

Aus dem Institut für Arbeitsmedizin  
der Westfälischen Wilhelms-Universität Münster  
-Direktorin: Univ.-Prof. Dr. Ute Witting-

---

---

**Evaluation eines projekt- und problemorientierten  
Lehransatzes im Fachgebiet Arbeitsmedizin**

INAUGURAL - DISSERTATION  
zur  
Erlangung des doctor medicinae

der Medizinischen Fakultät  
der Westfälischen Wilhelms-Universität Münster

vorgelegt von  
Sabine Woltering  
aus Rheine

**2007**

Gedruckt mit Genehmigung der  
Medizinischen Fakultät der  
Westfälischen Wilhelms-Universität Münster

Dekan: Univ.-Prof. Dr. med. Volker Arolt

1. Berichterstatter: Univ.- Prof. Dr. med. Ute Witting

2. Berichterstatter: Prof. Dr. rer. pol. R. Peter Nippert

Tag der mündlichen Prüfung: 12. September 2007

Der Mensch ist ein Wesen das lernt.  
Ein Leben lang.

Theo M. Liket

**Aus dem Universitätsklinikum Münster  
Institut für Arbeitsmedizin**

-Direktorin: Univ. -Prof. Dr. med. Ute Witting-  
Referentin: Univ. -Prof. Dr. med. Ute Witting  
Korreferent: Prof. Dr. rer. pol. R. Peter Nippert

**ZUSAMMENFASSUNG**

**Evaluation eines projekt- und problemorientierten Lehransatzes im Fach Arbeitsmedizin**

Sabine Woltering

**Ziel:** Seit dem Wintersemester 1999/2000 werden die Pflichtlehrveranstaltungen im Fachgebiet Arbeitsmedizin auf der methodischen Grundlage des projekt- und problemorientierten Lernens durchgeführt und kontinuierlich institutsintern evaluiert. Ziel dieser Studie ist es, die Evaluationsergebnisse auszuwerten, um den Erfolg des neuen Lehr- und Lernansatzes und insbesondere die Validität des Evaluationsverfahrens zu überprüfen.

**Methode:** Im Untersuchungszeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 haben insgesamt 1020 Studierende am „Praxisprojekt Arbeitsmedizin“ teilgenommen. 981 ausgefüllte Evaluationsbögen wurden in die Auswertung einbezogen. Der Fragebogen besteht aus insgesamt 25 Fragen (Items), strukturiert in Indikatoren für die didaktischen Elemente (Grundlagenseminare, Betriebsbegehungen, praktische Übungen, Gruppenarbeit) und für die Gesamtbewertung der Lehrveranstaltung. Die statistische Auswertung der Daten wurde mit SPSS Version 10.0 durchgeführt. Als multivariate Analysemethoden wurden die Regressionsanalyse und die Faktorenanalyse verwendet.

**Ergebnisse:** Durch die explorative und konfirmatorische Faktorenanalyse wurde belegt, dass das Evaluationsverfahren die methodischen Anforderungen an Reliabilität und Validität erfüllt. Im Erhebungszeitraum zeigte sich eine statistisch signifikante Verbesserung der Beurteilung über die Indikatoren der didaktischen Elemente und für die Gesamtbewertung der Lehrveranstaltung. Dies gilt insbesondere für die Items, die durch Faktorenanalyse dem Faktor „eigene Teilnahmemotivation und Lernerfolg“ und dem Faktor „Struktur, Didaktik und Lehrkompetenz“ zugeordnet wurden. Faktorenanalytisch konnte eine Korrelation zwischen dem Faktor „eigene Teilnahmemotivation und Lernerfolg“ und den Faktoren „Struktur, Didaktik und Lehrkompetenz“ und „Gruppenarbeit“ belegt werden. Zum Einfluss der Bewertung der Gruppenarbeit wird aufgezeigt, dass mehr als 2/3 der Studierenden, die die Gruppenarbeit positiv bewerteten, auch alle anderen evaluativen Kriterien signifikant positiver bewerteten als die Gruppe Derjenigen, die ihre Gruppenarbeit negativ beurteilten.

**Schlussfolgerungen:** Die kontinuierliche Evaluation und die konstruktive Kritik der Studierenden waren immer Anlass und Motivation für weitere strukturelle und prozedurale Verbesserungen, die wiederum zu einer verbesserten Bewertung der Lehrveranstaltung in den nachfolgenden Semestern führte. Insgesamt konnte gezeigt werden, dass kontinuierliche studentische Lehrevaluation auf der Grundlage eines validen Evaluationsverfahrens zur Qualität der Lehre in der Arbeitsmedizin beitragen kann.

**Tag der mündlichen Prüfung: 12. September 2007**

# Inhaltsverzeichnis

<b>1.</b>	<b>Einleitung</b>	<b>1</b>
<b>2.</b>	<b>Theoretischer Teil</b>	<b>6</b>
<b>2.1</b>	<b>Traditionelle und neue Lehr- und Lernformen</b>	<b>6</b>
2.1.1	Gruppenarbeit	7
2.1.2	Projekt-basiertes Lehren und Lernen	8
2.1.3	Problem-orientiertes Lehren und Lernen	12
<b>2.2</b>	<b>Grundlagen der Evaluation</b>	<b>16</b>
2.2.1	Definition von Evaluation	16
2.2.2	Historische Entwicklung	16
2.2.3	Standards, Ziele und Gegenstandsbereiche der Evaluation	18
2.2.4	Formen der Evaluation	19
2.2.5	Methoden der Evaluationsforschung	21
<b>2.3</b>	<b>Evaluation der Lehre</b>	<b>23</b>
2.3.1	Bedeutung der Evaluation der Lehre	23
2.3.2	Historische Entwicklung der Evaluation der Lehre	24
2.3.3	Formen und Ebenen der Evaluation der Lehre	25
2.3.4	Qualitätsentwicklung durch Evaluation der Lehre	26
2.3.5	Verfahren zur Evaluation von Lehrveranstaltungen	28
2.3.6	Validität studentischer Lehrevaluation	30
<b>3.</b>	<b>Empirischer Teil</b>	<b>33</b>
<b>3.1</b>	<b>Das „Praxisprojekt Arbeitsmedizin“: Lernziele , Struktur und Inhalt</b>	<b>33</b>
<b>3.2</b>	<b>Zielsetzung und Fragestellung</b>	<b>37</b>
<b>3.3</b>	<b>Methodik</b>	<b>38</b>
3.3.1	Kollektiv	38
3.3.2	Fragebogen	39
3.3.3	Statistik	41
<b>3.4</b>	<b>Ergebnisse</b>	<b>46</b>
<b>3.4.1</b>	<b>Häufigkeitsverteilungen</b>	<b>46</b>
3.4.1.1	Grundlagenseminar	47
3.4.1.2	Praktische Übungen	50
3.4.1.3	Betriebsbegehung	53
3.4.1.4	Gruppenarbeit	55
3.4.1.5	Gesamtbewertung	58

<b>3.4.2</b>	<b>Mittelwerte</b>	63
3.4.2.1	Grundlagenseminar	63
3.4.2.2	Praktische Übungen	63
3.4.2.3	Betriebsbegehung	64
3.4.2.4	Gruppenarbeit	65
3.4.2.5	Gesamtbewertung	65
<b>3.4.3</b>	<b>Der Einfluss der Bewertung der Gruppenarbeit auf die Gesamtbewertung der Lehrveranstaltung</b>	67
3.4.3.1	Grundlagenseminar	68
3.4.3.2	Praktische Übungen	70
3.4.3.3	Betriebsbegehung	73
3.4.3.4	Gruppenarbeit	75
3.4.3.5	Gesamtbewertung	76
3.4.3.6	Mittelwerte und t-Tests	80
<b>3.4.4</b>	<b>Vergleiche der Semesterhälften 1999 bis 2003</b>	83
<b>3.4.4.1</b>	<b>Häufigkeitsverteilungen</b>	83
3.4.4.1.1	Grundlagenseminar	83
3.4.4.1.2	Praktische Übungen	86
3.4.4.1.3	Betriebsbegehung	88
3.4.4.1.4	Gruppenarbeit	90
3.4.4.1.5	Gesamtbewertung	92
3.4.4.1.6	Mittelwerte und t-Tests	96
<b>3.4.5</b>	<b>Zeitliche Veränderungen von 1999 bis 2003</b>	97
3.4.5.1	Grundlagenseminar	98
3.4.5.2	Praktische Übungen	100
3.4.5.3	Betriebsbegehung	103
3.4.5.4	Gruppenarbeit	105
3.4.5.5	Gesamtbewertung	107
3.4.5.6	Mittelwerte und t-Tests	111
3.4.5.6.1	Grundlagenseminar	113
3.4.5.6.2	Praktische Übungen	114
3.4.5.6.3	Betriebsbegehung	114
3.4.5.6.4	Gruppenarbeit	114
3.4.5.6.5	Gesamtbewertung	114
<b>3.4.6</b>	<b>Regressionsanalysen</b>	115
3.4.6.1	Grundlagenseminar	116
3.4.6.2	Praktische Übungen	116
3.4.6.3	Betriebsbegehung	116
3.4.6.4	Gruppenarbeit	116
3.4.6.5	Gesamtbewertung	117

<b>3.4.7</b>	<b>Explorative Faktorenanalyse</b>	117
3.4.7.1	Grundlagenseminar	118
3.4.7.2	Praktische Übungen	119
3.4.7.3	Betriebsbegehung	120
3.4.7.4	Gruppenarbeit	120
3.4.7.5	Gesamtbewertung	121
3.4.7.6	Externe Validierung	121
<b>3.4.8</b>	<b>Konfirmatorische Faktorenanalyse</b>	126
<b>4.</b>	<b>Diskussion</b>	129
<b>4.1</b>	<b>Bewertung struktureller und didaktischer Elemente</b>	130
4.1.1	Kompetenz der Lehrenden	130
4.1.2	Betriebsbegehung	132
4.1.3	Praktische Übung	134
4.1.4	Gruppenarbeit	134
<b>4.2</b>	<b>Einfluss struktureller und didaktischer Elemente auf Teilnahmemotivation und persönlichen Lernerfolg</b>	137
<b>4.3</b>	<b>Bewertung des Evaluationsverfahrens</b>	139
<b>4.4</b>	<b>Kontinuierliche Evaluation als Beitrag zur Qualität der Lehre</b>	149
<b>4.5</b>	<b>Beantwortung der gestellten Fragen</b>	156
<b>5.</b>	<b>Zusammenfassung</b>	159
<b>6.</b>	<b>Literaturverzeichnis</b>	160
<b>7.</b>	<b>Danksagung</b>	169
<b>8.</b>	<b>Lebenslauf</b>	170
<b>9.</b>	<b>Anhang: Fragebogen</b>	

## 1. Einleitung

Bereits seit einigen Jahren haben immer mehr Hochschulen die Bedeutung der Qualitätssicherung erkannt und setzen Evaluationsverfahren ein, um die Qualität der Lehre zu verbessern. An den medizinischen Fakultäten soll die Evaluation der Lehre erstens einen Anreiz schaffen, die als defizitär geltende akademische Ausbildung zu verbessern; die Indikatoren des Lehrerfolgs reichen von subjektiven Einschätzungen der Studierenden und der Lehrenden über die Feststellung von Tatsachen-, Zusammenhangs- und Orientierungswissen bis zur erfolgreichen Lösung praxisnah gestellter Aufgaben. Zweitens erlangt die Lehre auch zunehmende Bedeutung als Output-Größe im Rahmen der leistungsorientierten Mittelvergabe an den und zwischen den Fakultäten (Fischer-Blum et al. 2001).

Das Hochschulrahmengesetz fordert eine Orientierung der staatlichen Hochschulfinanzierung unter anderem an den in der Lehre erbrachten Leistungen. Dazu werden weitergehende Informationen über die Lehrleistung, das heißt eine Vorstellung darüber, mit welchen Mitteln oder Maßnahmen welcher Lehrerfolg erreicht wurde, benötigt (Brähler et al. 1998)

Die „Qualität der Lehre“ ist ein aktuelles hochschulpolitisches Thema, deren Sicherung und Weiterentwicklung als zentrale Aufgaben der kommenden Jahre angesehen werden. Die Evaluation von Studiengängen, Fachbereichen und Hochschulen als Ganzes ist eine Aufgabe, der sich alle Beteiligten des Hochschulsystems widmen müssen (Rindermann 1996 a). Die Gründe dafür liegen im wachsenden öffentlichen Interesse, im zunehmenden Wettbewerb unter den Hochschulen und in den gesetzlichen Vorgaben der Hochschulen.

Vor einiger Zeit standen noch Strukturprobleme im Zentrum der hochschulpolitischen Diskussion, aktuell ist es die Lehre, die in den Blickpunkt der Hochschulreformdebatte rückt. Angesichts der Verschärfung des internationalen Wettbewerbs wird eine Steigerung der „Kundenorientierung“ proklamiert (Maudsley und Strivens 2000).

Um die Effektivität des Studiums (der Studierenden) zu erhöhen, wird vermehrt über andere Lehr- und Lernmethoden diskutiert (Albanese 2000, Wildt 2002).

Studierende werden als Experten zu den Lehrfähigkeiten ihrer Lehrenden befragt. Dabei kommt der Evaluation eine immer größere Bedeutung zu. Sie wird als ein Mittel zur Optimierung der Lehr- und Lernprozesse im gesamten Spektrum des Bildungswesens gesehen, ob auf der Ebene einzelner Unterrichtssituationen oder auf der Ebene ganzer Aus- und Weiterbildungsprogramme (Reissert 1994).

Die Evaluation der Lehre als Methode zur Qualitätssicherung ist inzwischen auch in der ärztlichen Ausbildung fest etabliert. In der neuen Ärztlichen Approbationsordnung (ÄAppO 2002), die am 1.10. 2003 in Kraft getreten ist, wird vorgeschrieben, dass Lehrveranstaltungen regelmäßig auf ihren Erfolg zu evaluieren und die Ergebnisse bekannt zu geben sind (§ 2, Abs. 9 ÄAppO 2002).

Das Fachgebiet Arbeitsmedizin ist seit 1970 in der Ärztlichen Approbationsordnung (ÄAppO 1970) als Pflichtlehr- und Prüfungsfach verankert. Bis zum Inkrafttreten der neuen Approbationsordnung am 01.10.2003 (ÄAppO 2002) war das Fach Arbeitsmedizin gemeinsam mit den Fächern Hygiene, Sozialmedizin und Rechtsmedizin im „Kursus des Ökologischen Stoffgebietes“ vertreten. In der neuen Approbationsordnung wurden die Fächer der klinischen Studienabschnitte neu geordnet. Danach ist für die ehemals im „Kursus des Ökologischen Stoffgebietes“ vertretenen Fächer folgende Zuordnung zur Erbringung benoteter Leistungsnachweise getroffen worden: „Arbeitsmedizin, Sozialmedizin“, „Hygiene, Mikrobiologie, Virologie“ und „Rechtsmedizin“. Neu ist, wie oben ausgeführt, auch die Vorschrift, Lehrveranstaltungen regelmäßig zu evaluieren (§ 2 Abs. 9 ÄAppO).

Die curriculare Lehre in der Arbeitsmedizin an der Medizinischen Fakultät der Universität Münster wird seit dem Sommersemester 1997 kontinuierlich durch eine studentische Lehrevaluation begleitet. Anlass für diese institutsintern initiierte Evaluation der Lehre war insbesondere eine Neustrukturierung des Pflichtlehrangebotes, das durch die Studierenden bewertet werden sollte. Bislang wurden die Lehrinhalte des Faches Arbeitsmedizin traditionell in Form von Vorlesungen und durch begleitende Seminare und Übungen in Gruppen angeboten. Exkursionen in Betriebe waren auf freiwilliger Basis möglich. Der Lehr- und Lernerfolg wurde durch eine Klausur (MC-Fragen) überprüft. Der Kurs musste wegen der hohen Studierendenzahl zweimal pro Semester durchgeführt werden.

Aufgrund einer Prüfungsordnung, die 1996 in Kraft trat (Westfälische Wilhelms-Universität Münster, Amtliche Bekanntmachungen 1996), wurde das Pflichtlehrangebot in der Arbeitsmedizin in Form von Seminaren und Übungen in Gruppen im Umfang von zwei Semesterwochenstunden festgelegt.

Um die bislang in Form von Vorlesungen angebotenen Lehrinhalte in die Pflichtveranstaltungen zu integrieren, wurden Grundlagenseminare eingeführt. Nach Vermittlung von Basiswissen über die gesetzlichen Grundlagen und ärztlichen Aufgaben des Faches unter Berücksichtigung von fachübergreifenden Aufgaben wurden darauf aufbauende Lehrinhalte in speziellen Seminaren mit Beiträgen der Studierenden exemplarisch und praxisnah bearbeitet und durch praktische Übungen ergänzt und vertieft. Exkursionen in Betriebe waren nach wie vor auf freiwilliger Basis möglich. Anstelle der bislang üblichen MC-Fragen-Klausur wurde die Überprüfung des Lehr- und Lernerfolges im Rahmen eines Abschlusskolloquiums mit mündlichen Prüfungen in Gruppen von je 4 Studierenden eingeführt.

Eine kontinuierliche institutsinterne Evaluation durch spezielle Projektgruppen der Studierenden mit Erarbeitung geeigneter Erhebungsinstrumente für die Evaluation des neu strukturierten arbeitsmedizinischen Lehrangebotes wurde, wie oben angeführt, seit dem Sommersemester 1997 fester Bestandteil des Kurses. Das Abschlusskolloquium mit mündlichen Prüfungen zu den Grundlagen und fachübergreifenden Inhalten des Faches erbrachte für die Lehrenden in unmittelbarer Rückkopplung nicht nur Erfahrungen über den Erfolg ihrer eigenen Bemühungen, sondern vor allem auch über das Lernverhalten der Studierenden.

Wenn sich herausstellte, dass bestimmte Lehrinhalte von den Studierenden nur unzureichend wahrgenommen wurden, war dies immer wieder Anlass, im darauf folgenden Semester die Aufmerksamkeit der Studierenden für bestimmte Themen durch in aller Regel praxisnahe Beispiele zu erhöhen. Dieser Prozess wurde intensiv begleitet und gefördert durch die Ergebnisse der kontinuierlichen studentischen Lehrevaluation. Dabei stellte sich u. a. heraus, dass die Exkursionen in Betriebe von Lehrenden und Lernenden unterschiedlich beurteilt wurden.

Das Angebot, auf freiwilliger Basis an Betriebsbegehungen teilzunehmen, wurde zwar begrüßt, aber auf Grund eines als zu hoch empfundenen Zeitaufwandes nur unzureichend angenommen. Jedoch bewerteten Studierende, die an den Betriebsbegehungen teilgenommen hatten, die Möglichkeit eines derartigen unmittelbaren Einblicks in die Praxis als sehr positiv.

Aufgrund der positiven Bewertung der Betriebsbegehungen durch Studierende, die freiwillig daran teilgenommen hatten, entwickelte sich die Idee, die Betriebsbegehung als zentrales Element in die Pflichtlehrveranstaltung zu integrieren. Dies ließ sich aber nur dann realisieren, wenn allen Studierenden in Kleingruppen eine Betriebsbegehung angeboten werden könnte. Die Umsetzung dieser Zielvorstellung wurde entscheidend dadurch gefördert, dass mittlerweile die Semesterstärke von ehemals mehr als 250 Studierenden auf weniger als 150 Studierende abgenommen hatte. Da der Kurs aufgrund struktureller und personeller Ressourcen nach wie vor zweimal im Semester mit jeweils acht Gruppen angeboten werden muss, können nunmehr Gruppengrößen von 8 – 10 Studierenden erreicht werden. Nachdem eine ausreichende Zahl geeigneter Betriebe für die Betriebsbegehungen gewonnen werden konnte, waren wesentliche Voraussetzungen für eine grundlegende Neustrukturierung des Lehrangebotes erfüllt. Unter der Leitidee „die Praxis des Faches ist der Betrieb“ wurde im Wintersemester 1999/2000 ein innovatives Lehrprojekt „Praxisprojekt Arbeitsmedizin“ auf der methodischen Grundlage des projekt- und problemorientierten Lehrens und Lernens entwickelt.

Jede Gruppe erhält ihren „eigenen“ Betrieb. Nach Vermittlung von Basiswissen durch Grundlagenseminare werden die Studierenden durch konkreten Praxisbezug im Rahmen der Betriebsbegehung dazu angeleitet, die Probleme in der Praxis eigenständig zu erkennen, die Arbeitsbedingungen zu beurteilen und die daraus resultierenden ärztlichen Aufgaben zu erarbeiten.

Durch diesen konkreten Praxisbezug in Betrieben unterschiedlicher Industrie- und Gewerbebranche werden wesentliche Inhalte der Arbeitsmedizin, die Aufgaben und Ziele der ärztlichen Tätigkeit sowie fachübergreifende Grundlagen für die ärztliche Tätigkeit in anderen Fachdisziplinen in Gruppenarbeit mit Dozentenunterstützung in 8 aufeinander folgenden Lernschritten erarbeitet.

Die Ergebnisse der Gruppenarbeit werden von den Mitgliedern jeder Gruppe in strukturierter Form in einer Projektpräsentation dargestellt, die dem Plenum vorgestellt wird. Durch die einzelnen Projektpräsentationen im Plenum werden alle Gruppenprojekte zum Gesamtprojekt „Praxisprojekt Arbeitsmedizin“ zusammengeführt. Der Lehr- und Lernerfolg wird wie bisher durch ein Abschlusskolloquium mit 4 Studierenden pro Gruppe überprüft. Anschließend beteiligen sich alle Studierenden an der institutsinternen Evaluation der Lehre.

Integraler Bestandteil des Lehrprojektes „Praxisprojekt Arbeitsmedizin“ ist der Arbeitsmedizinische Qualitätszirkel „Lehre“ mit allen an der Lehre Beteiligten, insbesondere auch mit den externen Partnern, d. h. mit den Betriebsärzten der beteiligten Betriebe.

Die studentische Evaluation des Praxisprojektes Arbeitsmedizin und deren Auswertung im Zeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 ist Thema der vorliegenden Arbeit. Vorangestellt wird ein theoretischer Teil über traditionelle und neue Lehr- und Lernformen sowie über Formen und Methoden der Evaluation. Im empirischen Teil werden zunächst Lehr- und Lernziele, Struktur und Inhalte des Praxisprojektes Arbeitsmedizin vorgestellt sowie Zielsetzung und Fragestellung dieser Arbeit formuliert. Daran anschließend werden Methoden und Ergebnisse dargestellt und in der Diskussion kritisch bewertet. Abschließend werden die wesentlichen Ergebnisse zusammengefasst und die zur Zielsetzung dieser Arbeit formulierten Fragen beantwortet.

## **2. Theoretischer Teil**

### **2.1 Traditionelle und neue Lehr- und Lernformen**

Zu den traditionellen Lehr- und Lernformen zählt insbesondere der Frontalunterricht, eine Unterrichtsform, bei der die Studierenden im Plenum unterrichtet werden. Die Aufgabe des Dozenten besteht darin, durch Erklärungen, Visualisierungen, Impulse und Fragen die Studierenden zum Denken, Darstellen, Begreifen und zum Wiederholen von Sachverhalten anzuleiten. Zu den traditionellen Formen des Lehrens zählt ebenfalls der Vortrag, bei dem der Dozent dem Plenum einen fachlichen Zusammenhang mündlich darstellt. Der Vortrag ist eine geplante und strukturierte Methode, die möglichst mit Visualisierungen verbunden sein sollte (Eitel 1994). Bei praktischen Übungen oder Experimenten wird erprobt, ob die Annahmen über den Ablauf eines Ereignisses und die Zusammenhänge der verschiedenen Wirkungsfaktoren bestätigt oder verworfen werden müssen. Praktische Übungen sind wissenschaftliche Versuche, weil sich in einer Versuchsanordnung und im Ablauf die natürlichen Gegebenheiten objektiv widerspiegeln müssen. In Experimenten sowie praktischen Übungen werden komplexe Phänomene in einer künstlichen Versuchsanordnung modellhaft vereinfacht so dargestellt, dass der Ablauf präzise beobachtbar, aber nicht verfälscht wird (Peterssen 1996, Tippelt 1999, Mattes 2004).

Traditionell gestaltete Lernumgebungen mit strukturorientiertem Vorgehen, strengen Fächergrenzen und Lernen als vorwiegend rezeptivem Prozess werden offensichtlich nicht mehr als geeignet angesehen, um Studierende auf die spätere berufliche Tätigkeit vorzubereiten (Eitel 1995, Habeck 1995, Siebert 1997). Gefragt ist neben Kreativität, Flexibilität und Problemlösungsfähigkeit vor allem die Entwicklung von persönlichen und sozialen Kompetenzen. Diese Kompetenzen entwickeln sich im Besonderen durch problembezogenes, reflektierendes Lernen (Dohmen 1996, Becker 1997).

### 2.1.1 Gruppenarbeit

Aus didaktischer Sicht kommt im handlungsorientierten Unterricht der Gruppenarbeit eine Schlüsselbedeutung zu, weil hier „Learning by doing“ stattfindet. Neu erworbene theoretische Kenntnisse werden umgehend praktisch angewendet (Pratt und Magill 1983, Gudjons 1998). Gute Gruppenarbeit deckt vier Ebenen des erweiterten Lernbegriffs ab. Die Studierenden erarbeiten sich Wissen, lernen Wichtiges von Unwichtigem zu unterscheiden, wenden das Gelernte kommunikativ an und trainieren soziale Fähigkeiten wie Teamgeist, Rücksichtnahme und Toleranz.

Zudem sticht eine Gruppe bei der Bearbeitung eines konkreten Problems jedes noch so begabte Individuum aus. Zu diesem Ergebnis kommen amerikanische Forscher in einer aktuellen Studie. Sie stellten fest, dass Gruppen von drei, vier oder fünf Personen besser abschnitten als die jeweils besten Einzelpersonen und schrieben diese Leistung der Fähigkeit der Menschen zu, zusammenzuarbeiten (Laughlin et al. 2006).

In den Phasen einer Gruppenarbeit zum Beispiel erarbeiten sich Studierende in Gruppengrößen zwischen drei und sechs Mitgliedern eine Aufgabenstellung eigenverantwortlich und kooperativ. Meistens werden die Arbeitsergebnisse so aufbereitet, dass sie nach Beendigung der Arbeitsphase vor einem Plenum präsentiert werden können. Gruppenarbeit wird in der Regel vom Dozenten geplant, belässt aber den Studierenden eigene Gestaltungsmöglichkeiten (Peterssen 1997). In ihrer Binnenstruktur lassen sich die Planungsphase, Erarbeitungsphase und Vorbereitungsphase für die Präsentation unterscheiden.

Während der Arbeitsphase tritt der Dozent zurück und beschränkt sich auf das Beobachten, Beraten und das Bewerten der Qualität des Arbeitsprozesses. Gruppenarbeit findet häufig in der Phase der Anwendung von neu Gelerntem statt. In der Regel erfolgt eine erste exemplarische Erarbeitung des neuen Lerngegenstands im Frontalunterricht. Auch in Übungs- und Wiederholungsphasen ist Gruppenarbeit sinnvoll.

Die Ziele können so weit gesteckt werden, dass Studierende sich völlig selbstständig neue Themengebiete erarbeiten und ihre Ergebnisse in Form einer Präsentation darstellen. Hier werden dann Gruppenarbeit und Präsentationen zu den Arbeitsformen, die am häufigsten zum Einsatz kommen (Mattes 2004).

Gruppenarbeit muss gelernt werden, damit sie erfolgreich verläuft. Die Voraussetzung für den Lernerfolg ist auch die Beherrschung elementarer Arbeitstechniken. Gruppenarbeit verlangt vom Dozenten Planungs- und Organisationskompetenz, von den Studierenden Einsicht, Übung und Routine. In der Planung muss die Gruppenarbeit im Gesamtablauf des Unterrichtes reflektiert werden. Zum Beispiel sollte überlegt werden, wie viel Anleitung im Frontalunterricht notwendig ist, damit die Gruppen produktiv arbeiten können. Erfolgreiche Gruppenarbeit hängt auch oft von der Formulierung der Arbeitsanweisungen ab. Sie sollten sich stets auch auf die Art der Präsentation beziehen (Mattes 2004).

### **2.1.2 Projekt-basiertes Lehren und Lernen**

Mit dem Begriff Projekt bezeichnet man ein Vorhaben, einen Plan. Im Projektunterricht stellen sich Lehrende und Lernende eine gemeinsame Aufgabe, die sie in einer vorgegebenen Zeit durchführen wollen. Das Verfahren bezieht sich auf den gesamten Ablauf des Unterrichtsverlaufes von der Phase der Planung bis zur Präsentation des Projektergebnisses. Projektunterricht ist keine einzelne Methode, sondern setzt sich aus einer Vielzahl von Methoden zusammen, wobei auch Phasen traditionellen Frontalunterrichtes ihren Platz haben können. Wichtigstes Merkmal ist die Verbindung von theoretischer Erkenntnis und praktischem Tun.

Am Ende steht in der Regel ein vorzeigbares Produkt, das einer Öffentlichkeit präsentiert werden kann. Projektunterricht ist eine offene Unterrichtsform. Es gibt innerhalb des Ablaufs Entscheidungsoffenheit und das Projektergebnis entwickelt sich sukzessiv aus dem unterrichtlichen Prozess heraus. Der Arbeits- und Lernprozess ist dabei ebenso wichtig wie das Handlungsergebnis oder Produkt, das am Ende des Projektes stehen soll (Meyer 1994).

Die zu Beginn dieses Jahrhunderts von John Dewey und William Heard Kilpatrick entwickelte Projektmethode wurde im Projektorientierten Unterricht weiterentwickelt (Geißler 1970, Bohnsack 1976, Frey 1982). Die Übergänge zwischen Projekt, Projektwoche, Vorhaben, Projektorientiertem Unterricht, Projektunterricht und Projektarbeit sind fließend (Stach 1978, Heller und Semmerling 1983).

Die ersten Formen des Projektunterrichts wurden schon im 16. Jahrhundert in Italien und Frankreich entwickelt und zunächst in der Architekturausbildung angewandt. Die Studierenden mussten gemeinsam Projekte erarbeiten. Besonderer Wert wurde dabei auf die Fähigkeit zur Kooperation der Studierenden gelegt: sie sollten eine gestellte Aufgabe innerhalb der Gemeinschaft lösen (Suin de Boutemard 1975, Frey 1996).

Im Zuge des wissenschaftlichen Austausches gelangte dieses methodische Konzept zu Beginn des 19. Jahrhunderts nach Deutschland, kurze Zeit später auch nach Nordamerika. Zwei Varianten des Projektunterrichts kristallisierten sich während der wissenschaftlichen Beschäftigung mit diesem Thema heraus, ein sozialkonservativ-technologischer Ansatz und eine sozialreformerisch-politische Variante (Kaiser 1977, Struck 1980).

Im reformerischen Ansatz etabliert sich die demokratische Komponente des Projektunterrichts endgültig, denn bei diesem Ansatz wird der Schwerpunkt darauf gelegt, bei den Lernenden durch Lernen am Objekt selbstständiges Denken zu fördern. Am „Teacher`s College“, einer pädagogischen Hochschule, wurde die Projektmethode in ihrer sozialreformerischen Variante gepflegt und durch den Philosophen John Dewey „zum Zentrum der amerikanischen Reformpädagogik“ (progressive education) geführt (Eitel 1996). Sein Schüler Kilpatrick entwickelte die Projektmethode weiter, deren Hauptmerkmal „Learning by doing“ war. Im Gegensatz zum herkömmlichen Frontal- oder Einzelunterricht ist der Projektunterricht gekennzeichnet durch Gruppenarbeit und Rollenwechsel zwischen Lehrenden und Lernenden (Geisler und Scholz 1976).

Das Aufgreifen von und Anknüpfen an Interessen des Lernenden, Handlungsorientierung und eine besondere Produktbezogenheit sind wesentliche Merkmale des Projektlernens. Dabei steht das Entwickeln einer Strategie zur Problemlösung im Mittelpunkt der Gruppenarbeit.

Zur Förderung eines Verständnisses von Verantwortung werden Zeitpläne und Organisationsabläufe innerhalb der Gruppe festgelegt. Dabei hat die Gruppe den Auftrag, alle Lernmittel und Fachinformationen selbst zu beschaffen. Auch fächerübergreifende Ansätze, außerinstitutionelle Expertise und interdisziplinäres Denken sind zentrale Elemente des Projektlernens. Zielfokussierung und Produktorientierung stehen in dieser Unterrichtsform an erster Stelle (Apel und Knoll 1997).

Die Qualität der Projekte kann am Arbeitsergebnis bemessen werden. Aus diesem Grunde sollte das Produkt im Vorfeld genau beschrieben werden, um im Verlauf der Projektarbeit immer wieder als Ziel fixiert werden zu können. Beweglichkeit im Denken und Verständnis für komplexe Zusammenhänge sollen so gefördert werden.

Projektarbeit hat auch eine starke soziale Komponente. Die Arbeit muss von der gesamten Gruppe bewältigt werden, die Gruppenteilnehmer lernen voneinander und miteinander (Bastian 1993). Als Abschluss der Projektarbeit dient die öffentliche Präsentation der Ergebnisse, wobei damit Austausch mit dem Betrachter und Kritik und Feedback von ihm erhofft wird. Mit der Veröffentlichung der Arbeitsergebnisse ist die Erwartung verbunden, dass eine Auseinandersetzung mit dem Thema über die Projektdauer hinaus gefördert wird.

Diese Lehrform erfordert ein neues Selbstverständnis der Lehrenden: Neben dem Erwerb pädagogischer Qualifikation muss der Dozent für die Projektarbeit seine Gruppenleiterkompetenz stärken, Verantwortung abgeben, über sich selbst reflektieren und sich selbst motivieren können. Aber auch der Lernende kann nicht mehr nur passiv rezipieren, sondern muss sich aktiv in den Gruppen- und Arbeitsprozess einbringen: Er muss sich selbst immer wieder motivieren, seine Zielvorstellungen mit denen der Gruppe abstimmen, muss sich auf eine Lernform einlassen, die ihm ein Höchstmaß an Eigenverantwortung abverlangt. Er muss Kritikfähigkeit, Toleranz und Rücksichtnahme entwickeln.

Nach Karl Frey (1996) beginnt das ideale Projekt mit der Projektinitiative: Demnach ist nicht festgelegt, ob die Projektgruppe selbst oder die Institution diese Arbeit anstößt. Dabei liefert die Initiative nur die Grundidee. Nach Frey kann jede Idee Ausgangspunkt für ein Projekt sein, „alles Denkbare ist auch machbar“.

Die Teilnehmer eines Projektes sind alle gleichberechtigt. Innerhalb eines vorher vereinbarten Rahmens geschieht die Auseinandersetzung mit der Initiative. Dabei können alle Teilnehmer eigene Wünsche und Vorstellungen einbringen. Nachdem eine Übereinkunft erreicht ist, wird eine Projektskizze angefertigt, in welcher Ziel und Umfang des Projektes beschrieben werden und ein zeitlicher Rahmen festgelegt wird. Erst dann können sich die Lernenden gemeinsam um die Entwicklung des Betätigungsbereiches kümmern (Gudjons 2001).

Hierzu entwickeln die Projektteilnehmer Ablaufpläne und üben fehlende Fertigkeiten ein. Bereits zu diesem Zeitpunkt wird die Bereitschaft zur Kooperation zur wichtigsten Eigenschaft des Projektteams. Nach diesen vorbereitenden Übungen kann zur Projektdurchführung übergegangen werden: Dies stellt den Hauptteil und somit auch wichtigsten Teil des Projektes dar und erfordert detaillierte Planung. Dabei sind alle Arbeitsformen möglich, insbesondere Arbeit in Kleingruppen.

Das Ende eines Projektes sollte im Vorfeld definiert werden: Das kann mit Fertigstellung des Produktes erreicht sein oder nach Abschluss eines bestimmten Zeitraumes, wobei die Teilnehmer mit einer Rückschau zum Anfang zurückkehren, indem sie Schlusstand mit Projektinitiative vergleichen. Vor dem Start eines Projektes sollten Fixpunkte definiert werden, die an bestimmten Stellen eine Unterbrechung der Arbeit zur gegenseitigen Information vorsehen.

Auf diese Art können sich die Teilnehmer den derzeitigen Stand im Hinblick auf das Gesamtkonzept vergegenwärtigen. Auch für die Metainteraktion muss im Vorfeld Raum geschaffen werden. Hier ist der Ort, an dem Konflikte und Beziehungsprobleme ausdiskutiert und Konfliktbewusstsein gefördert wird. Die Teilnehmer sollten gemeinsam versuchen Konflikte auszutragen, zu verkleinern oder zu lösen.

In der Definition der Projektmethode ist die Technik des Wissenserwerbs und der Wissensverankerung nicht festgelegt. Von daher sollte dieser Bereich mit einer anderen didaktischen Methode abgestützt werden. Da das Problemorientierte Lernen (POL) in den letzten Jahren als die Methode der Wahl in der Mediziner Ausbildung angesehen wird (Barrows und Tamblin 1980, Schmidt 1983, Barrows 1985, Pfaff 1997), wurde sie dahingehend analysiert, ob sie als zusätzliche methodische Komponente in die Projektarbeit integriert werden konnte.

Anstoß für die Überlegungen hinsichtlich einer neuen Studienmethode gab das Problem der raschen Veralterung von Wissen und die Schwierigkeit, in den bestehenden Kursstrukturen auf neue Entwicklungen reagieren und Innovationen einbringen zu können (Dörner 1987, Schmidt 1993)

### 2.1.3 Problem-orientiertes Lehren und Lernen

Problemorientiertes Lernen im Sinne des Problem-based-Learning-Konzepts (PBL) ist ein Ansatz, der seit dem Beginn des vorigen Jahrhunderts in Harvard an der juristischen Fakultät begonnen, und international in der Hochschulausbildung etabliert ist. Das Problem-based-Learning-Konzept in der Medizinerbildung (Schmidt et al. 1993, Vernon 1993, Armstrong 1997) wurde entwickelt, um isoliertes, disziplinäres Wissen in eine interdisziplinäre Ausbildung zu integrieren, die Lernenden in das selbstständige Problemlösen einzuführen und gleichzeitig die Anwendung von Wissen und Können bereits im Lernprozess zu fördern (Gudjons 1998).

Mit der Einrichtung der PBL-Kurse in der Medizin verfolgte die Mc-Master University das Ziel, die Zahl der Vorlesungen zu reduzieren. Gleichzeitig sollten die Grundlagenwissenschaften im Zusammenhang mit klinischen Problemen dargestellt werden, um die Studierenden so auf zu erwartende Situationen in der Praxis vorzubereiten (Kahlke et al. 2000, Norman und Schmidt 2000).

Es sollten kritisch und analytisch denkende Akademiker ausgebildet werden, mit deren Hilfe sich die Behandlung des Menschen effektiv und effizient gestalten ließe (Clade 1998, Finucare und Johnson 1998, Wilkie 2000). Auch die Novellierung der Ärztlichen Approbationsordnung von 2002 greift diesen methodischen Ansatz auf und fordert eine stärkere Ausrichtung des Lehrens und Lernens anhand konkreter Fälle von Patienten.

Problemorientiertes Lernen wurde 1957 auch von Heinrich Roth auf der Grundlage empirischer lerntheoretischer Forschung als Stufenkonzept entwickelt (Gräsel 1997). Nach dem Phasenschema von Roth, welches den Unterricht in die drei Teile „Einleitung, Hauptteil und Schluss“ aufteilt, vollzieht sich Problemorientiertes Lernen in sechs Stufen:

**Einleitung:**

1. Stufe der Motivation: Ein Lernwunsch erwacht.
2. Stufe der Schwierigkeiten: Widerstände im Lerngegenstand werden erkannt.

**Hauptteil:**

3. Stufe der Lösung: Entdeckung eines neuen Lösungsweges.
4. Stufe des Tuns und Ausführens: Der neue Lösungsweg wird ausprobiert, die neue Leistungsform aktiv vollzogen.

**Schluss:**

5. Stufe des Behaltens und Einübens: Die neue Leistungsform wird bewusst eingeübt.
6. Stufe der Bewährung im Leben und der Integration in die Persönlichkeitsstruktur.

Diese Phasen werden in der Mediziner Ausbildung genutzt, um den Wissenserwerb zu strukturieren. Zentrale Bedeutung erhalten in diesem Zusammenhang die Begriffe Problemdefinition, Anschlusswissen, Lernzieldefinition, Selbststudium und Integration des Wissens (Albanese und Mitchell 1993, Gräsel 1997).

Auf der Basis der Arbeiten von Barrows und Tamblyn (1980) sowie von Barrows (1985), entwickelten die Reformfakultäten weltweit ähnliche Schemata. In Maastricht als 7- Jump bekannt (Schmidt 1983), in Berlin als „POL-Bogen“ (Vernon 1995, Pfaff 1997), der um den Feedbackprozess erweitert wurde, ermöglicht diese didaktische Methode dem Lehrenden je nach Erfordernis ein aktives, direktives Eingreifen oder ein eher passives, beobachtendes Moderationsverhalten.

1971 wurde das problemorientierte Lernen an der Medizinischen Fakultät der niederländischen Universität Limburg in Maastricht eingeführt. Hier wurde auch die Problemlösungsstruktur „Siebensprung“ entwickelt, die inzwischen europaweit anerkannt ist (Hmelo 1998, Wilkie 2000).

Seit Anfang der 90er Jahre gibt es das problemorientierte Lernen mit zunehmender Tendenz auch an deutschen Hochschulen. Das Medizinstudium an der privaten Universität Witten-Herdecke ist seit 1992 problemorientiert organisiert. Zum Wintersemester 1999 wurde an der Berliner Humboldt-Universität ein POL-Reformstudien-gang Medizin eingeführt (Scheffner 2001).

Die Universität Hamburg hat seit dem Wintersemester 2001 bis zum Wintersemester 2003 mit einem Modellstudiengang die Möglichkeit zum problemorientierten Studium der Medizin für die ersten drei Studienjahre angeboten.

An der Medizinischen Fakultät der Westfälischen-Wilhelms Universität Münster findet Problemorientiertes Lernen seit dem Wintersemester 2000/2001 statt, um nur einige exemplarische Angebote zu nennen. Am Institut für Arbeitsmedizin wurde durch die Implementation des neuen Lehr- und Lernansatzes bereits im Wintersemester 1999/2000 problemorientiertes Lernen möglich.

Neben dem Erwerb von Fertigkeiten zum lebenslangen Lernen werden mit dem Einsatz des Problemorientierten Lernens verschiedene Ziele verfolgt. Die Motivation der Lernenden und Lehrenden, Wissenserwerb im Zusammenhang mit der späteren Anwendungssituation und die Erlangung einer wissenschaftlichen Denkweise der Lernenden soll gefördert werden (Robert Bosch Stiftung 1989).

Die Vermittlung des Wissens im Kontext der späteren Anwendungssituation geschieht unter der Vorstellung, dass das Wissen in der realen Situation später schneller aktiviert werden kann (Colliver 2000).

Moust und Mitarbeiter (1999) betonen neben dem Wissenserwerb den interaktiven Aspekt des Lernens beim POL. In Unterrichtsgruppen, die hinsichtlich Alter, Lebenserfahrung, Vorbildung, Geschlecht und gesellschaftlichen Einstellungen unterschiedlich sind, können die Gruppenmitglieder besonders gut voneinander lernen (Maudsley 1999).

Auch Engel (1999) erachtet die mit dem PoL angestrebten Kompetenzen wie logisches und analytisches Denken, begründetes Handeln und Entscheiden und nicht zuletzt die Fähigkeit zur Selbstevaluation als notwendig, um sich dem schnell verändernden Berufsleben anzupassen und an Veränderungen mitwirken zu können.

Neben dem Erwerb klinischer, psychosozialer, naturwissenschaftlicher und pflegerischer Kenntnisse sehen Harden und Mitarbeiter (2000) die Förderung und Entwicklung von persönlichen und sozialen Kompetenzen wie z. B. Selbstständigkeit, Teamarbeit, die Fähigkeit zur Entwicklung von Problemlösungsstrategien und Kommunikationstechniken als Ziele des problemorientierten Lernens.

Die von den verschiedenen Autoren aufgeführten Ziele des Problemorientierten Lernens sind zusammenfassend auf den Erwerb von Schlüsselqualifikationen sowie der Kompetenztrilogie aus Sach-, Sozial- und Selbstkompetenz zusammenzuführen.

Die Studierenden sollen sich nicht nur eine breite fachliche Wissensbasis erarbeiten, sondern gleichermaßen die Fähigkeit und Motivation zum lebenslangen Lernen erwerben (Kissel 1984, Maudsley 1999). Darüber hinaus spielt die Entwicklung einer reflektierten Persönlichkeit eine wichtige Rolle, die es schließlich ermöglicht, das Gelernte situationsabhängig anzuwenden, es nötigenfalls zu verändern oder neue Handlungsalternativen zu generieren (Habeck 1995, Mandl et al. 2001).

Das Konzept des problemorientierten Lernens soll demzufolge die Studierenden auf die bestehenden und zukünftigen Anforderungen ihres Berufes vorbereiten und sie befähigen, sich wandelnden Bedingungen anzupassen und Veränderungen verantwortlich mitzugestalten.

## **2.2 Grundlagen der Evaluation**

### **2.2.1 Definition von Evaluation**

„Jeder Mensch betreibt in seiner Erfahrungswelt ständig Evaluation, indem er auf der Grundlage kognitiver und affektiver Kriterien Gutes von Schlechtem und Sinnvolles von Sinnlosem trennt“ (Grüner 1993).

Es besteht eine große Spannweite der Bedeutung des Wortes oder der Methode Evaluation. Terminologisch geht der Begriff Evaluation auf das lateinische Verbum „valere“ zurück, was „wert sein“ bedeutet bzw. in Verbindung mit „e- valere“, „bewerten, einschätzen“ beinhaltet (Petschnig 1971). Laut Duden bedeutet Evaluation „Bewertung, Beurteilung“ (Duden 2005). Die Doppelbedeutung des englischen Wortes Evaluation, einerseits Auswertung, andererseits Bewertung/Beurteilung, lässt laut Allgäuer (1998) den Schluss zu, dass „die Aufgabe einer Evaluation nicht in einer bloßen Festschreibung beobachteter oder festgelegter Daten liegt, sondern dass der oder die Evaluierende auch eine Bewertung vornehmen muss und dadurch aufgefordert ist, Stellung zum Beobachteten bzw. Durchgeführten zu beziehen“.

### **2.2.2 Historische Entwicklung von Evaluation**

Wottawa und Thierau (1998) geben zur historischen Entwicklung von Evaluation folgende Übersicht:

Die ersten Evaluationsversuche technischer Art fanden bereits in der Urgesellschaft statt, indem empirisch gewonnene Kenntnisse über Materialeigenschaften bei der Herstellung von Arbeitsgeräten und Waffen verwendet wurden. Auch die Arbeitsteilung (Ackerbau und Viehzucht) wurde aufgrund von Erfahrungen nach Alter und Geschlecht der Arbeitenden vollzogen.

In der griechisch-römischen Antike forderte Aristoteles die empirische Nutzbestimmung zur Bewertung gesellschaftlich relevanter Maßnahmen, insbesondere von Staatsformen mit dem Zweck der optimalen Gestaltung der Gesellschaft.

Zunächst bleibt es bei dieser theoretischen Forderung ohne praktische Konsequenzen. Im Mittelalter kommt es zu einer weiterführenden theoretischen Systematisierung des aristotelischen Verständnisses von Nutzen durch Thomas von Aquin, und doch zeigt sich eine Innovationsschwäche und Unterbrechung der in der Antike begonnenen empirischen Wissenschaftsansätze.

Die Renaissance ist geprägt von einer Vielzahl neuer wissenschaftlicher Erkenntnisse, wie zum Beispiel dem heliozentrischen Weltbild von Kopernikus und Galilei. Der starke Aufschwung der empirischen Wissenschaft geht mit Innovations- und Evaluationsversuchen in den verschiedensten Gebieten einher.

Im 17. Jahrhundert bringt der enorme Aufschwung der Naturwissenschaften eine gezielte nutzorientierte Veränderung der technischen und ökonomischen Entwicklung mit sich. In engem Zusammenhang mit der europäischen Aufklärung kommt es zur Festigung und Institutionalisierung der empirischen Forschung. In der Philosophie dieser Zeit entwickelt sich der Utilitarismus, dessen ethisches Prinzip des Nutzens eine wesentliche geistige Wurzel moderner Evaluation darstellt (Wottawa und Thierau 1998).

Im Zeitalter der industriellen Revolution werden die empirischen Wissenschaften immer zielgerichteter nutzenorientiert zur Lösung gesellschaftlicher Probleme eingesetzt.

Das 20. Jahrhundert ist gekennzeichnet durch eine explosionsartige Vermehrung des Wissens in den empirischen Wissenschaften. Mit Fragen der Evaluation beschäftigen sich nun die verschiedensten Wissenschaftsdisziplinen, wie zum Beispiel die Wirtschaftswissenschaften, die Pädagogik und auch die Psychologie (Hager et al. 2000).

In den letzten zwanzig Jahren ist das Feld der Evaluation der stärkste Wachstumsfaktor innerhalb der Sozialwissenschaften geworden. In den USA begannen die Evaluationsbemühungen zu Beginn der 60er Jahre in Verbindung mit staatlichen Reformprogrammen. Mittlerweile ist die Evaluation fester Bestandteil der Implementierung von innovativen Programmen (Helmke und Krapp 1999).

Im deutschsprachigen Raum setzten die Evaluationsverfahren vorwiegend mit dem Ziel der effizienteren staatlichen Schul- und Bildungsplanung zum Ende der 60er Jahre ein. Die Entwicklungen begannen später als in den USA und verliefen in der inhaltlichen

Ausrichtung unstätiger. Die Evaluation ist nach Scholz (1995) ein entscheidendes Instrument der weiteren Qualitätsentwicklung in den verschiedensten Wissenschaftsdisziplinen.

### **2.2.3 Standards, Ziele und Gegenstandsbereiche der Evaluation**

Auf dem Gebiet der Evaluation gibt es eine Reihe zum Teil stark divergierender Ansätze. Der Grund hierfür liegt darin begründet, dass die vielen verschiedenen Formen und Modelle der Evaluation für spezifische Anwendungsbereiche entwickelt wurden (Richter 1994, Wesseler 1999). Es gibt jedoch allgemeine Standards für Evaluation, wie sie auch vom Joint Committee on Standards for Educational Evaluation (2006) veröffentlicht wurden. Die „Program Evaluation Standards“ sind von sechzehn amerikanischen Organisationen des Erziehungs- und Bildungsbereiches entwickelt worden. Sorgfältige Evaluationen von Programmen, Projekten und Materialien sollten vier grundlegende Eigenschaften aufweisen:

- |                     |                |
|---------------------|----------------|
| 1. Nützlichkeit     | 3. Korrektheit |
| 2. Durchführbarkeit | 4. Genauigkeit |

Nützlichkeitsstandards sollten sicherstellen, dass eine Evaluation sich an den Informationsbedürfnissen der vorgesehenen Evaluationsnutzer ausrichtet.

Durchführbarkeitsstandards sollten gewährleisten, dass eine Evaluation realistisch, gut durchdacht, diplomatisch und kostenbewusst sein wird.

Korrektheitsstandards sollten einbeziehen, dass eine Evaluation rechtlich und ethisch korrekt durchgeführt wird, und dem Wohlergehen der in die Evaluation Einbezogenen und auch der durch die Ergebnisse betroffenen Personen gebührende Aufmerksamkeit gewidmet wird.

Genauigkeitsstandards sollten sicherstellen, dass eine Evaluation über die Güte und die Verwendbarkeit des evaluierten Programms fachlich angemessene Informationen hervorbringen und vermitteln wird.

Nach Allgäuer (1998) gibt es *die* Evaluation nicht, denn jede Evaluation muss aufgrund der Interessenslage und der vorhandenen Kenntnisse und Erfahrungen neu konzipiert und initiiert werden. Dies bedeutet, dass jede Evaluation eine spezielle Aufgabe zu erfüllen hat, die von den Initiatoren im Evaluationskonzept festgelegt werden muss. Entscheidend für die Fragestellung eines Evaluationsvorganges sind das Interesse der Evaluatoren und die Zielsetzung, die damit verknüpft ist (Richter 1994).

Evaluation kann dazu dienen, die Wahrscheinlichkeit für die Auswahl der besonders guten Alternative zu erhöhen und für die Wahl der besonders schlechten zu verringern. Ein wesentlicher Schritt bei der Planung einer Evaluation ist es, die Gegenstände, d. h. die eigentlichen Objektbereiche der Evaluation, festzulegen. Erst muss geklärt werden, wer oder was zum Gegenstand der Evaluationstätigkeit wird, bevor über die Form der Evaluation, die Zielgruppen und einzusetzenden Methoden entschieden werden kann (Wottawa und Thierau 1998).

#### **2.2.4 Formen der Evaluation**

Eine wichtige Unterscheidung besteht zwischen interner und externer Evaluation. Beide Formen der Evaluation dienen der Qualitätssicherung und der Verbesserung von Lehre und Studium in einem Fach oder einem Fachbereich.

Interne Evaluationen werden durch das Fach selbst initiiert, dabei steht nicht allein die Person des Dozenten im Vordergrund, sondern der interaktive Ablauf einer Lehrveranstaltung (Marsch 1993, Beywl 1996). In ihrem Verlauf werden von Studierenden und Dozenten systematisch Daten und Informationen über Veranstaltungen und die Qualität von Veranstaltungen erhoben. Dabei ist es entscheidend, Profile der Stärken und Schwächen herauszuarbeiten und unter Berücksichtigung der Evaluationsziele zu bewerten, um dann durch gemeinsame Interpretationen Verbesserungen zu erzielen (Abs et al. 2000).

Externe Evaluationen, auch Peer-Evaluationen genannt, werden von externen Gutachtern erarbeitet. Diese führen ein- oder mehrfach vor Ort Interviews mit verschiedenen Funktionsträgern, Dozenten sowie Studierenden durch, um am Ende ein Gutachten über die Qualität eines Studienganges oder Fachgebietes zu erarbeiten.

Im Vordergrund von externen Evaluationen durch Peers stehen selten einzelne Veranstaltungen, sondern die Studienorganisation, die Prüfungsordnungen und die Relation von hauptamtlichen zu nebenamtlichen Beschäftigten (Abs et al. 2000). Externe Evaluation bezieht sich auf Studiengänge, aber auch auf größere Einheiten wie Hochschulen, Weiterbildungseinrichtungen und Schulen.

Sowohl interne als auch externe Evaluationsprozesse sind aber nur ergiebig, wenn sich ein Maßnahmenprogramm und Empfehlungen zur weiteren Qualitätssicherung anschließen, die dann auch in absehbarer Zeit realisiert werden (Angelo und Cross 1993).

Für die unmittelbare entwicklungsbegleitende Optimierung einer Bildungsmaßnahme ist die formative Evaluation zu nennen. Dieses Konzept bezüglich der Lehre erfordert, dass Erhebungen zur Qualitätsentwicklung nicht am Ende, sondern zu einem möglichst frühen Zeitpunkt der Veranstaltung stattfinden, um so durch sofortige Rückmeldung Maßnahmen zur Verbesserung einleiten zu können (Scholz 1995, Urban 2004). Formative Evaluationen beziehen sich auf die Sicherung der Qualität von Prozessen wie Interaktionsprozesse in Präsenzphasen und Ablaufprozesse von Bildungsveranstaltungen.

Die summative Evaluation ist der Sicherung der Qualität von Veranstaltungen und Prozessen zuzuordnen. Dabei geht es um eine abschließende Bewertung von Maßnahmen und eine genaue Beschreibung von beobachteten Effekten von Interventionen und Bildungsmaßnahmen. Im Vordergrund der Bewertung stehen Lehrmittel, Medien, Lernerfolg und auch die rückblickende Einschätzung einer Bildungsmaßnahme (Abs et al. 2000).

### **2.2.5 Methoden der Evaluationsforschung**

Die Evaluationsforschung hat nicht, wie andere Forschungsgebiete, die Überprüfung von Theorien zum Ziel, sondern die Bewertung von Handlungsalternativen. Als Grundlage aller wissenschaftlich gestützten Evaluationsbemühungen ist die zielgerichtete, systematische Sammlung, Aufbereitung und Bewertung von Daten zu sehen. Bei Evaluationsstudien steht die Verbesserung von Interventionsmaßnahmen in inhaltlicher und ökonomischer Sicht im Vordergrund (Bortz und Döring 2002).

Als anwendungsorientierte Forschung will die Evaluationsforschung zuverlässige Daten und Informationen liefern, um die Beurteilung von Programmen und Projekten im Bildungs-, Gesundheits- und Sozialwesen zu ermöglichen (Rossi 1988, Vedung 1999).

In der Evaluationsforschung finden sich daher folgende methodologische Forschungsansätze:

In der empirisch-analytischen Tradition steht die quantitative Erfassung und Beschreibung von Merkmalsausprägungen im Vordergrund. Es werden hauptsächlich standardisierte Messinstrumente wie Tests und Fragebögen verwendet. Sie sollten ein hohes Maß an Objektivität, Reliabilität und Validität aufweisen. Quantitative Methoden werden eher bei summativen Evaluationsmodellen, bei denen das technisch-funktionalistische Erkenntnisinteresse im Vordergrund steht, verwendet (Bortz 1995).

Die üblichen, am Markt erhältlichen Testverfahren sind nach Wottawa und Thierau (1998) aus nahe liegenden Gründen entweder für allgemeine Forschungszwecke oder für angewandt-diagnostische Fragen entwickelt worden, nicht aber für Evaluationszwecke.

Daraus ergeben sich folgende Nachteile:

- Geringe Änderungssensitivität: Die Items sind so ausgewählt, dass sie auch eine hohe Testwiederholungsreliabilität aufzeigen, und damit gerade nicht kurzfristig fluktuierende, leichter einer Intervention zugängliche Aspekte erfassen.
- Die Länge der Testverfahren ist auf die Bedürfnisse der Einzelfalldiagnostik abgestimmt, so dass sie für eine maßnahmebezogene Evaluation, die auf der Basis von Gruppendaten erfolgt, häufig zu aufwendig sind.
- Die Konstrukte, für die die jeweiligen Testverfahren Operationalisierungen darstellen, sind überwiegend nach Bedürfnissen der Forschung oder Diagnostik ausgerichtet. Sie decken oft nicht den Operationalisierungsbedarf von Evaluationsprojekten ab.
- Der Zusammenhang zwischen Testwert und Nutzen ist entweder überhaupt nicht bekannt oder dem Auftraggeber schwer vermittelbar.

Bei den qualitativen Methoden nach dem methodologischen Konzept der Handlungsforschung geht es laut Lamnek (1993) um die Herausarbeitung und die diskussionsfähige, deskriptive Darstellung von aktuellen Ergebnissen einer Maßnahme.

Die qualitative Erfassung von Merkmalen steht im Vordergrund. Methoden wie z.B. das Interview, die teilnehmende Beobachtung, Fragebögen mit zum Teil offenen Antwortmöglichkeiten, Stimmungsbarometer und Gruppendiskussionen werden hier eingesetzt. Die qualitativen Methoden finden eher bei formativen Evaluationsmodellen Verwendung. Der qualitative Ansatz dient nach Lamneck (1993) vor allem dazu, die Wirkung von Programmen aus der Sicht der Betroffenen zu verstehen.

## 2.3 Evaluation der Lehre

### 2.3.1 Bedeutung der Evaluation der Lehre

Worin besteht das große Interesse an der Evaluation universitärer Lehre? Von gesellschaftspolitischer Seite besteht eine zunehmende Bedeutung von Wissen in den Industriegesellschaften. Neben dem Produktionsfaktor Kapital entscheidet der Produktionsfaktor Wissen zunehmend über die Wettbewerbsfähigkeit eines Wirtschaftsstandortes oder Unternehmens (Wilke 1993). Zudem sind die teilweise dramatisch gestiegenen Kosten des universitären Ausbildungssystems zu nennen.

Die zunehmende Internationalisierung der Universitäten und das noch nicht gelöste Problem, wie Universitäten effektiv und effizient gesteuert werden können, sind ebenfalls anzuführen. Der Legitimationsdruck auf die Universitäten hat in Zeiten von Sparmaßnahmen stark zugenommen und wird durch die Konkurrenz der Fachhochschulen, die unter völlig anderen Ausgangsbedingungen als Universitäten eingerichtet wurden, noch verschärft (Altrichter 1997, Laske et al. 2000)

Der Ruf nach Evaluation universitärer Lehre wurde jedoch nicht nur aufgrund globaler gesellschaftspolitischer Entwicklungen an die Universitäten herangetragen, sondern auch universitätsinterne Bedingungen und Entwicklungen haben das Interesse an einer Qualitätssicherung in der Lehre geweckt und sogar noch gesteigert. Auf der einen Seite sind die steigenden Studierendenzahlen zu nennen, auf der anderen Seite die Geringschätzung von Lehre im Verhältnis zur Forschung (Sturm 1994, Enders und Teichler 1995).

Diskutiert werden in diesem Zusammenhang auch die Funktionen Kontrolle und Wettbewerb; denn wenn es gelingt, die Qualität der Leistungen einer Hochschule gültig, zuverlässig und umfassend zu messen, steht damit einerseits eine Kontrollmöglichkeit zur Verfügung, andererseits wird ein funktionierender Wettbewerb ermöglicht und angeregt. Der Wettbewerb erstreckt sich auf Reputation, Forschungsförderung und im Besonderen auf öffentliche Finanzmittel. Der Punkt leistungsorientierte Mittelverteilung ist im Grundsatz vom Wissenschaftsrat bereits in seiner Evaluationsempfehlung von 1996 erwähnt worden: „Evaluation dient als Grundlage einer leistungsbezogenen Mittelvergabe durch die Hochschulleitung und Ministerien.“

### **2.3.2 Historische Entwicklung der Evaluation der Lehre**

Während die universitäre Lehrevaluation in den USA eine lange Tradition hat - 1926 wurde der erste standardisierte Fragebogen veröffentlicht - gilt dies nicht für die deutschsprachigen Länder (McKeachie 1990). Flächendeckende Evaluationen wurden in den USA allerdings auch erst in den fünfziger und sechziger Jahren des vergangenen Jahrhunderts durchgeführt.

In der Folge boomte die Evaluationsforschung im Bereich universitärer Lehre. Bereits 1984 berichtet Marsh von über 1000 einschlägigen amerikanischen Studien. Dieser Unterschied zu den deutschsprachigen Ländern liegt sicherlich zu einem wesentlichen Teil in den unterschiedlichen Universitätsstrukturen begründet.

In den 70er Jahren wurde das Thema Evaluation von Hochschuldidaktikern wieder aufgegriffen (Marsh 1984), und es kam zur Entwicklung der ersten deutschsprachigen Fragebögen. Diese Pionierarbeiten wiesen allerdings eine Reihe methodischer Schwächen auf. Unter anderem wurden nur kleine, meist aus einer Studienrichtung stammende Stichproben von Studierenden befragt. Validität wurde postuliert, aber nicht geprüft.

In der Folge sank das Interesse an Lehrevaluation wieder und wurde erst Ende der 80er Jahre erneut wach, diesmal jedoch auf Seiten der Studierenden. Zwischen 1988 und 1993 wurden an österreichischen Hochschulen über 50 Lehrveranstaltungen bewertet, die jedoch hinsichtlich Qualität und Quantität beträchtlich variierten.

Parallel dazu traten der Staat bzw. die Hochschulverwaltung als Initiator und Förderer von Evaluation an Hochschulen auf. Während auf Seiten der Hochschuldidaktiker Kommunikation und Qualifikation im Zentrum des Interesses standen und für die Studierenden der Informationsaspekt zentral war, entdeckten Ministerien und Hochschulverwaltung Evaluation als mögliches Steuerungsinstrument.

In Deutschland wurden Mitte der 90er Jahre systematische Verfahren für die Evaluation von Studium und Lehre eingeführt. Die Hochschulrektorenkonferenz und der Wissenschaftsrat führten in den Jahren 1994 und 1995 zwei Pilotprojekte durch, um die Diskussion über Evaluationsverfahren auf nationaler Ebene auszuweiten.

Allgemein anerkannte Regeln für die Qualitätssicherung wurden herausgearbeitet und standardisierte Indikatoren der Leistungsmessung wurden begründet (Hochschulrektorenkonferenz 1995; Wissenschaftsrat 1996).

### **2.3.3 Formen und Ebenen der Evaluation der Lehre**

Evaluation verbindet sich heute auch mit Ansätzen wie Qualitätsmanagement und Organisationsentwicklung. Um Evaluationen als Verfahren der Qualitätskontrolle an deutschen Hochschulen einzuführen, hat es mehr als ein Jahrzehnt gedauert. Während dieser Zeit wurden eine Fülle von Konzepten und vielfältige Evaluationsansätze entwickelt. Besonders die Beurteilung von Lehrveranstaltungen durch Studierende (Reissert 1994), die durch Ländergesetze vorgeschriebenen Lehrberichte (Fischer – Blum 2002) und das Verfahren der internen und externen Evaluation (Reissert und Carstensen 1997) sind dabei besonders hervorzuheben.

Hochschuldidaktische Lehrevaluationen werden in verschiedenen Ebenen gesellschaftlicher Realität zugeordnet. Evaluationen auf der Mikroebene setzen die Studierenden in den Mittelpunkt des Interesses und setzen die Eingangsqualifikationen in eine Beziehung zu den Anforderungen einer Lehrveranstaltung. Auch Lernstrategien und das Messen des Lernerfolgs sind einer evaluativen Mikroebene zuzuordnen. Auf dieser Ebene wird auch die Qualität der Lehrpersonen durch Selbst- bzw. Fremdbewertung eingeschätzt. Grundsätzlich geht es auf dieser Ebene um die Sicherung der Qualität der Lehre.

Evaluationen auf der Mesoebene beschäftigen sich mit Fragen der Auswahl und Anordnung von Lerninhalten in einem Studiengang. Aber auch die Funktionalität von Bibliotheken, Studienberatungen und Prüfungsverfahren werden auf dieser Ebene untersucht. Solche Evaluationen sollten möglichst organisatorische Folgerungen nach sich ziehen.

Im Unterschied dazu beschäftigen sich Evaluationen auf der Makroebene mit dem Übergang der Studierenden in die berufliche Arbeitswelt. In Form von Übergangsstudien kann aufgezeigt werden, inwieweit die im Studium vermittelten Inhalte für berufliche Arbeitsfelder funktional sind.

Systematische Analysen können dann wiederum Rückwirkungen auf die Studieninhalte, Studienorganisationen und Studienpläne haben. Hochschuldidaktische Evaluationen sind insbesondere auf der Mikro- und Mesebene angeordnet (Abs et al. 2000).

#### **2.3.4 Qualitätsentwicklung durch Evaluation der Lehre**

Um Evaluation als Verfahren der Qualitätskontrolle an deutschen Hochschulen einzuführen, hat es etwas mehr als ein Jahrzehnt gedauert. Seitdem kann die Entwicklung der Evaluation als eine bemerkenswerte Erfolgsgeschichte angesehen werden (Hochschulrektorenkonferenz 1995, Wissenschaftsrat 1996). Es gibt inzwischen zahlreiche Verbände von Hochschulen, Evaluationsagenturen, -beratern und -experten, die sich professionell mit der Sicherung der Qualität in Studium und Lehre befassen und bereits Hunderte von Fächern evaluiert haben. Außerdem existieren zahlreiche Handbücher zu den verschiedenen Zielen, zu Methoden und zu einzelnen Verfahren der Evaluation, ebenso zu Teilaspekten wie der Akkreditierung, der Zielvereinbarung oder den Erfolgsbedingungen (z.B. Berendt et al. (Hrsg.) 2007: Neues Handbuch Hochschullehre).

Ein zentrales Thema der aufkommenden Qualitätsdiskussion wurde die Rechenschaftspflicht der Hochschulen durch Nachweis ihrer Leistungen in Lehre und Studium gegenüber einer breiteren Öffentlichkeit (Green 1994). In der ersten Hälfte der 90er Jahre wurde sie begleitet von mehreren Aktionsprogrammen, z.B. dem Aktionsprogramm „Qualität der Lehre“ 1990/91 des Landes Nordrhein- Westfalen, das die Verbesserung von Lehre und Studium in den Mittelpunkt rückte. Ebenfalls gab es vermehrt Tagungen der Wissenschafts- und Interessenorganisationen, die Empfehlungen und Vorschläge zur Verbesserung der Lehre entwickelten, aus denen sich im Laufe der Zeit die heute praktizierten Verfahren zur Evaluation und Berichterstattung über Lehre und Studium als Instrumente der Qualitätsförderung durchsetzten.

Welche Kriterien über gute Lehre werden nun den durchgeführten Evaluationen zugrunde gelegt? Was zeichnet „gute Lehre“ aus? Welche Merkmale charakterisieren gute Dozenten, welche zusätzlich gute Veranstaltungen?

Zur Beantwortung dieser Fragen können Lehr- und Ausbildungszielkriterien definiert werden (Krempkow 1998, Berendt 2000). Hierzu zählen unter anderem der Erwerb von

spezifischen oder allgemeinen Wissensinhalten, Kompetenzen und Einstellungen zum jeweiligen Fachgebiet (Diel 1996, Freed und Klugman 1997).

Bei „guten“ Dozenten in „guten“ Veranstaltungen sollte die Lehre den Studierenden ermöglichen, sich diesen Zielen anzunähern (Rindermann und Oubaid 1999). Merkmale guter Lehre können auch analysiert werden durch Beobachtung und Beschreibung von Veranstaltungen, die von als „gute“ Dozenten bekannten Personen gehalten werden (Helmke und Krapp 1999).

Ebenso lassen sich hochschuldidaktische Theorien heranziehen und Curricula sowie Lehrevaluationsbögen interpretieren. Der direkteste Weg, Kennzeichen guter Lehre zu finden, besteht jedoch darin, Dozenten und Studenten direkt zu befragen. Diese Personen, die tagtäglich gute oder schlechte Lehre, Vorträge und Veranstaltungen erleben, verfügen dank jahrelanger Erfahrungen über gute Kenntnisse der Hochschulpraxis, sie lehren selbst oder nahmen an Hochschulveranstaltungen unterschiedlicher Fachrichtungen teil (Rindermann 1999).

Zu den für erfolgreiche Lehre verantwortlichen Lehrmethoden zählen insbesondere solche, die sich durch Strukturierung und Klarheit auszeichnen. In vielen Studien und experimentellen Untersuchungen konnte ein inhaltlich bedeutsamer Zusammenhang zwischen klarer Strukturierung der Lehre und studentischem Lerneffekt festgestellt werden (Patton 1981, Dunkin und Barns 1986).

Zu den Merkmalen auf Seiten des Dozentenverhaltens gehören darüber hinaus auch verständliche Erklärungen und Darstellungen, rhetorische Kompetenz, Motivierung und die Darstellung der Relevanz oder der Theorie-Praxis-Bezüge. Verarbeitungstiefe, Zeit- und Schwierigkeitsmanagement, Engagement, Kooperativität sowie in Seminaren das Leiten von Diskussionen ist ein weiterer Aspekt der Handlungskompetenz (Regehr und Norman 1996).

Weitere Bedingungsvariablen des Erfolgs einer Veranstaltung stellen aber auch die studentischen Verhaltensweisen dar. Dazu zählen insbesondere die Vor- und Nachbereitung der Lehrinhalte, die veranstaltungsinterne Beteiligung sowie Diskussionen und Referate (Daniel 1995).

Zudem können Rahmenbedingungen, wie das gestellte Thema, die allgemeinen Anforderungen, Prüfungen, Pflichtmäßigkeit des Besuches, Veranstaltungstyp und Ähnliches als Kontextvariablen den Lehrerfolg wie auch die Veranstaltungsbewertung beeinflussen (Sturm 1994). Dozenten, Studierende und Kontextbedingungen wirken nicht nur direkt auf den Lehrerfolg, sondern üben über gegenseitige Einflussnahme auch indirekt Wirkung auf den Erfolg der Lehre aus.

Ein weiterer Auswertungsbefund aus Lehrevaluationen wird selten berücksichtigt, hat aber eine entscheidende Bedeutung für die Interpretation: der Einfluss der studentischen „Peers“ auf die Wahrnehmung und Beurteilung der Lehre, durch die anderen Teilnehmer. Wird eine Veranstaltung überwiegend von desinteressierten Studierenden besucht, werden alle Teilnehmer negativ beeinflusst. Im Gegensatz dazu findet sich ein deutlich positiver Effekt, wenn die überwiegende Mehrheit der Teilnehmer Interesse zeigt. Dies gilt für die Beurteilung der Lehrveranstaltung ebenso wie für die Selbsteinschätzung des eigenen Lernerfolgs oder die Bereitschaft, sich mit dem behandelten Lernstoff intensiver selbständig zu beschäftigen (Komrey 1994).

In Veranstaltungen mit negativer Grundstimmung hat die Lehrperson durch eigene Bemühungen kaum Möglichkeiten, das Interesse für die Lehrinhalte zu wecken. Eine genauere vergleichende Analyse über eine Vielzahl von Veranstaltungen hinweg zeigt aber auch, dass das von Studierenden mitgebrachte eigene Interesse am Lehrstoff und die Einsicht in den Nutzen für das weitere Studium oder für den späteren Beruf eine positive Grundstimmung hervorruft (Abrami und Cohen 1990). Diese Einsicht wiederum lässt sich im Kontext des Gesamtfachs durch ein in sich stimmiges und für die Studierenden nachvollziehbares Curriculum ebenso fördern, wie durch Studienberatung und Orientierungsangebote über die Berufspraxis.

### **2.3.5 Verfahren zur Evaluation von Lehrveranstaltungen**

Es gibt verschiedene Verfahren, Lehrveranstaltungen zu evaluieren. Methoden, wie das Interview, die teilnehmende Beobachtung, Fragebögen mit zum Teil offenen Antwortmöglichkeiten, Stimmungsbarometer und Gruppendiskussionen mit dem systematisch Urteile zum Beurteilungsgegenstand erhoben werden, finden im Rahmen von Evaluationsverfahren Verwendung.

Bei traditionellen Verfahren werden am Ende der Lehrveranstaltung Fragebögen ausgeteilt, von den Studierenden ausgefüllt und wieder abgegeben. Anschließend müssen die erhobenen Daten für die statistische Auswertung in Rechner eingegeben werden. Dieses Verfahren erfordert einen hohen organisatorischen und auch finanziellen Aufwand. Hinzu kommt, dass die Ergebnisse meist erst vorliegen, wenn das Semester längst beendet ist und so für Studenten kein unmittelbarer Zusammenhang zwischen Lehrveranstaltung und Auswertung der Lehrevaluation besteht.

Um eine hinreichend große Beteiligung der Studierenden bei der Beurteilung einzelner Lehrveranstaltungen, eine möglichst zeitnahe Auswertung der Befragung und eine schnelle Rückmeldung an Studierende und Dozenten zu erreichen, bietet das Internet neue Möglichkeiten.

An der Medizinischen Fakultät der Westfälischen Wilhelms-Universität Münster wird zur flächendeckenden Evaluation der curricularen Lehre seit dem Wintersemester 2001/2002 EVALuna eingesetzt. Hierbei handelt es sich um eine kontinuierliche Online-Evaluation zur Erfassung der studentischen Beurteilung von Lehrveranstaltungen.

Die Studierenden der Medizinischen Fakultät werden durch die Studienordnung und organisatorisch durch die Koppelung mit der Scheinvergabe der einzelnen medizinischen Fachrichtungen dazu verpflichtet, sich an der anonym durchgeführten Online-Evaluation zu beteiligen.

Um die Effektivität der Evaluation darzustellen, wird die Häufigkeit der Teilnahme der Studierenden an den jeweiligen Lehrveranstaltungen miterfasst und in drei Kategorien eingeteilt: „unter 50%“, „50 bis 75%“ und „über 75%“. Diese Angaben der Studierenden können aber nicht systematisch überprüft werden. Für die Evaluation der Lehre sollen die Studierenden dann eine bestimmte Lehrveranstaltung in ihrem gesamten Semesterumfang durch nur eine einzige Bewertung beurteilen. Hierzu wird auf einer Analogskala von „sehr gut“ bis „sehr schlecht“ mit einer Skala von 0 bis 100 ein „Schiebereglert“ bewegt, wobei die Ziffer 0 die beste und die Ziffer 100 die schlechteste Bewertung darstellt.

Darüber hinaus haben die Studierenden in der Online-Evaluation die Möglichkeit, freie Kommentare abzugeben und die jeweilige Lehrveranstaltung nach eigener Einschätzung zu kommentieren. Jeder Lehrende hat über ein persönliches Passwort Zugang zu EVALuna und damit Einsicht in seine persönlichen Lehrveranstaltungs-Bewertungen der Studierenden. Auf diesem Weg ist es den Dozenten auch möglich, über einen anonymen Link dem einzelnen Studierenden ein persönliches Feedback zu geben.

Einblick in die Gesamtevaluation und Überblick über die am besten und am schlechtesten beurteilten Lehrveranstaltungen hat als Einziger der Studiendekan der Medizinischen Fakultät der WWU Münster. Die Ergebnisse der studentischen Beurteilungen fließen neben quantitativen Faktoren wie Zahl und Größe der Lehrveranstaltungen in die leistungsbezogene Mittelvergabe der einzelnen Kliniken und Institute ein. Diese leistungsorientierte Mittelvergabe wird mittlerweile zunehmend eingesetzt, um Anreize zu Qualitätsverbesserung von Lehrveranstaltungen zu schaffen.

### **2.3.6 Validität studentischer Lehrevaluation**

Obwohl in der Regel gute Beurteilungen von Lehrenden und Lehrveranstaltungen durch Studierende erzielt werden, ist die Skepsis vieler in der Lehre Tätiger gegenüber studentischen Evaluationen groß (Kromrey 1996, Salzer und Nixon 1997). Selbst Gerichte kamen zu der Ansicht, dass Studenten „zu einer objektiven Beurteilung generell nicht in der Lage seien“ (Mitteilungen des deutschen Hochschulverbandes 1991).

Die Validität studentischer Veranstaltungsbeurteilungen wird von vielen angezweifelt (Greenwald 1997) „One needs not to talk with faculty very long to be aware of their concern about possible biases in student ratings“ (Kromrey 1995).

Von Dozenten wird vielfach die Vermutung gehegt, dass Studierende die Qualität von Lehrveranstaltungen auf Grund so genannter Biasvariablen nicht adäquat beurteilen können (Gold 1996). Biasvariable sind Einflussgrößen, die einen potentiell verzerrenden Effekt auf das zu evaluierende Zielkriterium haben.

Im Fall der Lehrveranstaltungsevaluierung sind dies Variablen, die grundsätzlich nichts mit Lehrqualität zu tun haben, jedoch Studierende systematisch bei der Bewertung

beeinflussen. Zu diesen potentiellen Verzerrungsfaktoren zählen nach Befragungen amerikanischer Dozenten Themenschwierigkeit, Benotungsstrenge, Popularität, Vorinteresse, Arbeitslast, Besuchszahl und Besuchsgrund (Marsh 1984). Deutsche Dozenten nennen an erster Stelle ebenso Themenschwierigkeit, dann Besuchsgrund und Interessantheit des Veranstaltungsstoffes als die Beurteilungen positiv oder negativ beeinflussende Faktoren (Rindermann 1996, Spiel 2001).

Solche urteilsverzerrenden Einflussgrößen lassen sich zu vier übergeordneten Bereichen zusammenfassen (Spiel und Gössler 2000):

- Merkmale der Lehrenden (z.B. Sympathie oder Attraktivität des Dozenten)
- Merkmale der Studierenden (z.B. Teilnahmemotivation)
- Umgebungscharakteristika (z.B. Raummerkmale)
- Charakteristika der Studienrichtung (z.B. Größe der Studienrichtung)

Das Interesse am Inhalt einer Lehrveranstaltung zum Beispiel kann in Vorinteresse als echter Bias und in im Laufe der Lehrveranstaltung gewecktes Interesse als Variable der Lehrwirksamkeit differenziert werden. Von daher stellt sich die Frage, ob es sich bei Interesse tatsächlich um einen Biaseffekt handelt oder eher um ein Qualitätsmerkmal (Rindermann 1997a).

Eine explizite Trennung zwischen Vorinteresse und dem in der Lehrveranstaltung geweckten Interesse durch unabhängige Items ist somit sinnvoll. Dabei ist allerdings die Zuverlässigkeit der Aussagekraft eingeschränkt, da bereits von vornherein Interessierte offensichtlich leichter zu begeistern sind als wenig Interessierte (Kromrey 1994). Nur Längsschnittuntersuchungen, die jedoch über die gesetzlichen Forderungen hinausgehen und flächendeckend an Universitäten auch aus Kostengründen kaum realisierbar sind, können hier zuverlässige Informationen liefern.

Nicht nur Einflüsse von Biasvariablen führten zur Kritik an der Validität studentischer Lehrveranstaltungsbewertung. Auch die Frage, ob bzw. inwieweit Studierende überhaupt in der Lage sind, die fachlichen Kompetenzen eines Dozenten zu beurteilen, wurde diskutiert. Vielfach wird der Standpunkt vertreten, dass vor allem in den Anfangssemestern eine entsprechende Bewertung nicht adäquat erfolgen kann (Kromrey 1996, Rindermann 2000a). Es ist aber dennoch anzunehmen, dass fachliche

Defizite wie Widersprüchlichkeiten im Vortrag oder ungenügende Beantwortung von Nachfragen auch Studienanfängern auffallen, die selbst noch keine bzw. wenig Fachkenntnisse erworben haben.

Ergebnisse studentischer Lehrevaluationen können das Lehrverhalten des einzelnen Dozenten oder aber auch die spezifische Teilnehmerzusammensetzung oder die Themeninteressantheit reflektieren. Diese Faktoren lassen sich analysieren, wenn Beurteilungen aus mehreren Veranstaltungen miteinander verglichen werden. (Rindermann 2000 b).

Zusammenfassend ist die Lehrbeurteilung prinzipiell das Resultat mehrerer Faktoren. Zu diesen Faktoren gehören der Dozent, die Studierenden, die Rahmenbedingungen und die sich daraus ergebenden Interaktionen.

### **3. Empirischer Teil**

#### **3.1 Das „Praxisprojekt Arbeitsmedizin“: Lernziele, Struktur und Inhalt**

Ziel bei der Entwicklung des Praxisprojektes Arbeitsmedizin war es, die Motivation und das Interesse der Studierenden für das Fach Arbeitsmedizin zu erhöhen und über die Vermittlung von theoretischen Grundlagen hinaus insbesondere die ärztlichen Aufgaben des Fachgebietes praxisbezogen zu vermitteln. Zudem sollte die Fähigkeit der Studierenden erhöht werden, angeeignetes Wissen fächerübergreifend anzuwenden.

Auf dem Hintergrund der beschriebenen Erkenntnisse zum Problemorientierten Lernen und Projektlernen wurden für das Fachgebiet Arbeitsmedizin Lehrmethoden zur Wissensverankerung gefunden, die durch Anschlusswissen und Aktivierung des selbstständigen Wissenserwerbs exemplarisches Lernen ermöglicht. Das Problemorientierte Lernen wurde für das Praxisprojekt Arbeitsmedizin angepasst. Im Gegensatz zur herkömmlichen Praxis wurde hier kein „paper case“ für die Gruppenarbeit zugrunde gelegt, sondern als „real case“ ein der Gruppe zugeordneter realer Betrieb.

Seit dem Wintersemester 1999/2000 werden die arbeitsmedizinischen Lehrinhalte in Form von Praxisprojekten erarbeitet. Der Kurs wird als Pflichtlehrveranstaltung zweimal pro Semester mit jeweils 8 Projektgruppen zu jeweils ca. 8 Studierenden durchgeführt.

In der Einführungsveranstaltung werden die Lehr- und Lernziele bekannt gegeben. Dabei wird zwischen speziellen und allgemeinen Lehr- und Lernzielen unterschieden. Zu den speziellen Zielen gehören die wesentlichen Grundlagen und Inhalte der Arbeitsmedizin, die Aufgaben und Ziele der ärztlichen Handlungen sowie die fächerübergreifenden Grundlagen für die ärztliche Tätigkeit in anderen Fachdisziplinen.

Das allgemeine Lehr- und Lernziel besteht darin zu lernen, ein Problem zu bearbeiten und in Teamarbeit zu lösen. Dazu sollen Ideen und Initiativen für eigenständige Problemlösungen entwickelt werden mit Delegation von Aufgaben innerhalb der Gruppe und mit regelmäßigem Informations- und Erfahrungsaustausch. Diese Ziele sollen durch projekt- und problemorientiertes Lernen in Gruppenarbeit mit Dozentenunterstützung erreicht werden.

Dazu werden praktische Beispiele im Rahmen einer Betriebsbegehung in Betrieben unterschiedlicher Branchen angeboten, z. B. Kunststoffverarbeitung, Metallverarbeitung, Produktion von Farben und Lacken, Papierverarbeitung, Fahrzeugbau, Textilherstellung und -verarbeitung, Fleischverarbeitung, Gießerei, Feuerwehr, Zahntechnik, Flughafen. Jede Gruppe erhält ihren „eigenen“ Betrieb zugeordnet und wird dort vor Ort während der Begehung von den zuständigen Sicherheitsfachkräften und Betriebsärzten betreut.

Durch diesen konkreten Praxisbezug werden wesentliche Inhalte des Fachgebietes mit den methodischen Grundlagen des projekt- und problemorientierten Lernens, wie ausführlich dargestellt wurde, exemplarisch erarbeitet.

Die Ergebnisse werden strukturiert und in einer Projektpräsentation dargestellt. Durch die Projektpräsentation im Plenum werden alle Gruppenprojekte zum Gesamtprojekt Arbeitsmedizin zusammengeführt. Dazu wurde inhaltlich und organisatorisch eine Grundstruktur mit aufeinander aufbauenden Lernschritten entwickelt (Witting 2002), die nachfolgend dargestellt wird.

### **Schritt 1: Plenum**

In Grundlagenseminaren wird Basiswissen vermittelt zum Fachgebiet Arbeitsmedizin, zu fachspezifischen und fächerübergreifenden ärztlichen Aufgaben, zum System der sozialen Sicherung, zu den Betrieben, zur jeweiligen Problemstellung sowie zur entsprechenden arbeitsmedizinischen Terminologie.

### **Schritt 2: Gruppenarbeit**

In diesem Lernschritt werden im Rahmen eines Brainstormings Vorkenntnisse und Vermutungen im Hinblick auf die Art des Betriebes, der Produkte und Arbeitsverfahren, auf mögliche arbeitsmedizinische und sicherheitstechnische Probleme und auf mögliche Gesundheitsgefährdungen zusammengetragen und in einer Problemliste strukturiert.

### **Schritt 3: Gruppenarbeit**

Der dritte Lernschritt baut auf dem Erkenntnisgewinn aus dem zweiten Schritt auf und beinhaltet die kritische Diskussion der Vorschläge und Ideen. Die Erstellung einer Checkliste ist an diesem Punkt von besonderer Bedeutung, um das Erkennen der je-

weiligen Problemfelder des Betriebes zu ermöglichen. Die jeweilige Informationsbeschaffung erfolgt dann während der sich anschließenden Betriebsbegehung anhand dieser Checkliste. Von besonderer Bedeutung sind dabei Informationen zu: Betriebsstruktur, Arbeitsorganisation, Produktionsverfahren, Organisation der arbeitsmedizinischen und sicherheitstechnischen Betreuung, möglichen Gesundheitsgefährdungen und Berufskrankheiten sowie zu den Aufgaben des Betriebsarztes. Dabei sollen insbesondere Arbeitsabläufe und Arbeitsbedingungen unter besonderer Berücksichtigung der Arbeitsschutzmaßnahmen beobachtet und registriert werden.

#### **Schritt 4: Gruppenarbeit**

In der Nachbereitung der Betriebsbegehung werden die erhaltenen Informationen zu den jeweiligen Problemfeldern systematisch geordnet.

#### **Schritt 5: Gruppenarbeit**

Dieser Lernschritt dient der Fokussierung auf wichtige Problemfelder sowie der Formulierung der Gruppenlernziele mit Festlegung des notwendigen Informationsbedarfs. Besonderes Augenmerk soll dabei auf Gesundheitsgefährdungen, Berufskrankheiten, Unfälle und Präventionsmaßnahmen gerichtet werden.

#### **Schritt 6: Gruppenarbeit**

In diesem Lernschritt werden die notwendigen Lerninhalte erarbeitet. Dabei wird die bestehende Gruppenstruktur aufgelöst, um durch Aufgabenteilung und Delegation die gesamte Informationsbeschaffung zu sichern. Hierzu werden Spezielle Seminare und praktische Übungen angeboten, in die die Gruppe Mitglieder delegiert.

In diesen Veranstaltungen werden unmittelbar projektbezogene Lehrinhalte zu Berufskrankheiten und arbeitsmedizinischen Untersuchungsverfahren vermittelt. Als Informationsquellen werden auch das Internet und die Präsenzbibliothek des Instituts genutzt.

#### **Schritt 7: Gruppenarbeit / Plenum**

Dieser Lernschritt dient dem Informationsaustausch der einzelnen Gruppenmitglieder, der Strukturierung des erarbeiteten Wissens und der Zusammenfassung und Darstellung der Ergebnisse in einer Präsentation. Mit der Vorstellung aller Gruppenprojekte im Plenum erfolgt dann die Zusammenführung der Lehr- und Lerninhalte zum Gesamtprojekt Arbeitsmedizin.

### **Schritt 8: Abschlusskolloquium / Evaluation**

Der Lehr- und Lernerfolg wird im Abschlusskolloquium in Form einer mündlichen Prüfung mit jeweils vier Studierenden pro Gruppe überprüft. Nach Teilnahme der Studierenden an der institutsinternen Evaluation wird das Abschlusstestament erteilt.

Integraler Bestandteil des Lehrprojektes „Praxisprojekt Arbeitsmedizin“ ist der arbeitsmedizinische Qualitätszirkel „Lehre“ mit allen an der Lehre Beteiligten, insbesondere auch mit den externen Partnern, d. h. mit den Betriebsärzten der beteiligten Betriebe.

Das Praxisprojekt Arbeitsmedizin mit dem neuen Lehr- und Lernkonzept wurde erstmals im Wintersemester 1999/2000 angeboten und seitdem kontinuierlich von allen Studierenden evaluiert. Die Auswertung der institutsinternen studentischen Evaluation bis einschließlich Sommersemester 2003 ist Gegenstand der vorliegenden Arbeit.

### 3.2 Zielsetzung und Fragestellung

Ziel dieser Studie ist es, die Ergebnisse der vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 kontinuierlich durchgeführten studentischen Evaluation auszuwerten, um

1. den Erfolg des neuen Lehr- und Lernansatzes und
2. die Validität des Evaluationsverfahrens zu überprüfen.

Dazu sollen Aussagen insbesondere zu folgenden Fragestellungen erarbeitet werden:

1. Welchen Einfluss hat das neu konzipierte didaktische und methodische Lehrkonzept auf die Teilnahmemotivation und den Lernerfolg der Studierenden?
2. Wie bewerten die Studierenden Struktur, Didaktik und Lehrkompetenz?
3. Wie bewerten die Studierenden den direkten Praxisbezug durch die Betriebsbegehung?
4. Welche Auswirkung hat die Arbeit in Kleingruppen auf die Gesamtbewertung der Lehrveranstaltung?
5. Entspricht die durchgeführte Evaluation den methodischen Ansprüchen an die Reliabilität und Validität von Evaluationsverfahren?
6. Kann die kontinuierliche studentische Lehrevaluation zur Qualität der Lehre in der Arbeitsmedizin beitragen?

### 3.3 Methodik

#### 3.3.1 Kollektiv

Die Pflichtlehrveranstaltungen des Fachgebiets Arbeitsmedizin werden wie einleitend dargestellt in Form eines Praxisprojektes durchgeführt. Die Teilnahme ist grundsätzlich nur Studierenden möglich, die durch das Institut für Ausbildung und Studienangelegenheiten (IfAS) in eine Klinikgruppe eingewiesen wurden.

Der Kurs ist dem 4. klinischen Semester zugeordnet und findet zweimal pro Semester statt mit jeweils 8 Projektgruppen zu jeweils ca. 8 Studierenden. Um diese Gruppengröße zu erreichen, werden die Klinikgruppen 1-20 in die Projektgruppen AH und IQ eingeteilt. In diese Gruppen werden auch Studierende aufgenommen, die z. B. aufgrund eines Auslandsstudiums den Kurs zu einem früheren oder auch späterem Zeitpunkt belegen. Die Zuordnung der Studierenden zu diesen neuen Gruppen wird in der Vorbesprechung und durch Aushang bekannt gegeben. Im Untersuchungszeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 haben 1020 Studierende den Kurs absolviert. 981 ausgefüllte Fragebögen liegen für die Auswertung der institutsinternen Evaluation vor. Auf eine Überprüfung der Geschlechtsvariablen wurde aufgrund von zum Teil fehlenden Angaben in den Studierendenlisten verzichtet.

Die Anzahl der Studierenden, die jeweils in der ersten und zweiten Semesterhälfte am Kurs teilgenommen haben, ist in Tabelle 1 aufgeführt.

Tabelle 1: „Semester numerisch“ Anzahl der Studierenden

Semester	1. Semesterhälfte	2. Semesterhälfte	gesamt	Fragebögen
WS 99/00	64	62	126	125
SS 00	61	71	132	123
WS 00/01	67	59	126	124
SS 01	69	63	132	122
WS 01/02	56	60	116	111
SS 02	86	44	130	122
WS 02/03	66	67	133	129
SS 03	66	59	125	125
		<b>Summe</b>	1020	981

### 3.3.2 Fragebogen

Der Fragebogen zur Evaluation des „Praxisprojekt Arbeitsmedizin“ bezieht sich auf die vier didaktischen Elemente der Lehrveranstaltung (die Grundlagenseminare, praktische Übungen, Betriebsbegehung und Gruppenarbeit) und eine Gesamtbeurteilung der Lehrveranstaltung. Für jeden dieser fünf Aspekte wurden vom Institut für Arbeitsmedizin der Universität Münster Fragen entwickelt, um die Beurteilung dieser Aspekte durch die Studierenden zu erfassen. Die Bewertungen werden nach der Skalierung: „trifft völlig zu“, „trifft überwiegend zu“, „trifft selten zu“ und „trifft überhaupt nicht zu“ vorgenommen, die mit Bewertungsziffern von „eins“ bis „vier“ unterlegt sind.

Die Beurteilung der Grundlagenseminare wurde über fünf Items operationalisiert. Zu den entsprechenden Itemformulierungen gehören:

- Die Medien wurden sinnvoll und effektiv eingesetzt
- Das Thema wurde durch praxisnahe Beispiele veranschaulicht
- Inhaltliche Fragen waren erwünscht
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet
- Es viel mir leicht aufmerksam zu bleiben.

Für die Beurteilung der praktischen Übungen wurden die folgenden Items verwendet:

- Mein Basiswissen war ausreichend um den Übungen zu folgen
- Der Lernstoff wurde im Dialog erarbeitet
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet
- Der Lehrinhalt wurde interessant präsentiert
- Es viel mir leicht aufmerksam zu bleiben.

Die Beurteilung der Betriebsbegehung wurde über die folgenden vier Items erfasst:

- Die Betriebsbegehung war inhaltlich gut vorbereitet
- Die Betriebsbegehung war organisatorisch gut vorbereitet
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet
- Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft.

Zur Erfassung der Bewertung der Gruppenarbeit wurden drei Items eingesetzt:

- Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes
- Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation
- Die Gruppenarbeit hat Spaß gemacht.

Abschließend werden die Studierenden nach ihrer Gesamteinschätzung durch die folgenden acht Items befragt:

- Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert
- Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen
- Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind
- Mein persönlicher Lernerfolg durch die Veranstaltung war groß
- Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen
- Die Lehrpersonen waren gut vorbereitet
- Ich hätte an der Veranstaltung auch freiwillig teilgenommen
- Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt.

### 3.3.3 Statistik

Im Untersuchungszeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 haben insgesamt 1020 Studierende die Lehrveranstaltung „Praxisprojekt Arbeitsmedizin“ der Universität Münster absolviert. 981 der abgegebenen Evaluationsbögen konnten in die Auswertung einbezogen werden. Die statistische Auswertung der Daten wurde mit SPSS Version 10.0 durchgeführt.

#### Mittelwert

Der arithmetische Mittelwert wird mit dem Statistikprogramm SPSS 10.0 nach

folgender Formel berechnet:  $\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$

#### Standardabweichung

Die Standardabweichung wird mit dem Statistikprogramm SPSS 10.0 nach folgender Formel berechnet:

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N}}$$

#### t-Test für unabhängige Stichproben

Der t-Test wird zur Prüfung auf Mittelwertsunterschiede zwischen zwei unabhängigen Stichproben eingesetzt. Dabei wird die Normalverteilung des zu untersuchenden Merkmals in der Grundgesamtheit vorausgesetzt und die t-Verteilung als Prüfverteilung herangezogen. Im ersten Analyseschritt wird geprüft, ob sich die Varianzen des Merkmals in beiden Stichproben signifikant unterscheiden oder nicht. Dies wird durch den F-Test untersucht:

$$F = \frac{s_{x_2}^2}{s_{x_1}^2}$$

Der empirische F-Wert wird dann mit dem theoretischen F-Wert verglichen. Ist der empirische F-Wert größer als der theoretische, ist von einer signifikanten Unterschiedlichkeit der Varianzen auszugehen. Ist der empirische F-Wert kleiner als der theoretische, unterscheiden sich die Varianzen des Merkmals in den beiden Stichproben nicht signifikant.

In Abhängigkeit davon berechnet sich die Prüfgröße t wie folgt:

$$\text{bei gleichen Varianzen: } t = \left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_{x_1}^2 + n_2 s_{x_2}^2}{n_1 + n_2} \cdot \frac{1}{n_1 n_2}}} \right|$$

Die Freiheitsgrade berechnen sich über  $n_1 + n_2 - 2$

$$\text{bei ungleichen Varianzen: } t = \left| \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_{x_1}^2}{n_1 - 1} + \frac{s_{x_2}^2}{n_2 - 1}}} \right|$$

Freiheitsgrade werden über folgende Formel bestimmt:  $df = \frac{\left[ \frac{s_{x_1}^2}{(n_1 - 1)} + \frac{s_{x_2}^2}{(n_2 - 1)} \right]^2}{\frac{\left[ \frac{s_{x_1}^2}{n_1 - 1} \right]^2}{n_1 + 1} + \frac{\left[ \frac{s_{x_2}^2}{n_2 - 1} \right]^2}{n_2 + 1}} - 2$

Ist die empirische Prüfgröße t größer als der dem gewünschten Signifikanzniveau entsprechende t - Wert der t - Verteilung, unterscheiden sich die Mittelwerte der beiden Stichproben signifikant.

## Multiple Regression

Die multiple Regression dient der Analyse von Beziehungen zwischen einer abhängigen und mehreren unabhängigen Variablen. Sie wird verwendet, um (a) Zusammenhänge quantitativ zu beschreiben und zu erklären und um (b) Werte der abhängigen Variablen zu prognostizieren. Das mathematisch zu lösende Gleichungssystem lautet:

$$\text{Regressionsgleichung } y = b_0 + b_1x_1 + b_2x_2 + \dots + b_jx_j + \dots + b_Jx_J$$

wobei folgende Zielfunktion zu erreichen ist:

$$\text{Zielfunktion: } \sum_{k=1}^K e_k^2 = \sum_{k=1}^K \left[ y_k - (b_0 + b_1x_{1k} + b_2x_{2k} + \dots + b_jx_{jk} + \dots + b_Jx_{Jk}) \right]^2 \longrightarrow \min$$

mit:

$e_k$  = Werte der Residualgröße ( $k=1, 2, \dots, K$ )

$y_k$  = Werte der abhängigen Variablen ( $k=1, 2, \dots, K$ )

$b_0$  = konstantes Glied

$b_j$  = Regressionskoeffizienten ( $j=1, 2, \dots, J$ )

$x_{jk}$  = Werte der unabhängigen Variablen ( $j=1, 2, \dots, J; k=1, 2, \dots, K$ )

$J$  = Zahl der unabhängigen Variablen

$K$  = Zahl der Beobachtungen

Die Regressionskoeffizienten geben dabei den marginalen Effekt der Änderung einer unabhängigen Variablen auf die abhängige Variable an. Sowohl für die einzelnen Regressionseffekte wie auch für das Gesamtmodell stehen entsprechende Signifikanztests zur Verfügung. Berechnet wurde die multiple Regression mit dem Software-Paket SPSS 10.0.

## Explorative Faktorenanalyse

Ziel der explorativen Faktorenanalyse ist es, die Gemeinsamkeiten mehrerer untereinander unabhängiger Variablen zu entdecken und auf einen gemeinsamen Faktor zurückzuführen.

Solche Faktoren werden als „hinter den Variablen“ stehende Größen betrachtet, die den korrelativen Zusammenhang zwischen den Variablen repräsentieren. Hohe Korrelationen zwischen den Variablen deuten darauf hin, dass (a) sich die Variablen gegenseitig bedingen oder (b) das Zustandekommen der Korrelationswerte durch einen oder mehrere hinter den zusammenhängenden Variablen stehende Faktoren bestimmt wird.

Grundannahme ist, dass sich jeder Beobachtungswert einer Ausgangsvariablen  $x_j$  als eine Linearkombination mehrerer (hypothetischer) Faktoren beschreiben lässt.

Fundamentaltheorem:  $x_{kj} = a_{j1}p_{k1} + a_{j2}p_{k2} + \dots + a_{jq}p_{kq}$

Die Faktorenladung ist dabei im mathematisch-statistischen Sinne eine Maßgröße für den Zusammenhang zwischen Variable und Faktor (Korrelationskoeffizient). Berechnen lässt sie sich mit dem Software-Paket SPSS 10.0

## Konfirmatorische Faktoren- und Kausalanalyse

Die konfirmatorische Faktoren- und Kausalanalyse ist eine Modellvariante aus dem Bereich der Strukturgleichungsmodelle. Strukturgleichungsmodelle haben konfirmatorischen Charakter, weil sie theoretisch fundierte Hypothesen empirisch überprüfen. Im Unterschied zu anderen kausalanalytischen Verfahren können mit Strukturgleichungsmodellen auch Beziehungen zwischen latenten, also nicht direkt beobachtbaren Variablen überprüft werden. Latente Variablen repräsentieren hypothetische Konstrukte, für deren abstrakte Inhalte sich nicht unmittelbar entscheiden lässt, ob der gemeinte Sachverhalt in der Realität vorliegt oder nicht. Ferner berücksichtigen Strukturgleichungsmodelle unsystematische und systematische Messfehler explizit, indem auch diese geschätzt werden.

Im Strukturmodell werden die Beziehungen zwischen den latenten Variablen und hypothetischen Konstrukten dargestellt.

Strukturgleichung:  $\mathbf{h} = \mathbf{B}\mathbf{h} + \mathbf{\Gamma}\mathbf{x} + \mathbf{z}$

mit:

$\eta$  = latente endogene Variablen

$\mathbf{B}$  = (m x m)-Matrix der postulierten Beziehungen zwischen  $\eta$ -Variablen

$\mathbf{\Gamma}$  = (m x n)-Matrix der postulierten Beziehungen zwischen  $\eta$ - und  $\xi$ -Variablen

$\xi$  = latente exogene Variablen

$\zeta$  = Residualvariable für latente endogene Variablen

Ferner wird für jedes hypothetische Konstrukt ein Messmodell formuliert. Die Zuverlässigkeit der Schätzungen steht so zur Gesamtbeurteilung des Modells zur Verfügung. Je nach Ergebnis werden Messmodelle empirische Indikatoren für die latenten Variablen enthalten.

Messmodell der latenten exogenen Variablen:  $X = \Lambda_x \mathbf{x} + \mathbf{d}$

mit:

$X$  = Indikatorenvariablen der latenten exogenen Variablen

$\Lambda_x$  = (q x n)-Matrix, die die Koeffizienten der Pfade zwischen x- und  $\xi$ -Variablen repräsentiert

$\xi$  = latente exogene Variablen

$\delta$  = Residuen der exogenen Variablen

Messmodell der latenten endogenen Variablen:  $Y = \Lambda_y \mathbf{h} + \mathbf{e}$

mit:

$Y$  = Indikatorenvariablen der latenten endogenen Variablen

$\Lambda_y$  = (p x m)-Matrix, die die Koeffizienten der Pfade zwischen y- und  $\eta$ -Variablen repräsentiert

$\eta$  = latente endogene Variablen

$\varepsilon$  = Residuen der endogenen Variablen

Diese Matrixgleichungen lassen sich unter Verwendung der empirischen Kovarianzen zwischen den im Modell berücksichtigten manifesten Variablen schätzen. Ziel der Parameterschätzung ist es, die Differenz zwischen der modelltheoretischen Varianz-Kovarianzmatrix ( $S$ ) und der empirischen Varianz-Kovarianzmatrix der Stichprobe ( $S$ ) zu minimieren (Reinecke 2005, Backhaus et al. 2003,).

## **3.4 Ergebnisse**

### **3.4.1 Häufigkeitsverteilungen**

Der Fragebogen zur Evaluation der Lehrveranstaltung bezieht sich auf die vier didaktischen Elemente der Lehrveranstaltung (Grundlagenseminare, praktische Übungen, Betriebsbegehung, Gruppenarbeit) und eine Gesamtbeurteilung der Lehrveranstaltung. Für jeden dieser fünf Aspekte wurden vom Institut für Arbeitsmedizin der Universität Münster Fragen entwickelt, um die Beurteilung dieser Aspekte durch die Studierenden zu erfassen.

Im Folgenden werden zunächst für jeden Teilaspekt die Itemformulierungen zusammen mit den prozentualen Häufigkeitsverteilungen für die Gesamtstichprobe vorgestellt. Die absoluten Häufigkeitsverteilungen werden aufgrund der ordinalen Antwortskalierung in Säulendiagrammen dargestellt, denn die Abstände zwischen den Antwortkategorien sind im metrischen Sinne nicht eindeutig definiert. Liniendiagramme würden suggerieren, dass die Abstände zwischen den einzelnen Antwortkategorien linear und kontinuierlich sind, was aber bei ordinalem Skalenniveau nicht der Fall ist.

### 3.4.1.1 Grundlagenseminare

Die Beurteilung der Grundlagenseminare wurde über fünf Items operationalisiert. Die entsprechenden Itemformulierungen und die deskriptiven Häufigkeitsverteilungen der Bewertungen sind in Tabelle 2 dargestellt.

Tabelle 2: Die Bewertungen der Grundlagenseminare: Itemformulierungen und prozentuale Häufigkeitsverteilung

Label	Itemformulierung	1 trifft völlig zu	2 trifft über- wiegend zu	3 trifft selten zu	4 trifft über- haupt nicht zu
MEDIEN	Die Medien wurden sinnvoll und effektiv eingesetzt	29,6 %	66,4 %	3,6 %	0,3 %
THEMA	Das Thema wurde durch praxisnahe Beispiel veranschaulicht	26,7 %	61,2 %	11,2 %	0,7 %
FRAGEN 1	Inhaltliche Fragen waren erwünscht	54,8 %	37,6 %	6,8 %	0,4 %
KOMPET 1	Inhaltliche Fragen wurden kompetent und verständlich beantwortet	44,4 %	49,7 %	5,2 %	0,1 %
AUFMERK1	Es fiel mir leicht aufmerksam zu bleiben	5,8 %	56,7 %	28,1 %	8,4 %

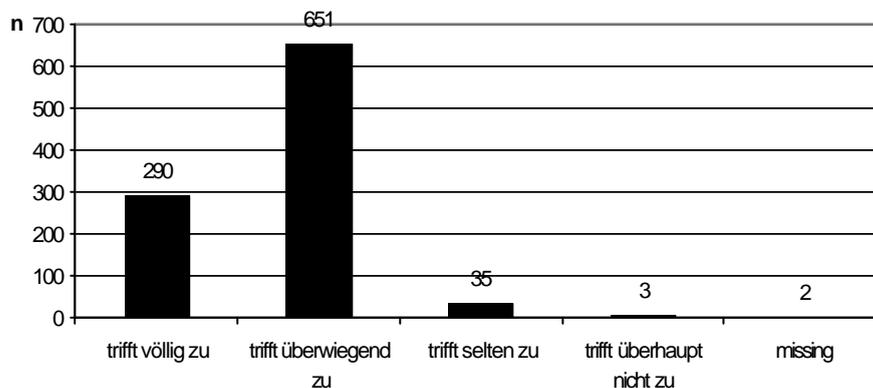


Abbildung 1: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Medien wurden sinnvoll und effektiv eingesetzt" (MEDIEN).

Mit großer Mehrheit (96,1%) sind die befragten Studierenden der Ansicht, die Medien seien überwiegend bzw. völlig sinnvoll und effektiv eingesetzt worden (MEDIEN) (vgl. Abb. 1).

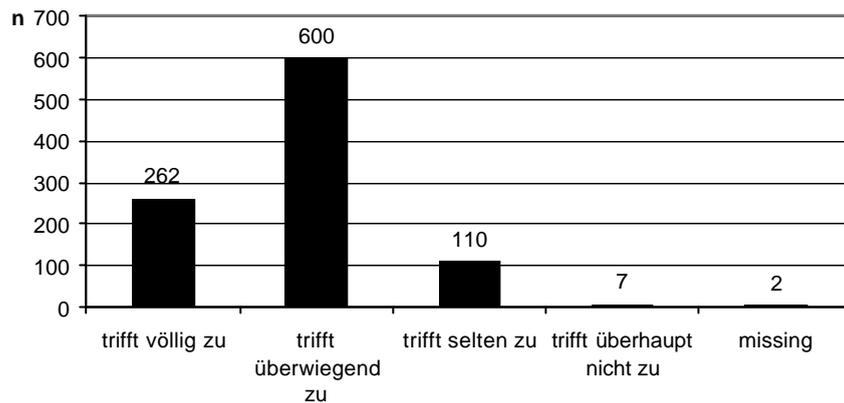


Abbildung 2: Absolute Häufigkeitsverteilung der Bewertungen des Items „Das Thema wurde durch praxisnahe Beispiele veranschaulicht“(THEMA).

88,1 Prozent der befragten Studierenden sind zudem der Ansicht, dass das Thema überwiegend bzw. völlig durch praxisnahe Beispiele veranschaulicht wurde (THEMA) (vgl. Abb. 2).

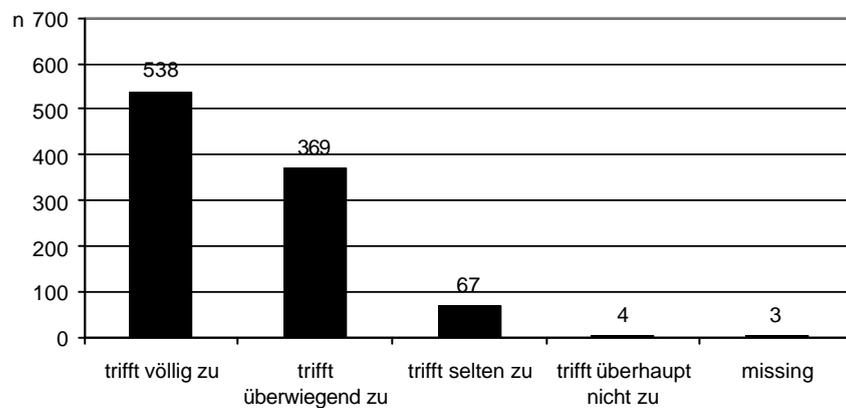


Abbildung 3: Absolute Häufigkeitsverteilung der Bewertungen des Items "Inhaltliche Fragen waren erwünscht" (FRAGEN).

Das inhaltliche Fragen erwünscht waren wird von fast allen Befragten (92,7%) als überwiegend oder völlig zutreffend beschrieben (FRAGEN) (vgl. Abb. 3).

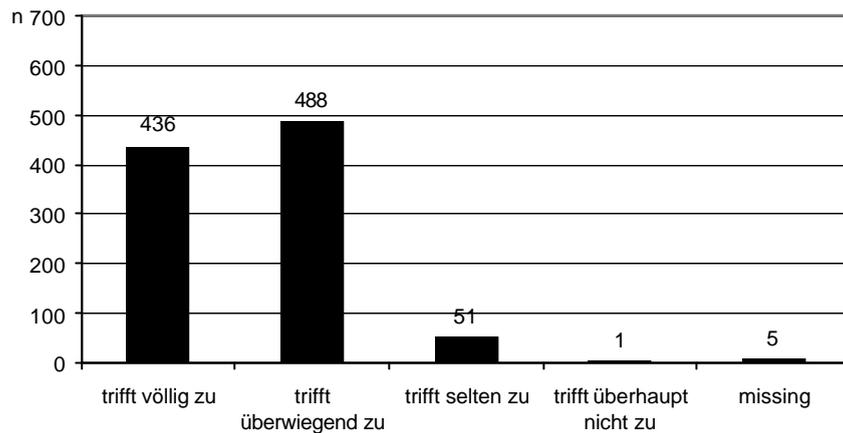


Abbildung 4: Absolute Häufigkeitsverteilung der Bewertungen des Items "Inhaltliche Fragen wurden kompetent und verständlich beantwortet" (KOMPET1).

Ferner sind fast alle Befragten (94,7%) der Ansicht, dass ihre inhaltlichen Fragen überwiegend oder völlig kompetent und verständlich beantwortet wurden (KOMPET1) (vgl. Abb. 4).

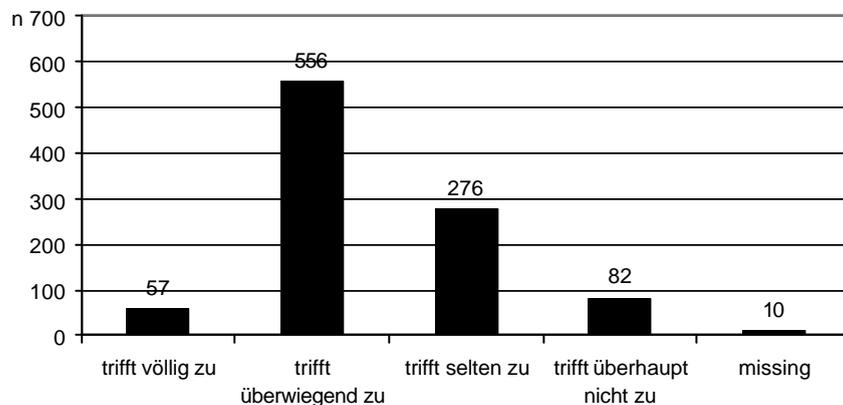


Abbildung 5: Absolute Häufigkeitsverteilung der Bewertungen des Items "Es fiel mir leicht, aufmerksam zu bleiben" (AUFMERK1).

Gegenüber dieser insgesamt sehr positiven Bewertung der Grundlagenseminare fiel es jedoch nur 62,5% der Befragten völlig oder überwiegend leicht, aufmerksam der Veranstaltung zu folgen (AUFMERK1) (vgl. Abb. 5).

### 3.4.1.2 Praktische Übungen

Für die Beurteilung der praktischen Übungen im Rahmen der Lehrveranstaltung wurden die fünf in Tabelle 3 dargestellten Items verwendet.

Tabelle 3: Die Bewertungen der praktischen Übungen: Itemformulierungen und prozentuale Häufigkeitsverteilung

Label	Itemformulierung	1 trifft völlig zu	2 trifft über- wiegend zu	3 trifft selten zu	4 trifft über- haupt nicht zu
BASISWIS	Mein Basiswissen war ausreichend, um den Übungen zu folgen	31,4 %	57,6 %	9,2 %	1,4 %
DIALOG	Der Lernstoff wurde im Dialog erarbeitet	24,5 %	52,7 %	20,6 %	1,3 %
KOMPET2	Inhaltliche Fragen wurde kompetent und verständlich beantwortet	43,8 %	51,5 %	3,5 %	0,5 %
LEHRINHA	Die Lehrinhalte wurden interessant präsentiert	23,8 %	58,6 %	14,0 %	2,3 %
AUFMERK2	Es fiel mir leicht aufmerksam zu bleiben	15,8 %	61,7 %	17,3 %	4,1 %

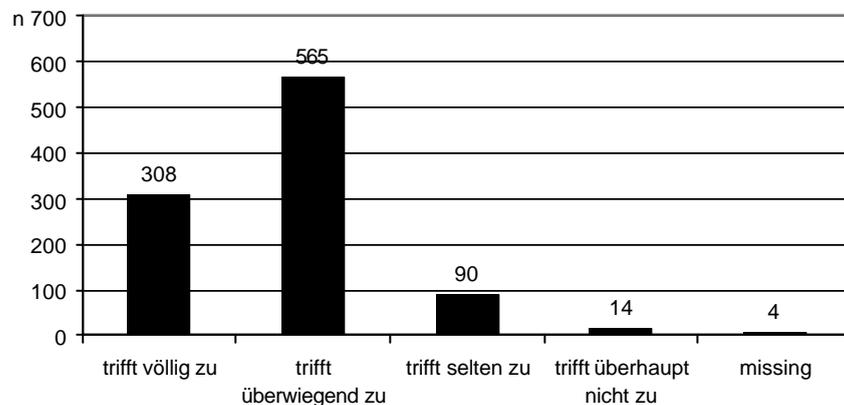


Abbildung 6: Absolute Häufigkeitsverteilung der Bewertungen des Items "Mein Basiswissen war ausreichend, um den Übungen zu folgen" (BASISWIS).

89,0% der Befragten gaben an, dass ihr Basiswissen überwiegend bzw. völlig ausreichend war, um den Übungen zu folgen (BASISWIS) (vgl. Abb. 6).

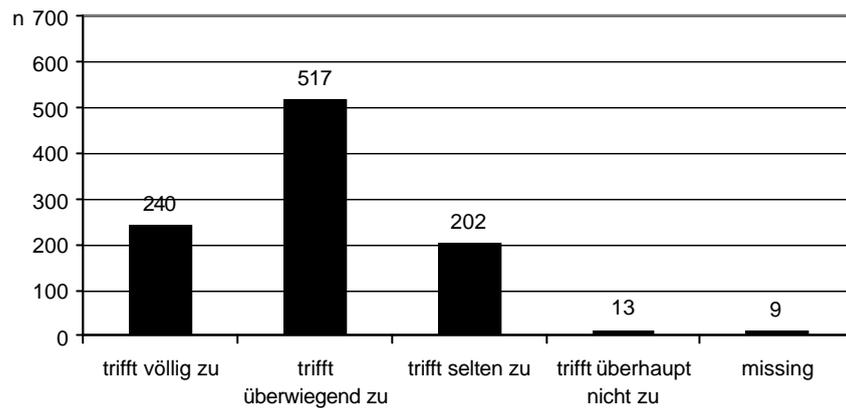


Abbildung 7: Absolute Häufigkeitsverteilung der Bewertungen des Items "Der Lernstoff wurde im Dialog erarbeitet" (DIALOG).

Ebenfalls die überwiegende Mehrheit (77,2%) der Befragten ist der Ansicht, dass der Lernstoff überwiegend bzw. völlig im Dialog erarbeitet wurde (DIALOG) (vgl. Abb. 7).

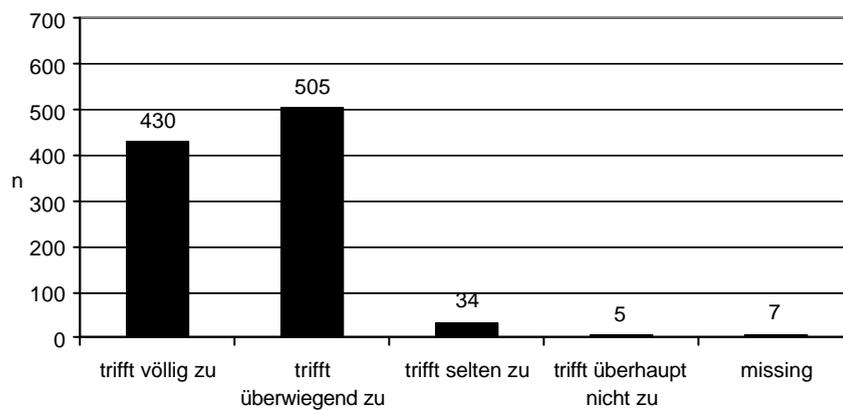


Abbildung 8: Absolute Häufigkeitsverteilung der Bewertungen des Items "Inhaltliche Fragen wurden kompetent und verständlich beantwortet" (KOMPET2).

95,3% der Studierenden sind überwiegend bzw. völlig der Ansicht, dass inhaltliche Fragen kompetent und verständlich beantwortet wurden (KOMPET2) (vgl. Abb. 8).

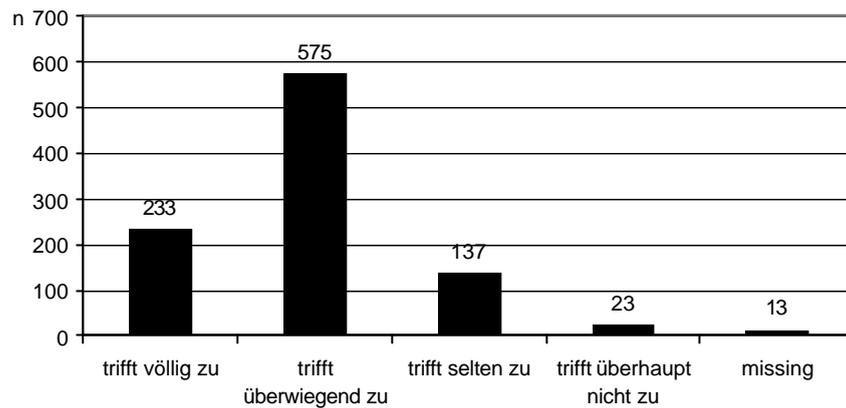


Abbildung 9: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Lehrinhalte wurden interessant präsentiert" (LEHRINHA).

Und auch für die Mehrheit (82,4%) der Befragten trifft es überwiegend, bzw. völlig zu, dass die Lehrinhalte interessant präsentiert wurden (LEHRINHA) (vgl. Abb. 9).

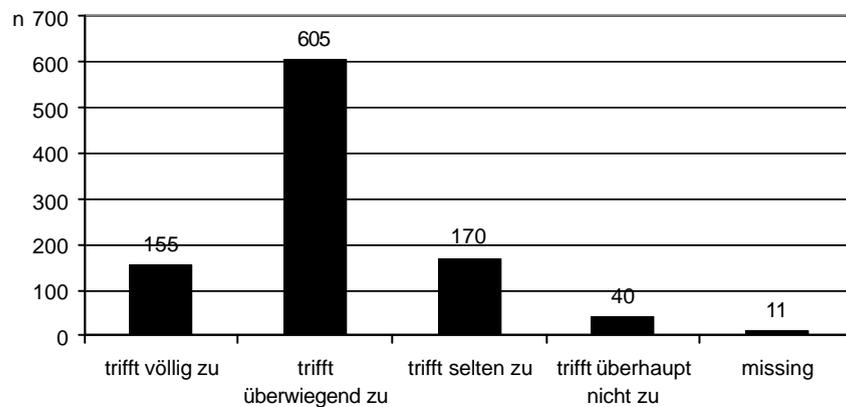


Abbildung 10: Absolute Häufigkeitsverteilung der Bewertungen des Items "Es fiel mir leicht aufmerksam zu bleiben" (AUFMERK2).

77,5% der Studierenden fiel es im Rahmen der praktischen Übungen überwiegend oder völlig leicht, aufmerksam zu bleiben (AUFMERK2) (vgl. Abb. 10). Dieser Anteil ist 15% höher als beim vergleichbaren Item zur Beurteilung der Grundlagenseminare.

### 3.4.1.3 Betriebsbegehung

Die Beurteilung der Betriebsbegehung wurde über die vier in Tabelle 4 dargestellten Items erfasst.

Tabelle 4: Die Bewertungen der Betriebsbegehung: Itemformulierungen und prozentuale Häufigkeitsverteilung

Label	Itemformulierung	1 trifft völlig zu	2 trifft über- wiegend zu	3 trifft selten zu	4 trifft über- haupt nicht zu
INHALT1	Die Betriebsbegehung war inhaltlich gut vorbereitet ...	36,1 %	41,0 %	10,6 %	2,7 %
ORGANIST	... und organisatorisch gut vorbereitet	53,0 %	36,5 %	6,9 %	2,3 %
KOMPET3	Inhaltliche Fragen wurden kompetent und verständlich beantwortet	41,8 %	41,8 %	13,0 %	2,5 %
EINBLICK	Die Betriebsbegehung hat mir einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft	48,0 %	40,0 %	9,1 %	2,3 %

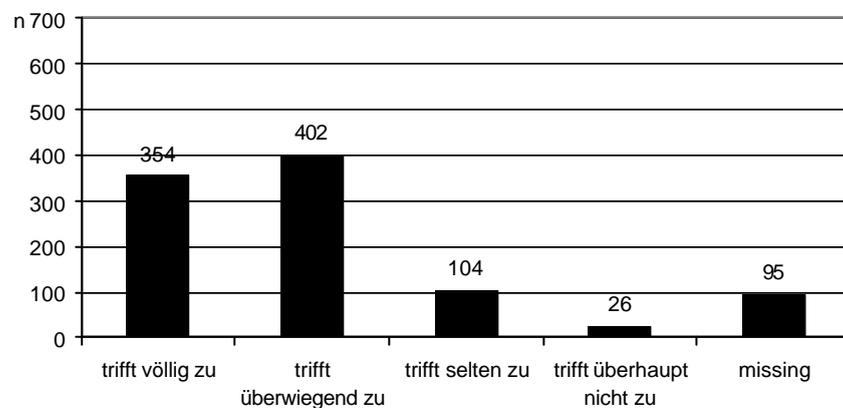


Abbildung 11: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Betriebsbegehung war inhaltlich gut vorbereitet" (INHALT1).

77,1% der Befragten sind überwiegend oder völlig der Ansicht, dass die Betriebsbegehung inhaltlich gut vorbereitet war (INHALT1) (vgl. Abb. 11).

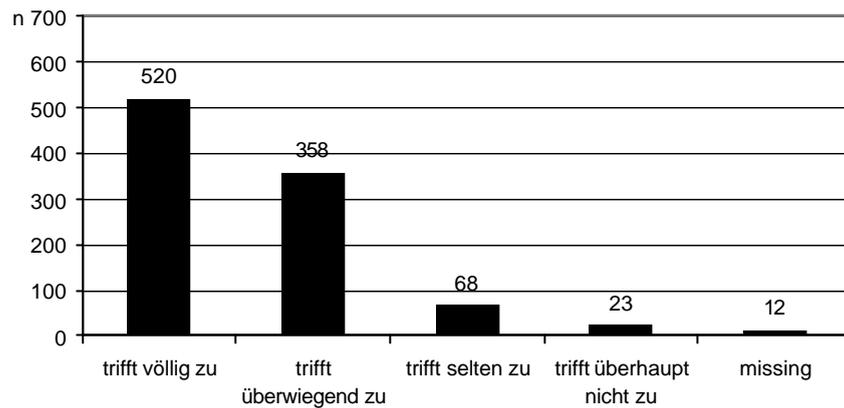


Abbildung 12: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Betriebsbegehung war organisatorisch gut vorbereitet" (ORGANIST).

89,5% der Befragten sind überwiegend oder völlig der Ansicht, dass die Betriebsbegehung organisatorisch gut vorbereitet war (ORGANIST) (vgl. Abb. 12).

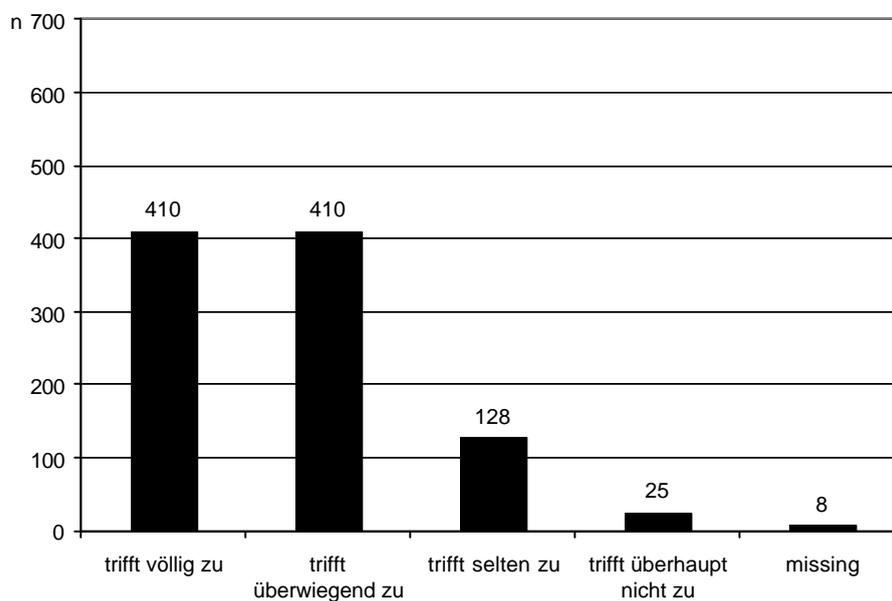


Abbildung 13: Absolute Häufigkeitsverteilung der Bewertungen des Items "Inhaltliche Fragen wurden kompetent und verständlich beantwortet" (KOMPET3).

83,6% der befragten Studierenden meinten, dass inhaltliche Fragen überwiegend oder völlig kompetent und verständlich beantwortet wurden (KOMPET3) (vgl. Abb. 13).

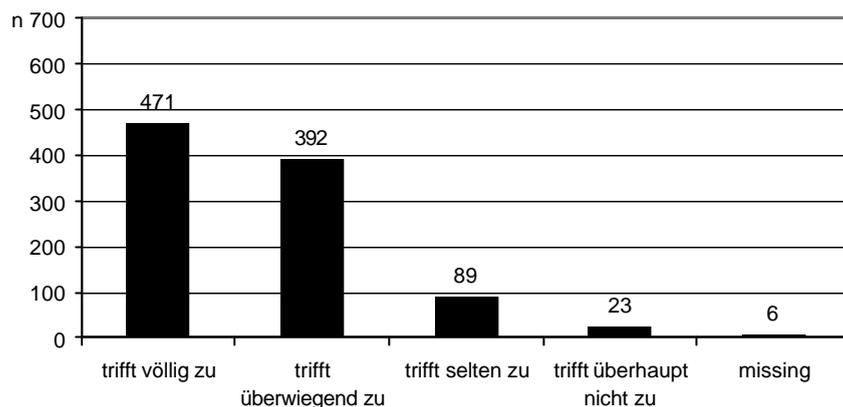


Abbildung 14: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Betriebsbegehung hat mir einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft" (EINBLICK).

Für die überwiegenden Mehrheit (88,0%) der befragten Studierenden hat die Betriebsbegehung nach ihrer Ansicht einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft (EINBLICK) (vgl. Abb. 14).

### 3.4.1.4 Gruppenarbeit

Zur Erfassung der Bewertungen der Gruppenarbeit durch die Studierenden wurden die drei in Tabelle 5 dargestellten Items eingesetzt.

Tabelle 5: Die Bewertungen der Gruppenarbeit: Itemformulierungen und prozentuale Häufigkeitsverteilung

Label	Itemformulierung	1 trifft völlig zu	2 trifft über- wiegend zu	3 trifft selten zu	4 trifft über- haupt nicht zu
DISKURSI	Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes	16,2 %	52,6 %	27,5 %	3,5 %
UNTERSTÜ	Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation	27,3 %	44,3 %	21,7 %	6,4 %
SPAß	Die Gruppenarbeit hat Spaß gemacht	26,3 %	51,6 %	17,1 %	4,8 %

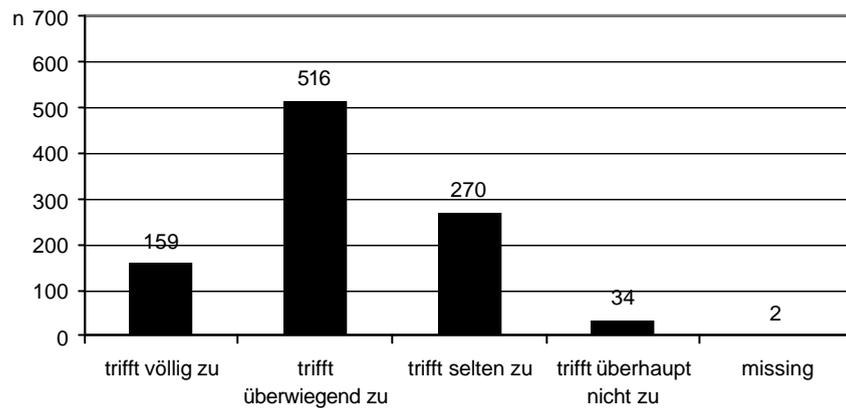


Abbildung 15: Absolute Häufigkeitsverteilung der Bewertungen des Items: "Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes" (DISKURSI).

Bei 68,8% der Befragten hat die Diskussion in der Gruppe völlig oder überwiegend zu einem vertieften Verständnis des Lerninhaltes geführt (DISKURSI) (vgl. Abb. 15).

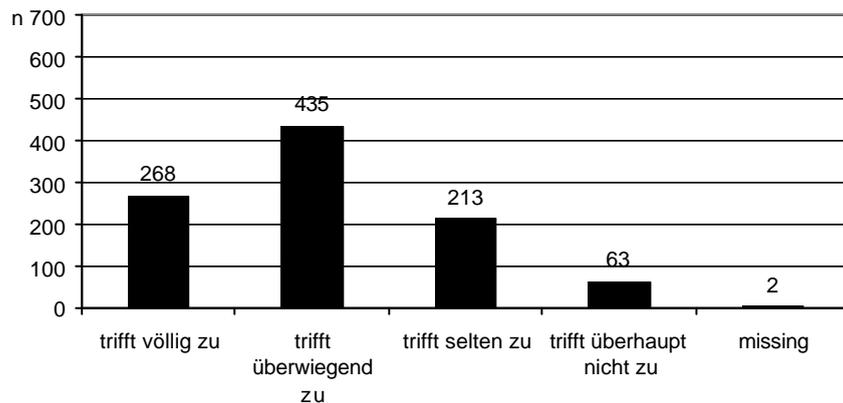


Abbildung 16: Absolute Häufigkeitsverteilung der Bewertungen des Items: "Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation" (UNTERSTÜ).

Die Gruppenarbeit mit Dozentenunterstützung war für 71,6% der Studierenden eine völlig oder überwiegend wichtige Hilfe für die Projektpräsentation (UNTERSTÜ) (vgl. Abb. 16).

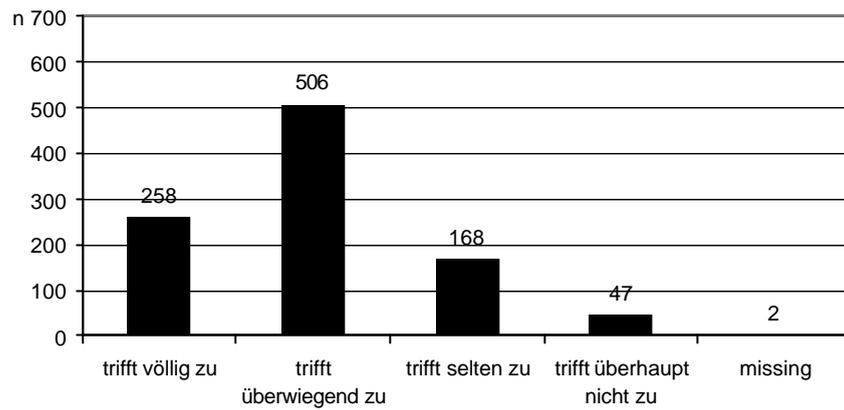


Abbildung 17: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Gruppenarbeit hat Spaß gemacht" (SPAß).

77,9% aller Befragten hat die Gruppenarbeit völlig oder überwiegend Spaß gemacht (SPAß) (vgl. Abb. 17).

### 3.4.1.5 Gesamtbewertung

Abschließend wurden die Studierenden in acht Items nach ihrer Gesamteinschätzung der Veranstaltung befragt. Die entsprechenden Itemformulierungen und prozentuale Häufigkeitsverteilungen sind in Tabelle 6 dargestellt.

Tabelle 6: Die Gesamtbewertung: Itemformulierungen und prozentuale Häufigkeitsverteilung

Label	Itemformulierung	1 trifft völlig zu	2 trifft über- wiegend zu	3 trifft selten zu	4 trifft über- haupt nicht zu
STRUKTUR	Die LV war insgesamt gut und systematisch strukturiert	33,5 %	58,5 %	7,0 %	0,7 %
VERNITTL	Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen	30,0 %	54,0 %	13,1 %	2,4 %
KENNTNIS	Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind	42,8 %	38,4 %	10,4 %	1,4 %
LERNERFO	Mein persönlicher Lernerfolg durch die Veranstaltung war groß	14,5 %	59,8 %	22,4 %	2,3 %
KOLLOQUI	Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen	59,2 %	34,5 %	4,2 %	1,3 %
VORBEREI	Die Lehrpersonen waren gut vorbereitet	66,1 %	31,6 %	1,5 %	0,2 %
FREIWILL	Ich hätte an der Veranstaltung auch freiwillig teilgenommen	6,9 %	38,6 %	39,7 %	13,8 %
INTERESS	Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt	11,2 %	39,0 %	37,0 %	11,6 %

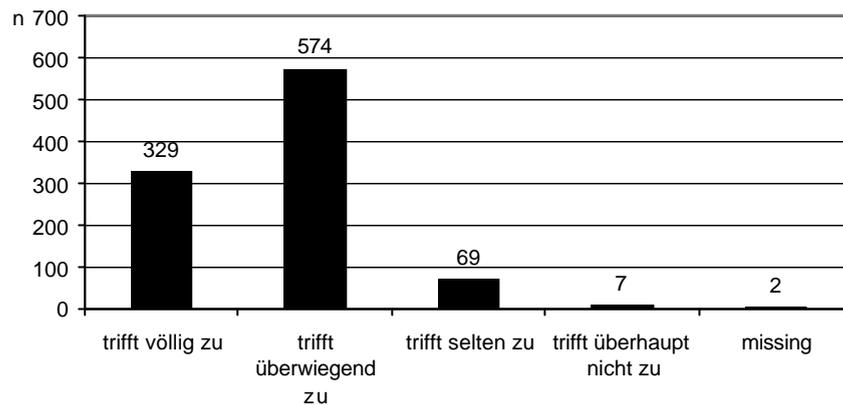


Abbildung 18: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert" (STRUKTUR).

92,0% der Befragten sind völlig oder überwiegend der Ansicht, die Lehrveranstaltung sei gut und systematisch strukturiert gewesen (STRUKTUR) (vgl. Abb. 18).

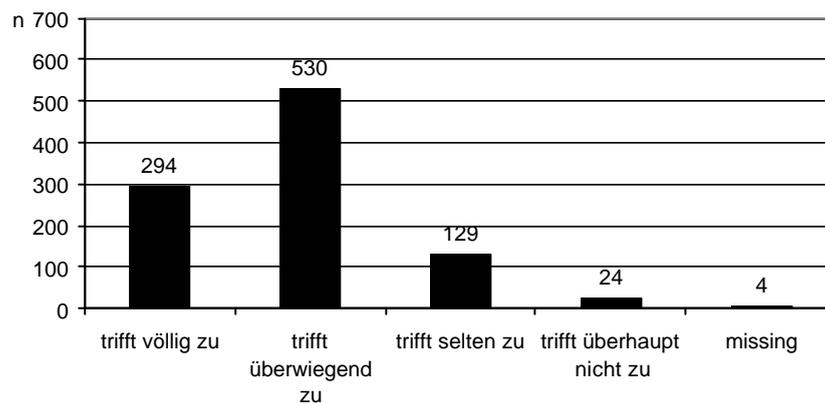


Abbildung 19: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen" (INHALT2).

Ebenfalls der überwiegenden Mehrheit der Befragten (84%) hat die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit völlig oder überwiegend gut gefallen (INHALT2) (vgl. Abb. 19).

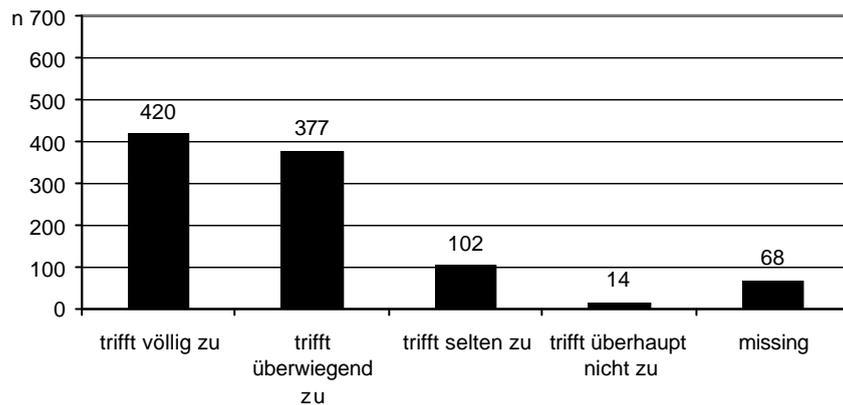


Abbildung 20: Absolute Häufigkeitsverteilung der Bewertungen des Items "Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Arbeit wichtig sind" (KENNTNIS).

81,2% der befragten Studierenden sind völlig oder überwiegend der Ansicht, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für ihre spätere ärztliche Tätigkeit wichtig sind (KENNTNIS) (vgl. Abb. 20). Jedoch beantworteten fast 7 Prozent der Befragten diese Frage nicht.

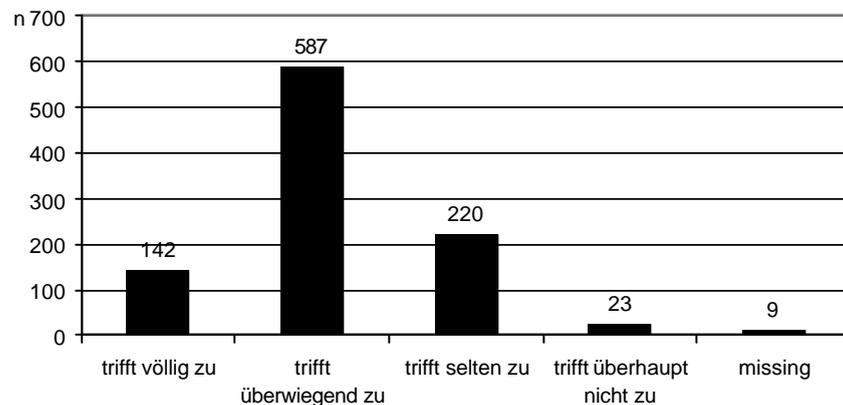


Abbildung 21: Absolute Häufigkeitsverteilung der Bewertungen des Items "Mein persönlicher Lernerfolg durch die Veranstaltung war groß" (LERNERFO).

Einen großen persönlichen Lernerfolg haben 74,3% der Befragten völlig oder überwiegend festgestellt (LERNERFO) (vgl. Abb. 21).

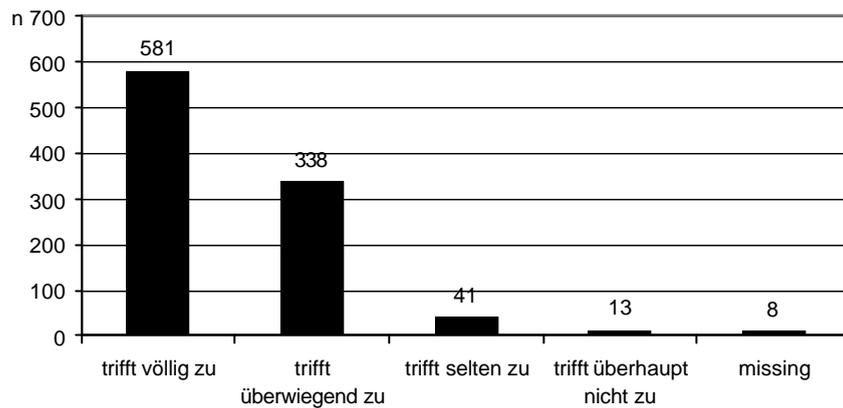


Abbildung 22: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Überprüfung des Lernerfolges durch das Abschlusskolloquium war dem Lernziel angemessen" (KOLLOQUI).

Fast alle befragten Studierenden (93,7%) fanden die Überprüfung des Lernerfolges durch das Abschlusskolloquium völlig oder überwiegend dem Lernziel angemessen (KOLLOQUI) (vgl. Abb. 22).

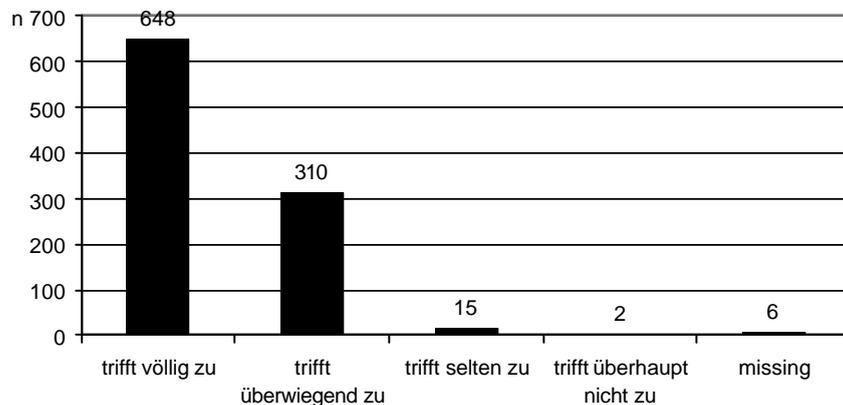


Abbildung 23: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Lehrpersonen waren gut vorbereitet" (VORBEREI).

Ebenso meinten fast alle Befragten (97,7%), dass die Lehrpersonen völlig oder überwiegend gut vorbereitet waren (VORBEREI) (vgl. Abb. 23).

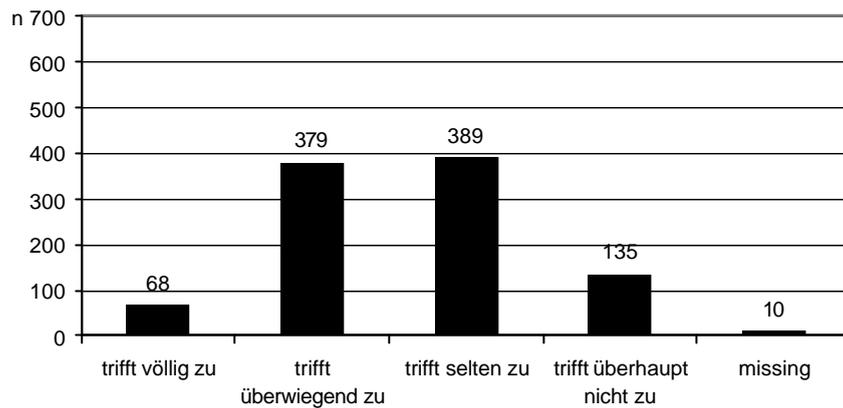


Abbildung 24: Absolute Häufigkeitsverteilung der Bewertungen des Items "Ich hätte an der Veranstaltung auch freiwillig teilgenommen" (FREIWILL).

Dennoch hätte nicht einmal die Hälfte der Studierenden (45,5%) völlig oder überwiegend freiwillig an der Veranstaltung teilgenommen (FREIWILL) (vgl. Abb. 24).

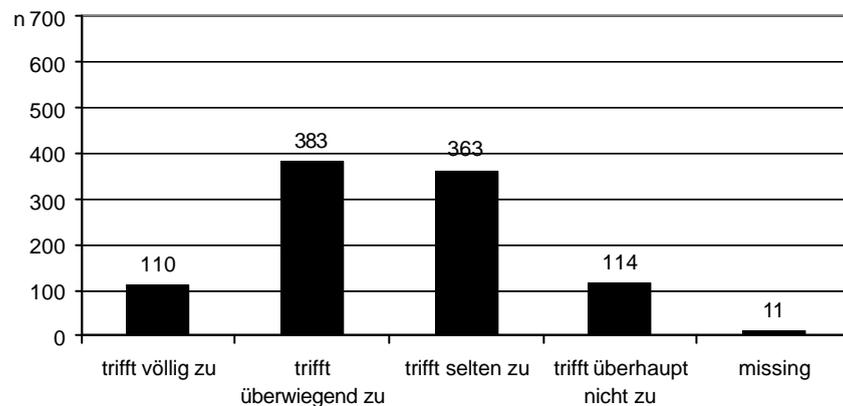


Abbildung 25: Absolute Häufigkeitsverteilung der Bewertungen des Items "Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt" (INTERES).

Und nur bei 50,2% hat die Veranstaltung völlig oder überwiegend ein Interesse am Fach Arbeitsmedizin geweckt (INTERES) (vgl. Abb. 25).

### 3.4.2 Mittelwerte

In diesem Abschnitt werden die deskriptiven Verteilungsparameter zur zusammenfassenden Beschreibung der Häufigkeitsverteilungen für alle Items der einzelnen Skalen vorgestellt. Die Bewertungen werden nach der Skalierung: „trifft völlig zu“, „trifft überwiegend zu“, „trifft selten zu“ und „trifft überhaupt nicht zu“ vorgenommen, die mit Bewertungsziffern von „eins“ bis „vier“ unterlegt sind.

#### 3.4.2.1 Grundlagenseminar

In Tabelle 7 sind die deskriptiven Verteilungsparameter für die Items zur Beurteilung der Grundlagenseminare dargestellt. Insgesamt sind die vorhandenen Mittelwertsunterschiede nur relativ und schmälern nicht die durchweg positive Beurteilung der Grundlagenseminare durch die Studierenden. Am positivsten wird an den Grundlagenseminaren der Umgang mit inhaltlichen Fragen beurteilt (FRAGEN). Am schlechtesten beurteilen die Befragten ihre Aufmerksamkeit in diesen Seminaren (AUFMERK1).

Tabelle 7: Die Beurteilung der Grundlagenseminare:

Deskriptive Verteilungsparameter

	x	s	M	Missing
MEDIEN	1,75	0,53	2	0,2 %
THEMA	1,86	0,62	2	0,2 %
FRAGEN	1,53	0,64	1	0,3 %
KOMPET1	1,61	0,59	2	0,5 %
AUFMERK1	2,39	0,73	2	1,0 %

x = arithmetischer Mittelwert; s = Standardabweichung, M = Median

#### 3.4.2.2 Praktische Übungen

In Tabelle 8 sind die deskriptiven Verteilungsparameter für die Items zur Beurteilung der praktischen Übungen dargestellt. Wiederum sind die arithmetischen Mittelwertsunterschiede nur relativ und stellen die insgesamt positive Beurteilung der praktischen Übungen durch die Befragten nicht in Frage. Am positivsten an den praktischen Übungen finden die Studierenden, dass der Lernstoff im Dialog erarbeitet wurde (DIALOG). Am negativsten bewerten sie auch hier ihre eigene Aufmerksamkeit (AUFMERK2).

Tabelle 8: Die Beurteilung der praktischen Übungen:

Deskriptive Verteilungsparameter

	x	s	M	Missing
BASISWIS	1,81	0,65	2	0,4 %
DIALOG	1,59	0,71	2	0,9 %
KOMPET2	1,60	0,58	2	0,7 %
LEHRINHA	1,95	0,69	2	1,3 %
AUFMERK2	2,10	0,70	2	1,1 %
x = arithmetischer Mittelwert; s = Standardabweichung, M = Median				

### 3.4.2.3 Betriebsbegehung

Die deskriptiven Verteilungsparameter der Items zur Beurteilung der Betriebsbegehung sind in Tabelle 9 dargestellt. Auch hier sind die Unterschiede in den arithmetischen Mittelwerten nur marginaler Natur und beeinträchtigen die insgesamt positive Beurteilung der Betriebsbegehung durch die Studierenden nicht. Am positivsten an der Betriebsbegehung empfanden die befragten Studierenden die organisatorische Vorbereitung (ORGANIST). Am negativsten beurteilen sie die inhaltliche Vorbereitung (INHALT1).

Tabelle 9: Die Beurteilung der Betriebsbegehung:

Deskriptive Verteilungsparameter

	x	s	M	Missing
INHALT1	1,78	0,76	2	9,7 %
ORGANIST	1,58	0,73	1	1,2 %
KOMPET3	1,76	0,77	2	0,8 %
EINBLICK	1,66	0,74	2	0,6 %
x = arithmetischer Mittelwert; s = Standardabweichung, M = Median				

### 3.4.2.4 Gruppenarbeit

Die Zusammenfassung der Beurteilung der Gruppenarbeit über deskriptive Verteilungsparameter findet sich in Tabelle 10. Hier liegen die arithmetischen Mittelwerte der drei Items sehr nah beieinander.

Relativ am positivsten wird hier der Spaß an der Gruppenarbeit (SPAß) und am negativsten die Diskussion in der Gruppe zur Förderung eines tieferen Verständnisses der Lerninhalte (DISKURSI) gesehen.

Tabelle 10: Die Beurteilung der Gruppenarbeit:

Deskriptive Verteilungsparameter

	x	s	M	Missing
DISKURSI	2,18	0,74	2	0,2 %
UNTERSTÜ	2,07	0,86	2	0,2 %
SPAß	2,00	0,79	2	0,2 %
x = arithmetischer Mittelwert; s = Standardabweichung, M = Median				

### 3.4.2.5 Gesamtbewertung

Die deskriptiven Verteilungsparameter für die Items zur Gesamtbewertung der Lehrveranstaltung sind in Tabelle 11 zusammengefasst. Hier zeigen sich trotz der insgesamt positiven Gesamtbeurteilung deutliche Unterschiede zwischen den verschiedenen Bewertungsaspekten. Am positivsten wird die Vorbereitung der Lehrpersonen (VORBEREI) und die Überprüfung des Lernerfolges durch das Abschlusskolloquium (KOLLOQUI) beurteilt. Als weniger positiv wird die Freiwilligkeit zur Teilnahme an der Lehrveranstaltung (FREIWILL) bewertet sowie das Wecken des Interesses am Fach Arbeitsmedizin (INTERESS).

Tabelle 11: Die Gesamtbewertung der Lehrveranstaltung:

Deskriptive Verteilungsparameter

	x	s	M	Missing
STRUKTUR	1,75	0,61	2	0,2 %
INHALT2	1,88	0,76	2	0,4 %
KENNTNIS	1,68	0,73	2	6,9 %
LERNERFO	2,13	0,67	2	0,9 %
KOLLOQUI	1,47	0,64	1	0,8 %
VORBEREI	1,35	0,52	1	0,6 %
FREIWILL	2,61	0,81	3	1,0 %
INTERESS	2,50	0,84	2	1,1 %

x = arithmetischer Mittelwert; s = Standardabweichung, M = Median

### **3.4.3 Einfluss der Bewertung der Gruppenarbeit auf die Gesamtbewertung der Lehrveranstaltung**

In diesem Abschnitt wird untersucht, welche Auswirkungen die Bewertung der Gruppenarbeit auf die Beurteilung der Grundlagenseminare, der praktischen Übungen, der Betriebsbegehung und auf die Gesamtbeurteilung der Lehrveranstaltung haben.

Die Gruppenarbeit ist ein zentrales didaktisches Element, das von den Studierenden unterschiedlich bewertet und damit auch zu unterschiedlichen Bewertungen der Lehrveranstaltung insgesamt führen kann. So gibt es Studierende, die frontale Unterrichtsformen favorisieren, andere hingegen erfahren durch Gruppenarbeit größere Motivation und erzielen Lernerfolge. Deshalb werden im Folgenden die Gruppe der Studierenden, die die Gruppenarbeit positiv und diejenigen, die sie negativ bewertet haben hinsichtlich ihrer Beurteilung der Lehrveranstaltung anhand der hier verwendeten Items miteinander verglichen.

Bei 68,8% der Befragten hat die Diskussion in der Gruppe völlig oder überwiegend zu einem vertieften Verständnis des Lerninhaltes geführt (DISKURSI). Die Gruppenarbeit mit Dozentenunterstützung war für 71,6% der Studierenden eine völlig oder überwiegend wichtige Hilfe für die Projektpräsentation (UNTERSTÜ). 77,9% aller Befragten hat die Gruppenarbeit völlig oder überwiegend Spaß gemacht (SPAß).

Der überwiegenden Mehrheit der Befragten (84%) hat die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit völlig oder überwiegend gut gefallen.

Die drei Items zur Bewertung der Gruppenarbeit wurden zunächst unter Verwendung der Regressionsmethode zu einer Index-Variablen zusammengefasst. Anschließend wurden durch Median-Splitt zwei Gruppen gebildet: Studierende mit numerisch niedrigen Werte in dieser Index-Variablen wurden der Gruppe zugeordnet, die die Gruppenarbeit positiv bewerten. Studierende mit numerisch hohen Werten wurden entsprechend der Gruppe zugeordnet, die die Gruppenarbeit negativ bewerten. Damit wurde die kritische Grenze, die zwischen positiven und negativen Bewertungen der Gruppenarbeit differenziert, relativ zur Häufigkeitsverteilung und nicht absolut definiert. Dadurch sind beide Gruppen gleich groß. In den folgenden Abbildungen sind die Häufigkeitsverteilungen für sämtliche Bewertungsitems getrennt nach diesen beiden Gruppen dargestellt.

### 3.4.3.1 Grundlagenseminar

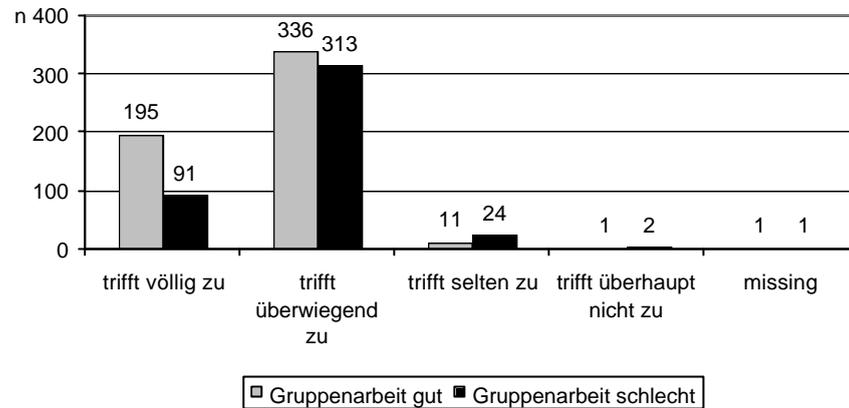


Abbildung 26: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Medien wurden sinnvoll und effektiv eingesetzt“ (MEDIEN) in Abhängigkeit von der Bewertung der Gruppenarbeit

Die Befragten, die die Gruppenarbeit positiv einschätzen, sind auch eher der Meinung, dass die Medien im Grundlagenseminar sinnvoll und effektiv eingesetzt wurden (MEDIEN) (vgl. Abb. 26).

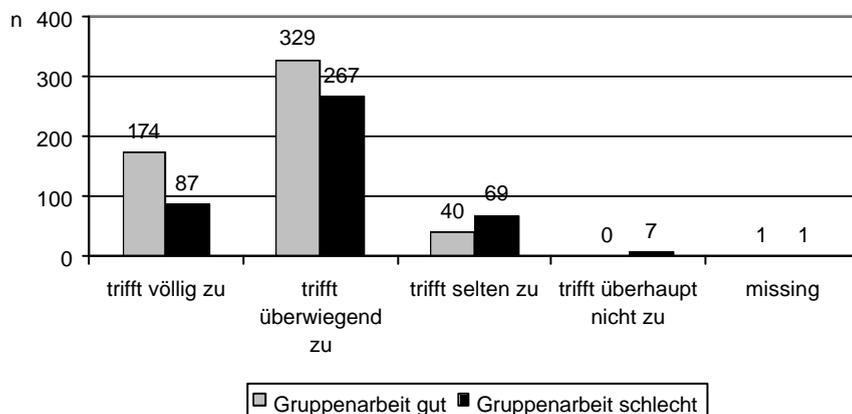


Abbildung 27: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Das Thema wurde durch praxisnahe Beispiele veranschaulicht“ (THEMA) in Abhängigkeit von der Bewertung der Gruppenarbeit

Diejenigen, die die Gruppenarbeit positiv bewerten, sind eher der Ansicht, dass im Grundlagenseminar das Thema durch praxisnahe Beispiele veranschaulicht wurde (THEMA) (vgl. Abb. 27).

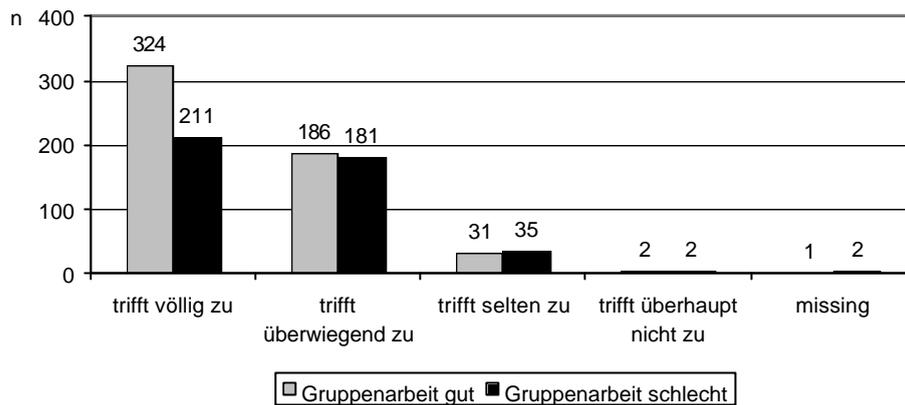


Abbildung 28: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen waren erwünscht“ (FRAGEN) in Abhängigkeit von der Bewertung der Gruppenarbeit

Wird die Gruppenarbeit positiv bewertet, sind die Befragten auch eher der Ansicht, dass inhaltliche Fragen im Grundlagenseminar erwünscht waren (FRAGEN) (vgl. Abb. 28).

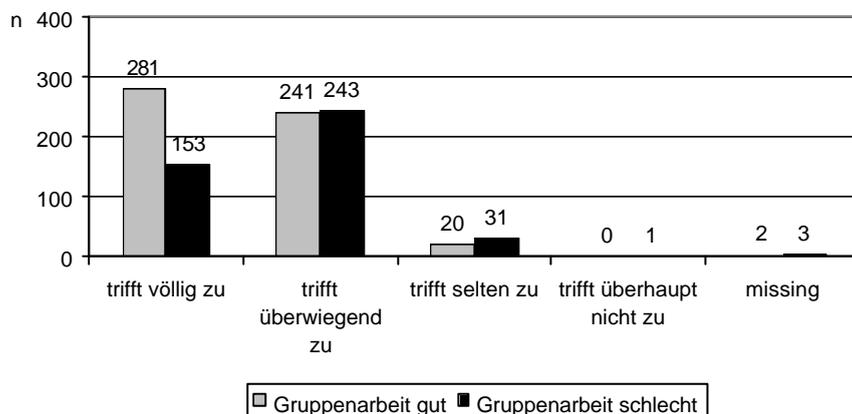


Abbildung 29: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET1) in Abhängigkeit von der Bewertung der Gruppenarbeit

Die Befragten sind eher der Ansicht, dass inhaltliche Fragen kompetent und verständlich beantwortet wurden (KOMPET1), wenn sie auch die Gruppenarbeit positiv bewerten (vgl. Abb. 29).

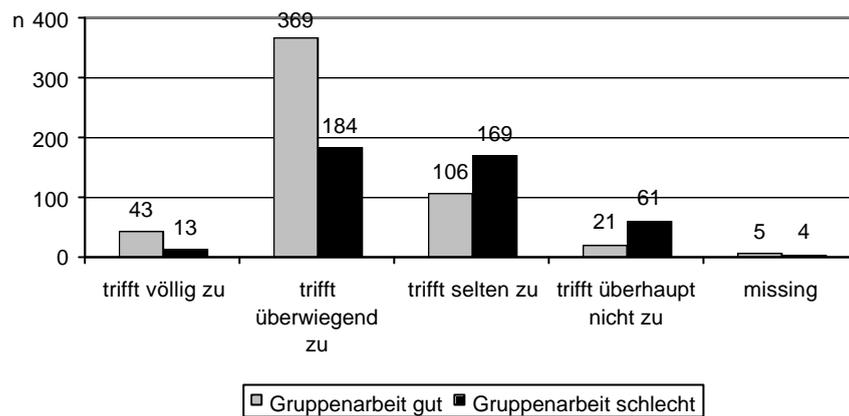


Abbildung 30: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Es fiel mir leicht, aufmerksam zu bleiben“ (AUFMERK1) in Abhängigkeit von der Bewertung der Gruppenarbeit

Die Aufmerksamkeit der Studierenden (AUFMERK1) ist deskriptiv deutlich höher, wenn sie die Gruppenarbeit positiv bewerten (vgl. Abb. 30).

### 3.4.3.2 Praktische Übungen

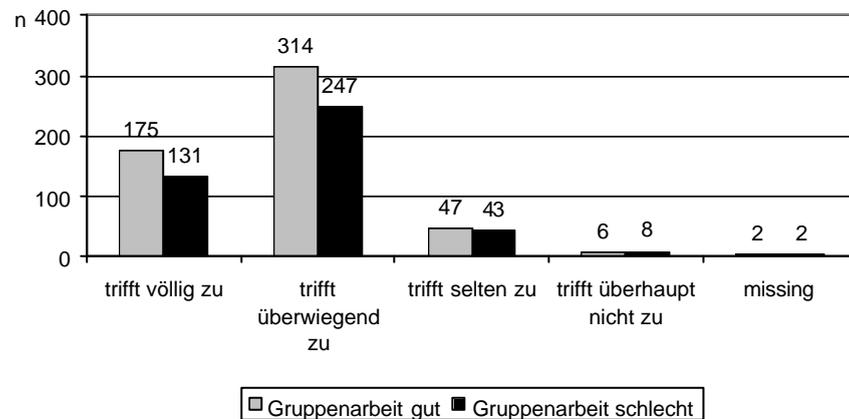


Abbildung 31: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Mein Basiswissen war ausreichend, um den Übungen zu folgen“ (BASISWIS) in Abhängigkeit von der Bewertung der Gruppenarbeit

Wenn sie die Gruppenarbeit positiv bewerten, dann sind die Studierenden auch häufiger der Ansicht, dass ihr Basiswissen ausreichend war, um den Übungen zu folgen (BASISWIS) (vgl. Abb. 31).

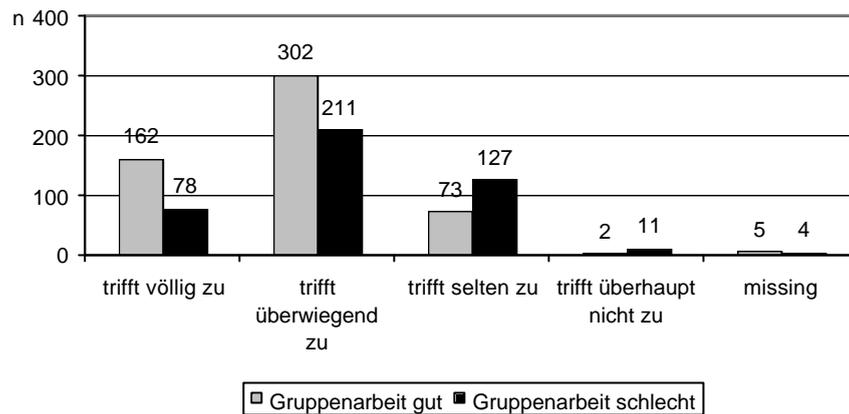


Abbildung 32: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Der Lernstoff wurde im Dialog erarbeitet“ (DIALOG) in Abhängigkeit von der Bewertung der Gruppenarbeit

Je eher die Studierenden der Ansicht zustimmen, dass der Lernstoff in den praktischen Übungen im Dialog erarbeitet wurde (DIALOG), desto eher beurteilen sie auch die Gruppenarbeit positiv (vgl. Abb. 32).

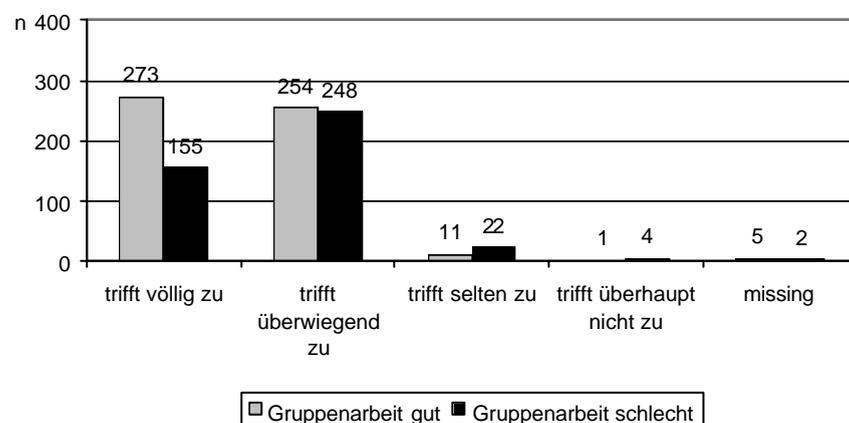


Abbildung 33: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET2) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv beurteilen, sind eher der Ansicht, dass inhaltliche Fragen kompetent und verständlich beantwortet wurden (KOMPET2) (vgl. Abb. 33).

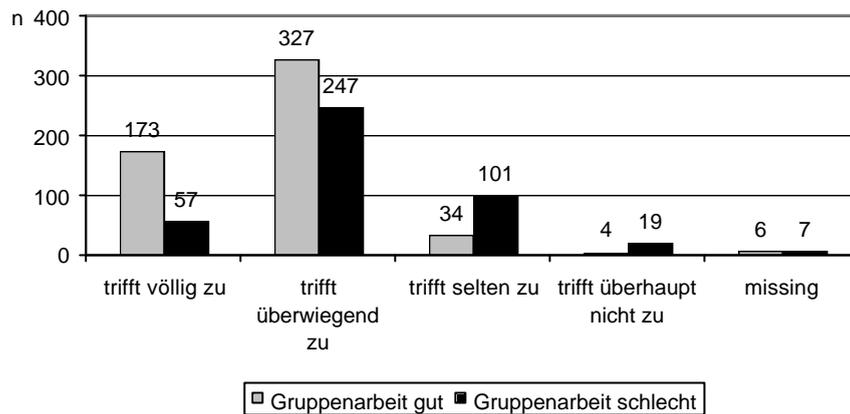


Abbildung 34: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrinhalte wurden interessant präsentiert“ (LEHRINHA) in Abhängigkeit von der Bewertung der Gruppenarbeit

Wenn Studierende die Gruppenarbeit positiv bewerten, dann sind sie auch eher der Ansicht, dass die Lehrinhalte interessant präsentiert wurden (LEHRINHA) (vgl. Abb. 34).

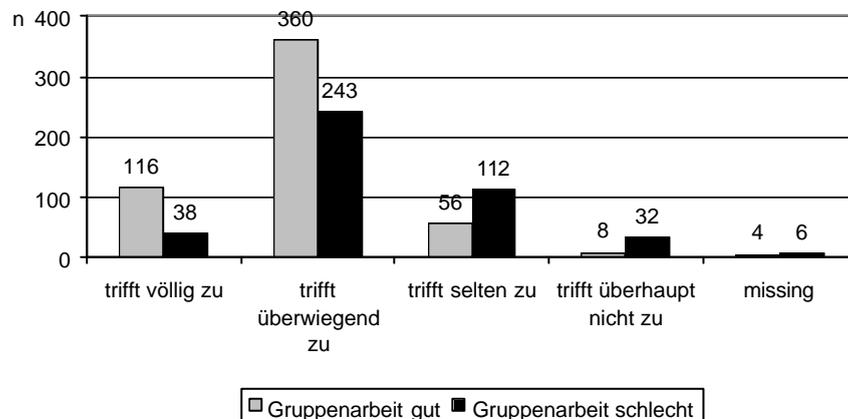


Abbildung 35: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Es fiel mir leicht aufmerksam zu bleiben“ (AUFMERK2) in Abhängigkeit von der Bewertung der Gruppenarbeit

Beurteilen Studierende die Gruppenarbeit positiv, dann fällt es ihnen in den praktischen Übungen leichter, aufmerksam zu bleiben (AUFMERK2) (vgl. Abb. 35).

### 3.4.3.3 Betriebsbegehung

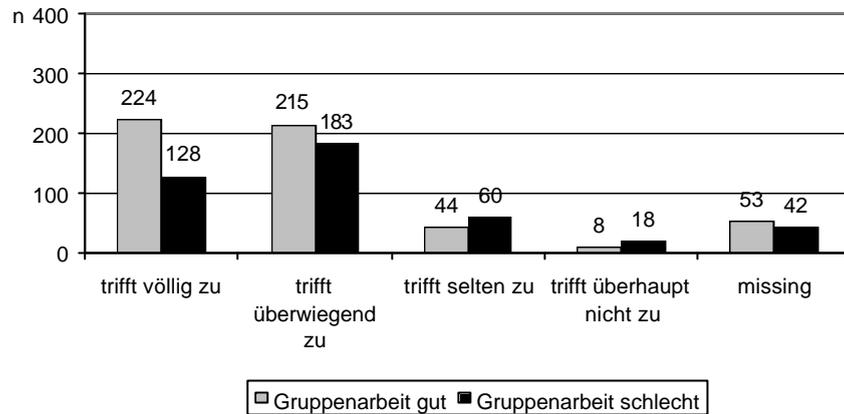


Abbildung 36: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung war inhaltlich gut vorbereitet.“ (INHALT1) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv bewerten, sind eher der Ansicht, dass die Betriebsbegehung inhaltlich gut vorbereitet war (INHALT1) (vgl. Abb. 36).

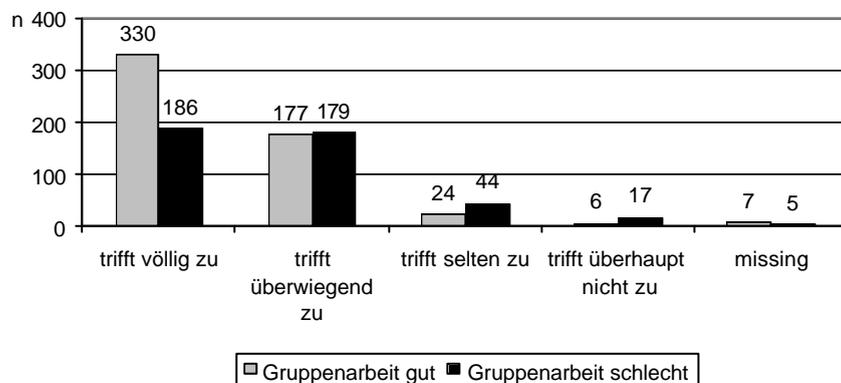


Abbildung 37: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung war organisatorisch gut vorbereitet.“ (ORGANIST) in Abhängigkeit von der Bewertung der Gruppenarbeit

Ebenso sind Studierende, die die Gruppenarbeit positiv bewerten, eher der Ansicht, dass die Betriebsbegehung organisatorisch gut vorbereitet war (ORGANIST)(vgl. Abb. 37).

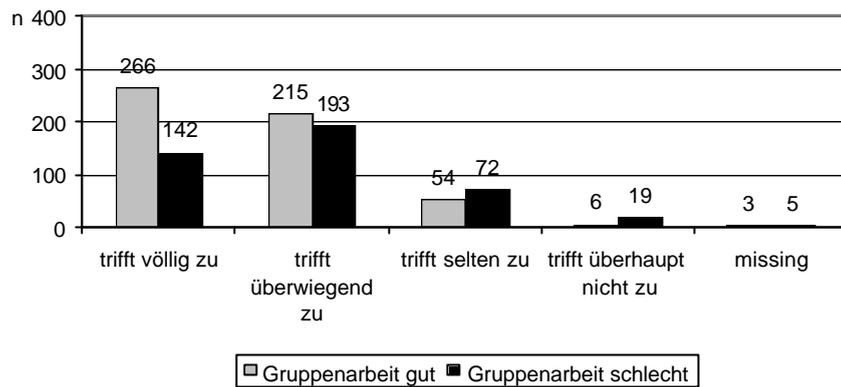


Abbildung 38: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET3) in Abhängigkeit von der Bewertung der Gruppenarbeit

Auch sind Studierende, die die Gruppenarbeit positiv beurteilen, eher der Ansicht, dass inhaltliche Fragen im Rahmen der Betriebsbegehung kompetent und verständlich beantwortet wurden (KOMPET3) (vgl. Abb. 38).

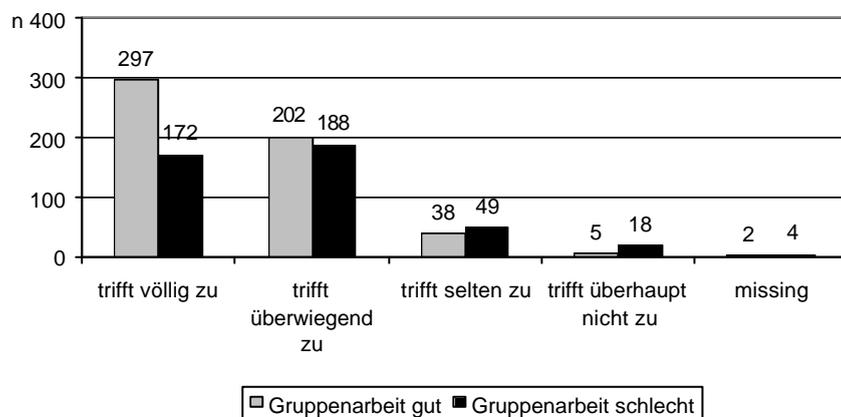


Abbildung 39: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung hat mir einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft“ (EINBLICK) in Abhängigkeit von der Bewertung der Gruppenarbeit

Fast plausibel erscheint es dann, dass Studierende, die die Gruppenarbeit positiv bewerten, auch eher der Ansicht sind, dass die Betriebsbegehung ihnen einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft hat (EINBLICK) (vgl. Abb. 39).

### 3.4.3.4 Gruppenarbeit

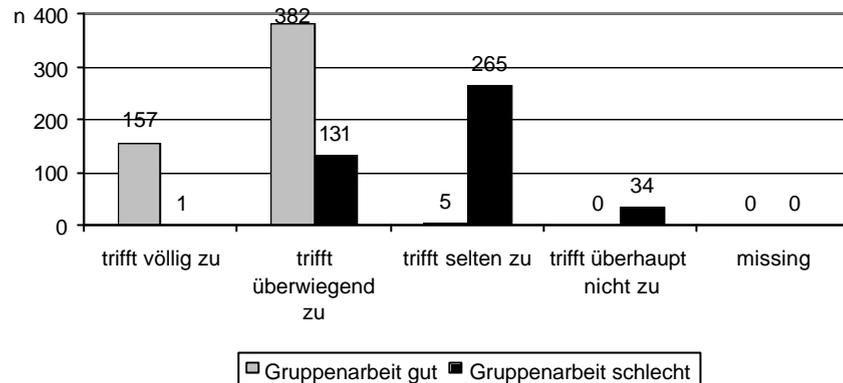


Abbildung 40: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes“ (DISKURSI) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv beurteilen, sind eher der Ansicht, dass die Diskussion in der Gruppe zu einem vertieften Verständnis des Lerninhaltes geführt habe (DISKURSI) (vgl. Abb. 40).

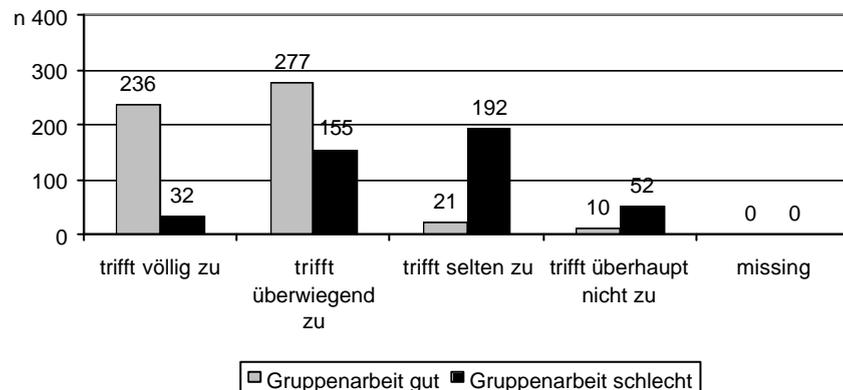


Abbildung 41: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation“ (UNTERSTÜ) in Abhängigkeit von der Bewertung der Gruppenarbeit

Ebenso sind Studierende, die die Gruppenarbeit positiv bewerten, auch eher der Ansicht, dass die Gruppenarbeit mit Dozentenunterstützung eine wichtige Hilfe für die Projektpräsentation war (UNTERSTÜ) (vgl. Abb. 41).

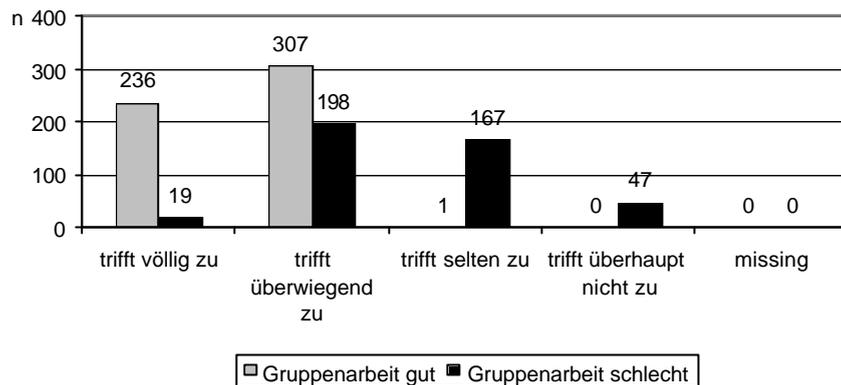


Abbildung 42: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Gruppenarbeit hat mir Spaß gemacht“ (SPAß) in Abhängigkeit von der Bewertung der Gruppenarbeit

Diejenigen Befragten, denen die Gruppenarbeit Spaß gemacht hat, bewerten auch die Gruppenarbeit insgesamt eher als positiv als Studierende, denen die Gruppenarbeit keinen Spaß gemacht hat (SPAß) (vgl. Abb. 42).

### 3.4.3.5 Gesamtbewertung

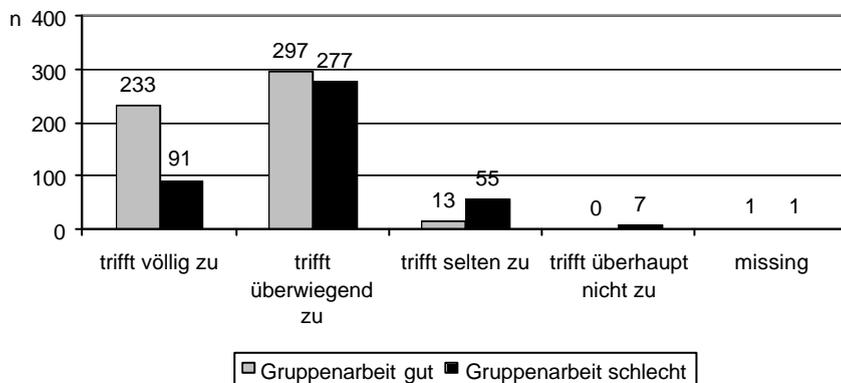


Abbildung 43: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert“ (STRUKTUR) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv bewerten, sind eher der Ansicht, dass die Lehrveranstaltung insgesamt gut und systematisch strukturiert war (STRUKTUR) (vgl. Abb. 43).

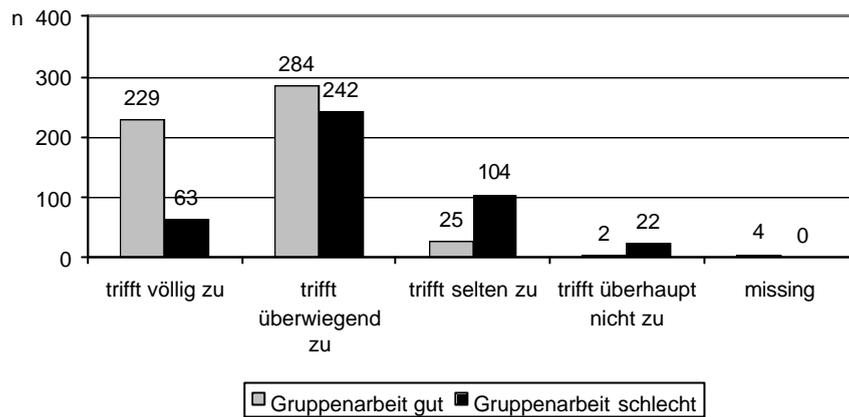


Abbildung 44: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen“ (INHALT2) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv bewerten, sind eher der Ansicht, dass die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit ihnen gut gefallen habe (INHALT2) (vgl. Abb. 44).

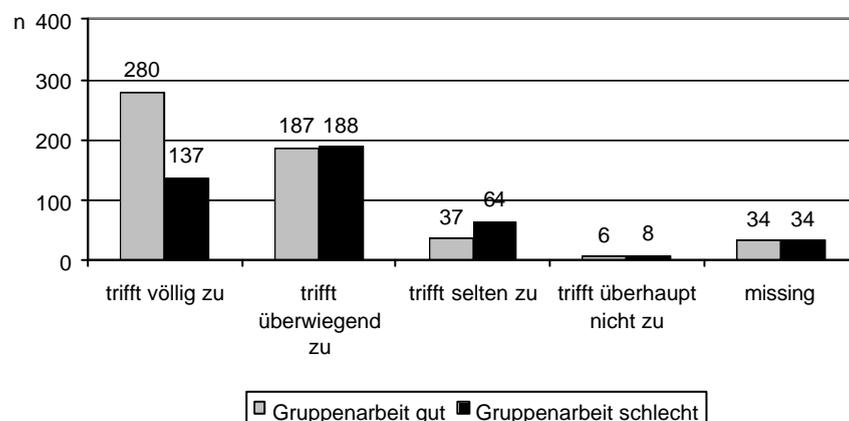


Abbildung 45: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Arbeit wichtig sind“ (KENNTNIS) in Abhängigkeit von der Bewertung der Gruppenarbeit

Ebenso sind Studierende, die die Gruppenarbeit positiv beurteilen, auch eher der Ansicht, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für ihre spätere ärztliche Arbeit wichtig seien (KENNTNIS) (vgl. Abb. 45).

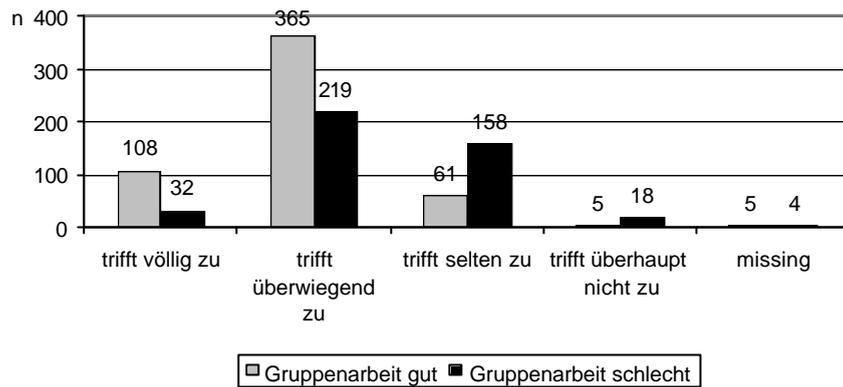


Abbildung 46: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Mein persönlicher Lernerfolg durch die Veranstaltung war groß“ (LERNERFO) in Abhängigkeit von der Bewertung der Gruppenarbeit

Der persönliche Lernerfolg durch die Veranstaltung (LERNERFO) wird von Studierenden, die die Gruppenarbeit positiv bewerten, höher eingeschätzt als von Studierenden, die die Gruppenarbeit negativ beurteilen (vgl. Abb. 46).

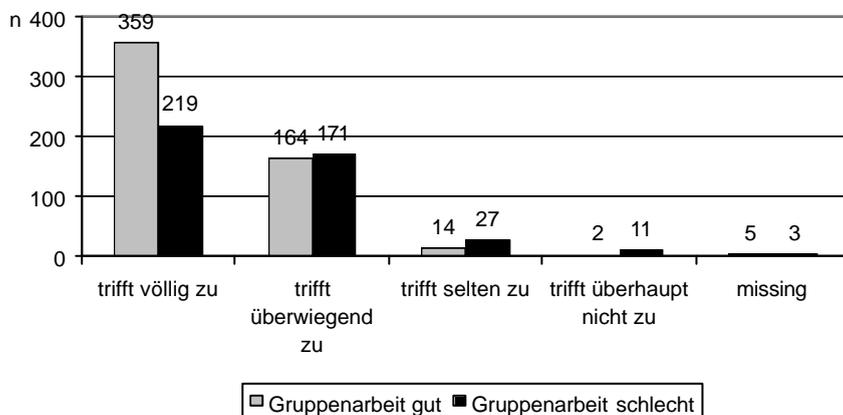


Abbildung 47: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Überprüfung des Lernerfolges durch das Abschlusskolloquium war dem Lernziel angemessen“ (KOLLOQUI) in Abhängigkeit von der Bewertung der Gruppenarbeit

Je eher Studierende die Gruppenarbeit positiv bewerten, desto eher sind sie auch der Ansicht, dass die Überprüfung des Lernerfolges durch ein Abschlusskolloquium dem Lernziel angemessen sei (KOLLOQUI) (vgl. Abb. 47).

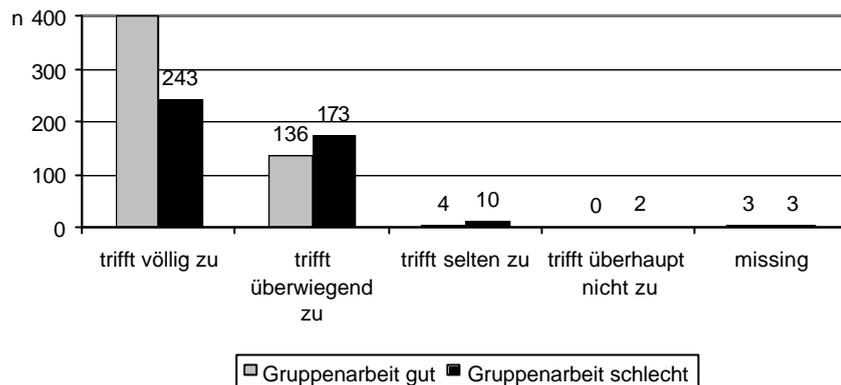


Abbildung 48: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrpersonen waren gut vorbereitet“ (VORBEREI) in Abhängigkeit von der Bewertung der Gruppenarbeit

Diejenigen, die die Gruppenarbeit positiv bewerten, sind auch eher der Ansicht, dass die Lehrpersonen gut vorbereitet waren (VORBEREI) (vgl. Abb. 48).

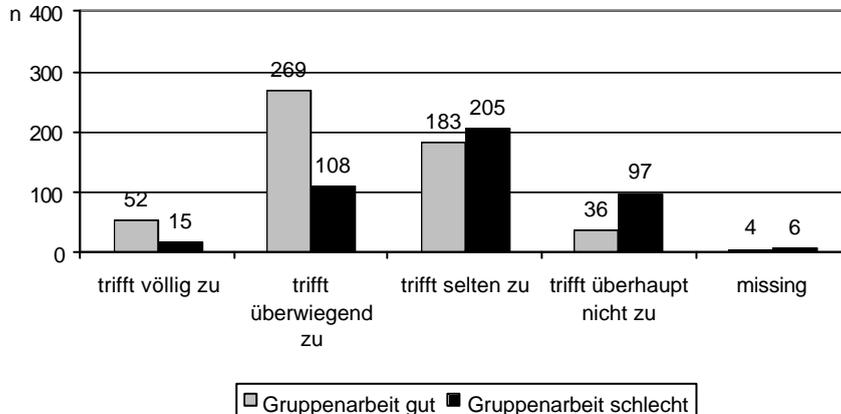


Abbildung 49: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Ich hätte an der Veranstaltung auch freiwillig teilgenommen“ (FREIWILL) in Abhängigkeit von der Bewertung der Gruppenarbeit

Studierende, die die Gruppenarbeit positiv beurteilen, hätten eher auch freiwillig an der Veranstaltung teilgenommen (FREIWILL) als Studierende, die die Gruppenarbeit insgesamt negativ beurteilen (vgl. Abb. 49).

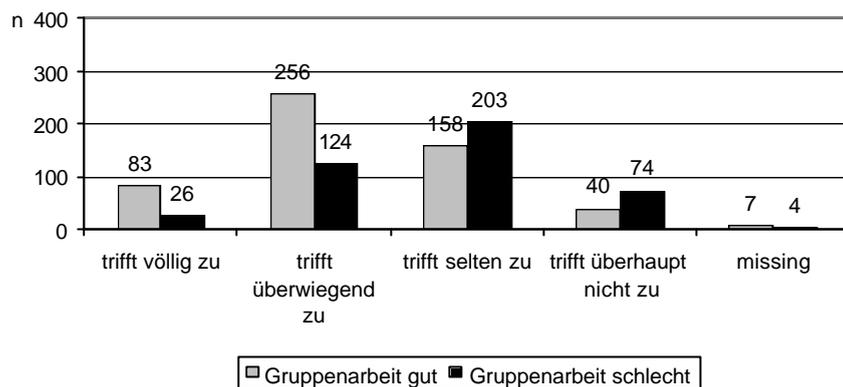


Abbildung 50: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt“ (INTERESS) in Abhängigkeit von der Bewertung der Gruppenarbeit

Bei denjenigen Studierenden, die die Gruppenarbeit positiv bewerten, hat die Veranstaltung ein deutlich höheres Interesse für das Fach Arbeitsmedizin geweckt als bei denjenigen Studierenden, die die Gruppenarbeit negativ bewerten (vgl. Abb. 50).

#### 3.4.3.6 Mittelwerte und t-Tests

Diese Vergleiche zwischen Studierenden, die die Gruppenarbeit positiv bzw. negativ beurteilen zeigen auf deskriptiver Ebene deutliche Unterschiede in der Bewertung der Lehrveranstaltung insgesamt und ihrer Teilaspekte. Inwieweit es sich dabei nur um deskriptive oder aber sogar um statistisch signifikante und damit inhaltlich interpretierbare Gruppenunterschiede handelt, wurde für alle Items anhand von t-Tests (Mittelwertsvergleichen) zwischen den beiden Gruppen untersucht. Das Signifikanzniveau liegt bei 5 Prozent. Die Ergebnisse dieser t-Tests sind in Tabelle 12 dargestellt.

Tabelle 12: Bewertung der Gruppenarbeit und Beurteilung der Lehrveranstaltung: Gruppenspezifische Verteilungsparameter und Ergebnisse des t-Tests.

	Label		Gruppen- arbeit gut	*	Gruppen- arbeit schlecht		Label		Gruppen- arbeit gut		Gruppen- arbeit schlecht
Grundlagenseminar	MEDIEN	M	1,66	*	1,85	Praktische Übungen	BASISWIS	M	1,79		1,83
		s	0,52		0,51			s	0,64		0,67
	THEMA	M	1,75	*	1,99		DIALOG	M	1,84	*	2,17
		s	0,58		0,65			s	0,65		0,75
	FRAGEN	M	1,47	*	1,60		KOMPET2	M	1,52	*	1,71
s		0,62		0,66	s	0,55			0,60		
KOMPET1	M	1,52	*	1,72	LEHRINHA	M	1,76	*	2,19		
	s	0,57		0,60		s	0,60		0,72		
AUFMERK1	M	2,19	*	2,65	AUFMERK2	M	1,92	*	2,32		
	s	0,63		0,76		s	0,61		0,74		
Gesamtbeurteilung	STRUKTUR	M	1,59	*	1,95	Betriebsbegehung	INHALT1	M	1,67	*	1,92
		s	0,54		0,63			s	0,71		0,81
	INHALT2	M	1,63	*	2,20		ORGANIST	M	1,45	*	1,75
		s	0,59		0,74			s	0,64		0,80
	KENNTNIS	M	1,55	*	1,86		KOMPET3	M	1,63	*	1,92
s		0,68		0,75	s	0,71			0,82		
LERNERFO	M	1,93	*	2,38	EINBLICK	M	1,54	*	1,80		
	s	0,59		0,69		s	0,67		0,80		
Gesamtbeurteilung	KOLLOQUI	M	1,37	*	1,60	Gruppenarbeit	DISKURSI	M	1,72	*	2,77
		s	0,55		0,72			s	0,47		0,58
	VORBEREI	M	1,27	*	1,46		UNTERSTÜ	M	1,64	*	2,61
		s	0,46		0,57			s	0,65		0,79
	FREIWILL	M	2,38	*	2,90		SPAß	M	1,57	*	2,56
s		0,75		0,78	s	0,50			0,74		
INTERESS	M	2,29	*	2,76							
	s	0,82		0,81							

\* signifikanter Mittelwertsunterschied; M = arith. Mittelwert; s = Standardabweichung

Fast durchgängig über alle Indikatoren der Lehrveranstaltung beurteilt die Gruppe derjenigen, die die Gruppenarbeit positiv bewerten, alle übrigen evaluativen Kriterien signifikant positiver als die Gruppe derjenigen, die die Gruppenarbeit negativ bewerten. Lediglich beim Item Basiswissen (BASISWIS) besteht kein signifikanter Unterschied zwischen beiden Gruppen. Dies erscheint plausibel, da das vorher bereits vorhandene Basiswissen bei allen Studierenden zu Beginn der Lehrveranstaltung nicht systematisch in Abhängigkeit von der Lehrveranstaltung und den dabei eingesetzten Methoden variiert.

Grundsätzlich ist die Lehrveranstaltung so konzipiert, dass sie an das mitgebrachte Basis- und Vorwissen anknüpft. Ob die Veranstaltung selbst zum Erfolg wird, ist somit von der eingesetzten Lehrform und deren Akzeptanz durch die Studierenden abhängig. Es zeigt sich anhand dieser kontinuierlichen Evaluation, dass die Studierenden die Gruppenarbeit als Unterrichtsform favorisieren, darüber hinaus hohe Motivation, Akzeptanz und persönliche Lernerfolge erzielen.

### 3.4.4 Vergleiche der Semesterhälften 1999 bis 2003

Die Lehrveranstaltung wurde zweimal pro Semester angeboten, einmal in der ersten Semesterhälfte und einmal in der zweiten Semesterhälfte. Da in der zweiten Semesterhälfte für die Studierenden die Prüfungsbelastung in den übrigen Lehrveranstaltungen ansteigt, stellt sich die Frage, welchen Einfluss diese zusätzlichen Stressoren auf die Bewertung der arbeitsmedizinischen Lehrveranstaltung haben.

#### 3.4.4.1 Häufigkeitsverteilungen

In diesem Abschnitt werden die absoluten Häufigkeitsverteilungen je Item für die Befragten aus den beiden Semesterhälften gegenübergestellt.

##### 3.4.4.1.1 Grundlagenseminar

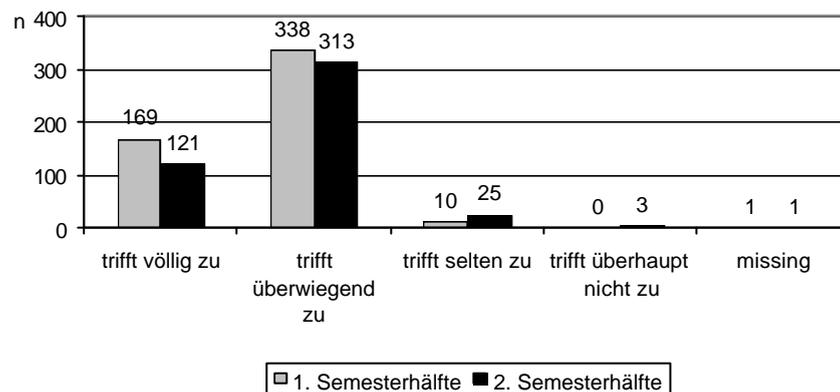


Abbildung 51: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Medien wurden sinnvoll und effektiv eingesetzt“ (MEDIEN) nach Befragten je Semesterhälfte

Die Befragten aus der ersten Semesterhälfte sind eher der Meinung, dass die Medien sinnvoll und effektiv eingesetzt wurden (MEDIEN) als Befragte aus der zweiten Semesterhälfte (vgl. Abb. 51).

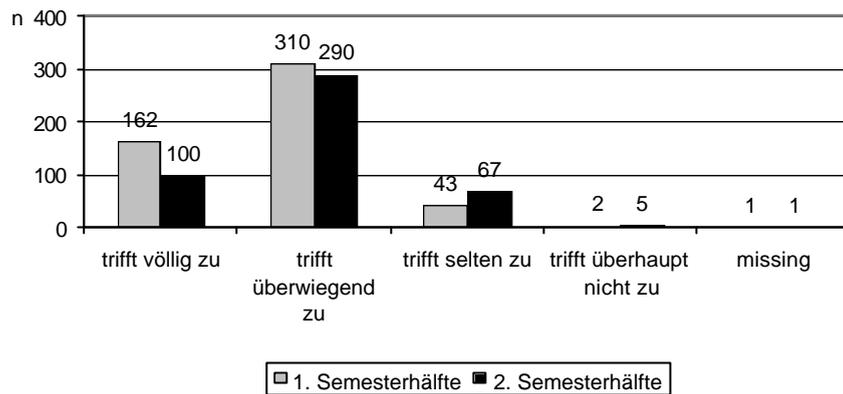


Abbildung 52: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Das Thema wurde durch praxisnahe Beispiele veranschaulicht“ (THEMA) nach Befragten je Semesterhälfte

In der ersten Semesterhälfte Befragte sind eher der Meinung, dass das Thema durch praxisnahe Beispiele veranschaulicht wurde (THEMA) als Befragte aus der zweiten Semesterhälfte (vgl. Abb. 52).

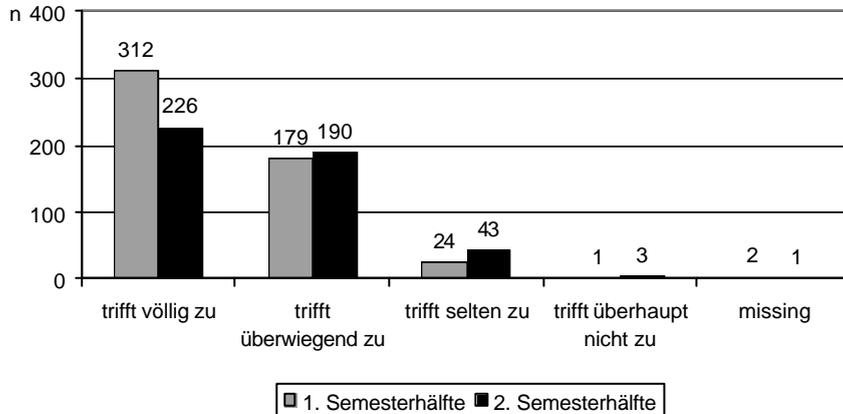


Abbildung 53: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen waren erwünscht“ (FRAGEN) nach Befragten je Semesterhälfte

Die Befragten aus der ersten Semesterhälfte sind eher der Ansicht, dass inhaltliche Fragen erwünscht waren (FRAGEN) als Befragte aus der zweiten Semesterhälfte (vgl. Abb. 53).

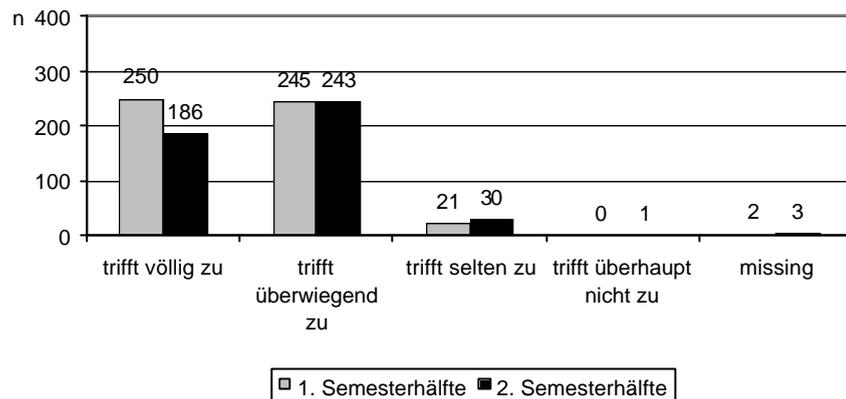


Abbildung 54: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET1) nach Befragten je Semesterhälfte

Die Befragten aus der zweiten Semesterhälfte sind seltener der Meinung, dass inhaltliche Fragen im Grundlagenseminar kompetent und verständlich beantwortet wurden (KOMPET1) als Befragte aus der ersten Semesterhälfte (vgl. Abb. 54).

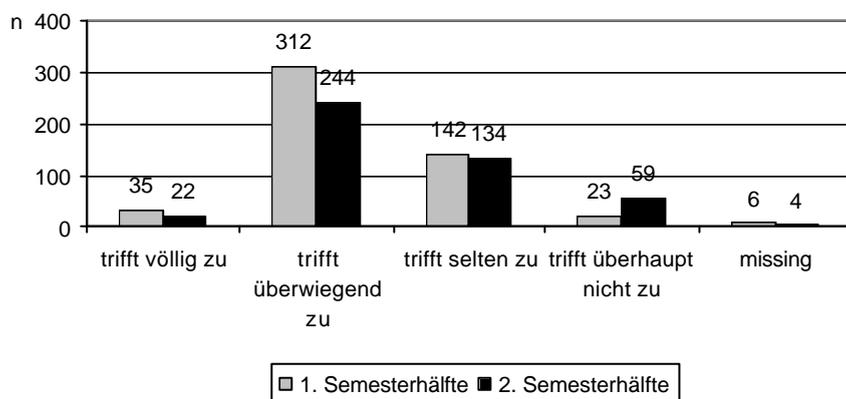


Abbildung 55: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Es fiel mir leicht aufmerksam zu bleiben“ (AUFMERK1) nach Befragten je Semesterhälfte

Und schließlich fiel es den Befragten aus der ersten Semesterhälfte im Grundlagenseminar leichter, aufmerksam zu bleiben (AUFMERK1) als Befragten aus der zweiten Semesterhälfte (vgl. Abb. 55).

### 3.4.4.1.2 Praktische Übungen

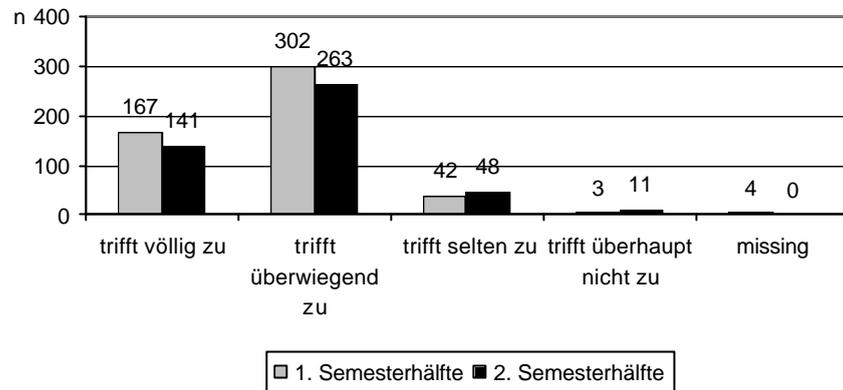


Abbildung 56: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Mein Basiswissen war ausreichend, um den Übungen zu folgen“ (BASISWIS) nach Befragten je Semesterhälfte

Die Befragten aus der ersten Semesterhälfte sind eher der Ansicht, dass ihr Basiswissen ausreichend war, um den Übungen zu folgen (BASISWIS), als die Befragten aus der zweiten Semesterhälfte (vgl. Abb. 56).

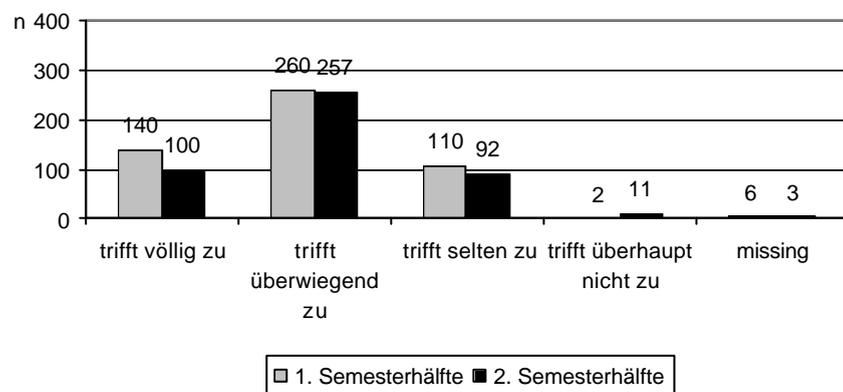


Abbildung 57: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Der Lernstoff wurde im Dialog erarbeitet“ (DIALOG) nach Befragten je Semesterhälfte

Dass der Lernstoff im Dialog erarbeitet wurde (DIALOG), wird eher von den Befragten aus der ersten Semesterhälfte als von den Befragten aus der zweiten Semesterhälfte wahrgenommen (vgl. Abb. 57).

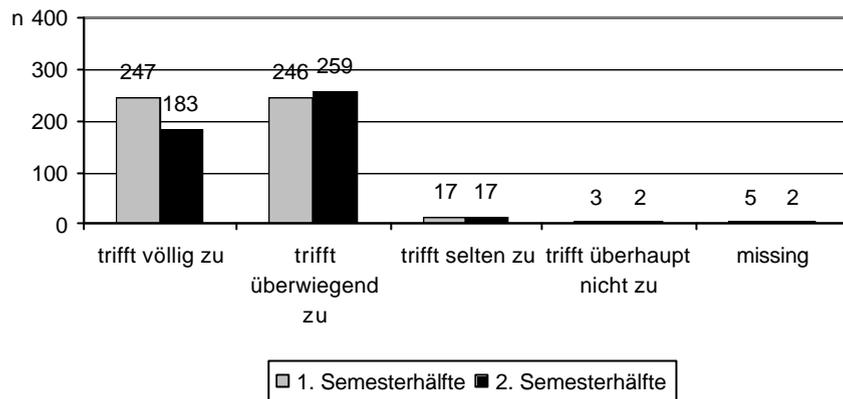


Abbildung 58: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET2) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte sind eher der Meinung, dass inhaltliche Fragen in den praktischen Übungen kompetent und verständlich beantwortet wurden (KOMPET2) als Befragte aus der zweiten Semesterhälfte (vgl. Abb. 58).

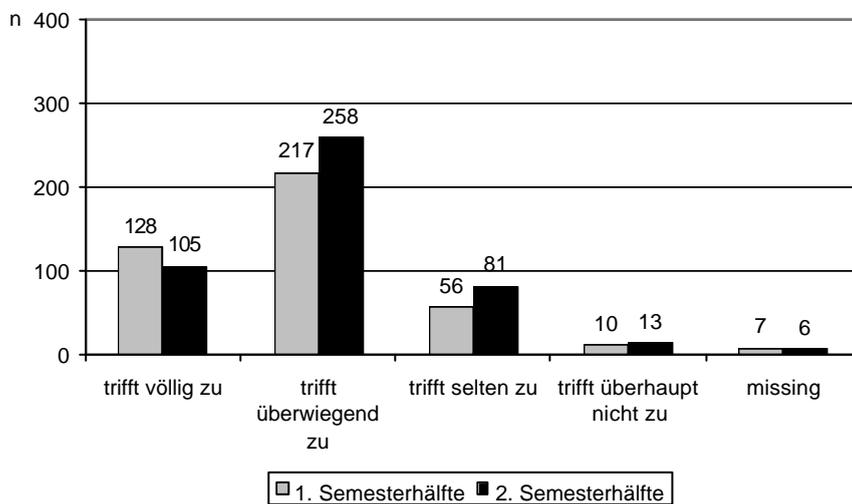


Abbildung 59: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrinhalte wurden interessant präsentiert“ (LEHRINHA) nach Befragten je Semesterhälfte

Die Befragten aus der zweiten Semesterhälfte sind seltener als die Befragten der ersten Semesterhälfte der Ansicht, dass die Lehrinhalte in den praktischen Übungen interessant präsentiert wurden (LEHRINHA) (vgl. Abb. 59).

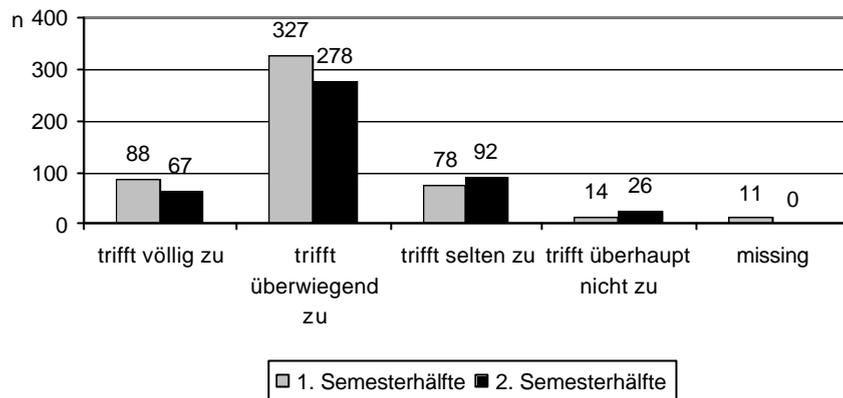


Abbildung 60: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Es fiel mir leicht, aufmerksam zu bleiben“ (AUFMERK2) nach Befragten je Semesterhälfte

Den Befragten aus der ersten Semesterhälfte fiel leichter, in den praktischen Übungen aufmerksam zu bleiben (AUFMERK2) als den Befragten aus der zweiten Semesterhälfte (vgl. Abb. 60).

### 3.4.4.1.3 Betriebsbegehung

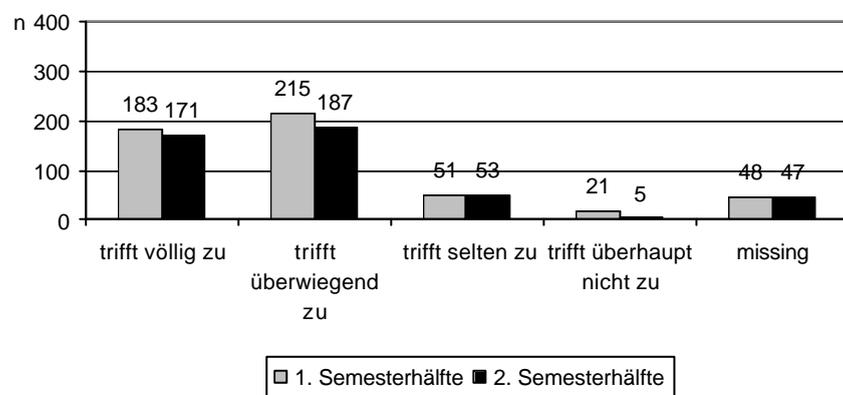


Abbildung 61: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung war inhaltlich gut vorbereitet“ (INHALT1) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte sind eher der Ansicht, dass die Betriebsbegehung inhaltlich gut vorbereitet war (INHALT1) als die Befragten aus der zweiten Semesterhälfte (vgl. Abb. 61).

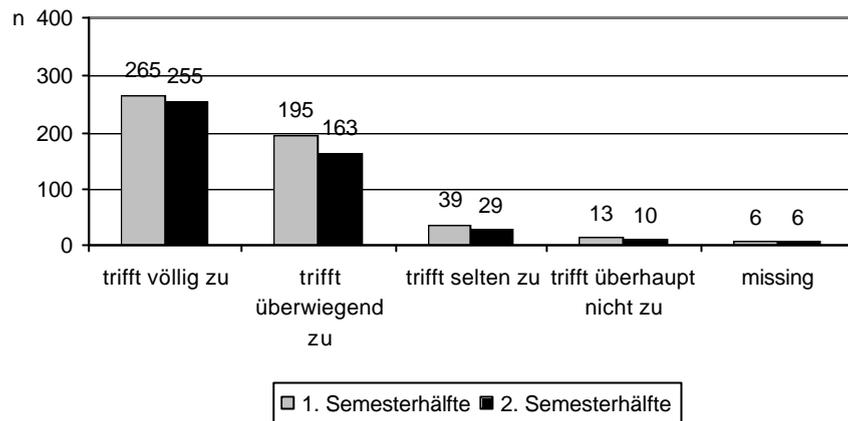


Abbildung 62: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung war organisatorisch gut vorbereitet“ (ORGANIST) nach Befragten je Semesterhälfte

Ebenfalls sind die Befragten der ersten Semesterhälfte auch eher als die Befragten der zweiten Semesterhälfte der Ansicht, dass die Betriebsbegehung organisatorisch gut vorbereitet war (Organist)(vgl. Abb. 62).

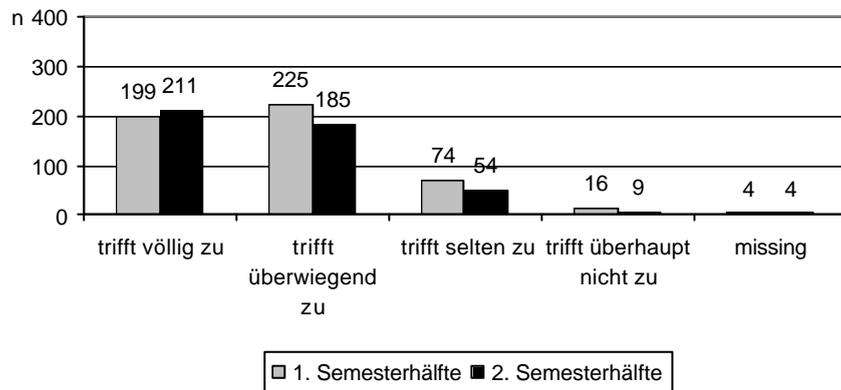


Abbildung 63: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET3) nach Befragten je Semesterhälfte

Nahezu keine Unterschiede zwischen den Befragten beider Semesterhälften bestehen hinsichtlich der Einschätzung, inwieweit inhaltliche Fragen kompetent und verständlich beantwortet wurden (KOMPET3) (vgl. Abb. 63).

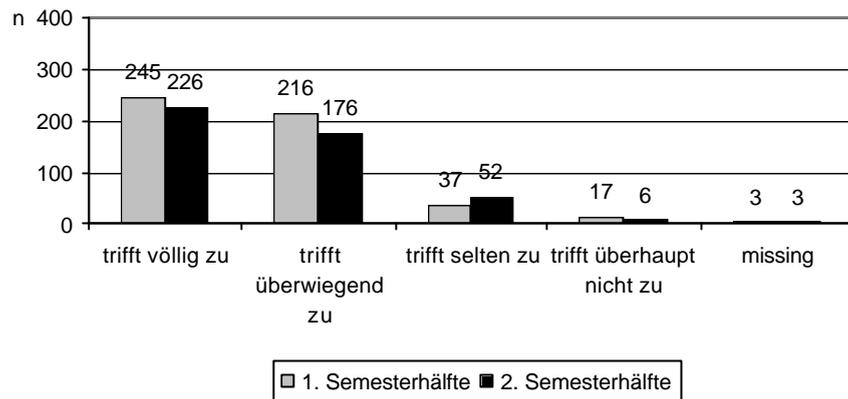


Abbildung 64: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Betriebsbegehung hat mir einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft“ (EINBLICK) nach Befragten je Semesterhälfte

Dass die Betriebsbegehung einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz vermittelt habe (EINBLICK), wird eher von den Befragten der ersten Semesterhälfte so eingeschätzt, als von den Befragten der zweiten Semesterhälfte (vgl. Abb. 64).

#### 3.4.4.1.4 Gruppenarbeit

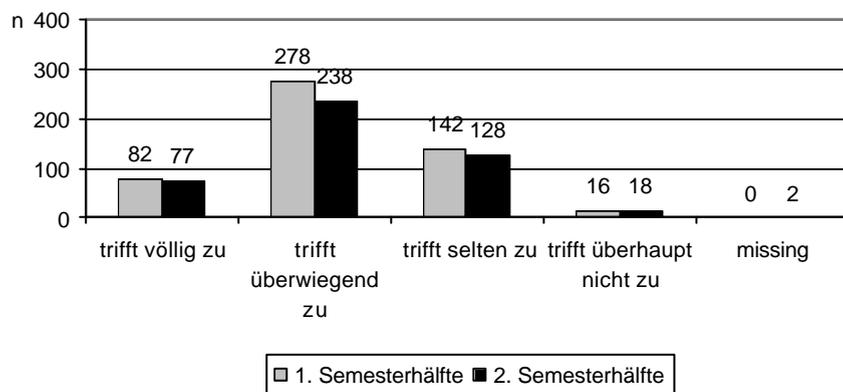


Abbildung 65: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes“ (DISKURSI) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte sind eher der Meinung, dass die Diskussion in der Gruppe zu einem vertieften Verständnis des Lerninhaltes geführt habe (DISKURSI) als die Befragten der zweiten Semesterhälfte (vgl. Abb. 65).

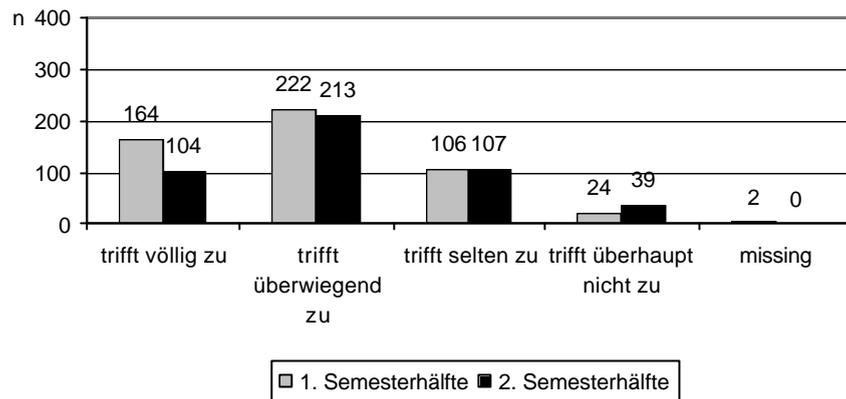


Abbildung 66: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation“ (UNTERSTÜ) nach Befragten je Semesterhälfte

Ebenso sind die Befragten der ersten Semesterhälfte eher der Ansicht, dass die Gruppenarbeit mit Dozentenunterstützung eine wichtige Hilfe für die Projektpräsentation gewesen sei (UNTERSTÜ) als die Befragten aus der zweiten Semesterhälfte (vgl. Abb. 66).

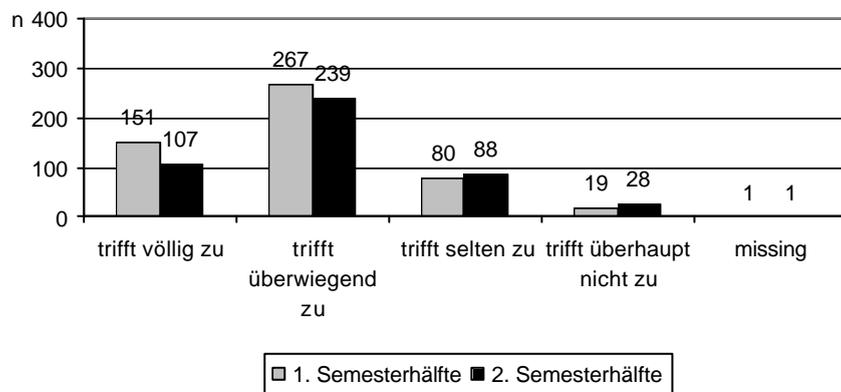


Abbildung 67: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Gruppenarbeit hat mir Spaß gemacht“ (SPAß) nach Befragten je Semesterhälfte

Den Befragten aus der ersten Semesterhälfte hat die Gruppenarbeit mehr Spaß gemacht (SPAß) als den Befragten aus der zweiten Semesterhälfte (vgl. Abb. 67).

### 3.4.4.1.5 Gesamtbewertung

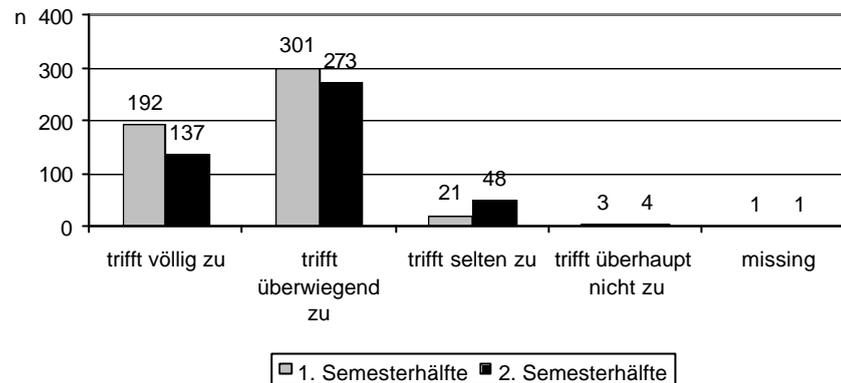


Abbildung 68: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert“ (STRUKTUR) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte beurteilen die Lehrveranstaltung insgesamt eher als gut und systematisch strukturiert (STRUKTUR), als die Befragten aus der zweiten Semesterhälfte (vgl. Abb. 68).

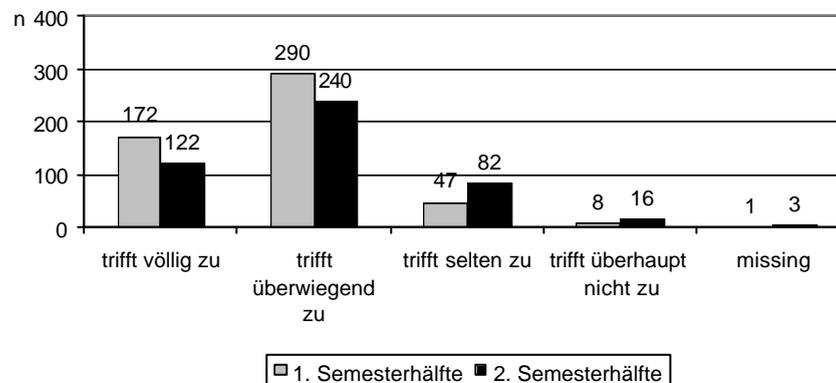


Abbildung 69: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen“ (INHALT2) nach Befragten je Semesterhälfte

Insgesamt hat den Befragten aus der ersten Semesterhälfte auch die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit (INHALT2) besser gefallen als den Befragten der zweiten Semesterhälfte (vgl. Abb. 69).

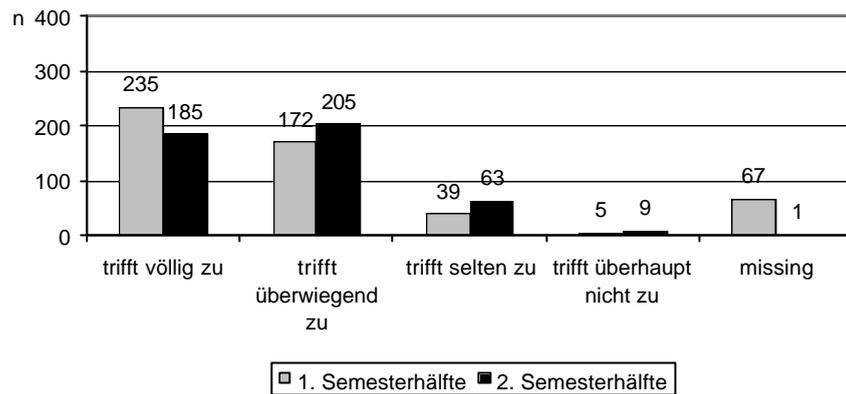


Abbildung 70: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind“ (KENNTNIS) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte sind eher der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für ihre spätere ärztliche Tätigkeit wichtig seien (KENNTNIS), als die Befragten der zweiten Semesterhälfte (vgl. Abb. 70).

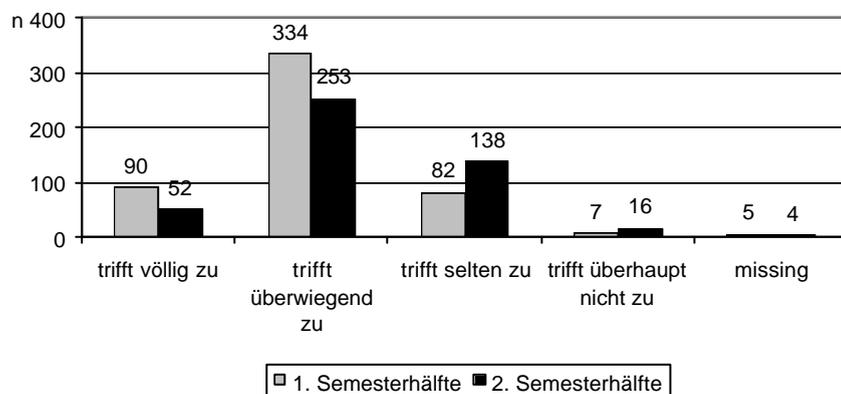


Abbildung 71: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Mein persönlicher Lernerfolg durch die Veranstaltung war groß“ (LERNERFO) nach Befragten je Semesterhälfte

Dass der persönliche Lernerfolg durch die Veranstaltung groß war (LERNERFO) wird eher von den Befragten der ersten Semesterhälfte als zutreffend beschrieben, als von den Befragten der zweiten Semesterhälfte (vgl. Abb. 71).

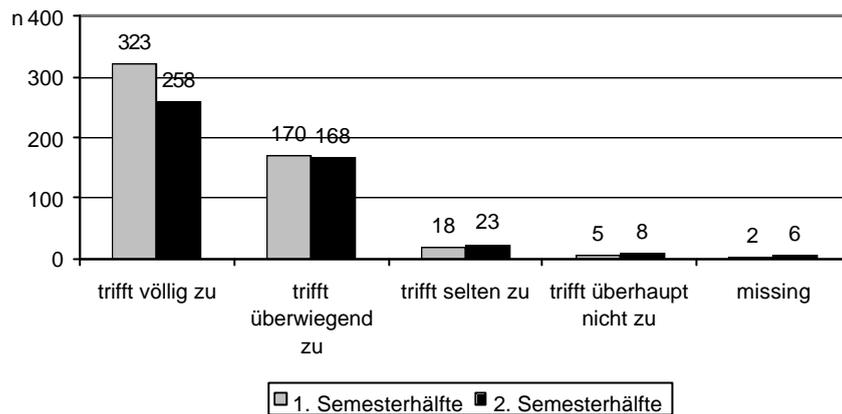


Abbildung 72: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Überprüfung des Lernerfolges durch das Abschlusskolloquium war dem Lernziel angemessen“ (KOLLOQUI) nach Befragten je Semesterhälfte

Ebenso fanden eher die Befragten der ersten Semesterhälfte die Überprüfung des Lernerfolges durch ein Abschlusskolloquium angemessen (KOLLOQUI), als die Befragten der zweiten Semesterhälfte (vgl. Abb. 72).

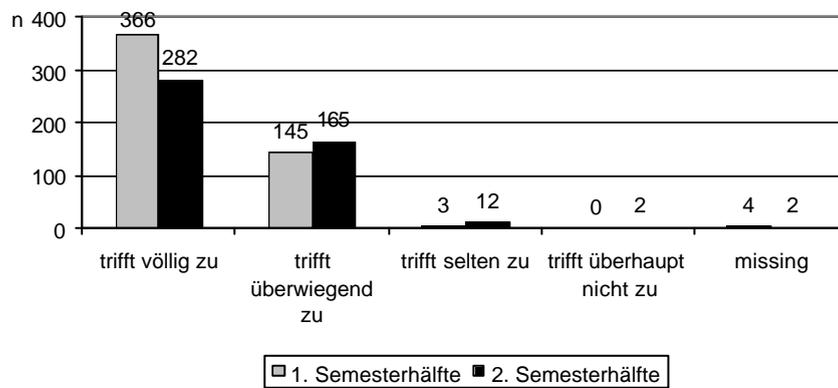


Abbildung 73: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Lehrpersonen waren gut vorbereitet“ (VORBEREI) nach Befragten je Semesterhälfte

Die Befragten der ersten Semesterhälfte sind eher der Meinung, dass die Lehrpersonen gut vorbereitet waren (VORBEREI), als die Befragten der zweiten Semesterhälfte (vgl. Abb. 73).

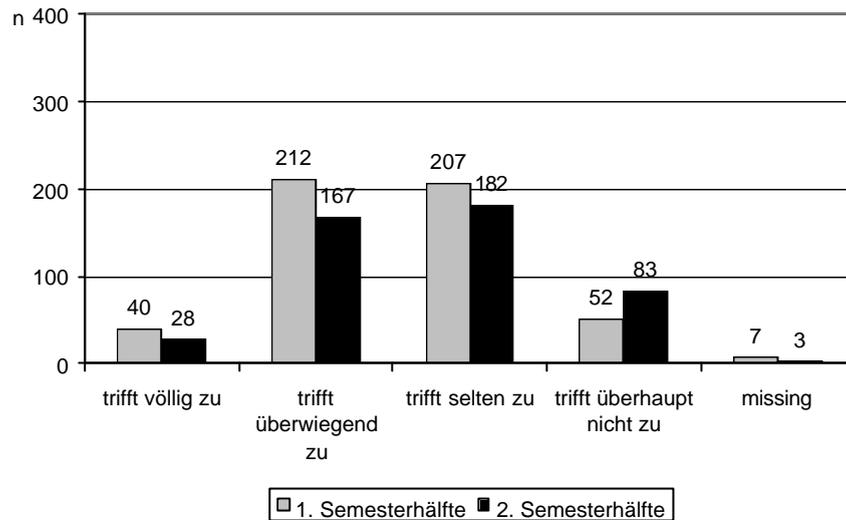


Abbildung 74: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Ich hätte an der Veranstaltung auch freiwillig teilgenommen“ (FREIWILL) nach Befragten je Semesterhälfte

Die Befragten aus der ersten Semesterhälfte wären eher bereit gewesen, auch freiwillig an der Veranstaltung teilzunehmen (FREIWILL), als die Befragten der zweiten Semesterhälfte (vgl. Abb. 74).

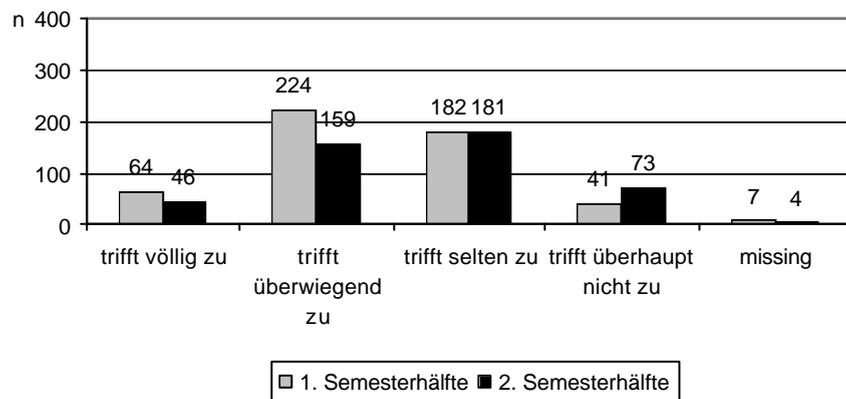


Abbildung 75: Absolute Häufigkeitsverteilungen der Bewertungen des Items „Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt“ (INTERESS) nach Befragten je Semesterhälfte

Bei den Befragten der ersten Semesterhälfte hat die Veranstaltung eher das Interesse an der Arbeitsmedizin geweckt (INTERESS) als bei Befragten der zweiten Semesterhälfte (vgl. Abb. 75).

### 3.4.4.1.6 Mittelwerte und t-Tests

Im Folgenden werden die deskriptiven Bewertungsunterschiede zwischen den Studierenden der ersten und der zweiten Semesterhälfte unter Anwendung des t - Testes daraufhin überprüft, inwieweit sie statistische Signifikanz besitzen und damit auch inhaltlich interpretierbar sind. Das Signifikanzniveau liegt bei 5 Prozent.

Die Aufteilung der Studierenden in die erste und zweite Semesterhälfte ist aus organisatorischen Gründen vorgegeben. In Tabelle 13 sind die Vergleichsergebnisse des t Tests über alle Beurteilungselemente dargestellt.

Tabelle 13: Vergleich der Bewertung nach Teilnahme in der 1. und 2. Semesterhälften

	Label		1. Semester -hälfte	*	2. Semester -hälfte		Label		1. Semester -hälfte		2. Semester -hälfte
Grundlagenseminar	MEDIEN	M	1,69	*	1,81	Praktische Übungen	BASISWIS	M	1,77		1,85
		s	0,50		0,55			s	0,61		0,69
	THEMA	M	1,78	*	1,95		DIALOG	M	1,95		2,03
		s	0,60		0,63			s	0,71		0,72
	FRAGEN	M	1,45	*	1,62		KOMPET2	M	1,56	*	1,65
	s	0,59		0,68		s	0,59		0,57		
	KOMPET1	M	1,56	*	1,67	LEHRINHA	M	1,90	*	2,00	
	s	0,57		0,61		s	0,65		0,72		
	AUFMERK1	M	2,30	*	2,50	AUFMERK2	M	2,04	*	2,17	
	s	0,66		0,78		s	0,66		0,74		
Gesamtbeurteilung	STRUKTUR	M	1,68	*	1,82	Betriebsbegehung	INHALT1	M	1,81		1,74
		s	0,58		0,64			s	0,80		0,72
	INHALT2	M	1,79	*	1,98		ORGANIST	M	1,61		1,55
		s	0,66		0,76			s	0,74		0,71
	KENNTNIS	M	1,59	*	1,77	KOMPET3	M	1,82	*	1,70	
		s	0,69		0,75		s	0,79		0,75	
	LERNERFO	M	2,01	*	2,26	EINBLICK	M	1,66		1,65	
		s	0,62		0,70		s	0,75		0,73	
Gruppenarbeit	KOLLOQUI	M	1,43	*	1,52	DISKURSI	M	2,18		2,19	
		s	0,61		0,68		s	0,72		0,75	
	VORBEREI	M	1,29	*	1,42	UNTERSTÜ	M	1,98	*	2,17	
		s	0,47		0,57		s	0,84		0,87	
FREIWILL	M	2,53	*	2,70	SPAß	M	1,94	*	2,08		
	s	0,78		0,83		s	0,77		0,81		
INTERESS	M	2,39	*	2,61							
	s	0,81		0,87							

\* signifikanter Mittelwertsunterschied; M = arith. Mittelwert; s = Standardabweichung

Die Vergleichsergebnisse des t-Tests über alle Beurteilungselemente zeigen, dass die Lehrveranstaltung „insgesamt“ sowie in den Teilaspekten „Grundlagenseminar“, praktische Übungen“ und „Gruppenarbeit“ von den Studierenden in der zweiten Semesterhälfte in fast allen Bewertungsaspekten signifikant schlechter beurteilt werden als von denen der ersten Semesterhälfte. Lediglich in der Beurteilung der Betriebsbegehung zeigen sich keine systematisch signifikanten Unterschiede in der Beurteilung. Deren Bewertung fällt in der zweiten Semesterhälfte deskriptiv, aber nicht statistisch signifikant sogar positiver aus als in der ersten Semesterhälfte. Dies liegt vermutlich in ihrem praktischen Charakter als Ausgleich zu den Prüfungsvorbereitungen in den anderen Fächern.

Die signifikant schlechtere Bewertung in der zweiten Semesterhälfte kann zum Teil sicherlich auf zusätzliche Stressoren aus anderen Lehrveranstaltungen zurückgeführt werden (z. B. Prüfungsbelastungen zum Ende des Semesters). Ob dies die alleinige Ursache ist, kann anhand der vorliegenden Daten nicht entschieden werden.

#### **3.4.5 Zeitliche Veränderungen von 1999 bis 2003**

Der Untersuchungszeitraum reicht vom Wintersemester 1999 bis zum Sommersemester 2003, also über eine Zeitspanne von vier Jahren. In diesem Zeitraum können z. B. Dozenten gewechselt haben, die inhaltliche und didaktische Konzeption verändert worden sein, aber auch Veränderungen auf Seiten der Studierenden stattgefunden haben. Deshalb lohnt es, den Verlauf der Lehrveranstaltungsbewertung näher zu betrachten. In den folgenden Abbildungen sind deshalb zunächst die Mittelwertverläufe der einzelnen Beurteilungsaspekte über den Untersuchungszeitraum dargestellt. Liniendiagramme wurden gewählt, um die zeitliche Entwicklung besser als Kontinuum darstellen zu können.

### 3.4.5.1 Grundlagenseminar

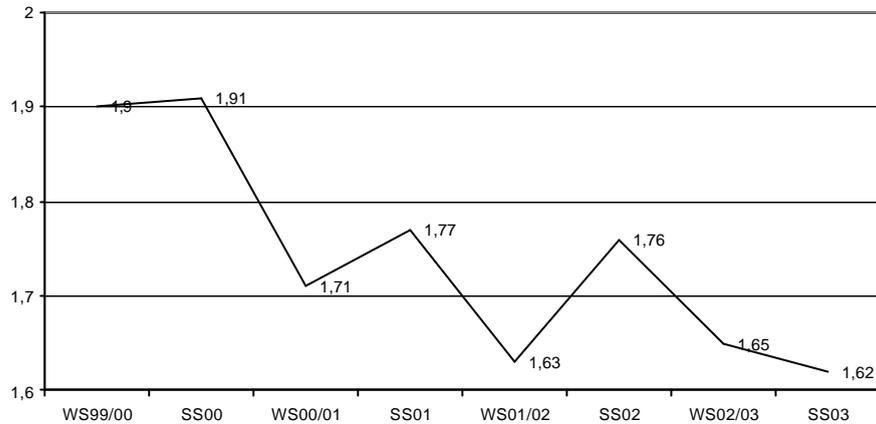


Abbildung 76: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Medien wurden sinnvoll und effektiv eingesetzt“ (MEDIEN)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass die Medien sinnvoll und effektiv eingesetzt wurden (MEDIEN) (vgl. Abb. 76).

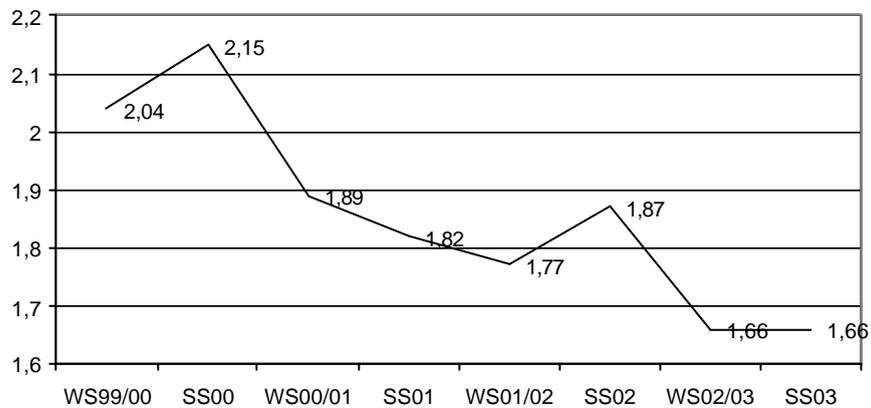


Abbildung 77: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Das Thema wurde durch praxisnahe Beispiele veranschaulicht“ (THEMA)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass das Thema durch praxisnahe Beispiele veranschaulicht wurde (THEMA) (vgl. Abb. 77).

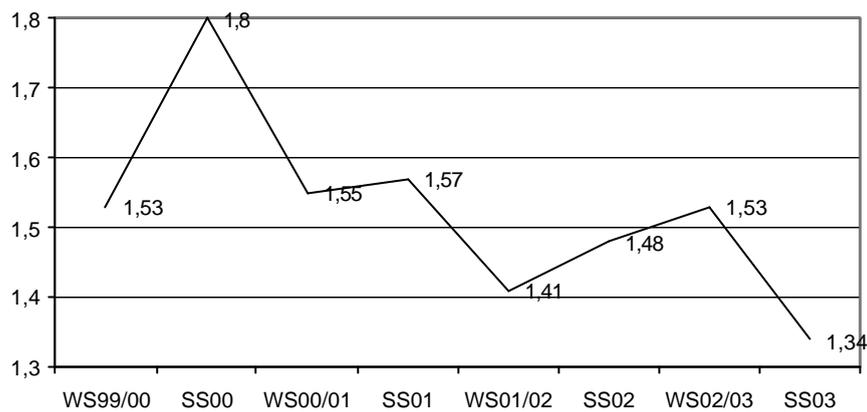


Abbildung 78: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Inhaltliche Fragen waren erwünscht“ (FRAGEN)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass inhaltliche Fragen erwünscht waren (FRAGEN) (vgl. Abb. 78).

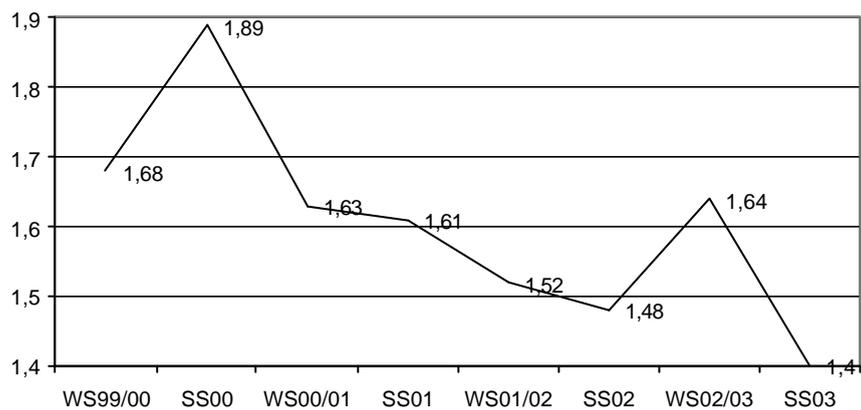


Abbildung 79: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET1)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass inhaltliche Fragen kompetent und verständlich beantwortet wurden (KOMPET1) (vgl. Abb. 79).

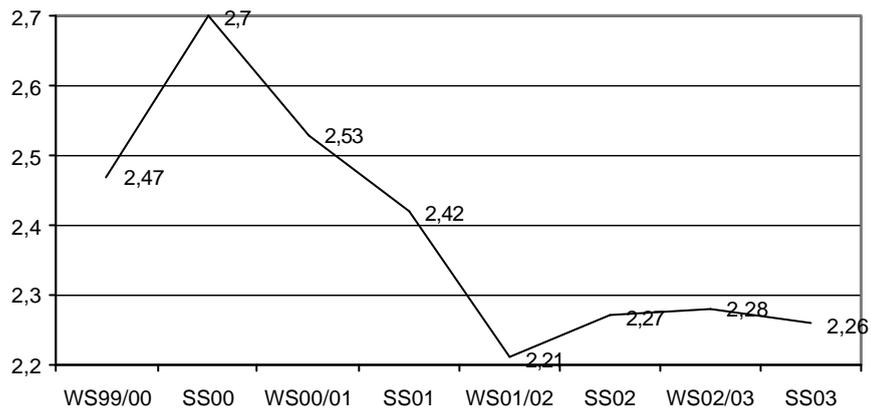


Abbildung 80: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Es fiel mir leicht aufmerksam zu bleiben“ (AUFMERK1)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass es ihnen leicht fiel aufmerksam zu bleiben (AUFMERK1) (vgl. Abb. 80).

### 3.4.5.2 Praktische Übungen

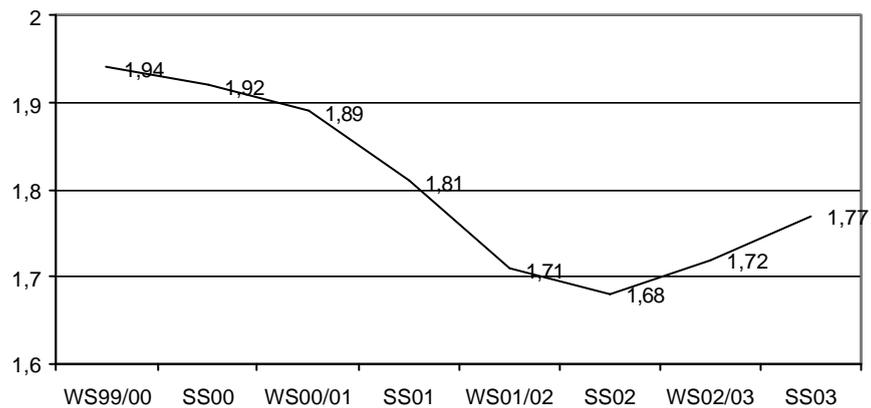


Abbildung 81: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Mein Basiswissen war ausreichend, um den Übungen zu folgen“ (BASISWIS)

Insgesamt verbessert sich über den Untersuchungszeitraum die Einschätzung der Befragten, dass ihr Basiswissen ausreichend war, um den Übungen zu folgen (BASISWIS) (vgl. Abb. 81). Der Verlauf der Kurve zeigt jedoch auch, dass sich diese Einschätzung in den letzten Semestern wieder etwas verschlechtert hat.

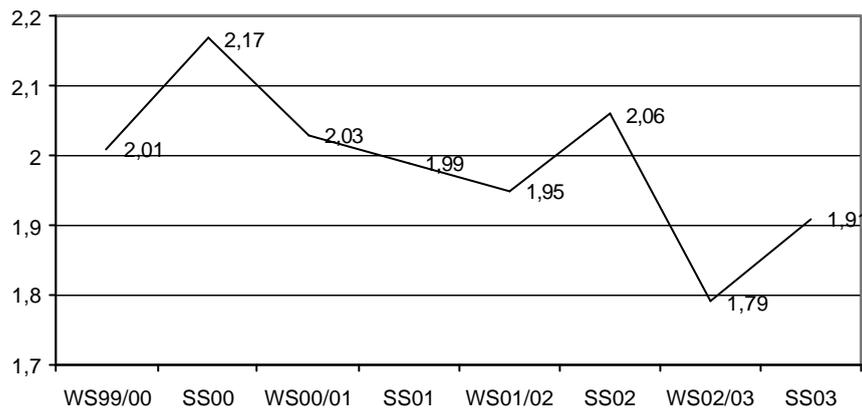


Abbildung 82: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Der Lernstoff wurde im Dialog erarbeitet“ (DIALOG)

Über den Untersuchungszeitraum verbessert sich die Einschätzung der Befragten, dass der Lernstoff im Dialog erarbeitet wurde (DIALOG) (vgl. Abb. 82). Diese Einschätzungsverbesserung ist jedoch über den Zeitraum nicht kontinuierlich. Vielmehr fällt auf, dass insbesondere in den beiden letzten Sommersemestern des Untersuchungszeitraumes die Bewertungen tendenziell schlechter ausfallen als in den jeweils vorangegangenen Wintersemestern.

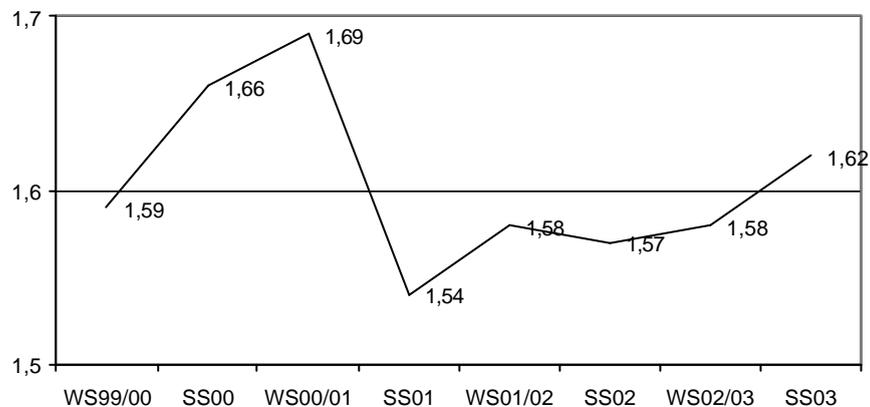


Abbildung 83: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET2)

Insgesamt zeitlich recht stabil ist die Einschätzung der befragten Studierenden, dass inhaltliche Fragen im Rahmen der praktischen Übungen kompetent und verständlich beantwortet wurden (KOMPET2) (vgl. Abb. 83). Zugleich zeigt aber der Kurvenverlauf auch, dass nach einer Verschlechterung der Einschätzung in den ersten untersuchten Semestern und einer deutlichen Verbesserung der Einschätzung im Sommersemester 2001 wiederum tendenziell eine Verschlechterung eintritt.

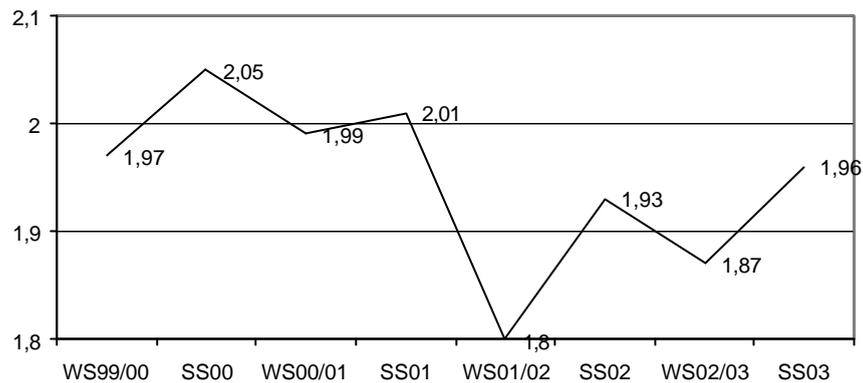


Abbildung 84: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Lehrinhalte wurden interessant präsentiert“ (LEHRINHA)

Ebenfalls recht zeitstabil ist die Einschätzung der Studierenden, dass die Lehrinhalte interessant präsentiert wurden (LEHRINHA) (vgl. Abb. 84). Eine zwischenzeitlich deutliche Verbesserung der Bewertung fand zum Wintersemester 2001/2002 statt, die aber im Gesamtverlauf der Kurve den Anschein eines Ausreißers in der durchschnittlichen Bewertung darstellen könnte.

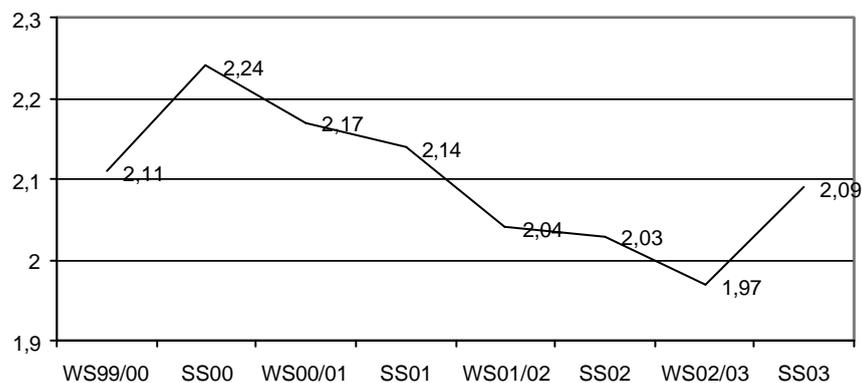


Abbildung 85: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Es fiel mir leicht, aufmerksam zu bleiben“ (AUFMERK2)

Insgesamt ist die Beurteilung ihrer Aufmerksamkeit bei den praktischen Übungen bei den Studierenden über den Untersuchungszeitraum hin konstant (AUFMERK2) (vgl. Abb. 85). In der großen Tendenz der Verlaufskurve ist aber über viele Semester hinweg eine stetige Verbesserung dieser Einschätzung festzustellen, die nur durch die Endpunkte des Untersuchungszeitraumes relativiert wird.

### 3.4.5.3 Betriebsbegehung

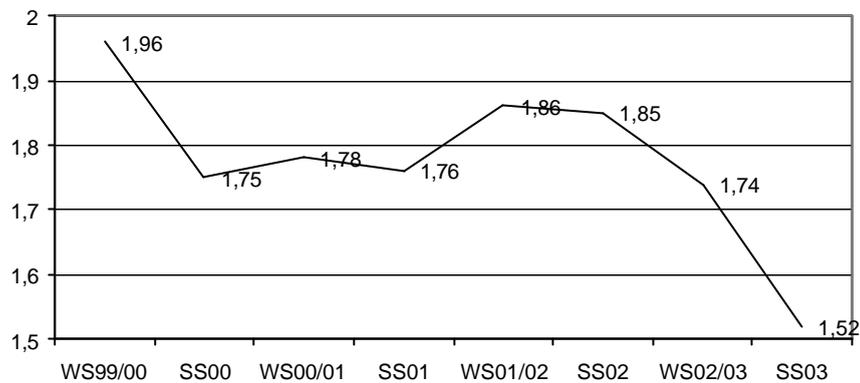


Abbildung 86: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Betriebsbegehung war inhaltlich gut vorbereitet“ (INHALT1)

In der großen Tendenz wird die Qualität der inhaltlichen Vorbereitung der Betriebsbegehung von Studierenden über den Untersuchungszeitraum hinweg immer besser eingeschätzt (INHALT1) (vgl. Abb. 86). Diese Verbesserung fand aber in erster Linie nur zwischen dem Wintersemester 1999/2000 zum Sommersemester 2000 und zwischen dem Wintersemester 2002/2003 zum Sommersemester 2003 statt.

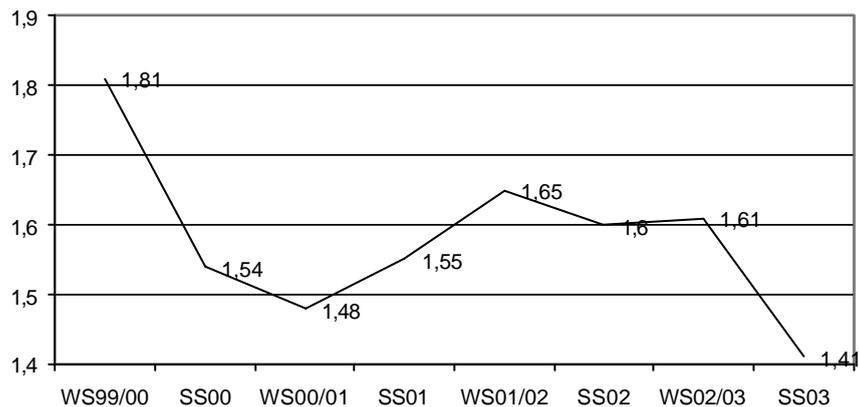


Abbildung 87: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Betriebsbegehung war organisatorisch gut vorbereitet“ (ORGANIST)

Für die zeitliche Entwicklung der Beurteilung der organisatorischen Vorbereitung der Betriebsbegehung (ORGANIST) (vgl. Abb. 87) gilt im wesentlichen das gleiche wie für die inhaltliche Vorbereitung (s.o.).

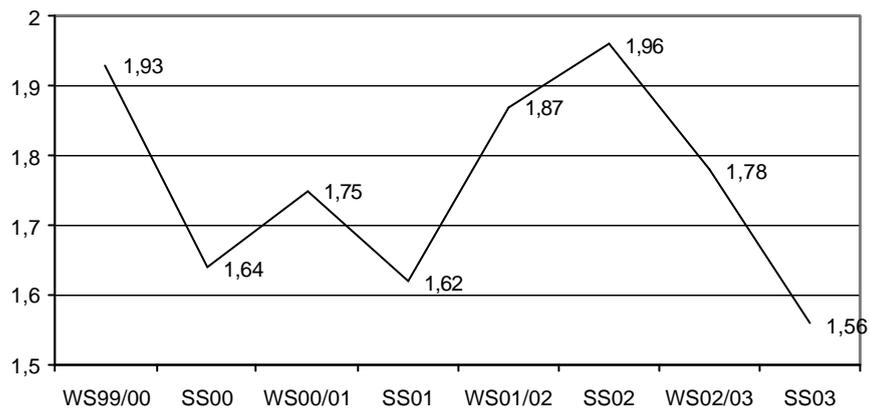


Abbildung 88: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Inhaltliche Fragen wurden kompetent und verständlich beantwortet“ (KOMPET3)

Der zeitliche Verlauf der Beurteilung, inwieweit inhaltliche Fragen im Rahmen der Betriebsbegehung kompetent und verständlich beantwortet wurden, zeigt einen uneinheitlichen Verlauf (KOMPET3) (vgl. Abb. 88). Insgesamt verbessert sich die Beurteilung über den Untersuchungszeitraum. Zugleich besteht eine deutliche Variation zwischen den verschiedenen Erhebungszeitpunkten.

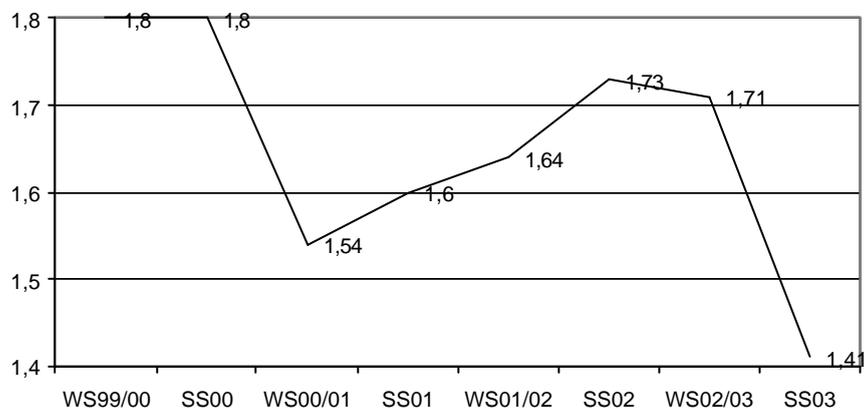


Abbildung 89: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Betriebsbegehung hat mir einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft“ (EINBLICK)

Über den Untersuchungszeitraum sind die Studierenden immer stärker der Ansicht, dass die Betriebsbegehung ihnen einen Eindruck über mögliche Gesundheitsgefahren am Arbeitsplatz vermittelt habe (EINBLICK) (vgl. Abb. 89).

### 3.4.5.4 Gruppenarbeit

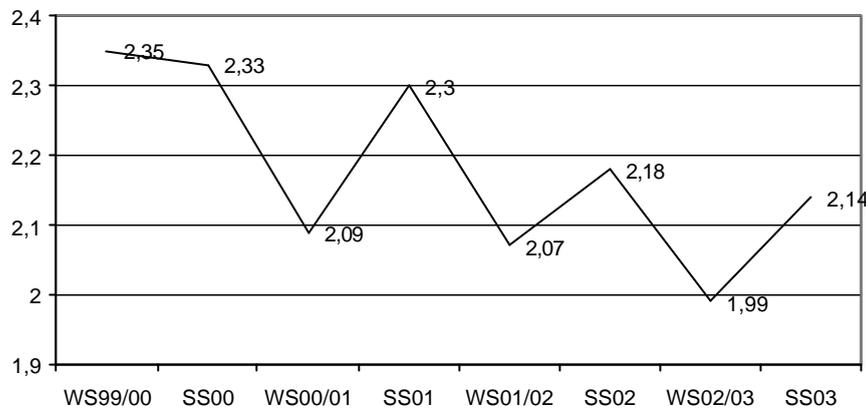


Abbildung 90: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes“ (DISKURSI)

Insgesamt verbessert sich die Einschätzung, dass die Diskussion in der Gruppe zu einem vertieften Verständnis des Lerninhaltes geführt habe über den Untersuchungszeitraum (DISKURSI) (vgl. Abb. 90). Auffällig ist jedoch, dass die Bewertung in den Sommersemestern deutlich schlechter ausfällt als in den Wintersemestern.

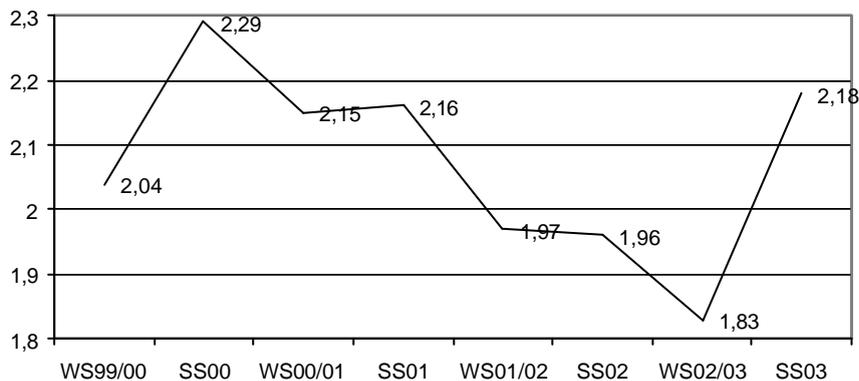


Abbildung 91: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation“ (UNTERSTÜ)

Die Einschätzung der Studierenden, die Gruppenarbeit mit Dozentenunterstützung wäre eine wichtige Hilfe für die Projektpräsentation gewesen (UNTERSTÜ) (vgl. Abb. 91) verbessert sich vom Sommersemester 2000 bis zum Wintersemester 2002/2003

kontinuierlich. Dagegen fällt die Bewertung für das Sommersemester 2003 wieder deutlich schlechter aus.

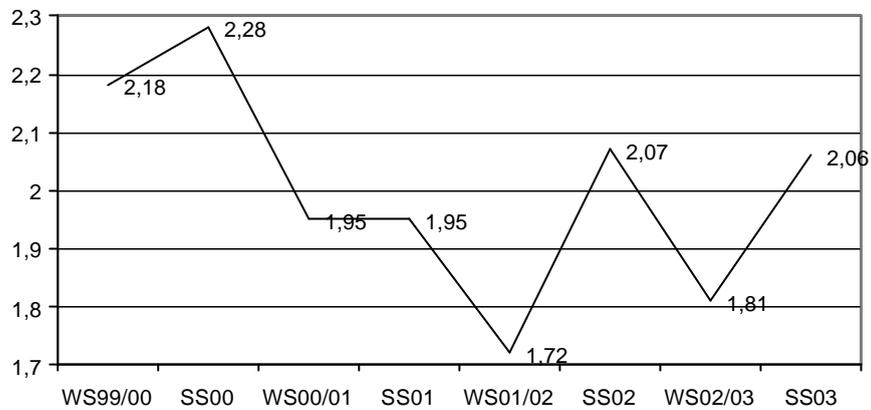


Abbildung 92: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Gruppenarbeit hat mir Spaß gemacht“ (SPAß)

Insgesamt haben die Studierenden mit Ausnahme der Sommersemester 2000, 2002 und 2003 zunehmend mehr Spaß an der Gruppenarbeit gehabt (SPAß) (vgl. Abb. 92).

Es fällt auf, dass die Items „Diskussion“ und „Spaß“ in den Wintersemestern zum Teil sehr viel positiver als in den Sommersemestern bewertet werden. Möglicherweise werden in Sommersemestern Rahmenbedingungen wahrgenommen, die eine positive Einstellung zur Gruppenarbeit erschweren.

### 3.4.5.5 Gesamtbewertung

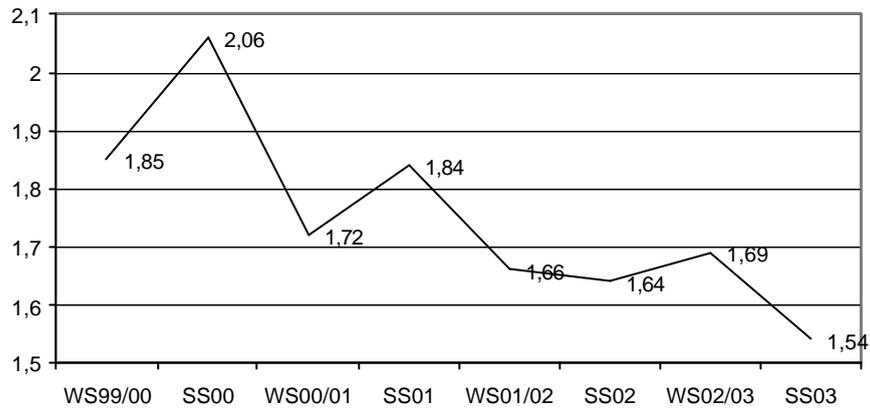


Abbildung 93: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert“ (STRUKTUR)

Insgesamt verbessert sich über den Untersuchungszeitraum die Einschätzung der Studierenden, die Lehrveranstaltung sei insgesamt gut und systematisch strukturiert gewesen (STRUKTUR) (vgl. Abb. 93).

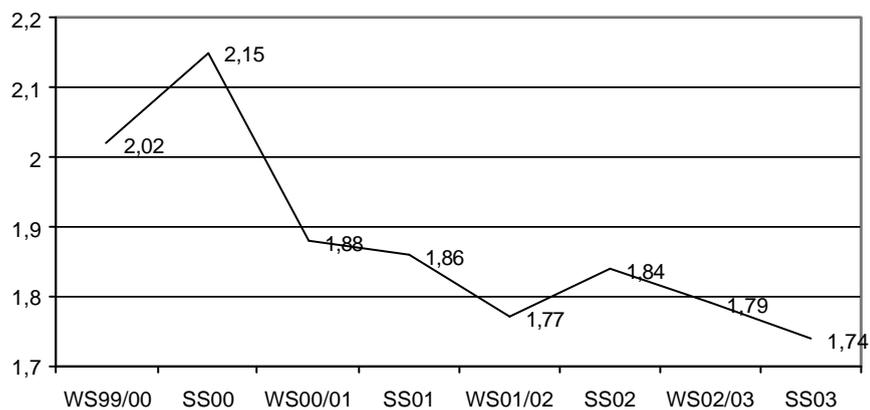


Abbildung 94: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen“ (INHALT2)

Über den Untersuchungszeitraum hat den Studierenden die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit zunehmend besser gefallen (INHALT2) (vgl. Abb. 94).

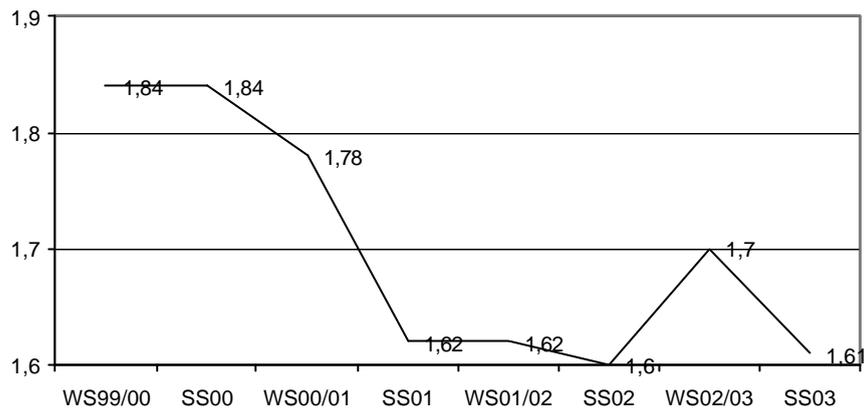


Abbildung 95: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind“ (KENNTNIS)

Insgesamt verbessert sich über den Untersuchungszeitraum die Einschätzung der Studierenden, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für ihre spätere ärztliche Tätigkeit wichtig seien (KENNTNIS) (vgl. Abb. 95).

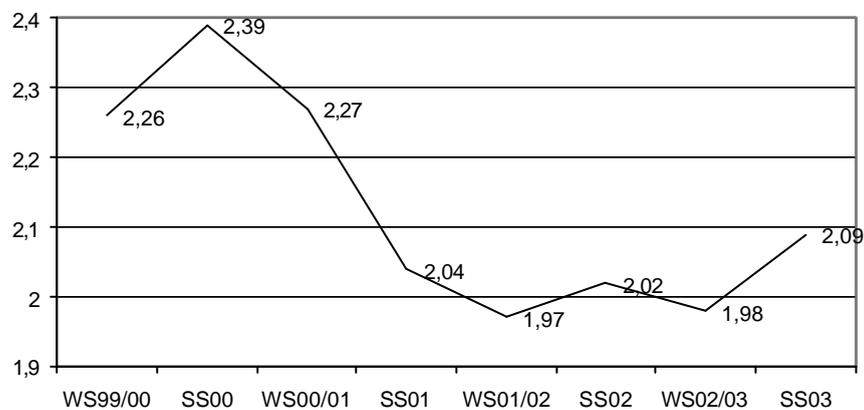


Abbildung 96: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Mein persönlicher Lernerfolg durch die Veranstaltung war groß“ (LERNERFO)

Insgesamt verbessert sich über den Untersuchungszeitraum der persönliche Lernerfolg durch die Veranstaltung (LERNERFO) (vgl. Abb. 96).

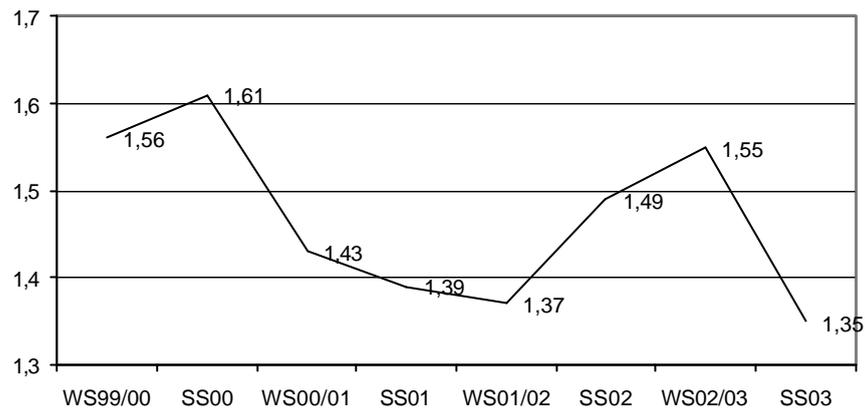


Abbildung 97: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Überprüfung des Lernerfolges durch das Abschlusskolloquium war dem Lernziel angemessen“ (KOLLOQUI)

Die Einschätzung, dass die Überprüfung des Lernerfolges durch das Abschlusskolloquium dem Lernziel angemessen sei, verbessert sich insgesamt über den Untersuchungszeitraum (KOLLOQUI) (vgl. Abb. 97). Zugleich zeigt der Verlauf der Kurve aber auch, dass es sich bei dieser Einschätzungsverbesserung nicht um eine kontinuierliche Entwicklung handelt.

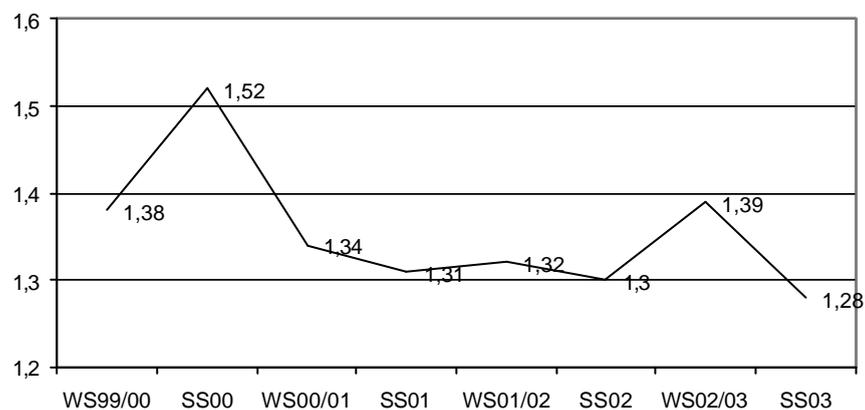


Abbildung 98: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Lehrpersonen waren gut vorbereitet“ (VORBEREI)

Tendenziell verbessert sich die Einschätzung, die Lehrpersonen seien gut vorbereitet gewesen über den Untersuchungszeitraum (VORBEREI) (vgl. Abb. 98). Insgesamt betrachtet handelt es sich dabei jedoch um eine relative zeitstabile Einschätzung.

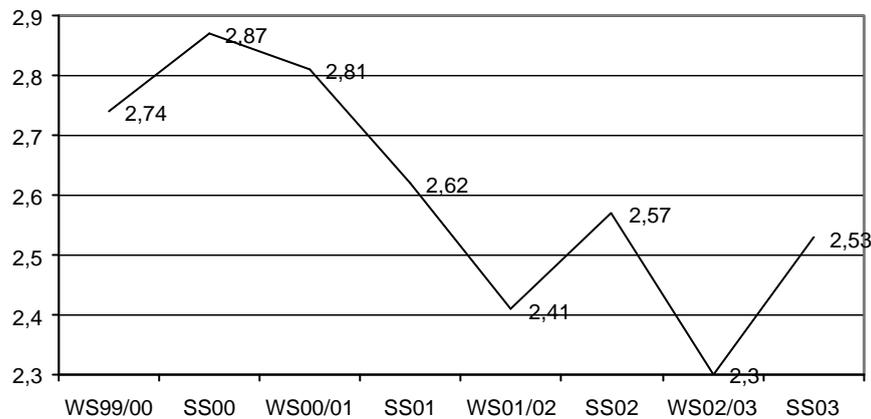


Abbildung 99: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Ich hätte an der Veranstaltung auch freiwillig teilgenommen“ (FREIWILL)

Die Bereitschaft der Studierenden, an der Veranstaltung freiwillig teilzunehmen, nimmt insgesamt über den Untersuchungszeitraum zu (FREIWILL) (vgl. Abb. 99). Tendenziell zeigt sich aber auch, dass die Freiwilligkeit in den Wintersemestern höher ist als in den Sommersemestern ist.

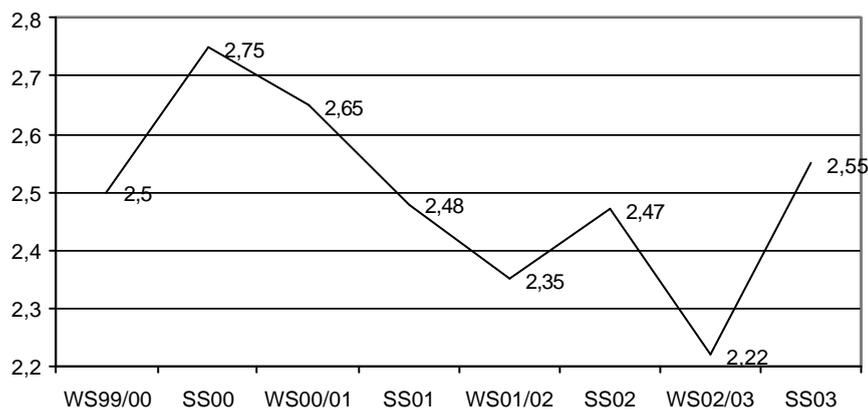


Abbildung 100: Zeitlicher Verlauf der Bewertungsmittelwerte für das Item „Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt“ (INTERESS)

Über den gesamten Untersuchungszeitraum hat die Veranstaltung zu keiner wesentlichen Veränderung des Interesses der Studierenden an der Arbeitsmedizin geführt (INTERESS) (vgl. Abb. 100).

Andererseits hat sich das Interesse vom Sommersemester 2000 bis zum Wintersemester 2002/2003 recht stetig verbessert, und ist erst im Sommersemester 2003 wieder auf das Niveau des Wintersemesters 1999/2000 gesunken.

In Abbildung 101 ist der zeitliche Verlauf der Mittelwerte über alle Items dargestellt. Hier zeigt sich, dass die Lehrveranstaltung insbesondere im Vergleich zwischen dem ersten und letzten Erhebungszeitpunkt deutlich positiver bewertet wurde.

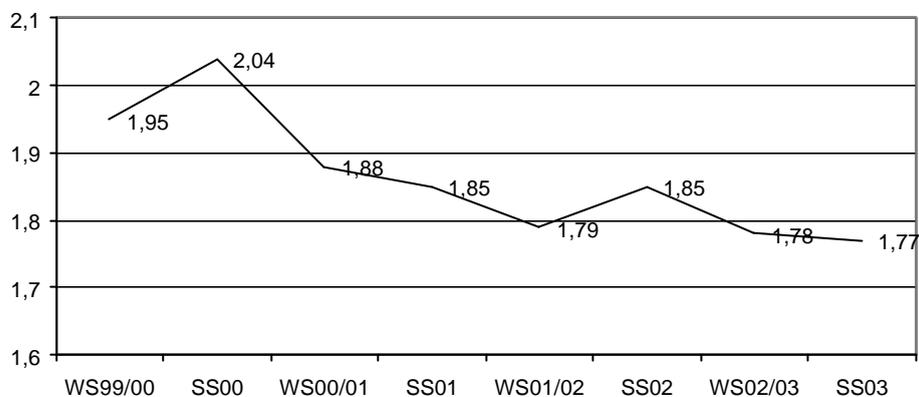


Abbildung 101: Zeitlicher Verlauf der Bewertungsmittelwerte über alle 25 Items des Fragebogens

### 3.4.5.6 Mittelwerte und t-Tests

Tabelle 14 zeigt die Mittelwerte und Standardabweichungen der Bewertungsitems getrennt für jedes Semester. Ferner sind dort signifikante Veränderungen zwischen Semestern markiert, die sich im t-Test zwischen den beiden aufeinander folgenden Semestern ergeben haben. Das Signifikanzniveau liegt bei 5 Prozent. Und schließlich enthält die Tabelle in der letzten Spalte die Ergebnisse des t - Testes zwischen Wintersemester 1999/2000 und Sommersemester 2003, um signifikante Veränderungen zwischen dem ersten und dem letzten Erhebungszeitpunkt aufzuzeigen.

Tabelle 14: Zeitliche Veränderungen vom WS 99/00 bis einschließlich SS 03

		Semester									WS99/00 - SS03
		WS 99/0 0	SS 00	WS 00/01	SS 01	WS 01/02	SS 02	WS 02/03	SS 03		
Grundlagenseminar	MEDIEN	M s	1,90 0,55	1,91 0,50	* 1,71 0,49	1,77 0,53	* 1,63 0,52	1,76 0,46	1,65 0,58	1,62 0,52	*
	THEMA	M s	2,04 0,57	2,15 0,64	* 1,89 0,53	1,82 0,56	1,77 0,60	1,87 0,63	* 1,66 0,68	1,66 0,61	*
	FRAGEN	M s	1,53 0,62	* 1,80 0,74	* 1,55 0,68	1,57 0,57	* 1,41 0,55	1,48 0,66	1,53 0,64	* 1,34 0,57	*
	KOMPET1	M s	1,68 0,60	* 1,89 0,67	* 1,63 0,59	1,61 0,52	1,52 0,52	1,48 0,58	* 1,64 0,60	* 1,40 0,51	*
	AUFMERK1	M s	2,47 0,75	* 2,70 0,81	2,53 0,70	2,42 0,64	* 2,21 0,69	2,27 0,72	2,28 0,68	2,26 0,68	*
praktische Übungen	BASISWIS	M s	1,94 0,64	1,92 0,67	1,89 0,69	1,81 0,63	1,71 0,64	1,68 0,63	1,72 0,64	1,77 0,64	*
	DIALOG	M s	2,01 0,66	2,17 0,84	2,03 0,68	1,99 0,68	1,95 0,73	2,06 0,74	* 1,79 0,67	1,91 0,65	
	KOMPET2	M s	1,59 0,64	1,66 0,57	1,69 0,61	1,54 0,55	1,58 0,56	1,57 0,59	1,58 0,54	1,62 0,60	
	LEHRINHA	M s	1,97 0,71	2,05 0,76	1,99 0,64	2,01 0,64	* 1,80 0,60	1,93 0,69	1,87 0,73	1,96 0,70	
	AUFMERK2	M s	2,11 0,71	2,24 0,75	2,17 0,66	2,14 0,67	2,04 0,67	2,03 0,66	1,97 0,68	2,09 0,77	
Betriebsbegehung	INHALT1	M s	1,96 0,88	1,75 0,80	1,78 0,74	1,76 0,69	1,86 0,78	1,85 0,78	1,74 0,76	* 1,52 0,60	*
	ORGANIST	M s	1,81 0,88	* 1,54 0,76	1,48 0,61	1,55 0,65	1,65 0,66	1,60 0,72	1,61 0,79	* 1,41 0,64	*
	KOMPET3	M s	1,93 0,91	* 1,64 0,73	1,75 0,76	1,62 0,70	* 1,87 0,75	1,96 0,81	1,78 0,77	* 1,56 0,65	*
	EINBLICK	M s	1,80 0,84	1,80 0,80	* 1,54 0,66	1,60 0,72	1,64 0,64	1,73 0,77	1,71 0,78	* 1,41 0,61	*
Gruppenarbeit	DISKURSI	M s	2,35 0,75	2,33 0,75	* 2,19 0,72	* 2,30 0,66	* 2,07 0,69	2,18 0,75	* 1,99 0,72	2,14 0,77	*
	UNTERSTÜ	M s	2,04 0,88	* 2,29 0,83	2,15 0,89	2,16 0,86	1,97 0,85	1,96 0,73	1,83 0,87	* 2,18 0,91	
	SPAß	M s	2,18 0,86	2,28 0,82	* 1,95 0,71	1,95 0,74	* 1,72 0,63	* 2,07 0,77	* 1,81 0,83	* 2,06 0,81	

Tabelle 14 Fortsetzung

Gesamtbeurteilung	STRUKTUR	M	1,85 *	2,06 *	1,72	1,84 *	1,66	1,64 *	1,69	1,54	*
		s	0,61	0,64	0,55	0,55	0,51	0,62	0,65	0,59	
	INHALT2	M	2,02	2,15 *	1,88	1,86	1,77	1,84	1,79	1,74	*
		s	0,82	0,70	0,69	0,66	0,60	0,73	0,72	0,73	
	KENNTNIS	M	1,84	1,84	1,78	1,62	1,62	1,60	1,70	1,61	*
		s	0,73	0,79	0,70	0,70	0,72	0,70	0,78	0,70	
	LERNERFO	M	2,26	2,39	2,27	* 2,04	1,97	2,02	1,98	2,09	*
		s	0,66	0,72	0,65	0,55	0,60	0,64	0,74	0,66	
KOLLOQUI	M	1,56	1,61 *	1,43	1,39	1,37	1,49	1,55 *	1,35	*	
	s	0,64	0,69	0,54	0,55	0,50	0,76	0,72	0,64		
VORBEREI	M	1,38	1,52 *	1,34	1,31	1,32	1,30	1,39	1,28		
	s	0,56	0,55	0,48	0,47	0,50	0,48	0,60	0,49		
FREIWILL	M	2,74	2,87	2,81	* 2,62 *	2,41	2,57 *	2,30 *	2,53	*	
	s	0,82	0,78	0,80	0,71	0,85	0,76	0,77	0,84		
INTERESS	M	2,50 *	2,75	2,65	2,48	2,35	2,47 *	2,22 *	2,55		
	s	0,83	0,81	0,79	0,83	0,77	0,83	0,90	0,89		

\* signifikanter Mittelwertsunterschied zwischen den beiden Erhebungszeitpunkten

letzte Spalte signifikante Mittelwertsunterschiede zwischen erstem und letztem Erhebungszeitpunkt

M = arithmetischer Mittelwert; s = Standardabweichung

### 3.4.5.6.1 Grundlagenseminar

Von 1999 bis 2003 verbessert sich die Bewertung des Medieneinsatzes (MEDIEN) und der Veranschaulichung des Themas durch praxisnahe Beispiele (THEMA) signifikant. Auch die Einschätzungen, ob inhaltliche Fragen erwünscht waren (FRAGEN) und die inhaltlichen Fragen kompetent und verständlich beantwortet wurden (KOMPET1), verbessern sich über den gesamten Betrachtungszeitraum signifikant. Diese Veränderungen sind teilweise sogar zwischen zwei aufeinander folgenden Semestern signifikant. Dass der Anteil der Befragten, denen es leicht fiel, aufmerksam zu bleiben (AUFMERK1), in diesem Zeitraum ebenfalls signifikant zunahm, verwundert vor diesem Hintergrund dann nicht.

#### **3.4.5.6.2 Praktische Übungen**

Die Bewertung der praktischen Übungen verändert sich über den Untersuchungszeitraum nicht signifikant, obwohl sich das Verhältnis zwischen Lehrveranstaltung und Basiswissen signifikant verbessert hat. So haben im Sommersemester 2003 signifikant mehr Studierende ihr Basiswissen als ausreichend bezeichnet, als noch im Wintersemester 1999.

Inwieweit dies auf konzeptionelle und inhaltliche Veränderungen der Übungen oder auf ein verändertes Vorwissen der Studierenden zurückzuführen ist, kann hier allerdings nicht geklärt werden.

#### **3.4.5.6.3 Betriebsbegehung**

Über den gesamten Betrachtungszeitraum verbessert sich die Beurteilung der Betriebsbegehung bei allen vier Items signifikant.

#### **3.4.5.6.4 Gruppenarbeit**

Die Gruppenarbeit wird über den ganzen Zeitraum relativ konstant bewertet. Zwischen einzelnen Messzeitpunkten jedoch zeigen sich unsystematische, aber signifikante Veränderungen der Beurteilungen, die mal besser, aber auch wieder mal schlechter werden.

#### **3.4.5.6.5 Gesamtbewertung**

Die Gesamtbeurteilung der Lehrveranstaltung verbessert sich bei fast allen Indikatoren von 1999 bis 2003 signifikant. Diese Veränderungen fanden jedoch nicht sprunghaft zwischen zwei Erhebungszeitpunkten statt, sondern zeigen sich erst in der Gesamtbeurteilung. Lediglich die Bewertung der Strukturierung der Lehrveranstaltung weist eine Reihe signifikanter Veränderungen zwischen je zwei Messzeitpunkten auf. Für die Verbesserung der Gesamtbeurteilung über die Zeit können verschiedene Gründe verantwortlich sein: Erstens können diese Veränderungen in Modifikationen des strukturellen und didaktischen Konzeptes der Lehrveranstaltung begründet sein. Zweitens kann sich eine Konkurrenzsituation unter den Studierenden entwickelt oder weiterentwickelt haben. Diese könnte dazu geführt haben, dass Studierende effektiver studieren als noch vor einigen Jahren und deshalb an Lehrveranstaltungen aufmerksamer, konzentrierter und motivierter teilnehmen.

### 3.4.6 Regressionsanalysen

In einer weiteren Analyse wurde das Semester als unabhängige und die Bewertungsitems als abhängige Items regressionsanalytisch untersucht. Fragestellung dabei ist der langzeitige Effekt der Modifikation von Teilbereichen der Lehrveranstaltung über die Zeit auf die Bewertung der Studierenden. Untersucht wird dabei aber nur, inwieweit solche Modifikationen einen linearen Effekt auf die Bewertungen haben. Die entsprechenden Ergebnisse sind in Tabelle 15 dargestellt.

Tabelle 15: Erklärte Varianzen, standardisierte Regressionseffekte und Signifikanzniveaus der Effekte des Semesters auf die Bewertungsitems

	Label	r <sup>2</sup>	p <sub>a</sub>	r	p <sub>b</sub>		Label	r <sup>2</sup>	p <sub>a</sub>	r	p <sub>b</sub>
Grundlagenseminar	MEDIEN	0,027	0,000	-0,168	0,000	Praktische Übungen	BASISWIS	0,014	0,000	-0,122	0,000
	THEMA	0,051	0,000	-0,229	0,000		DIALOG	0,009	0,002	-0,098	0,002
	FRAGEN	0,016	0,000	-0,130	0,000		KOMPET2	0,000	0,488	-0,022	0,488
	KOMPET1	0,029	0,000	-0,173	0,000		LEHRINHA	0,002	0,097	-0,053	0,097
	AUFMERK1	0,029	0,000	-0,172	0,000		AUFMERK2	0,005	0,015	-0,078	0,015
Gesamtbeurteilung	STRUKTUR	0,038	0,000	-0,198	0,000	Betriebsbegehung	INHALT1	0,009	0,003	-0,100	0,003
	INHALT2	0,022	0,000	-0,151	0,000		ORGANIST	0,004	0,022	-0,074	0,022
	KENNTNIS	0,006	0,010	-0,085	0,010		KOMPET3	0,000	0,270	-0,035	0,270
	LERNERFO	0,027	0,000	-0,166	0,000		EINBLICK	0,008	0,003	-0,095	0,003
	KOLLOQUI	0,004	0,035	-0,067	0,035	Gruppenarbeit	DISKURSI	0,013	0,000	-0,116	0,000
	VORBEREI	0,005	0,018	-0,076	0,018		UNTERSTÜ	0,004	0,035	-0,068	0,035
	FREIWILL	0,031	0,000	-0,180	0,000		SPAß	0,010	0,001	-0,105	0,001
	INTERESS	0,008	0,003	-0,095	0,003						

r<sup>2</sup> = erklärte Varianz; p<sub>a</sub> = Signifikanzniveau für r<sup>2</sup>; r = standardisierter Regressionskoeffizient; p<sub>b</sub> = Signifikanzniveau für r

### **3.4.6.1 Grundlagenseminar**

Sämtliche Bewertungsaspekte des Grundlagenseminars verbessern sich signifikant über den Untersuchungszeitraum. Alle Regressionskoeffizienten weisen negative Vorzeichen auf. Dies ist inhaltlich gleichbedeutend mit einer Verbesserung der Bewertung. Der Einfluss des Untersuchungszeitraums auf die Items variiert zwischen 1,6% und 5,1%.

### **3.4.6.2 Praktische Übungen**

Bei drei Aspekten Basiswissen, Dialog, Aufmerksamkeit (BASISWIS, DIALOG, AUFMERK2) hat der Untersuchungszeitraum einen signifikanten bewertungsverbessernden Effekt. Der Einfluss des Untersuchungszeitraums variiert dabei zwischen 0,5% und 1,4%. Bei den verbleibenden zwei Items KOMPET2 und LEHRINHA ergibt sich kein signifikanter Zusammenhang zwischen Untersuchungszeitraum und Bewertung. Im Hinblick auf die Vorzeichen der Regressionskoeffizienten zeigt sich tendenziell aber die gleiche Entwicklungsrichtung wie bei den drei übrigen Items.

### **3.4.6.3 Betriebsbegehung**

Die Bewertung der Betriebsbegehung verbessert sich bei drei Items (INHALT1, ORGANIST, EINBLICK) über die Zeit signifikant. Der Einfluss des Untersuchungszeitraums liegt dabei zwischen 0,4% und 0,9%. Bei dem vierten Item (KOMPET3) hat der Untersuchungszeitraum zwar keinen signifikanten Effekt, das Vorzeichen des Regressionskoeffizienten weist aber in die gleiche Richtung wie bei den übrigen Items und damit auf eine Verbesserung der Bewertung hin.

### **3.4.6.4 Gruppenarbeit**

Auf die drei Indikatoren zur Bewertung der Gruppenarbeit hat der Untersuchungszeitraum einen signifikanten Einfluss. So verbessert sich die Bewertung über die Zeit. Der Einfluss der Zeit auf die Bewertung der Gruppenarbeit variiert je nach Item zwischen 0,4% und 1,3%.

### **3.4.6.5 Gesamtbewertung**

Auf alle acht Indikatoren zur Gesamtbeurteilung der Lehrveranstaltung hat der Untersuchungszeitraum einen signifikanten Effekt. In allen acht Aspekten verbessert sich die Bewertung der Lehrveranstaltung insgesamt über die Zeit. Dabei beträgt der Einfluss der Zeit auf die Bewertung zwischen 0,4% und 3,8%.

Zusammenfassend zeigt diese Analyse somit, dass sich die Bewertung der Lehrveranstaltung insgesamt sowie in ihren Teilaspekten über den Untersuchungszeitraum fast durchgängig verbessert hat.

### **3.4.7 Explorative Faktorenanalyse**

Das Ziel der Faktorenanalyse ist es, aus einer Zahl von Variablen die voneinander unabhängigen Einflussfaktoren herauszukristallisieren und damit einen Beitrag zu liefern, die von untereinander unabhängigen Beschreibungs- und Erklärungsvariablen zu finden. Dieses Verfahren ermöglicht somit durch die Variablenreduktion eine Erleichterung empirischer Arbeit.

Aus messtheoretischer Sicht ist die Messung von Einstellungen über einzelne Items immer mit Messfehlern behaftet. Um diese Messfehler messtheoretisch in den Griff zu bekommen, bietet sich alternativ die Messung einer Einstellung über mehrere Items an. Die Gemeinsamkeiten des Antwortverhaltens auf diese Items können dann als eine präzisere Annäherung an den wahren Wert betrachtet werden.

Sind solche Gemeinsamkeiten vorhanden, können die Items zu einer Skala zusammengefasst werden. Aus diesem Grunde ist es sinnvoll zu prüfen, ob die verwendeten Items empirische Gemeinsamkeiten – systematische Korrelationen – aufweisen und somit als verschiedene Indikatoren des gleichen theoretischen Einstellungskonstruktes betrachtet werden können.

Dabei stellt sich die Frage, inwieweit z. B. die fünf Indikatoren zur Bewertung der Grundlagenseminare das gleiche messen oder jedes für sich die Bewertung eines spezifischen Teilbereiches der Grundlagenseminare. Ein geeignetes statistisches Verfahren zur Untersuchung dieser Fragestellung ist die Faktorenanalyse.

Dabei unterscheidet man zwischen explorativen (= strukturentdeckenden) und konfirmatorischen (= strukturprüfenden) Faktorenanalysen. Während die explorative Faktorenanalyse nach Gemeinsamkeiten im Antwortverhalten (Korrelationen) sucht, werden bei der konfirmatorischen Faktorenanalyse messtheoretisch a-priori postulierte Gemeinsamkeiten auf ihre empirische Gültigkeit geprüft. Im Folgenden werden zunächst die Ergebnisse der explorativen Faktorenanalysen vorgestellt, anschließend die Ergebnisse der konfirmatorischen Prüfung.

Alle fünf Skalen wurden zunächst mittels explorativer Faktorenanalysen auf ihre Dimensionalität und damit auf ihre interne Validität überprüft. Die Berechnungen wurden mit dem Statistikprogramm SPSS 10.0 durchgeführt.

Da das Antwortverhalten auf alle Items naturgemäß mit Messfehlern behaftet ist, wurde die Hauptachsenanalyse als faktorenanalytisches Verfahren ausgewählt. Sie berücksichtigt im Gegensatz zur Hauptkomponentenanalyse die Annahme von Messfehlern (Backhaus et al. 2003).

Die Zahl der Faktoren wurde jeweils über das Eigenwertkriterium bestimmt. Anhand solcher Analysen wird geprüft, inwieweit die jeweiligen Items mit welcher Qualität die entsprechende theoretisch postulierte Dimension abbilden. Deshalb wird jede dieser fünf Skalen zunächst getrennt analysiert, um die Frage der internen Validität zu klären. Anschließend werden sämtliche Items einer simultanen Faktorenanalyse unterzogen. Damit wird zum einen untersucht, inwieweit die Skalen auch gegeneinander trennscharf sind. Zum anderen liefert die Analyse der Faktorenkorrelationen Hinweise auf die externe Validität der Skalen.

#### **3.4.7.1 Grundlagenseminare**

Die Faktorenanalyse ergab eine eindimensionale Lösung. Der extrahierte Faktor hat einen Eigenwert von 1,738 und durch den Faktor werden 34,77% der Varianz der Items erklärt. Die Faktorenladungen sind mit Werten über .5 ausreichend hoch (vgl. Tab. 16). Das Cronbachs Alpha dieser Skala beträgt 0,72 und ist damit auch ausreichend hoch. Aufgrund dieses Ergebnisses ist von einer ausreichend hohen internen Validität der Skala auszugehen.

Tabelle 16: Faktorenladungen der Dimension „Grundlagenseminare“

Itemlabel	Faktorenladung
MEDIEN	0,52
THEMA	0,57
FRAGEN	0,59
KOMPET1	0,64
AUFMERK1	0,62

### 3.4.7.2 Praktische Übungen

Auch für die Skala „praktische Übungen“ ergab die Faktorenanalyse eine eindimensionale Lösung (vgl. Tab. 17). Der Eigenwert des Faktors beträgt 1,876 und es werden 37,51% der Itemvarianzen durch den Faktor erklärt. Bis auf ein Item liegen die Faktorenladungen über 0,5. Lediglich das Item zum Basiswissen (BASISWIS) weist eine sehr geringe Faktorenladung auf, denn Faktorenladungen unter 0,3 werden aus Konvention als sehr gering betrachtet (Backhaus et al. 2003).

Aus inhaltlicher Sicht ist dieses Ergebnis plausibel, weil sich das Item zum Basiswissen (BASISWIS) auf eine individuelle Voraussetzung zur erfolgreichen Teilnahme an den praktischen Übungen bezieht, während die übrigen Items sich stärker auf die inhaltliche und didaktische Gestaltung der Veranstaltung beziehen. Die Reliabilitätsanalyse dieser Skala unter Berücksichtigung aller Items ergibt ein Cronbachs Alpha von 0,72. Dieser Wert erhöht sich auf 0,75, wenn das Item zum Basiswissen (BASISWIS) nicht berücksichtigt wird. Da sich bei Verzicht auf dieses Item die Reliabilität der Skala erhöht und zudem die Faktorenladung dieses Items sehr gering ist, wird es in den folgenden Analysen nicht weiter berücksichtigt.

Tabelle 17: Faktorenladungen der Dimension „praktische Übungen“

Itemlabel	Faktorenladung
BASISWIS	0,29
DIALOG	0,57
KOMPET2	0,56
LEHRINHA	0,76
AUFMERK2	0,76

Eine Faktorenanalyse ohne Berücksichtigung des Items zum Basiswissen (BASISWIS) kommt zu dem in Tabelle 18 dargestellten Ergebnis. Der Eigenwert beträgt nun 1,807 und der Anteil der erklärten Varianz erhöht sich auf 45,17%. Ebenso sind nun alle Faktorenladungen ausreichend hoch (vgl. Tab. 18). Wie oben bereits angedeutet, erhöht sich Cronbachs Alpha für diese Skala auf 0,75.

Tabelle 18. Faktorenladungen der Dimension „praktische Übungen“ ohne das Item Basiswissen

Itemlabel	Faktorenladung
DIALOG	0,54
KOMPET2	0,54
LEHRINHA	0,81
AUFMERK2	0,75

### 3.4.7.3 Betriebsbegehung

Für die vier Items zur Beurteilung der Betriebsbegehung ergibt die Faktorenanalyse ebenfalls eine eindimensionale Lösung. Der Eigenwert des Faktors beträgt hier 2,316 und der durch den Faktor erklärte Varianzanteil der Items liegt bei 57,89%. Die Faktorenladungen der vier Items sind ausreichend hoch (vgl. Tab. 19). Das Reliabilitätsmaß Cronbachs Alpha beträgt 0,84 und ist damit ausreichend hoch.

Tabelle 19: Faktorenladungen der Dimension „Betriebsbegehung“

Itemlabel	Faktorenladung
INHALT1	0,83
ORGANIST	0,77
KOMPET3	0,77
EINBLICK	0,67

### 3.4.7.4 Gruppenarbeit

Drei Items wurden zur Beurteilung der Gruppenarbeit verwendet. Die eindimensionale Lösung basiert auf einem Eigenwert des Faktors von 1,452 und einem erklärten Varianzanteil von 48,39%. Die Faktorenladungen und damit die interne Validität sind ausreichend hoch (vgl. Tab. 20). Cronbachs Alpha beträgt 0,72. Die Reliabilität der Skala ist damit ebenfalls ausreichend hoch.

Tabelle 20: Faktorenladungen der Dimension „Gruppenarbeit“

Itemlabel	Faktorenladung
DISKURSI	0,75
UNTERSTÜ	0,58
SPAß	0,75

### 3.4.7.5 Gesamtbewertung

Die Skala zur Erfassung der Gesamtbeurteilung der Lehrveranstaltung umfasst acht Items. Sie ist faktorenanalytisch eindimensional. Der Eigenwert des Faktors beträgt 2,965 und dieser Faktor erklärt 37,06% der Gesamtvarianz. Die Faktorenladungen der Items sind in Tabelle 21 dargestellt. Alle Faktorenladungen sind ausreichend hoch. Cronbachs Alpha für diese Skala beträgt 0,82. Beide Befunde deuten auf eine ausreichend hohe Reliabilität und interne Validität der Skala hin.

Tabelle 21: Faktorenladungen der Dimension „Gesamtbeurteilung“

Itemlabel	Faktorenladung
STRUKTUR	0,61
INHALT2	0,66
KENNTNIS	0,58
LERNERFO	0,75
KOLLOQUI	0,50
VORBEREI	0,52
FREIWILL	0,62
INTERESSE	0,61

### 3.4.7.6 Externe Validierung

Die externe Validität der fünf Skalen wurde ebenfalls faktorenanalytisch untersucht. Ausgehend von der grundsätzlichen Annahme, dass die fünf Skalen unterschiedliche Dimensionen bzw. Aspekte der Lehrveranstaltung trennscharf messen, müssten sich bei einer entsprechenden Faktorenanalyse diese fünf Dimensionen als fünf Faktoren zeigen.

Die Berechnung einer solchen Analyse ohne Berücksichtigung des Items zum Basiswissen (BASISWIS) ergibt jedoch nach dem Kaiserkriterium nur zwei Faktoren mit Eigenwerten größer als 1. Für den ersten Faktor beträgt der Eigenwert 6,836 und er erklärt 28,48% der Varianz. Der Eigenwert des zweiten Faktors beträgt 1,755 und er erklärt weitere 7,31% der Varianz. Beide Faktoren korrelieren positiv mit 0,40. Die entsprechenden Faktorenladungen sind in Tabelle 22 dargestellt.

Tabelle 22: Faktorenladungen aller Items für die zweifaktorielle Lösung

	Itemlabel	Faktor 1	Faktor 2		Itemlabel	Faktor 1	Faktor 2
Grundlagen- seminar	MEDIEN	<b>0,49</b>	0,00	praktische Übungen	DIALOG	<b>0,50</b>	-0,00
	THEMA	<b>0,52</b>	-0,00		KOMPET2	<b>0,51</b>	0,00
	FRAGEN	<b>0,42</b>	0,01		LEHRINHA	<b>0,66</b>	-0,01
	KOMPET1	<b>0,46</b>	0,01		AUFMERK2	<b>0,67</b>	-0,12
	AUFMERK1	<b>0,70</b>	-0,01				
Gruppen- arbeit	DISKURSI	<b>0,47</b>	0,00	Betriebs- begehung	INHALT1	-0,00	<b>0,85</b>
	UNTERSTÜ	<b>0,43</b>	0,01		ORGANIST	-0,00	<b>0,77</b>
	SPAß	<b>0,52</b>	0,01		KOMPET3	0,00	<b>0,73</b>
Gesamt- beurteilung	STRUKTUR	<b>0,63</b>	0,00		EINBLICK	0,15	<b>0,58</b>
	INHALT2	<b>0,66</b>	0,00				
	KENNTNIS	<b>0,51</b>	0,00				
	LERNERFO	<b>0,72</b>	-0,01				
	KOLLOQUI	<b>0,44</b>	0,01				
	VORBEREI	<b>0,52</b>	0,01				
	FREIWILL	<b>0,60</b>	-0,00				
	INTERESS	<b>0,60</b>	-0,01				

fett = höchste Faktorenladung des Items

Dieser Befund widerspricht zunächst den messtheoretischen Annahmen. Alle Items mit Ausnahme der Items zur Evaluation der Betriebsbegehung laden auf dem ersten Faktor. Der zweite Faktor bildet inhaltlich die Bewertung der Betriebsbegehung ab. Aus messtheoretischer Sicht ist dies jedoch nicht verwunderlich. Erstens wurden zur Evaluation der verschiedenen Teilaspekte der Lehrveranstaltung semantisch identische Indikatoren verwendet (z.B. AUFMERK1, AUFMERK2, KOMPET1, KOMPET2, KOMPET3).

Da die in diesen Items abgefragten Aspekte eher genereller Natur sind, ist ihr Stellenwert in den faktorenanalytischen Befunden durch die Mehrfachverwendung überhöht. Zweitens berücksichtigt diese Analyse die Gesamtbeurteilung der Lehrveranstaltung nicht als Aggregat der verschiedenen Skalen, sondern als eigenständige Operationalisierung.

Da sich die Gesamtbeurteilung der Lehrveranstaltung aber sowohl aus der explizit gemessenen Gesamtbeurteilung wie auch aus den Bewertungen der Teilaspekte (Grundlagenseminar, Gruppenarbeit, praktische Übungen) zusammensetzt, ist bei dieser analytischen Vorgehensweise die Trennschärfe der Skalen empirisch nicht mehr gegeben.

Die Betriebsbegehung fällt aus diesem Schema deshalb etwas heraus, weil sie nicht so sehr nur durch die Dozenten und Studierenden strukturiert ist, sondern externe Einflüsse der Betriebsumgebung hinzukommen, die nicht in der Handlungsautonomie der Dozenten und Studierenden liegen. Analysiert man den kompletten Itempool faktorenanalytisch unter der Vorgabe, fünf Faktoren zu bilden, so führt dies zu dem in Tabelle 23 dargestellten Befund.

Tabelle 23: Faktorenladungen aller Items für die fünffaktorielle Lösung

	Itemlabel	FAKTOR 1	FAKTOR 2	FAKTOR 3	FAKTOR 4	FAKTOR 5
Grundlagen-seminare	MEDIEN	0,13	0,00	<b>0,39</b>	-0,00	0,01
	THEMA	0,18	-0,01	<b>0,44</b>	0,01	0,11
	FRAGEN	0,00	0,01	<b>0,63</b>	-0,00	-0,01
	KOMPET1	-0,01	0,00	<b>0,72</b>	0,00	0,00
	AUFMERK1	<b>0,40</b>	-0,00	0,19	-0,18	0,01
Praktische Übungen	DIALOG	-0,00	0,00	0,16	<b>-0,40</b>	0,13
	KOMPET2	-0,01	0,01	<b>0,44</b>	-0,37	-0,00
	LEHRINHA	0,01	0,00	-0,00	<b>-0,69</b>	0,10
Betriebsbegehung	AUFMERK2	0,21	0,00	-0,01	<b>-0,74</b>	0,00
	INHALT1	0,00	<b>0,87</b>	-0,00	-0,00	-0,00
	ORGANIST	-0,00	<b>0,77</b>	0,00	0,00	0,00
	KOMPET3	-0,00	<b>0,76</b>	-0,01	-0,01	-0,00
Gruppenarbeit	EINBLICK	0,01	<b>0,58</b>	0,01	0,00	0,01
	DISKURSI	-0,01	0,00	-0,00	-0,01	<b>0,72</b>
	UNTERSTÜ	0,01	0,01	0,00	0,00	<b>0,51</b>
Gesamtbewertung	SPAß	-0,00	0,00	-0,01	-0,01	<b>0,76</b>
	STRUKTUR	0,16	0,00	<b>0,45</b>	-0,00	0,17
	INHALT2	0,29	0,00	0,16	-0,00	<b>0,36</b>
	KENNTNIS	<b>0,52</b>	0,01	0,01	-0,00	-0,00
	LERNERFO	<b>0,67</b>	0,00	0,01	-0,01	0,01
	KOLLOQUI	0,17	0,01	<b>0,26</b>	-0,00	0,01
	VORBEREI	0,12	0,01	<b>0,38</b>	-0,01	0,01
	FREIWILL	<b>0,62</b>	0,00	-0,00	-0,01	0,00
INTERESS	<b>0,67</b>	0,00	-0,00	-0,00	-0,00	
Faktorenkorrelationen						
	FAKTOR 1	1,00				
	FAKTOR 2	0,30	1,00			
	FAKTOR 3	0,53	0,35	1,00		
	FAKTOR 4	-0,49	-0,21	-0,44	1,00	
	FAKTOR 5	0,59	0,34	0,38	-0,44	1,00
	Eigenwert	6,94	1,78	1,00	0,66	0,56
	Erklärte Varianz	28,9 %	7,4 %	4,2 %	2,8 %	2,4 %

fett = höchste Faktorenladung des Items

Wie oben bereits dargestellt, erfüllen nur zwei der fünf Faktoren das Kaiserkriterium mit Eigenwerten über 1 und sind somit im strengen Sinne interpretierbar. Die Items zur Evaluation der Grundlagenseminare laden mit einer Ausnahme (AUFMERK1) in diesem Modell auf einem Faktor (Faktor 3). Das Item zur kompetenten Beantwortung von Fragen (KOMPET1) weist dabei die höchste, das Item zum Medieneinsatz (MEDIEN) die geringste Faktorenladung auf.

Ähnlich sind die Ergebnisse für die vier Items zur Beurteilung der praktischen Übungen. Drei dieser Items (DIALOG, LEHRINHA, AUFMERK2) laden auf einem Faktor (Faktor 4), wobei das Item zur „Erarbeitung der Lehrinhalte im Dialog“ (DIALOG) die niedrigste und das Item zur „Aufmerksamkeit“ (AUFMERK2) die höchste Faktorenladung aufweist. Das Item zur „kompetenten Beantwortung von Fragen“ (KOMPET2) jedoch differenziert entgegen den messtheoretischen Annahmen nicht eindeutig zwischen den Faktoren, da es auf zwei Faktoren (Faktor 3 und Faktor 4) fast gleich hohe Faktorenladungen aufweist.

Die vier Items zur Beurteilung der Betriebsbegehung weisen eine deutliche eindimensionale Struktur auf. Sie laden alle auf dem Faktor 2 und ihre Faktorenladungen sind ausreichend hoch. Damit sind sie im Hinblick auf die hier analysierte Faktorenstruktur trennscharf. Gleiches gilt für die drei Items zur Evaluation der Gruppenarbeit (Faktor 5). Die Skala für die Gesamtbeurteilung der Seminare besitzt faktorenanalytisch keine eindeutige Struktur. Die Faktorenladungen der Items verteilen sich auf drei Faktoren.

Aufgrund der Zuordnung der Items zu den einzelnen Faktoren lassen sich diese wie folgt charakterisieren: Faktor 1 umfasst Items zur eigenen Teilnahmemotivation und zum Lernerfolg. Faktor 2 beschreibt die Evaluation der Betriebsbegehung, Faktor 3 die Struktur und die didaktischen Methoden der Veranstaltung insgesamt. Faktor 4 bezieht sich auf die praktischen Übungen und Faktor 5 auf die Evaluation der Gruppenarbeit. Wie die Korrelation der Faktoren zeigt (vgl. Tabelle 23), hängen Motivation und Lernerfolg (Faktor 1) besonders mit den didaktischen Methoden (Faktor 3) und der Gruppenarbeit (Faktor 5) zusammen. Die Betriebsbegehung hingegen wird relativ unabhängig von den anderen Dimensionen bewertet. Damit ist auch hier die Betriebsbegehung relativ unabhängig von den übrigen Bewertungen der Lehrveranstaltung. Zur Begründung wurde bereits oben auf die Fremdstrukturierung der Betriebssituation hingewiesen.

### 3.4.8 Konfirmatorische Faktorenanalysen

Diese explorative Analyse der faktoriellen Struktur aller Items berücksichtigt zwar grundsätzlich die Annahme von Messfehlern. Insbesondere systematische Messfehler werden nicht explizit berücksichtigt. Gerade diese jedoch können weitere Hinweise auf messtheoretische Probleme der Items und deren Trennschärfe liefern (Backhaus et al 2003). Deshalb wurden die Items einer weiteren Analyse unterzogen

Tabelle 24: Signifikante Faktorenladungen des konfirmatorischen Messmodells

Itemlabel	Grundlagenseminar (GRUNDLA)	praktische Übung (PRAKTÜB)	Betriebsbegehung (BETRIEB)	Gruppenarbeit (GRUPARB)	Gesamtbeurteilung (GESABEUR)	erklärte Varianz
MEDIEN	0,54					0,29
THEMA	0,55					0,30
FRAGEN	0,48					0,23
KOMPET1	0,51					0,26
AUFMERK1	0,72					0,52
DIALOG		0,51				0,26
KOMPET2		0,51				0,26
LEHRINHA		0,82				0,67
AUFMERK2		0,78				0,61
INHALT1			0,85			0,73
ORGANIST			0,76			0,58
KOMPET3			0,74			0,55
EINBLICK			0,56		0,18	0,43
DISKURSI				0,69		0,47
UNTERSTÜ				0,60		0,36
SPAß				0,75		0,56
STRUKTUR					0,63	0,40
INHALT2					0,71	0,50
KENNTNIS					0,49	0,24
LERNERFO					0,70	0,48
KOLLOQUI					0,46	0,22
VORBEREI					0,57	0,33
FREIWILL					0,57	0,32
INTERESS					0,55	0,30
Faktorenkorrelationen (Subdiagonale) und –varianzen (Diagonale)						
Grundlagenseminar	0,09					
Prakt. Übung	0,69	0,31				
Betriebsbegehung	0,35	0,24	0,30			
Gruppenarbeit	0,58	0,57	0,34	0,33		
Gesamtbeurteilung	0,89	0,74	0,38	0,72	0,15	

Tabelle 25: Residuen (Diagonale) und Residualkorrelationen (Subdiagonale) des konfirmatorischen Messmodells

	Grundlagenseminar					praktische Übungen				Betriebsbegehung				Gruppenarbeit			Gesamtbeurteilung									
Item	MEDIEN	THEMA	FRAGEN	KOMPET1	AUFMERK1	DIALOG	KOMPET2	LEHRINHA	AUFMERK2	INHALT1	ORGANIST	KOMPET3	EINBLICK	DISKURSI	UNTERSTÜ	SPAß	STRUKTUR	INHALT2	KENNTNIS	LERNERFO	KOLLOQUI	VORBEREI	FREIWILL	INTERESS		
Grundlagenseminar																										
MEDIEN	.20																									
THEMA	.12	.28																								
FRAGEN			.32																							
KOMPET1			.34	.25																						
AUFMERK1					.26																					
praktische Übungen																										
DIALOG						.38																				
KOMPET2			.20	.27		.18	.24																			
LEHRINHA			-.13	-.15				.16																		
AUFMERK2		-.11			.28				.19																	
Betriebsbegehung																										
INHALT1										.15																
ORGANIST											.20															
KOMPET3												.25														
EINBLICK													.30													
Gruppenarbeit																										
DISKURSI						.11								.28												
UNTERSTÜ															.47											
SPAß																.26										
Gesamtbeurteilung																										
STRUKTUR	.19	.12															.22									
INHALT2																	.14	.24								
KENNTNIS																			.37							
LERNERFO							.09												.22	.23						
KOLLOQUI																	.09	-.12	.08		.32					
VORBEREI					.13	-.17	.16														.17	.18				
FREIWILL																							.45			
INTERESS																			.15	.13				.31	.50	

Im Unterschied zum explorativen Modell (vgl. Tabelle 23) weisen die Items im konfirmatorischen Modell eine eindeutige Faktorenstruktur im Sinne der theoretischen Zuordnung im Fragebogen auf (vgl. Tabelle 24). Lediglich das Item zum „Einblick in Gesundheitsgefahren am Arbeitsplatz“ (EINBLICK) besitzt noch eine Faktorenfremdla-

dung auf dem Faktor ‚Gesamtbeurteilung‘. Alle Items weisen befriedigend hohe Faktorenladungen auf und sind trennscharf zwischen Faktoren. Für die Faktorenladungen bedeutet dies, dass Fremdladungen wie im explorativen Modell hier über die Schätzung systematischer Messfehler modelliert werden. Diese sind numerisch jedoch nicht besonders hoch und damit aus inhaltlicher Sicht zu vernachlässigen (vgl. Tabelle 25). Es sei deshalb hier lediglich auf den Zusammenhang zwischen der Bereitschaft zur freiwilligen Teilnahme an der Lehrveranstaltung und dem geweckten Interesse an Arbeitsmedizin hingewiesen (FREIWILL, INTERESS), der nicht auf die Gesamtbeurteilung der Veranstaltung zurückzuführen ist. Vielmehr zeigt sich hier auch ein grundsätzlicher Zusammenhang von Motivation, Interesse und Freiwilligkeit.

Die explorative Analyse hat Hinweise darauf geliefert, dass sich die Items auch nach anderen als im Fragebogen vorgegebenen Zuordnungen inhaltlich gruppieren lassen. Die konfirmatorische Analyse zeigt hingegen, dass auch eine valide Interpretation der Items im Sinne ihrer theoretischen a-priori-Zuordnung empirisch haltbar ist. Dennoch verweist die Unterschiedlichkeit der Befunde darauf, dass eine Überarbeitung der Items und Skalen sinnvoll sein könnte.

## 4. Diskussion

Die ausführlich dargestellten Ergebnisse der studentischen Lehrevaluation im Zeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 werden in ihren wesentlichen Teilaspekten zusammengefasst und diskutiert. Dabei wird der eingangs formulierten Zielsetzung gefolgt:

1. den Erfolg des neuen Lehr- und Lernansatzes und
2. die Validität des Evaluationsverfahrens zu überprüfen

Danach wird die Diskussion in folgende Kapitel gegliedert:

- Bewertung struktureller und didaktischer Elemente (Kompetenz der Lehrenden, Betriebsbegehung, praktische Übungen, Gruppenarbeit)
- Einfluss dieser Bewertung auf Teilnahmemotivation und Lernerfolg
- Bewertung des Evaluationsverfahrens

Abschließend wird diskutiert, ob die kontinuierliche Evaluation als Beitrag zur Qualität der Lehre in der Arbeitsmedizin bewertet werden kann.

## **4.1 Bewertung struktureller und didaktischer Elemente**

### **4.1.1 Kompetenz der Lehrenden**

Zu den für erfolgreiche Lehre verantwortlichen Lehrmethoden zählen nach Dunkin und Barns (1986) insbesondere solche, die sich durch Strukturierung und Klarheit auszeichnen. In vielen Studien und experimentellen Untersuchungen konnte ein inhaltlich bedeutsamer Zusammenhang zwischen klarer Strukturierung der Lehre und studentischem Lerneffekt festgestellt werden (Patton 1981, Dunkin und Barns 1986).

Auch Kromrey (1996) sieht einen inhaltlich bedeutsamen Zusammenhang zwischen klarer Strukturierung der Lehre und studentischem Lerneffekt. Zu den weiteren Merkmalen auf Seiten des Dozentenverhaltens gehören darüber hinaus auch verständliche Erklärungen und Darstellungen, rhetorische Kompetenz, Motivierung und die Darstellung der Relevanz oder der Theorie-Praxis-Bezüge. Verarbeitungstiefe, Zeit- und Schwierigkeitsmanagement, Engagement, Kooperativität sowie in Seminaren das Leiten von Diskussionen ist ein weiterer Aspekt der Handlungskompetenz (Regehr und Norman 1996).

Nach Schmidt und Moust (1995) wird von einem betreuenden Dozenten zudem erwartet, dass er einerseits auf die Studenten eingehen, sie motivieren und ermutigen kann, und andererseits, dass er fundiertes Wissen in Bezug auf die behandelten Themen besitzt. Die Studierenden sollen sich demnach nicht nur eine breite fachliche Wissensbasis erarbeiten, sondern gleichermaßen die Fähigkeit und Motivation zum lebenslangen Lernen erwerben (Kissel 1984, Maudsley 1999).

Darüber hinaus spielt die Entwicklung einer reflektierten Persönlichkeit eine wichtige Rolle, die es schließlich ermöglicht, das Gelernte situationsabhängig anzuwenden, es nötigenfalls zu verändern oder neue Handlungsalternativen zu generieren (Habeck 1995, Mandl et al. 2001).

Für die Handlungskompetenz der Lehrenden unter Einbeziehung der didaktischen Elemente und für die Methodik des Praxisprojektes Arbeitsmedizin konnten folgende wichtige Erkenntnisse gewonnen werden:

Mit großer Mehrheit (96,1%) sind die befragten Studierenden der Meinung, die Medien seien überwiegend bzw. völlig sinnvoll und effektiv eingesetzt worden. 88,1 Prozent der befragten Studierenden sind zudem der Überzeugung, dass das Thema überwiegend bzw. völlig durch praxisnahe Beispiele veranschaulicht wurde.

Das inhaltliche Fragen erwünscht waren wird von fast allen Befragten (92,7%) als überwiegend oder völlig zutreffend beschrieben. Ferner sind fast alle Befragten (94,7%) der Ansicht, dass ihre inhaltlichen Fragen überwiegend oder völlig kompetent und verständlich beantwortet wurden. 95,3% der Studierenden sind überwiegend bzw. völlig der Meinung, dass auch die inhaltlichen Fragen während der praktischen Übungen kompetent und verständlich beantwortet wurden.

Der Hauptteil (92%) der Studierenden sind völlig oder überwiegend der Ansicht, die Lehrveranstaltung sei gut und systematisch strukturiert gewesen. Fast alle befragten Studierenden (93,7%) waren ebenfalls der Ansicht, dass auch die Überprüfung des Lernerfolges durch das Abschlusskolloquium dem Lernziel angemessen war.

Ebenso sind fast alle Befragten (97,7%) der Auffassung, dass die Lehrpersonen völlig oder überwiegend gut vorbereitet waren.

Nach Rindermann 1997 hängt aber die Generalisierbarkeit studentischer Beurteilungen über Dozenten von spezifischen Untersuchungsbedingungen ab. Dazu zählen zum einen die Lehrmethoden des Dozenten, aber auch veranstaltungsspezifische Kriterien. Zu diesen veranstaltungsspezifischen Kriterien gehört die Möglichkeit der Studierenden sich zum Beispiel in Form von Referaten am Unterrichtsgeschehen zu beteiligen.

Im Rahmen des „Praxisprojektes Arbeitsmedizin“ geschieht dies zum Beispiel durch die Beteiligung der Studierenden an den speziellen Seminaren in Form eines Kurzreferates über ein projektbezogenes Thema.

Darüber hinaus beteiligen sich die Studierenden durch eine Präsentation ihres Projektes im Plenum aktiv am Unterrichtsgeschehen. Somit sind die studentischen Einschätzungen der Dozentenvariablen nach Rindermann nicht ausschließlich auf das Lehrverhalten zurückzuführen.

### 4.1.2 Betriebsbegehung

Die Vermittlung von Wissen im Kontext eines konkreten Praxisbezuges geschieht unter der Vorstellung, dass das Wissen in anderen realen Situationen später schneller aktiviert werden kann (Colliver 2000). Um diese Aktivierung herbeizuführen wurde im Rahmen des Praxisprojektes eine reale Problemsituation in Form eines Betriebes als Real Case gewählt. Es stehen hier somit komplexe und authentische Lernbedingungen zur Verfügung.

Insgesamt zeigen die Bewertungen der Betriebsbegehung, welchen Stellenwert der konkrete Praxisbezug bzw. das didaktische Element „Betriebsbegehung“ für die Studierenden hat.

77,1% der Befragten sind überwiegend oder völlig der Ansicht, dass die Betriebsbegehung inhaltlich gut vorbereitet war. Auch die gute vorbereitende Organisation der Betriebsbegehung spiegelt sich in den Beurteilungen der Studierenden wieder.

Die meisten (89,5%) der Befragten sind ebenfalls der Meinung, dass die Betriebsbegehung auch organisatorisch gut geplant war.

Die überwiegende Mehrheit (83,6%) der befragten Studierenden waren der Ansicht, dass inhaltliche Fragen während der Begehung überwiegend oder völlig kompetent und verständlich beantwortet wurden.

Für den persönlichen Lernerfolg ergab sich bei 88% der befragten Studierenden, dass die Betriebsbegehung ihnen einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft hat.

Da der Betrieb im „Praxisprojekt Arbeitsmedizin“ eine zentrale Position einnimmt und als Schlüssel zum Wissenserwerb dient, ist der zunächst hoch erscheinende inhaltliche und organisatorische Aufwand durchaus als gerechtfertigt zu betrachten.

Im Vergleich dazu zeigen sich an medizinischen Fakultäten in Großbritannien unterschiedliche Sichtweisen der Studierenden hinsichtlich einer Betriebsbegehung (Grime et al. 2003 und 2006).

Aufgrund des zu hoch bewerteten Zeitaufwandes im Verhältnis zum erlebten Nutzen für die Lehrinhalte stehen viele Studierende in London einer Begehung eher ablehnend gegenüber. Die Betriebsbegehung ist aber nicht wie im Praxisprojekt Arbeitsmedizin der Universität Münster als zentrales Element in das Curriculum eingebunden, sondern findet als zusätzliche Lehrveranstaltung statt.

Von den Lehrenden in Großbritannien wird diese Unterrichtsform aber für die Entwicklung des Interesses für das Fach Arbeitsmedizin ebenfalls als wichtig erachtet, und wird nach den erhobenen Evaluationsergebnissen weiter konzipiert. In Birmingham wird die Arbeitsplatzbegehung für die medizinische Ausbildung und spätere ärztliche Tätigkeit von den Lehrenden und Studierenden als bedeutsam empfunden, findet aber dort zunächst noch in geringem zeitlichem Umfang statt (Wynn et al 2003).

Auch in Singapore findet die Vermittlung von arbeitsmedizinischen Lehrinhalten in Kleingruppen und in Form von konkreten Fällen statt, und ist zudem eingebunden in allgemeinmedizinische Lehrinhalte und die Ausbildungsinhalte von Public Health (Koh et. al. 1995). Auch diese Form der Unterrichtsgestaltung findet bei den Studierenden großen Anklang.

Ein wesentliches Ziel des Praxisprojektes Arbeitsmedizin ist es, neben fachspezifischen Grundlagen auch fächerübergreifende Inhalte für die ärztliche Tätigkeit in anderen Fachgebieten zu vermitteln. Hierzu gehören insbesondere grundlegende Kenntnisse über Arbeitsbedingungen und gesundheitliche Belastungen am Arbeitsplatz, die jeder Arzt nach entsprechender Anamnese in diagnostische und therapeutische Maßnahmen einbeziehen muss.

Im Praxisprojekt Arbeitsmedizin werden diese Kenntnisse durch konkreten Praxisbezug im Rahmen einer Betriebsbegehung, die das zentrale Element der Lehrveranstaltung darstellt, exemplarisch vermittelt.

Dadurch ist es offensichtlich gelungen, die Studierenden auch zum Erwerb fächerübergreifender Grundlagenkenntnisse zu motivieren: über 80% der Studierenden gaben an, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für ihre spätere ärztliche Tätigkeit wichtig seien.

### **4.1.3 Praktische Übung**

Die überwiegende Mehrheit (77,2%) der Befragten ist der Auffassung, dass der Lernstoff der praktischen Übung überwiegend bzw. völlig im Dialog mit dem Dozenten erarbeitet wurde. Und auch für die Mehrheit (82,4%) der Befragten trifft es überwiegend bzw. völlig zu, dass die Lehrinhalte bei den praktischen Übungen interessant präsentiert wurden.

Auch durch die Bewertung der Aufrechterhaltung der Konzentration, lässt sich der Erfolg der Didaktik und Methodik der Lehrenden des Fachgebietes Arbeitsmedizin belegen. 77,5% der Studierenden fiel es im Rahmen der praktischen Übungen überwiegend oder völlig leicht, aufmerksam zu bleiben.

### **4.1.4 Gruppenarbeit**

Die Gruppenarbeit hat im Rahmen des projekt- und problemorientierten Lehransatzes des Fachgebietes Arbeitsmedizin eine besondere Bedeutung.

Aus didaktischer Sicht kommt hier im handlungsorientierten Unterricht, der Gruppenarbeit eine Schlüsselbedeutung zu, weil hier „Learning bei Doing“ stattfindet. Neu erworbene theoretische Kenntnisse werden umgehend praktisch angewendet (Pratt und Magill 1983, Gudjons 1998). In den Phasen einer Gruppenarbeit zum Beispiel erarbeiten sich Studierende in Gruppengrößen zwischen drei und sechs Mitgliedern eine Aufgabenstellung eigenverantwortlich und kooperativ. Meistens werden die Arbeitsergebnisse so aufbereitet, dass sie nach Beendigung der Arbeitsphase vor einem Plenum präsentiert werden können.

Gruppenarbeit wird in der Regel vom Dozenten geplant, belässt aber den Studierenden eigene Gestaltungsmöglichkeiten (Peterssen 1997). In ihrer Binnenstruktur lassen sich die Planungsphase, Erarbeitungsphase und Vorbereitungsphase für die Präsentation unterscheiden. Während der Arbeitsphase tritt der Dozent zurück und beschränkt sich auf das Beobachten, Beraten und das Bewerten der Qualität des Arbeitsprozesses.

Diese Arbeit erfordert ein neues Selbstverständnis der Lehrenden: Neben dem Erwerb pädagogischer Qualifikation muss der Dozent für die Projektarbeit seine Gruppenleiterkompetenzen stärken, Verantwortung abgeben, über sich selbst reflektieren und sich selbst motivieren können.

Aber auch der Lernende kann nicht mehr nur passiv rezipieren, sondern muss sich aktiv in den Gruppen- und Arbeitsprozess einbringen: Er muss sich selbst immer wieder motivieren, seine Zielvorstellungen mit denen der Gruppe abzustimmen, muss sich auf eine Lernform einlassen, die ihm ein Höchstmaß an Eigenverantwortung abverlangt. Er muss Kritikfähigkeit, Toleranz und Rücksichtnahme erwerben.

Gruppenarbeit muss gelernt werden, damit sie erfolgreich verläuft, die Voraussetzung für den Lernerfolg ist auch die Beherrschung elementarer Arbeitstechniken. Gruppenarbeit verlangt vom Dozenten Planungs- und Organisationskompetenz, von den Studierenden Einsicht, Übung und Routine. In der Planung muss sie im Gesamtablauf des Unterrichtes reflektiert werden. Zum Beispiel sollte überlegt werden, wie viel Anleitung im Frontalunterricht notwendig ist, damit die Gruppen produktiv arbeiten können.

In Unterrichtsgruppen, die hinsichtlich Alter, Lebenserfahrung, Vorbildung, Geschlecht und gesellschaftlichen Einstellungen unterschiedlich sind, können die Gruppenmitglieder besonders gut voneinander lernen (Maudsley 1999, Harden und Crosby 2000).

Gute Gruppenarbeit deckt vier Ebenen des erweiterten Lernbegriffs ab. Die Studierenden erarbeiten sich Wissen, lernen Wichtiges von Unwichtigem zu unterscheiden, wenden das Gelernte kommunikativ an und trainieren soziale Fähigkeiten wie Teamgeist, Rücksichtnahme und Toleranz.

Zudem sticht eine Gruppe bei der Bearbeitung eines konkreten Problems jedes noch so begabte Individuum aus. Zu diesem Ergebnis kommen amerikanische Forscher in einer aktuellen Studie. Sie stellten fest, dass Gruppen von drei, vier oder fünf Personen besser abschnitten als die jeweils besten Einzelpersonen und schreiben diese Leistung der Fähigkeit der Menschen zu, zusammenzuarbeiten (Laughlin et al.2006).

Auch im Rahmen des projekt- und problemorientierten Lehransatzes des Fachgebietes Arbeitsmedizin hat die Gruppenarbeit diese besondere Bedeutung erlangt.

Die besondere Bedeutung der Gruppenarbeit zeigt sich auch in der Evaluation des Praxisprojektes Arbeitsmedizin. Bei 68,8% der Befragten hat die Diskussion in der Gruppe völlig oder überwiegend zu einem vertieften Verständnis des Lerninhaltes geführt. Die Gruppenarbeit mit Dozentenunterstützung war für 71.6% der Studierenden eine völlig oder überwiegend wichtige Hilfe für die Projektpräsentation. 77,9% aller Befragten hat die Gruppenarbeit völlig oder überwiegend Spaß gemacht. Die Akzeptanz der Studierenden wird auch dadurch belegt, dass über 80% der Befragten die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit völlig oder überwiegend gut gefallen hat.

Die Auswertung der Evaluation des Praxisprojektes zeigte, dass fast durchgängig über alle Indikatoren der Lehrveranstaltung die Gruppe derjenigen, die die Gruppenarbeit positiv bewerten, alle übrigen evaluativen Aspekte signifikant positiver beurteilten, als die Gruppe derjenigen, die die Gruppenarbeit negativ einschätzen. Lediglich beim Item Basiswissen (BASISWIS) besteht kein signifikanter Unterschied zwischen beiden Gruppen.

Einen Einfluss der studentischen „Peers“ auf die Wahrnehmung und Beurteilung der Lehre durch die anderen Gruppenteilnehmer beschreibt auch Kromrey (1994). Einen deutlich positiven Effekt zeigt sich demnach, wenn die überwiegende Mehrheit der Gruppenmitglieder Interesse zeigt. Dies gilt für die Beurteilung der Lehrveranstaltung ebenso wie für die Selbsteinschätzung des eigenen Lernerfolgs oder die Bereitschaft, sich mit dem behandelten Lernstoff intensiver selbstständig zu beschäftigen. Studentische Verhaltensweisen stellen somit weitere Bedingungsvariablen für den Erfolg einer Gruppenarbeit dar.

## **4.2 Einfluss didaktischer und struktureller Elemente auf Teilnahmemotivation und persönlichen Lernerfolg**

Zu den für erfolgreiche Lehre verantwortlichen Lehrmethoden zählen insbesondere solche, die sich durch Strukturierung und Klarheit auszeichnen. In vielen Studien und experimentellen Untersuchungen konnte ein inhaltlich bedeutsamer Zusammenhang zwischen klarer Strukturierung der Lehre und studentischem Lerneffekt festgestellt werden (Patton 1981, Dunkin und Barns 1986).

Zu den Merkmalen auf Seiten des Dozentenverhaltens gehören darüber hinaus auch verständliche Erklärungen und Darstellungen, rhetorische Kompetenz, Motivierung und die Darstellung der Relevanz oder der Theorie-Praxis-Bezüge. Verarbeitungstiefe, Zeit- und Schwierigkeitsmanagement, Engagement, Kooperativität sowie in Seminaren das Leiten von Diskussionen sind weitere Aspekte der Handlungskompetenz (Regehr und Norman 1996).

Weitere Bedingungsvariablen für den persönlichen Lernerfolg stellen aber auch die studentischen Verhaltensweisen dar. Dazu zählen insbesondere die Vor- und Nachbereitung der Lehrinhalte, die veranstaltungsinterne Beteiligung sowie Diskussionen und Referate (Daniel 1995).

Zudem können Rahmenbedingungen, wie das gestellte Thema, die allgemeinen Anforderungen, Prüfungen, Pflichtmäßigkeit des Besuches, Veranstaltungstyp und Ähnliches als Kontextvariablen den Lehrerfolg wie auch die Veranstaltungsbewertung beeinflussen (Sturm 1994, Scholz 1995). Dozenten, Studierende und Kontextbedingungen wirken nicht nur direkt auf den Lehrerfolg, sondern üben über gegenseitige Einflussnahme auch indirekt Wirkung auf den Erfolg der Lehre aus.

Für das Praxisprojekt Arbeitsmedizin konnte anhand der explorativen Faktorenanalyse belegt werden, dass definierte Prozessfaktoren wie Klarheit und Strukturiertheit in Verbindung mit Engagement und teilnehmerorientierter Didaktik und Methodik, das Interesse und die Motivation der Studierenden fördern.

Die Korrelation der analysierten Faktoren untereinander zeigt, dass die eigene Teilnahmemotivation und der Lernerfolg (Faktor 1) besonders mit der Struktur, Didaktik und Lehrkompetenz (Faktor 3) und der Gruppenarbeit (Faktor 5) zusammenhängen

(Vergl. Kap. 3.4.7.6., Seite 124, Tabelle 23). Eine ausführliche Darstellung der Ergebnisse der Auswertung wird im Kapitel 3.4.7 aufgezeigt.

Zu einem weiteren Aspekt, der das Interesse und die Motivation für das Fachgebiet Arbeitsmedizin verdeutlicht, gehört die Einschätzung des persönlichen Lernerfolges. Mit der Frage des persönlichen Wissenszuwachses und der Zufriedenheit mit einer Lehrveranstaltung beschäftigen sich auch Rosemann und Schweer (1996).

Hierbei zeigte sich, dass Studierende mit wachsender Zufriedenheit auch ihren eigenen Lernerfolg positiver einschätzten. Rosemann und Mitarbeiter (1996) ziehen die geäußerte Studienzufriedenheit als Indikator für gute Lehre in Betracht. Auch Spiel und Gössler (2000) sowie die Arbeitsgruppe von Westermann (1996) sehen in der Zufriedenheit mit einer Lehrveranstaltung ebenfalls einen wesentlichen Indikator für gelungene Lehre.

Kromrey (1996) bezeichnet die Förderung von Interesse und Motivation der Studierenden, sich mit den Lehrinhalten auseinanderzusetzen, als die eigentlichen Kriterien für gute Lehre.

Im Untersuchungszeitraum hat sich die Bewertung der Lehrveranstaltung „Praxisprojekt Arbeitsmedizin“ durch die Studierenden kontinuierlich verbessert. Der persönliche Lernerfolg der Studierenden wurde statistisch signifikant immer besser eingeschätzt.

Einen großen persönlichen Lernerfolg durch das „Praxisprojekt Arbeitsmedizin“ haben 74,3% der Befragten völlig oder überwiegend festgestellt. Insbesondere dieses Item dokumentiert den Erfolg des neuen Lehr- und Lernkonzeptes.

### 4.3 Bewertung des Evaluationsverfahrens

Die studentische Lehrevaluation wird als nützlich betrachtet für die Hochschulleitungen, für Studierende in ihrer Kurswahl und als Feedback für Dozenten. Den starken Forderungen nach Lehrevaluation stehen kritische Argumente gegenüber, die hinterfragen, ob die gängige Lehrevaluation überhaupt relevante Dimensionen der Lehrqualität erfasst.

Tatsächlich zeigt die umfangreiche internationale Forschung zu Lehrevaluationsdaten, dass studentische Lehrveranstaltungsbewertungen systematischen Verzerrungen unterliegen, beispielsweise im Hinblick auf folgende Bias-Variablen (Kromrey 1994, Rindermann 1997):

- **Teilnahmegrund:**  
Pflichtveranstaltungen werden negativer bewertet als Wahlveranstaltungen.
- **Themenbeliebtheit:**  
Veranstaltungen mit populären Themen werden positiver bewertet als Veranstaltungen mit unpopulären Themen.
- **Studierverhalten:**  
Veranstaltungen, in denen Studierende als Kollektiv aktiver agieren, werden positiver bewertet als Veranstaltungen mit eher passivem Studierverhalten.

Marsh und Roche (1997) vertreten die Überzeugung, dass studentische Beurteilungen multidimensional, reliabel und stabil sowie relativ valide hinsichtlich vieler Indikatoren effektiver Lehre (z.B. Lerntests, Fremdbeurteilungen) sind und relativ unbeeinflusst von einer Vielzahl von potentiellen Verzerrungsvariablen (z.B. Geschlecht, Intelligenz) bleiben.

Zur Bestimmung der Reliabilität und Zuverlässigkeit, der Gütekriterien einer Messmethode studentischer Beurteilungen, können mehrere statistische Verfahren herangezogen werden, die unterschiedliche Implikationen für die Messgenauigkeit nach sich ziehen (Miles 2000). Die Generalisierbarkeit studentischer Beurteilungen über verschiedene Veranstaltungen oder Dozenten hängt von spezifischen Untersuchungsbedingungen ab.

Dazu zählen die Ähnlichkeit der Themen, die Veranstaltungsformen, die Lehrmethoden der Dozenten und die gewählten Skalen (Rindermann 1997).

Evaluation führt zu einer kontinuierlichen Verlaufs- und Qualitätskontrolle und damit auch zur Qualitätsentwicklung. Als Grundfragen der Qualitätsevaluation werden nach Mittag und Hager (2000) folgende drei Fragen angeführt:

- Bedeutsamkeit: Tun wir das Wichtige?
- Wirksamkeit: Tun wir es effektiv, wirksam?
- Effizienz: Tun wir es mit vernünftigem Aufwand?

Die Bedeutung der Evaluation des „Praxisprojektes „Arbeitsmedizin“ war (und ist), im Sinne der oben genannten Fragen die Ergebnisse der Evaluation als relevante Informationen für die Lehrenden zu betrachten. Dabei war es für die Lehrenden entscheidend, Profile der Stärken und Schwächen der Lehrveranstaltung herauszuarbeiten und unter Berücksichtigung der Evaluationsziele zu bewerten. Zu diesen Zielen gehörten insbesondere die Erfolgskontrolle, aber auch die Erhöhung der Akzeptanz und das Interesse der Studierenden für das Fachgebiet Arbeitsmedizin.

Die Evaluation durch die Studierenden schafft damit eine Basis für den kontinuierlichen Veränderungsprozess in der Lehre im Fachgebiet Arbeitsmedizin. Diese Evaluation dient damit als ein Beitrag zur Optimierung eines neuen Lehr- und Lehransatzes.

Nach Altrichter (1997) ist „Selbstevaluation ein unverzichtbares Merkmal professioneller Berufstätigkeit“. Ein Evaluationsverfahren der Selbstevaluation ist zweifellos die Lehrveranstaltungskritik durch die Studierenden.

Damit die Selbstevaluation auch tatsächlich zu den gewünschten Zielen führt, sind einige wichtige Punkte zu beachten. Dazu gehört eine hinreichende Anzahl von Studierenden, die konkrete Lehrveranstaltungen der Lehrenden beurteilen. Die Auswertung der Befragung muss in einem möglichst engen Zusammenhang zu der konkreten Lehrveranstaltung erfolgen. Zeitreihen, um die Entwicklung in den Lehrveranstaltungen messen zu können sind vorteilhaft, denn die über mehrere Einsatzperioden gewonnenen Einschätzungen ermöglichen auch den Vergleich der Ergebnisse und zeigen Verbesserungsaktivitäten durch veränderte Werte deutlich auf.

Die Evaluation des „Praxisprojektes Arbeitsmedizin“ berücksichtigt diese Voraussetzungen. Im Untersuchungszeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 haben insgesamt 1020 Studierende den Kurs absolviert. 981 ausgefüllte Evaluationsbögen wurden in die Auswertung einbezogen.

Die Ergebnisse der Lehrveranstaltungsevaluation wurden von den Lehrenden selbst rezipiert und reflektiert. In einer Befragung von Lehrenden in den USA zeigte sich ebenfalls, dass immerhin 45% regelmäßig ihre Lehrveranstaltungen auf der Basis des studentischen Feedbacks modifizierten (Allgäuer 1998).

Es erhebt sich nun die Frage, ob das hier verwendete Evaluationsverfahren des Fragebogens den methodischen Ansprüchen an Validität, Reliabilität und Objektivität genügt. Ein geeignetes statistisches Verfahren zur Untersuchung dieser Fragestellung ist die Faktorenanalyse. Das Ziel der Faktorenanalyse ist, aus einer Zahl von Variablen die voneinander unabhängigen Einflussfaktoren herauszukristallisieren und damit einen Beitrag zu liefern, die von untereinander unabhängigen Beschreibungs- und Erklärungsvariablen zu finden. Dabei unterscheidet man zwischen explorativen und konfirmatorischen Faktorenanalysen. Im Hinblick auf die Überprüfung der Reliabilität wurde der Reliabilitätskoeffizient Cronbachs Alpha als gebräuchlichster Koeffizient verwendet. Die Überprüfung der fünf Skalen: Grundlagenseminar, Praktische Übungen, Betriebsbegehung, Gruppenarbeit und die Gesamtbewertung zeigte nach der explorativen Faktorenanalyse und der Reliabilitätsanalyse folgendes Ergebnis:

Skala Grundlagenseminare: Eigenwert des extrahierten Faktors 1,738

34,77% Varianz der Items

Faktorladungen über 0,5

Cronbachs Alpha= 0,72

Aufgrund dieses Ergebnisses ist von einer ausreichend hohen internen Validität und Reliabilität der Skala auszugehen.

Skala Praktische Übungen : Eigenwert des extrahierten Faktors 1,876

37,51% Varianz der Items

Faktorladungen über 0,5

Ausnahme Basiswissen unter 0,3

Cronbachs Alpha= 0,72

Eine Faktorenanalyse ohne das Item Basiswissen zeigt folgendes Ergebnis:

Eigenwert des Extrahierten Faktors 1,807

45,17% Varianz der Items

Faktorladungen über 0,5

Cronbachs Alpha = 0,75

Das Item Basiswissen bezieht sich auf die individuellen Voraussetzungen der Studierenden für die erfolgreiche Teilnahme an den praktischen Übungen. Daher ist dieses Ergebnis aus inhaltlicher Sicht plausibel.

Skala Betriebsbegehung: Eigenwert des extrahierten Faktors 2,316

57,89% Varianz der Items

Faktorladungen über 0,6

Cronbachs Alpha = 0,84

Aufgrund dieses Ergebnisses ist hier ebenfalls von einer ausreichend hohen internen Validität und Reliabilität der Skala auszugehen.

Skala Gruppenarbeit: Eigenwert des extrahierten Faktors 1,452

48,39% Varianz der Items

Faktorladung über 0,5

Cronbachs Alpha = 0,72

Aufgrund dieses Ergebnisses ist ebenfalls von einer ausreichend hohen internen Validität und Reliabilität der Skala auszugehen.

Skala Gesamtbewertung: Eigenwert des extrahierten Faktors 2,965

37,06% Varianz der Items

Faktorladung gleich und größer 0,5

Cronbachs Alpha 0,82

Aufgrund dieses Ergebnisses ist auch bei dieser Skala von einer ausreichend hohen internen Validität und Reliabilität auszugehen.

Damit ist der Fragebogen im Hinblick auf die verwendeten Skalen und den dazugehörigen Items aufgrund dieser Ergebnisse als intern valide und reliabel zu bezeichnen.

Die externe Validität der Skalen Grundlagenseminar, Praktische Übungen, Betriebsbegehung, Gruppenarbeit und Gesamtbewertung wurde ebenfalls mit der Faktorenanalyse untersucht.

Durch die grundsätzliche Annahme, dass die genannten Skalen die Dimensionen der Lehrveranstaltung „Praxisprojekt Arbeitsmedizin“ trennscharf messen, müssten sich diese fünf Skalen auch als fünf Faktoren zeigen. Die Analyse zeigt aber nach dem Kaiserkriterium nur zwei Faktoren mit Eigenwerten größer als 1.

Faktor 1: Eigenwert 6,836

28,48% Varianz

Auf dem ersten Faktor laden alle Items mit Ausnahme der Items zur Evaluation der Betriebsbegehung.

Faktor 2: Eigenwert 1,755

7,31% Varianz

Der zweite Faktor bildet inhaltlich die Bewertung der Betriebsbegehung ab.

Die Ursache liegt darin begründet, dass zur Evaluation der verschiedenen Teilaspekte der Lehrveranstaltung semantisch identische Indikatoren verwendet wurden (z.B. INHALT 1, INHALT 2, KOMPET 1, KOMPET 2, KOMPET 3). Zudem berücksichtigt die Analyse die Gesamtbeurteilung des „Praxisprojektes Arbeitsmedizin“ nicht als Aggregat der verschiedenen Skalen, sondern als eigenständige Operationalisierung.

Die Trennschärfe der Skalen ist empirisch nicht mehr gegeben, da sich die Gesamtbeurteilung der Lehrveranstaltung sowohl aus der gemessenen Gesamtbeurteilung wie auch aus den Bewertungen der Teilaspekte (Grundlagenseminare, praktische Übungen, Betriebsbegehung, Gruppenarbeit) zusammensetzt.

Die Betriebsbegehung fällt aus diesem Schema etwas heraus, weil sie nicht nur durch die Dozenten und Studenten strukturiert ist, sondern auch externe Einflüsse hinzukommen, die nicht in der Handlungsautonomie der Dozenten und Studierenden liegen.

Die weitere Analyse des kompletten Itempools erfolgte unter der Vorgabe von fünf Skalen, fünf Faktoren zu bilden.

Diese ermittelten fünf Faktoren lassen sich aufgrund der Zuordnung der Items folgendermaßen charakterisieren:

Das Item der Grundlagenseminare:

- Es fiel mir leicht aufmerksam zu bleiben.

Die folgenden Items der Gesamtbewertung:

- Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind.
- Mein persönlicher Lernerfolg durch die Veranstaltung war groß.
- Ich hätte an der Veranstaltung auch freiwillig teilgenommen.
- Die Veranstaltung hat mein Interesse für das Fachgebiet Arbeitsmedizin geweckt.

laden auf einem Faktor, dem Faktor 1 und lassen sich aufgrund dieser Zuordnung als - Eigene Teilnahmemotivation und Lernerfolg - charakterisieren.

#### Die gesamten Items der Betriebsbegehung:

- Die Betriebsbegehung war inhaltlich gut vorbereitet
- Die Betriebsbegehung war organisatorisch gut vorbereitet.
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet.
- Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft.

laden auf einem Faktor, dem Faktor 2, und lassen sich wie in dem verwendeten Fragebogen als - Betriebsbegehung - kennzeichnen.

#### Die Items der Grundlagenseminare:

- Die Medien wurden sinnvoll und effektiv eingesetzt.
- Das Thema wurde durch praxisnahe Beispiele veranschaulicht.
- Inhaltliche Fragen waren erwünscht
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet.

#### Das Item der praktischen Übung:

- inhaltliche Fragen wurden kompetent und verständlich beantwortet

#### Die Items der Gesamtbewertung:

- Die LV war insgesamt gut und systematisch strukturiert.
- Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen.
- Die Lehrpersonen waren gut vorbereitet.

laden alle auf einem Faktor, dem Faktor 3 und lassen sich aufgrund dieser Zuordnung als - Struktur, Didaktik und Lehrkompetenz - zusammenfassen.

#### Die Items der praktischen Übung:

- Der Lernstoff wurde im Dialog erarbeitet.
- Der Lehrinhalt wurde interessant präsentiert.
- Es fiel mir leicht aufmerksam zu bleiben.

laden auf einem Faktor, dem Faktor 4 und werden als - praktische Übungen - gekennzeichnet.

Die Items der Gruppenarbeit:

- Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes.
- Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation
- Die Gruppenarbeit hat Spaß gemacht.

Das Item der Gesamtbewertung:

- Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen.

laden auf einem Faktor, dem Faktor 5 und können aufgrund dieser Zuordnung als - Gruppenarbeit - gedeutet werden.

Zusammenfassend ergeben sich daraus also folgende Zuordnungen:

Faktor 1: Eigene Teilnahmemotivation und Lernerfolg

Faktor 2: Betriebsbegehung

Faktor 3: Struktur, Didaktik und Lehrkompetenz

Faktor 4: Praktische Übungen

Faktor 5: Gruppenarbeit

Dabei laden auf dem Faktor 3 die Items zur Evaluation der Grundlagenseminare mit Ausnahme des Items AUFMERK1, welches sich auf Faktor 1 abbildet. Ähnlich sind die Ergebnisse für die vier Items zur Beurteilung der praktischen Übungen, mit Ausnahme von KOMPET2 laden die Items auf Faktor 4. Die Skala für die Gesamtbewertung besitzt faktorenanalytisch auch keine eindeutige Struktur, denn die Faktorenladungen der Items verteilen sich auf drei Faktoren: Faktor 1, Faktor 3 und Faktor 5. Die vier Items zur Beurteilung der Betriebsbegehung laden alle auf dem Faktor 2 und weisen eine deutliche eindimensionale Struktur auf. Die Faktorenladungen sind ausreichend hoch. Damit ist hier eine ausreichende Trennschärfe gegeben. Das gleiche gilt für die drei Items zur Evaluation der Gruppenarbeit (Faktor 5) (vgl: Kap. 3.4.7.6., Seite 124, Tabelle 23)

Das Kaiserkriterium mit Eigenwerten über 1 wird wie oben bereits dargestellt von zwei der fünf Faktoren erfüllt, die damit im strengeren Sinne interpretierbar sind. Diese zwei Faktoren sind Faktor 1 und Faktor 2.

Faktor 1: Eigenwert 6,836

28,48% Varianz

Faktor 2: Eigenwert 1,755

7,31% Varianz

Die Korrelation der Faktoren untereinander zeigt, dass die Teilnahmemotivation und der Lernerfolg (Faktor 1) besonders mit der Struktur, Didaktik und Lehrkompetenz (Faktor 3) und der Gruppenarbeit (Faktor 5) zusammenhängen (Kapitel 3.4.7.6, S. 124 Tabelle 23).

Auch hier wird die Betriebsbegehung relativ unabhängig von den anderen Dimensionen bewertet. Damit ist die Betriebsbegehung von den übrigen Bewertungen der Lehrveranstaltung vergleichsweise unabhängig.

Zusammenfassend lässt sich somit darstellen, dass die explorative Faktorenanalyse das Ergebnis erbracht hat, dass der Fragebogen intern valide und reliabel ist, aber Hinweise darauf geliefert hat, dass sich die Items auch nach anderen als im Fragebogen vorgegebenen Zuordnungen inhaltlich gruppieren lassen. Eine Überarbeitung der Items und Skalen wäre in diesem Sinne sinnvoll.

Die Items wurden einer weiteren Analyse im Rahmen eines Strukturgleichungsmodells unterzogen. Diese Strukturgleichungsmodelle haben konfirmatorischen Charakter, weil sie vorgegebene strukturelle Beziehungen zwischen den Items überprüfen. Alle Items weisen nach der Analyse befriedigend hohe Faktorenladungen auf und sind trennscharf zwischen den Faktoren (Kapitel 3.4.8, S. 126, Tabelle 24).

Im Unterschied zum explorativen Modell weisen die Items im konfirmatorischen Modell eine eindeutige Faktorenstruktur im Sinne der theoretischen Zuordnung im Fragebogen auf. Nur das Item: „Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft“ besitzt noch eine Faktorenfremdladung auf dem Faktor „Gesamtbeurteilung“. Die konfirmatorische Analyse zeigte somit ebenfalls, dass auch eine valide Interpretation der Items im Sinne ihrer theoretischen a-priori-Zuordnung empirisch haltbar ist.

Forderungen in methodischer Hinsicht zur Validität des Evaluationsverfahrens stehen aber auch solche in sozialer Kompetenz gegenüber. Die Zielsetzung von Evaluation besteht ebenfalls darin, die Ergebnisse der Evaluation in einem Diskurs zwischen Lehrenden und Lernenden zu reflektieren, um dadurch die Qualität der Lehre und des Lernens zu erhöhen. Diese Auffassung vertritt auch Wesseler (1999), der eine Ergänzung zur Evaluation durch eine veranstaltungsinterne Besprechung zwischen Studierenden und Lehrenden zusätzlich für notwendig hält. In einer solchen Besprechung sollen die Resultate der Evaluation aktualisiert, neue Informationen gewonnen sowie mögliche Veränderungen erörtert und angestoßen werden, wie dies auch durch die institutsinterne Evaluation des „Praxisprojektes Arbeitsmedizin“ erfolgte.

#### **4.4 Kontinuierliche Evaluation als Beitrag zur Qualität der Lehre**

Nach Altrichter (1997) führt Evaluation zu einer kontinuierlichen Verlaufs- und Qualitätskontrolle und damit auch zur Qualitätsentwicklung. Es gibt verschiedene Ansätze, Qualität im Hinblick auf die Lehre zu beschreiben (Siehe Kap. 2.3.4, S. 26). Die Beschreibung von Qualität als „fitness for purpose“ wird von Abs und Mitarbeitern (2000) als die treffendste bezeichnet. Danach kann „Qualität der Lehre“ nur dann adäquat beurteilt werden, wenn vorher Ziel und Zweck der Lehre bestimmt wurden. Diese grundlegende Voraussetzung wird durch die Bestimmung von Ziel und Zweck der Lehrveranstaltung „Praxisprojekt Arbeitsmedizin“ erfüllt (siehe Kap. 3.1, S. 33).

Kriterien für die Qualitätsbeurteilung sind nicht nur im Hinblick auf die Studierenden ableitbar, sondern auch als wichtige Komponente auf Seiten der Lehrenden. Die arbeits- und institutionelle Zufriedenheit erhöht die Motivation und den Lernerfolg der Studierenden. Dieser Aspekt trifft ebenfalls auf die Lehrenden zu. Denn die von den Studierenden geäußerte Wertschätzung, die sich auch in der eigenen Teilnahmemotivation an der Lehrveranstaltung und dem Lernerfolg widerspiegelt, erhöht die Motivation der Lehrenden, insbesondere neue Lehr- und Lernansätze weiter zu entwickeln.

Nach Altrichter und Schratz (1992) dienen Evaluationsverfahren als geeignetes Instrument dazu, Hinweise zu gewinnen, in welchem Rahmen einer Lehrveranstaltung etwas verbesserungsbedürftig ist, und wie es verbessert werden kann. Für diese Art der kontinuierlichen Verlaufs- und Qualitätskontrolle und damit auch der Qualitätsentwicklung müssen nicht nur Lehr- und Lernziele (s. o.) definiert werden, sondern auch die Ziele der studentischen Lehrevaluation.

Ziel der kontinuierlichen durchgeführten Evaluation war es, den Erfolg des neuen Lehr- und Lernansatzes zu überprüfen. Der Fragebogen für die studentische Lehrevaluation wurde durch die Lehrenden entwickelt. Ob dieses Erhebungsinstrument den methodischen Anforderungen an Validität und Reliabilität genügt, wurde in dieser Studie ebenfalls überprüft. Durch die explorative und konfirmatorische Faktorenanalyse wurde belegt, dass das Evaluationsverfahren die methodischen Anforderungen an Reliabilität und Validität erfüllt (siehe Kap. 4.3, S. 139).

Die Evaluation zur Überprüfung des Erfolgs des neuen Lehr- und Lernansatzes im Zeitraum vom Wintersemester 1999/2000 bis Sommersemester 2003 zeigte im Vergleich zwischen dem ersten und letzten Erhebungszeitpunkt signifikante Verbesserungen der studentischen Bewertung auf. Auf der Grundlage von insgesamt 981 ausgewerteten Fragebögen im Zeitraum vom Wintersemester 1999/2000 bis Sommersemester 2003 werden die wesentlichen Ergebnisse in Tabelle 26 zusammenfassend dargestellt.

Tabelle 26 : Zusammenfassende Darstellung der Ergebnisse der Evaluation und statistisch signifikante Verbesserungen zwischen dem ersten (WS 1999/2000) und dem letzten (SS 2003) Untersuchungszeitpunkt.

	Trifft zu/trifft überwiegend zu Anteil der Studierenden in %	Statistisch signifikante Verbesserungen (p=0,05)
<b><u>Grundlagenseminare</u></b>		
Die Medien wurden sinnvoll und effektiv eingesetzt	96,1	*
Das Thema wurde durch praxisnahe Beispiele veranschaulicht	88,1	*
Inhaltliche Fragen waren erwünscht	92,7	*
Inhaltliche Fragen wurden kompetent und verständlich beantwortet	94,7	*
Es fiel mir leicht aufmerksam zu bleiben	62,5	*
<b><u>Praktische Übungen</u></b>		
Mein Basiswissen war ausreichend um den Übungen zu folgen	89,0	*
Der Lernstoff wurde im Dialog erarbeitet	77,2	
inhaltliche Fragen wurden kompetent und verständlich beantwortet	95,3	
Der Lehrinhalt wurde interessant präsentiert	82,4	
Es fiel mir leicht aufmerksam zu bleiben	77,5	

Tabelle 26 Fortsetzung

	Trifft zu/trifft überwiegend zu Anteil der Studierenden in %	Statistisch signifikante Verbesserungen (p=0,05)
<b><u>Betriebsbegehung</u></b>		
Die Betriebsbegehung war inhaltlich gut vorbereitet	77,1	*
Die Betriebsbegehung war organisatorisch gut vorbereitet	89,5	*
Inhaltliche Fragen wurden kompetent und verständlich beantwortet	83,6	*
Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft.	88,0	*
<b><u>Gruppenarbeit</u></b>		
Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes	68,8	*
Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation.	71,6	
Die Gruppenarbeit hat Spaß gemacht.	77,9	

Tabelle 26 Fortsetzung

	Trifft zu/trifft überwiegend zu Anteil der Studierenden in %	Statistisch signifikante Verbesserungen (p=0,05)
<b><u>Gesamtbewertung</u></b>		
Die LV war insgesamt gut und systematisch strukturiert.	92,0	*
Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen.	84,0	*
Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind.	81,2	*
Mein persönlicher Lernerfolg durch die Veranstaltung war groß.	74,3	*
Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen.	93,7	*
Die Lehrpersonen waren gut vorbereitet	97,7	
Ich hätte an der Veranstaltung auch freiwillig teilgenommen.	45,5	*
Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt.	50,2	

Wie aus Tabelle 26 hervorgeht, weisen 17 der 25 Items signifikante Verbesserungen auf. Bei den verbleibenden Items bewegen sich die Bewertungen auf gleich bleibend hohem Niveau.

Durch die explorative Faktorenanalyse wurden die Items des Fragebogens aufgrund der Korrelation untereinander fünf neuen Faktoren zugeordnet: Faktor 1: „Eigene

Teilnahmemotivation und Lernerfolg“, Faktor 2: „Betriebsbegehung“, Faktor 3: „Struktur, Didaktik und Lehrkompetenz“, Faktor 4: „Praktische Übungen“, Faktor 5: „Gruppenarbeit“

Im Folgenden wird zusammengefasst, inwieweit sich die Bewertungen der den Faktoren zugeordneten Items vom ersten bis zum letzten Erhebungszeitpunkt verändert haben.

- **Eigene Teilnahmemotivation und Lernerfolg**

Zu diesem Aspekt, der in der Faktorenanalyse als Faktor 1 bezeichnet wurde, gehören folgende Items:

- Es fiel mir leicht aufmerksam zu bleiben (AUFMERK1)
- Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind (KENNTNIS)
- Mein persönlicher Lernerfolg durch die Veranstaltung war groß (LERNERFO)
- Ich hätte an der Veranstaltung auch freiwillig teilgenommen (FREIWILL)
- Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt (INTERESS)

Im Untersuchungszeitraum wurden mit Ausnahme des Items „Interesse“ alle anderen Items signifikant positiver bewertet. Das Item „Interesse“ blieb mit ca. 50% auf gleichbleibend gutem Niveau.

- **Betriebsbegehung**

Zu diesem Aspekt, der in der Faktorenanalyse als Faktor 2 bezeichnet wurde, gehören folgende Items:

- Die Betriebsbegehung war inhaltlich gut vorbereitet (INHALT1)
- Die Betriebsbegehung war organisatorisch gut vorbereitet (ORGANIST)
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet (KOMPET3)
- Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft (EINBLICK)

Signifikant positivere Bewertungen zeigten sich in allen Items der Betriebsbegehung. Alle Items wurden von den Studierenden zu 80 bis 90% positiv bewertet.

- **Struktur, Didaktik und Lehrkompetenz**

Zu diesem Aspekt, der durch die Faktorenanalyse als Faktor 3 bezeichnet wurde, gehören folgende Items:

- Die Medien wurden sinnvoll und effektiv eingesetzt (MEDIEN)
- Das Thema wurde durch praxisnahe Beispiele veranschaulicht (THEMA)
- Inhaltliche Fragen waren erwünscht (FRAGEN)
- Inhaltliche Fragen wurden kompetent und verständlich beantwortet (KOMPET1, KOMPET2)
- Die Lehrveranstaltung war insgesamt gut und systematisch strukturiert (STRUKTUR)
- Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen (KOLLOQUI)
- Die Lehrpersonen waren gut vorbereitet (VORBEREI)

Im Untersuchungszeitraum wurden mit Ausnahme der Items „Kompetenz“ und die „Vorbereitung der Lehrenden“ alle anderen Items signifikant positiver bewertet.

Die „kompetente und verständliche Beantwortung von inhaltlichen Fragen“ sowie die „Vorbereitung der Lehrpersonen“ wurden von mehr als 95% der Studierenden auf gleich bleibend hohem positiven Niveau bewertet.

- **Praktische Übungen**

Zu diesem Aspekt, der in der Faktorenanalyse als Faktor 4 bezeichnet wurde, gehören folgende Items:

- Der Lernstoff wurde im Dialog erarbeitet (DIALOG)
- Der Lehrinhalt wurde interessant präsentiert (LEHRINHA)
- Es viel mir leicht aufmerksam zu bleiben (AUFMERK2)

Im Untersuchungszeitraum zeigten sich keine signifikanten Veränderungen. Die drei Items wurden von ca. 80% der Studierenden auf gleich bleibend hohem Niveau positiv bewertet.

- **Gruppenarbeit**

Zu diesem Aspekt, der in der Faktorenanalyse als Faktor 5 bezeichnet wurde, gehören folgende Items:

- Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes (DISKURSI)
- Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation (UNTERSTÜ)
- Die Gruppenarbeit hat Spaß gemacht (SPAß)
- Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen (INHALT2)

Im Untersuchungszeitraum wurden die Items „Diskussion“ und „Inhalt“ signifikant positiver bewertet. Die Items „Dozentenunterstützung“ und „Spaß“ wurden von 70 bis 80% der Studierenden auf gleich bleibend hohem Niveau positiv bewertet. Darüber hinaus beurteilen insgesamt mehr als 2/3 der Studierenden, die die Gruppenarbeit positiv bewerten, auch alle anderen evaluativen Kriterien signifikant positiver als die Gruppe derjenigen, die ihre Gruppenarbeit negativ bewerten.

Durch die explorative Faktorenanalyse konnte zudem gezeigt werden, dass der Faktor „eigene Teilnahmemotivation und Lernerfolg“ mit den Faktoren „Struktur, Didaktik und Lehrkompetenz“ und „Gruppenarbeit“ korreliert.

Durch die kontinuierliche Durchführung und Auswertung der studentischen Lehrevaluation vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 wurde eine periodische Rückmeldung Semester für Semester erreicht. Insgesamt zeigen die Ergebnisse dieser Studie, dass durch studentische Lehrevaluation auf der Grundlage zuvor definierter Zielkriterien aussagekräftige Bewertungen erreicht wurden. Die aus diesen Bewertungen ersichtliche Zunahme von Teilnahmemotivation und Lernerfolg motivierten die Lehrenden zunehmend dazu, die erprobten Lehr- und Lernansätze in der curricularen Lehre des Fachgebietes weiter zu entwickeln.

Zusammenfassend lässt sich somit belegen, dass durch die kontinuierliche Durchführung und Auswertung der studentischen Lehrevaluation immer wieder strukturelle und prozedurale Verbesserungen der Lehrveranstaltung erreicht wurden.

Dies führte in den nachfolgenden Semestern zu einer zunehmend besseren Bewertung. Auf der Grundlage eines validen Evaluationsverfahrens kann kontinuierliche studentische Lehrevaluation somit zur Qualität der Lehre beitragen.

#### **4.5 Beantwortung der gestellten Fragen**

Ziel dieser Studie war es, die Ergebnisse der vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 kontinuierlich durchgeführten studentischen Evaluation auszuwerten, um

1. den Erfolg des neuen Lehr- und Lernansatzes und
2. die Validität des Evaluationsverfahrens zu überprüfen.

Nach den Ergebnissen dieser Studie können die eingangs gestellten Fragen wie folgt beantwortet werden:

##### **1. Welchen Einfluss hat das neu konzipierte didaktische und methodische Lehrkonzept auf die Teilnahmemotivation und den Lernerfolg der Studierenden?**

Das neu konzipierte didaktische und methodische Lehrkonzept wirkt sich positiv auf die Teilnahmemotivation und den Lernerfolg der Studierenden aus. Vier der fünf Items, die durch Faktorenanalyse dem Faktor „Eigene Teilnahmemotivation und Lernerfolg“ zugeordnet wurden, sind im Untersuchungszeitraum signifikant positiver bewertet worden: Die Aufmerksamkeit in den Grundlagenseminaren sowie die Wichtigkeit der Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für die spätere ärztliche Tätigkeit wurden signifikant positiver bewertet. Dies gilt auch für die Einschätzung des persönlichen Lernerfolgs sowie für die Bereitschaft zur freiwilligen Teilnahme an der Lehrveranstaltung.

Das Interesse für das Fachgebiet Arbeitsmedizin bewegt sich vom ersten bis zum letzten Untersuchungszeitpunkt auf gleich bleibend gutem Niveau.

Besonders hervorzuheben ist, dass nach dem Ergebnis der Faktorenanalyse der Faktor „Eigene Teilnahmemotivation und Lernerfolg“ mit den Faktoren „Struktur, Didaktik und Lehrkompetenz“ und „Gruppenarbeit“ korreliert.

## **2. Wie bewerten die Studierenden Struktur, Didaktik und Lehrkompetenz?**

Struktur, Didaktik und Lehrkompetenz wurden zum Ende des Untersuchungszeitraums positiver bewertet. Sechs der sieben Items, die durch Faktorenanalyse dem Faktor „Struktur, Didaktik und Lehrkompetenz“ zugeordnet wurden, zeigten signifikante Verbesserungen der Bewertung. Dies gilt für die Gesamtstruktur der Lehrveranstaltung, die Veranschaulichung der arbeitsmedizinischen Themengebiete anhand von praxisnahen Beispielen und den Dialog mit den Studierenden in den angebotenen Seminaren. Ebenso werden Medieneinsatz und das Prüfungsverfahren signifikant positiver bewertet. Hervorzuheben ist, dass die Vorbereitung der Lehrenden kontinuierlich auf gleich bleibend hohem Niveau bewertet wurde.

## **3. Wie bewerten die Studierenden den direkten Praxisbezug durch die Betriebsbegehung?**

Die Betriebsbegehung und der damit verbundene Einblick über Gesundheitsgefahren am Arbeitsplatz wurden von den Studierenden zum Ende des Untersuchungszeitraums signifikant positiver bewertet als zu Beginn der Untersuchung.

Dies gilt für alle vier Items des Faktors „Betriebsbegehung“: die inhaltliche und organisatorische Vorbereitung der Begehung, der Dialog mit den Studierenden und der gewonnene Einblick in mögliche Gesundheitsgefahren am Arbeitsplatz wurden signifikant positiver bewertet.

## **4. Welche Auswirkung hat die Arbeit in Kleingruppen auf die Gesamtbewertung der Lehrveranstaltung?**

Die Bewertung der Gruppenarbeit hatte einen erheblichen Einfluss auf die Gesamtbewertung der Lehrveranstaltung.

Fast durchgängig über alle Indikatoren der Lehrveranstaltung beurteilt die Gruppe derjenigen, die die Gruppenarbeit positiv bewerten, alle übrigen evaluativen Kriterien signifikant positiver als die Gruppe derjenigen, die die Gruppenarbeit negativ bewerten.

Der Anteil der Studierenden, die die Gruppenarbeit nach den Items Diskussion, Dozentenunterstützung und Spaß positiv bewerten, liegt zwischen 70 und 80 %.

Die Akzeptanz der Studierenden wird auch dadurch belegt, dass über 80% der Befragten die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit völlig oder überwiegend gut gefallen hat.

## **5. Entspricht die durchgeführte Evaluation den methodischen Ansprüchen an die Reliabilität und Validität von Evaluationsverfahren?**

Durch die explorative Faktorenanalyse wurden die fünf Skalen auf die interne Validität überprüft. Die Eigenwerte, Faktorladungen, Cronbachs Alpha aller fünf Skalen erreichen geeignete Werte. Damit ist die interne Validität gegeben. Bei der Überprüfung der externen Validität zeigt sich, dass sich 2 Faktoren mit Eigenwerten größer als eins abbilden. Die Gründe hierfür liegen in der Verwendung von semantisch Identischen Fragen und in dem Element Gesamtbewertung, dass sich aus den Bewertungen der Teilaspekte und der gemessenen Gesamtbeurteilung zusammensetzt. Die Trennschärfe der Skalen ist somit empirisch nicht mehr gegeben.

Die Betriebsbegehung als zentrales Element des Praxisprojektes Arbeitsmedizin bildet sich mit allen Items auf einem Faktor ab, und ist somit trennscharf. Unter der Vorgabe fünf Faktoren zu bilden zeigte sich bei der Faktorenanalyse, dass auch eine andere Zuordnung der verwendeten Fragen möglich ist. Durch die konfirmatorische Faktorenladung wird belegt, dass aber auch eine eindeutige Faktorenstruktur im Sinne der Zuordnung des verwendeten Fragebogens möglich und eine valide Interpretation der Items empirisch haltbar ist.

## **6. Kann kontinuierliche studentische Lehrevaluation zur Qualität der Lehre in der Arbeitsmedizin beitragen?**

Durch die kontinuierliche Durchführung und Auswertung der studentischen Lehrevaluation vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 wurde eine periodische Rückmeldung Semester für Semester erreicht. Die konstruktive Kritik der Studierenden waren immer Anlass und Motivation für weitere strukturelle und prozedurale Verbesserungen, die wiederum zu einer verbesserten Bewertung der Lehrveranstaltung in den nachfolgenden Semestern führte. Die daraus ersichtliche Zunahme von Teilnahmemotivation und Lernerfolg der Studierenden motivierten die Lehrenden zunehmend dazu, die erprobten Lehr- und Lernansätze in der curricularen Lehre des Fachgebietes weiter zu entwickeln. Auf der Grundlage eines validen Evaluationsverfahrens kann somit kontinuierliche studentische Lehrevaluation zur Qualität der Lehre in der Arbeitsmedizin beitragen.

## 5. ZUSAMMENFASSUNG

**Ziel:** Seit dem Wintersemester 1999/2000 werden die Pflichtlehrveranstaltungen im Fachgebiet Arbeitsmedizin auf der methodischen Grundlage des projekt- und problemorientierten Lernens durchgeführt und kontinuierlich institutsintern evaluiert. Ziel dieser Studie ist es, die Evaluationsergebnisse auszuwerten, um den Erfolg des neuen Lehr- und Lernansatzes und insbesondere die Validität des Evaluationsverfahrens zu überprüfen.

**Methode:** Im Untersuchungszeitraum vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 haben insgesamt 1020 Studierende am „Praxisprojekt Arbeitsmedizin“ teilgenommen. 981 ausgefüllte Evaluationsbögen wurden in die Auswertung einbezogen. Der Fragebogen besteht aus insgesamt 25 Fragen (Items), strukturiert in Indikatoren für die didaktischen Elemente (Grundlagenseminare, Betriebsbegehungen, praktische Übungen, Gruppenarbeit) und für die Gesamtbewertung der Lehrveranstaltung. Die statistische Auswertung der Daten wurde mit SPSS Version 10.0 durchgeführt. Als multivariate Analysemethoden wurden die Regressionsanalyse und die Faktorenanalyse verwendet.

**Ergebnisse:** Durch die explorative und konfirmatorische Faktorenanalyse wurde belegt, dass das Evaluationsverfahren die methodischen Anforderungen an Reliabilität und Validität erfüllt. Im Erhebungszeitraum zeigte sich eine statistisch signifikante Verbesserung der Beurteilung über die Indikatoren der didaktischen Elemente und für die Gesamtbewertung der Lehrveranstaltung. Dies gilt insbesondere für die Items, die durch Faktorenanalyse dem Faktor „eigene Teilnahmemotivation und Lernerfolg“ und dem Faktor „Struktur, Didaktik und Lehrkompetenz“ zugeordnet wurden. Faktorenanalytisch konnte eine Korrelation zwischen dem Faktor „eigene Teilnahmemotivation und Lernerfolg“ und den Faktoren „Struktur, Didaktik und Lehrkompetenz“ und „Gruppenarbeit“ belegt werden. Zum Einfluss der Bewertung der Gruppenarbeit wird aufgezeigt, dass mehr als 2/3 der Studierenden, die die Gruppenarbeit positiv bewerten, auch alle anderen evaluativen Kriterien signifikant positiver bewerten als die Gruppe Derjenigen, die ihre Gruppenarbeit negativ beurteilen.

**Schlussfolgerungen:** Die kontinuierliche Evaluation und die konstruktive Kritik der Studierenden waren immer Anlass und Motivation für weitere strukturelle und prozedurale Verbesserungen, die wiederum zu einer verbesserten Bewertung der Lehrveranstaltung in den nachfolgenden Semestern führte. Insgesamt konnte gezeigt werden, dass kontinuierliche studentische Lehrevaluation auf der Grundlage eines validen Evaluationsverfahrens zur Qualität der Lehre in der Arbeitsmedizin beitragen kann.

## 6. Literaturverzeichnis

1. ÄAppO (1970) Approbationsordnung für Ärzte. BGBl 1970 Teil 1:1458
2. ÄAppO (2002) Approbationsordnung für Ärzte. BGBl 2002 Teil 1:2405
3. Abrami P, Cohen P (1990) Validity of student ratings of instruction: What we know and what we do not? *J Educ Psych* 82: 219-231
4. Abs HJ, Raether W, Tippelt R (2000) *Besser Lehren*. Deutscher Studienverlag, Weinheim
5. Albanese M (2000) Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. *Med Educ* 34:729-738
6. Albanese MA, Mitchell S (1993) Problem-based learning: A review of literature on its outcomes and implementation issues. *Acad Med* 68:52-81
7. Allgäuer R (1998) *Evaluation macht uns stark! Zur Unverzichtbarkeit von Praxisforschung im schulischen Alltag*. 2. Auflage, Lang, Bern Berlin Frankfurt
8. Altrichter H (1997) *Hochschulen auf dem Prüfstand*. Studienverlag, Innsbruck, Wien
9. Altrichter H, Schratz M (1992) *Evaluation: Impulse für Innovation?* Studien-Verlag, Innsbruck
10. Angelo T, Cross K (1993) *Classroom Assessment Techniques. A handbook for college teachers*. Jossey-Bass, San Francisco
11. Apel HJ, Knoll M (1997) *Aus Projekten lernen. Grundlegung und Anregungen*. Oldenburg, München
12. Armstrong EG (1997) A hybrid model of problem-based learning. In: Boud D, Felletti G (Hrsg.) *The Challenge of Problem Based Learning*. Kogan Page, London, S. 137-150
13. Backhaus K, Erichson B, Plinke W, Weiber R (2003) *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. 10. neu bearbeitete und erweiterte Auflage, Springer, Berlin
14. Barrows HS (1985) *How to design a problem-based curriculum for the preclinical years*. Springer Pub. Co, New York
15. Barrows HS, Tamblyn RM (1980) *Problem-based learning. An approach to medical education*. Springer Pub. Co, New York
16. Bastian J (1993) Freie Arbeit und Projektunterricht. *Pädagogik* 10:6-9
17. Becker GE (1997) *Planung von Unterricht*. Beltz, Weinheim

18. Berendt B (2000) Was ist gute Hochschullehre. Zeitschrift für Pädagogik 41:247-260
19. Berendt B, Voss HP, Wildt J (Hrsg.) (2007) Neues Handbuch Hochschullehre. Raabe, Berlin
20. Beywl W (1996) Evaluation – Controlling-Qualitätsmanagement in der Weiterbildung. Bundesinstitut für Berufsbildung, Berlin, Bonn
21. Bohnsack F (1976) Erziehung zur Demokratie. John Deweys Pädagogik und ihre Bedeutung für die Reform unserer Schule. Otto Meyer, Ravensburg
22. Bortz J (1995) Lehrbuch der empirischen Forschung. Springer-Verlag, Berlin Heidelberg New York
23. Bortz J, Döring N (2002) Forschungsmethoden und Evaluation. Springer-Verlag, Berlin
24. Brähler E, Wittig U, Beckert C (1998). Der Studienerfolg an Medizinischen Fakultäten - wie viele Studienanfänger bestehen nach vier Semestern die ärztliche schriftliche Vorprüfung? Gesundheitswesen 60:317-321
25. Clade H (1998) Reform des Medizinstudiums – Mehr Praxisbezug angesagt. Deutsches Ärzteblatt 95:57-58
26. Colliver JA (2000) Effectiveness of problem-based learning curricula: Research and theory. Acad Med 75:259-266
27. Daniel HD (1995) Bewertung der Lehre aus Sicht der Studierenden und Absolventen. In: Müller-Böling D (Hrsg.) Qualitätssicherung in Hochschulen. Verlag Bertelsmann Stiftung, Gütersloh
28. Diel J (1996) Studentische Evaluation von Hochschulveranstaltungen. In: Zeitschrift für Pädagogische Psychologie 10:167-170
29. Dohmen G (1996) Das lebenslange Lernen – Leitlinien einer modernen Bildungspolitik. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, Bonn
30. Dörner D (1987) Problemlösen als Informationsverarbeitung, 3. Aufl., Kohlhammer, Stuttgart
31. Duden Fremdwörter, 2005, 8. Auflage, S: 296
32. Dunkin M, Barnes J (1986) Research on teaching in higher education. In: Wittrock M (Hrsg.) Handbook of research on teaching. Macmillan, New York, S. 754-777
33. Eitel F (1994) Wege zur Reform der Unterrichtsorganisation. Medizinische Ausbildung 1:122-137
34. Eitel F (1995) Neuere Entwicklungen in den Fakultäten zur Verbesserung der Lehrqualität. Medizinische Ausbildung 12:2-12

35. Eitel F (1996) Wer kennt John Dewey? Medizinische Ausbildung 13:37-41
36. Enders J, Teichler U (1995) Der Hochschullehrerberuf im internationalen Vergleich. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie, Bonn
37. Engel C (1997) Not just a method but a way of learning. In: Boud D, Feletti G (Hrsg.) The challenge of Problem-based-Learning. 2. Auflage, Kogan Page, London
38. Finucare P, Johnson S (1998) Problem-based learning: its rationale and efficacy. Med Educ 168:445-448
39. Fischer-Blum K (2002) Zielvereinbarungen als „Follow up“ der Evaluation von Studium und Lehre im Nordverbund. In: Hochschulrektorenkonferenz (HRK) (Hrsg.) Evaluation, was nun? Erfahrungen mit der Umsetzung von Evaluationsergebnissen (Beiträge zur Hochschulpolitik 1/2002) Projekt Qualitätssicherung, Bonn, S. 83-91
40. Fischer-Blum K, Mittag S, Sundermann E, Buß S, Habel E, Winter M (2001) Hochschulevaluation. Zu Einheit und Vielfalt in der Evaluation von Studium und Lehre – Bericht des Arbeitskreises Hochschulevaluation. In: Deutsche Gesellschaft für Evaluation e.V. (Hrsg.) Evaluation – Reformmotor oder Reformbremse? Köln, S. 51-79
41. Freed J., Klugmann M (1997) Quality principles and practices in higher education. Different Questions for different times. Oryx Press, Phoenix, Arizona
42. Frey K (1982) Die Projektmethode. Beltz, Weinheim Basel
43. Frey K (1996) Die Projektmethode, Beltz, Weinheim
44. Geisler W, Scholz G (1976) Projektorientierter Unterricht. Beltz, Weinheim Basel
45. Geißler G (1970) Das Problem der Unterrichtsmethode, Beltz, Weinheim Basel
46. Gold A (1996) Können Studierende die Qualität der Lehre beurteilen? Zeitschrift für Pädagogische Psychologie 10:147-150
47. Gräsel C (1997) Problemorientiertes Lernen. Hogrefe, Göttingen
48. Green D (1994) What is Quality in Higher Education? University Press, Buckingham
49. Greenwald A (1997) Validity concerns and usefulness of student ratings of instruction. Am Psychol 52:1182-1186
50. Grime P, Williams S, Nicholson S (2003) Reports of new ideas in medical education. A new undergraduate teaching session in occupational medicine. Med Educ 37:1033-1034

51. Grime P, Williams S, Nicholson S (2006) Medical students' evaluation of a teaching session in occupational medicine: the value of a workplace visit. *Occup Med* 56:110-114
52. Grüner H (1993) Evaluation und Evaluationsforschung im Bildungswesen. *Pädagogische Rundschau* 47:29-52
53. Gudjons H (1998) Didaktik zum Anfassen. Lehrer/in - Persönlichkeit und lebendiger Unterricht. Klinkhardt, Bad Heilbronn
54. Gudjons H (2001) Handlungsorientiert lehren und lernen – Schüleraktivierung, Selbsttätigkeit, Projektarbeit. 6. Auflage, Klinkhard, Bad Heilbronn
55. Habeck D (1995) Überlegungen zu einem Modellstundenplan, *Medizinische Ausbildung* 12:42-45
56. Hager W, Patry J, Brezing H (2000) Evaluation psychologischer Interventionsmaßnahmen, Huber Verlag, Bern
57. Harden RM, Crosby J, Davis MH, Howie PW, Struthers AD (2000) Task-based learning: The answer to integration and problem-based learning in the clinical years. *J Med Educ* 34:391-397
58. Heller A, Semmerling R (1983) *Das PROWO\_Buch. Leben, Lernen, Arbeiten in Projekten und Projektwochen.* Scriptor, Königsstein
59. Helmke A, Krapp A (1999) Lehren und Lernen in der Hochschule. *Zeitschrift für Pädagogik* 45:19-24
60. Hmelo CE (1998) Cognitive Consequences of Problem-Based Learning for the Early Development of Medical Expertise. *Teach Learn Med* 10:92-100
61. Hochschulrektorenkonferenz HRK (Hrsg.) (1995) Zur Evaluation im Hochschulbereich unter besonderer Berücksichtigung der Lehre. Entschließung des 176. Plenums der HRK vom 3. Juli 1995, Bonn. Zitiert nach: Berendt B, Voss HP, Wildt J (Hrsg.) *Neues Handbuch Hochschullehre*, 9. Lieferung 2002, Kap. I 2.2
62. Joint Committee on Standards for Educational (2006) *The program evaluation standards: how to assess evaluations of educational programs.* Thousand Oaks, California
63. Kahlke W, Kaie A, Kaiser H, Kratzert R, Schöne A (2000) Problemorientiertes Lernen eine Chance für die Fakultäten. *Deutsches Ärzteblatt* 97:1745-1749
64. Kaiser A (1977) *Projektstudium und Projektarbeit in der Schule*, Klinkhardt, Bad Heilbronn
65. Kissel D (1984) *Von der Pädagogik in der Medizin.* Hippokrates, Stuttgart
66. Koh D, Chia K, Jeyaratnam J, Chia S, Singh J (1995) Case studies in occupational medicine for medical undergraduate training. *National University of Singapore. Occup Med* 45:27-30

67. Krempkow R (1998) Ist „gute“ Lehre messbar? Das Hochschulwesen 46:195-199
68. Kromrey H (1994) Wie erkennt man „gute Lehre?“ Was studentische Vorlesungsbefragungen (nicht) aussagen. Empirische Pädagogik 8:153-168
69. Kromrey H (1995) Evaluation der Lehre durch Umfrageforschung? Methodische Fallstricke bei der Messung von Lehrqualität durch Befragung von Vorlesungsteilnehmern. In: Mohler P (Hrsg) Hochschullehre. Ihre Evaluation als Herausforderung an die Empirische Sozialforschung Waxmann, Münster, S. 105-128
70. Kromrey H (1996) Qualitätsverbesserung in Lehre und Studium statt so genannter Lehrevaluation. Zeitschrift für Pädagogische Psychologie 10:153-166
71. Lamnek S (1993) Qualitative Sozialforschung. Methoden und Techniken. Band 2, Beltz, Weinheim
72. Laske S, Scheytt T, Meister-Scheytt C, Scharmer C (2000) Universität und Gesellschaft - Schriften zur Universitätsentwicklung. Band 1, Rainer Hamp, München
73. Laughlin PR, Hatch EC, Silver JS, Boh L (2006) Groups perform better than the best individuals on letters-to-numbers problems: Effects of group size. J Pers Soc Psychol 90:644-651
74. Mandl H, Putz RV, Peter K, Höfling S (2001) Lernmodelle der Zukunft am Beispiel der Medizin. Hans Seidel Stiftung München
75. Marsch H (1993) Multidimensional Student's Evaluations of Teaching Effectiveness. A Profile Analysis. Journal of Higher Education 64:1-18
76. Marsh H (1984) Students' evaluations of university teaching: Dimensionality, reliability, validity, potential bias and utility. J Educ Psychol 76:707-754
77. Marsh H, Roche L (1997) Making students' evaluations of teaching effectiveness effective. The critical issues of validity, bias and utility. Am Psychol 52:1187-1197
78. Mattes W (2004) Methoden für den Unterricht. Schöningh, Paderborn
79. Maudsley G (1999) Do we all mean the same thing by "problem-based learning?" A review of the concepts and a formulation of the ground rules. Acad Med 74:178-185
80. Maudsley G, Strivens J (2000) Promoting professional knowledge, experiential learning and critical thinking for medical students. Med Educ 34:535-544
81. McKeachie WJ (1990) Research of college teaching: The historical background. J Educ Psychol 82:189-200
82. Meyer H (1994) Unterrichtsmethoden. Cornelsen Verlag, Frankfurt am Main
83. Miles J (2000) Statistical software. Br J Math Stat Psychol 53:145-153

84. Mittag W, Hager W (2000) Ein Rahmenkonzept zur Evaluation psychologischer Interventionsmaßnahmen. In: Patry J, Brezing H (Hrsg.) Evaluation psychologischer Interventionsmaßnahmen Standards und Kriterien. Huber, Bern: S.102-128
85. Mitteilungen des deutschen Hochschulverbandes (1991) 6:288
86. Moust J, Bouhouijs P, Schmidt H (1999) Problemorientiertes Lernen, Ullstein Medical, Wiesbaden
87. Norman GR, Schmidt HG (2000) Effectiveness of problem-based learning curricula: theory, practice and paper darts. Med Educ 34:721-728
88. Patton M (1981) Creative Evaluation. Beverly Hills, London
89. Peterssen WH (1996) Handbuch Unterrichtsplanung. Grundfragen, Modelle, Stufen, Dimensionen. Ehrenwirth, München
90. Peterssen WH (1997) Kleines Methoden-Lexikon. Oldenburg, München
91. Petschnig M (1971) Der kleine Stowasser. Lateinisch deutsches Schulwörterbuch. Hölder-Verlag, Wien
92. Pfaff M (1997) Problemorientiertes Lernen. Anleitung mit 20 Fallbeispielen. Chapman & Hall, Weinheim
93. Pratt D, Magill MK (1983) Educational contracts: A basis for effective clinical teaching. J Med Educ 58:462-467
94. Regehr G, Norman GR (1996) Issues in cognitive psychology: Implications for professional education. Acad Med 71:988-1001
95. Reinecke J (2005) Strukturgleichungsmodelle, Oldenburg, München
96. Reissert R (1994) Handbuch „Berichtssystem zu Studium und Lehre“ – Beschreibung eines Verfahrens zur Selbstevaluation. HIS - GmbH Selbstverlag, Hannover
97. Reissert R, Carstensen D (1997) Praxis der internen und externen Evaluation – Handbuch zum Verfahren. HIS - GmbH Selbstverlag, Hannover
98. Richter R (1994) Qualitätssorge in der Lehre. Leitfaden für die studentische Lehrevaluation. Luchterhand, Neuwied Berlin
99. Rindermann H (1996) Untersuchungen zur Brauchbarkeit studentischer Lehrevaluationen: Analysen der Validität und zu Auswirkungen ihres Einsatzes anhand des Heidelberger Inventars zur Lehrveranstaltungs-Evaluation (HILVE), Empirische Pädagogik, Landau
100. Rindermann H (1996 a) Qualitative Urteile über die Qualität der Lehre durch Studierende? Zeitschrift für Pädagogische Psychologie 10:151-152

101. Rindermann H (1997 ) Die studentische Beurteilung von Lehrveranstaltungen: Forschungsstand und Implikationen für den Einsatz von Lehrevaluationen. In: Jäger R, Trost G, Lehmann R (Hrsg.) Tests und Trends, 11. Jahrbuch der Pädagogischen Diagnostik. Beltz, Weinheim, S. 12-53
102. Rindermann H (1997a) Generalisierbarkeit studentischer Veranstaltungsbeurteilungen. *Psychologie* 44:216-234
103. Rindermann H (1999) Was zeichnet gute Lehre aus? Ergebnisse einer offenen Befragung von Studierenden und Lehrenden nach Merkmalen guter Dozenten und Veranstaltungen. *Zeitschrift für Hochschuldidaktik* 23:136-156
104. Rindermann H (2000a) Das Selbstobjektivierungsproblem im akademischen Milieu (I): Forschung an Hochschulen über die Qualität der Lehre an Hochschulen. *Das Hochschulwesen* 48:74-82
105. Rindermann H (2000b) Das Selbstobjektivierungsproblem im akademischen Milieu (II): Zusammenhang zwischen wissenschaftlicher Methodik und dem Einfluss sozialer Hintergrundvariablen in der Lehrevaluationsforschung. *Das Hochschulwesen* 48:117-123
106. Rindermann H, Oubaid V (1999) Auswahl von Studienanfängern durch Universitäten - Kriterien, Verfahren und Prognostizierbarkeit des Studienerfolgs. *Zeitschrift für Differentielle und Diagnostische Psychologie* 20:172-191
107. Robert Bosch Stiftung (1989) Das Arztbild der Zukunft. Analysen künftiger Anforderungen an den Arzt, Konsequenzen für die Ausbildung und Wege zur Reform; Abschlußbericht des Arbeitskreises Mediziner Ausbildung der Robert Bosch Stiftung - Murrhardter Kreis. Bleicher, Gerlingen
108. Rosemann B, Schweer MW (1996). Evaluation universitärer Lehre und Wissenszuwachs bei Studierenden. *Zeitschrift für pädagogische Psychologie* 10:175-180
109. Rossi P (1988) *Programm Evaluation: Einführung in die Methoden angewandter Sozialforschung*. Enke-Verlag, Stuttgart
110. Salzer M, Nixon C, (1997) Validating Quality Indicators. Quality as relationship between structure, process and outcome. *Eval Rev* 21:292-309
111. Scheffner D (2001) Reformstudiengang Medizin in Berlin: Who has to learn? In: Mandl H, Putz RV, Peter K, Höfling S (Hrsg.) *Lernmodelle der Zukunft am Beispiel der Medizin*. Hans Seidel Stiftung, München, S. 53-61
112. Schmidt H, Moust J (1995) What makes a tutor effective? A structural-equations modelling approach to learning in problem-based curricula. *Acad Med* 70:708-714
113. Schmidt HG (1983) Problem-based learning: rationale and description. *Med Educ* 17:11-16
114. Schmidt HG (1993) Foundations of problem based learning: Some explanatory notes. *J Med Educ* 27:422-432

115. Schmidt HG, Moust JH, Boon L (1993) Influence of tutors' subject matter expertise on student effort and achievement in problem-based learning. *Acad Med* 68:784-791
116. Scholz O (1995) Evaluation von Lehrveranstaltungen durch Studenten. *Forschung und Lehre* 9:497-501
117. Siebert H (1997) *Didaktisches Handeln in der Erwachsenenbildung*. Luchterhand, Neuwied
118. Spiel C (2001) Der differentielle Einfluss von Biasvariablen auf studentische Lehrveranstaltungsbewertungen. In: Engel U (Hrsg.) *Zur Qualitätsbewertung von Studium und Lehre*. Campus, Frankfurt am Main, S. 61-82
119. Spiel C, Gössler M (2000) Zum Einfluss von Biasvariablen auf die Bewertung universitärer Lehre durch Studierende. *Zeitschrift für Pädagogische Psychologie* 14:38-47
120. Stach R (1978) *Projektorientierter Unterricht*. Henn, Kastellaun
121. Struck P (1980) *Projektunterricht*. Kohlhammer, Stuttgart
122. Sturm M (1994) „Die Reise ins Ich.“ *Selbstevaluation im Hochschulunterricht*. Studien-Verlag, Innsbruck
123. Suin de Boutemard B (1975) *Projektunterricht und soziale Handlungsperformanz*. Fink, München
124. Tippelt R (1999) *Handbuch der Erwachsenenbildung*. Leske und Budrich, Opladen
125. Urban W (2004) *Evaluation des tertiären Bildungssystems*. Studien Verlag, Innsbruck
126. Vedung E (1999) *Evaluation im öffentlichen Sektor*. Böhlau-Verlag, Wien Köln Graz
127. Vernon DT (1995) Attitudes and opinions of faculty tutors about problem-based learning. *Acad Med* 70:216-223
128. Vernon DT, Blake RL (1993) Does problem-based learning work? A meta-analysis of evaluative research. *Acad Med* 68:550-563
129. Wesseler M (1999) Evaluation und Evaluationsforschung. In: Tippelt R (Hrsg.) *Handbuch Erwachsenenbildung*. 2. Aufl., Leske & Budrich, Opladen, S. 671-687
130. Westermann R, Heise E, Spies K und Trautwein U (1996). Identifikation und Erfassung von Komponenten der Studienzufriedenheit. *Psychologie in Erziehung und Unterricht* 43:1-22
131. Westfälische Wilhelms-Universität Münster (1996) *Amtliche Bekanntmachungen Jahrgang 1996 Nr. 8*

132. Wildt J (2002) Ein hochschuldidaktischer Blick auf Lehren und Lernen. Eine kurze Einführung in die Hochschuldidaktik. In: Berendt B, Voss HP, Wildt J (Hrsg.) Neues Handbuch Hochschullehre. Raabe, Berlin, Kap. A 1.1
133. Wilke H (1993) Systemlogik und kontextuelle Einbindung der Ökonomie in Hochdifferenzierten Gesellschaften. In: Heitger B, Schmitz C, Gester P (Hrsg.) 2. Jahrbuch für systemisches Denken und Handeln im Management. Auer, Heidelberg, S. 49-62
134. Wilkie K (2000) The nature of Problem-based-Learning. In: Glen S, Wilkie K (Hrsg.) Problem-based-Learning in Nursing – A new model for a new context? Macmillan Press, Houndmills, S. 37-43
135. Wissenschaftsrat (1996) Empfehlungen zur Stärkung der Lehre in den Hochschulen durch Evaluation. Wissenschaftsrat, Berlin
136. Witting U (2002) Problemorientiertes Lernen am Beispiel der Arbeitsmedizin in Münster. In: Nowak D, Raul G (Hrsg.) Dokumentationsband über die 42. Jahrestagung der Deutschen Gesellschaft für Arbeitsmedizin und Umweltmedizin e.V. in München vom 10.-13. April 2002. Rindt-Druck, Fulda, S. 554-558
137. Wottawa H, Thierau H (1998) Lehrbuch Evaluation. 2. Aufl., Hans Huber, Bern
138. Wynn P, Aw T, Williams N, Harrington M (2003) Teaching of occupational medicine to undergraduates in UK schools of medicine. Occup Med 53:349-353

## **7. Danksagung**

Frau Univ.-Prof. Dr. med. U. Witting, Direktorin des Institutes für Arbeitsmedizin der Westfälischen Wilhelms-Universität, danke ich herzlich für die freundliche Überlassung des Dissertationsthemas und die hervorragende Betreuung bei der Erstellung der Dissertation.

Mein besonderer Dank gilt ebenfalls Herrn Prof. Dr. N. Binding für seine ausgezeichnete Betreuung bei der Erstellung dieser Dissertation.

Den weiteren Mitarbeitern des Instituts danke ich für deren freundliches Entgegenkommen und die vielseitige Unterstützung.

## 8. Lebenslauf

## **9. Anhang: Fragebogen**

Beigefügt ist der Fragebogen, der vom Wintersemester 1999/2000 bis einschließlich Sommersemester 2003 jeweils den Studierenden der Gruppen A- H der ersten Semesterhälfte und den Studierenden der Gruppen I - Q in der zweiten Semesterhälfte vorgelegt wurde.



WESTFÄLISCHE  
WILHELMS-UNIVERSITÄT  
MÜNSTER

## Institut für Arbeitsmedizin

DIREKTORIN: UNIV.-PROF. DR. MED. UTE WITTING

### Fragebogen zur Evaluation des Praxisprojektes Arbeitsmedizin im SS 2003

#### Gruppen A-H

##### Grundlagenseminare

	trifft völlig	überwiegend	selten	überhaupt nicht zu
Die Medien wurden sinnvoll und effektiv eingesetzt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das Thema wurde durch praxisnahe Beispiele veranschaulicht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhaltliche Fragen waren erwünscht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhaltliche Fragen wurden kompetent und verständlich beantwortet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es fiel mir leicht aufmerksam zu bleiben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

##### Praktische Übungen

Mein Basiswissen war ausreichend um den Übungen zu folgen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Lernstoff wurde im Dialog erarbeitet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhaltliche Fragen wurden kompetent und verständlich beantwortet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Der Lehrinhalt wurde interessant präsentiert.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es fiel mir leicht aufmerksam zu bleiben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

##### Betriebsbegehung

Welchen Betrieb haben Sie besichtigt (bitte eintragen):

-----

Die Betriebsbegehung war inhaltlich gut vorbereitet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Betriebsbegehung war organisatorisch gut vorbereitet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhaltliche Fragen wurden kompetent und verständlich beantwortet.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Betriebsbegehung hat mir einen Einblick über mögliche Gesundheitsgefahren am Arbeitsplatz verschafft.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**Gruppenarbeit**

trifft  
völlig    überwiegend    selten    überhaupt nicht  
zu

Die Diskussion in der Gruppe führte zu einem vertieften Verständnis des Lerninhaltes.

Die Gruppenarbeit mit Dozentenunterstützung war eine wichtige Hilfe für die Projektpräsentation.

Die Gruppenarbeit hat Spass gemacht.

**Gesamtbewertung**

Die LV war insgesamt gut und systematisch strukturiert.

Die Vermittlung wesentlicher Inhalte des Fachgebietes im Rahmen eines Praxisprojektes mit Gruppenarbeit hat mir gut gefallen.

Ich bin der Meinung, dass Kenntnisse über gesundheitliche Belastungen am Arbeitsplatz für meine spätere ärztliche Tätigkeit wichtig sind.

Mein persönlicher Lernerfolg durch die Veranstaltung war gross.

Die Überprüfung des Lernerfolgs durch das Abschlusskolloquium war dem Lernziel angemessen.

Die Lehrenden waren gut vorbereitet.

Ich hätte an der Veranstaltung auch freiwillig teilgenommen.

Die Veranstaltung hat mein Interesse für das Fach Arbeitsmedizin geweckt.

Was hat mir besonders gut gefallen?

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

Was sollte meiner Meinung nach geändert werden?

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

(sollte der Platz nicht ausreichen, bitte benutzen Sie die Rückseite)

**Viele Dank für Ihre Teilnahme**