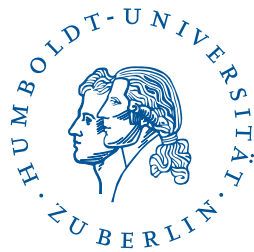


Bewertung von Lehrveranstaltungen mit Hilfe der Evaluationsdaten

Eine Masterarbeit von

Dipl.-Math. Melanie Reichelt (500969)

Erstgutachter: Prof. Dr. Wolfgang Härdle
Zweitgutachter: PD Dr. Marlene Müller
Betreuer: Dr. Sigbert Klinke



zur Erlangung des Grades
Master of Science in Statistic

Humboldt-Universität zu Berlin
Studiengang Statistik
Fachbereich Wirtschaftswissenschaften
Spandauer Str. 1 D-10178 Berlin

01. August 2007

Erklärung

Ich versichere: Ich habe die Masterarbeit selbständig verfasst. Andere als die angegebenen Hilfsmittel und Quellen habe ich nicht benutzt. Zitate wurden kenntlich gemacht.

Mir ist bekannt: Bei Verwendung von Inhalten aus dem Internet habe ich diese zu kennzeichnen und einen Ausdruck davon mit Datum sowie der Internet-Adresse (URL) als Anhang der Masterarbeit beizufügen.

Melanie Reichelt

Berlin, den

Inhaltsverzeichnis

1	Einleitung	9
2	Übersicht über die Daten	12
2.1	Verwendete Daten	12
2.2	Die fehlenden Daten	18
3	Statistische Methoden	26
3.1	Faktorenanalyse	26
3.1.1	Explorative Faktorenanalyse	27
3.1.2	Konfirmatorische Faktorenanalyse	32
3.1.3	Faktorwerte	35
3.2	Analyse der Faktorwerte	37
3.3	Clusteranalyse	39
4	Faktorenanalyse mit den Daten der Evaluation	43
4.1	Explorative Faktorenanalyse	43
4.1.1	Anzahl der Faktoren	43

4.1.2	Teststatistiken	44
4.1.3	Interpretation der Faktorladungen	47
4.2	Konfirmatorische Faktorenanalyse	50
4.2.1	Teststatistiken	51
4.2.2	Ergebnisse	54
4.3	Ein-Faktor-Modell	57
5	Bewertung der Lehrveranstaltungen mit Hilfe der Faktorwerte	59
5.1	Faktorwerte der Ein-Faktor-Modells	59
5.1.1	Auswertung der Größe, Level, Tag und Uhrzeit der LV	61
5.1.2	Abhängigkeit vom Dozenten	65
5.1.3	Verteilung der Faktorwerte	67
5.1.4	Clusteranalyse	68
5.2	Faktorwerte des Vier-Faktor-Modells	70
5.2.1	Clusteranalyse	70
5.2.2	Auswertung der Faktoren	72
6	Zusammenfassung	80
	Literaturverzeichnis	83
A	Appendix	87

Abbildungsverzeichnis

2.1	Auswertung der Studenten nach Geschlecht	15
2.2	Auswertung nach Studiengang	16
2.3	Links: Prozentuale Verteilung der Kurse, Rechts: Prozentuale Verteilung der ausgewerteten Bögen, HS=Hauptstudium, GS=Grundstudium	17
2.4	Auswertung der Lehrkräfte nach Geschlecht	17
2.5	Links: Prozentuale Verteilung der Lehrveranstaltungen an den Wochen- tagen, Rechts: Prozentuale Verteilung der Lehrveranstaltung zu be- stimmten Tageszeit	18
2.6	Anteil der fehlenden Werte pro Beobachtung	20
2.7	Fehlende Werte pro Variable für 4 Datensätze	21
3.1	Beispiel Pfaddiagramm	33
5.1	Auswertung der Vorlesungen nach Grund-und Hauptstudium	62
5.2	Auswertung der Vorlesungen nach der Größe	63
5.3	Auswertung der Vorlesungen nach dem Wochentag	64
5.4	Auswertung der Vorlesungen nach der Uhrzeit	65

5.5	Übungen	66
5.6	Vorlesungen	66
5.7	Vorlesungen SS06	67
5.8	Uebungen SS06	67
5.9	Dichteschätzung der Faktorwerte, blau: schlechter Kurs, schwarz: guter Kurs	68
5.10	Dendrogramm der Mittelwerte der Faktorwerte der Vorlesungen SS06 .	69
5.11	Dendrogramm der Mittelwerte der Faktorwerte der Vorlesungen SS06 (4F)	71
5.12	Scatterplot der Mittelwerte der Faktorwerte des 1. und 2. Faktors für die Übungen SS06	73
5.13	Scatterplot der Mittelwerte der Faktorwerte des 1. und 3. Faktors für die Übungen SS06	73
5.14	Scatterplot der Mittelwerte der Faktorwerte des 1. und 4. Faktors für die Übungen SS06	74
5.15	Lineplot der Faktorwerte (Ordinate) der ersten 3 Faktoren (Abszisse) - Vorlesungen SS06	75
5.16	Lineplot der Faktorwerte (Ordinate) der ersten 3 Faktoren (Abszisse) - Übungen SS06	76
5.17	Konfirmatorisches Faktorenmodell zweiter Ordnung Übungen SS06 . . .	77
5.18	Konfirmatorisches Faktorenmodell zweiter Ordnung Vorlesungen SS06 .	78

5.19 Zusammenhang zwischen Faktor zweiter Ordnung Lehre und Faktor An- forderung der Übungen im SS06	79
A.1 Fragebogen für Vorlesungen	96
A.2 Fragebogen für Übungen	97

Tabellenverzeichnis

2.1	Prozentuale Verteilung bei den Variablen 4 und 5	22
4.1	Erklärte Varianz für vier bzw. fünf Faktoren	44
4.2	Screeplots für Übung SS05 (oben links), Vorlesung SS05 (oben rechts), Übung SS06 (unten links) und Vorlesung SS06 (unten rechts)	45
4.3	Teststatistiken	47
4.4	Faktorladungen, rot: Ladungen $\in [0.5, 0.7]$, blau : Ladungen $\in [0.7, 1]$. . .	48
4.5	Teststatistiken	54
4.6	Ergebnisse für das Konstrukt Lehrbefähigung	55
4.7	Ergebnisse für das Konstrukt Skripte	55
4.8	Ergebnisse für das Konstrukt Selbsteinschätzung	56
4.9	Ergebnisse für das Konstrukt Leistungsanforderung	56
4.10	Faktorladungen, rot: Ladungen $\in [0.5, 0.7]$, blau : Ladungen $\in [0.7, 1]$. . .	58
5.1	Faktorwerte des Ein-Faktor-Modells der best bewerteten Lehrveranstaltungen im SS06	60

5.2 Faktorwerte des Ein-Faktor-Modells der schlechter bewerteten Lehrveranstaltungen im SS06	61
A.1 Prozentuale Verteilung des Studienganges und Geschlechts	88
A.2 Prozentuale Verteilung der Uhrzeiten und Tage	88
A.3 Anteil der fehlenden Werte pro Variable	89
A.4 Zuordnung der Variablen	90
A.5 Faktorwerte der Vorlesungen SS06	91
A.6 Faktorwerte der Vorlesungen SS05	92
A.7 Faktorwerte der Übungen SS06	93
A.8 Faktorwerte der Übungen SS05	94
A.9 Korrelation zwischen den Faktoren für Übungen SS05	94
A.10 Korrelation zwischen den Faktoren für Übungen SS06	95
A.11 Korrelation zwischen den Faktoren für Vorlesungen SS05	95
A.12 Korrelation zwischen den Faktoren für Vorlesungen SS06	95

1 Einleitung

Ganz allgemein betrachtet, bedeutet Evaluation den Prozess der Beurteilung des Wertes eines Produkts, Prozesses oder eines Programms. [Ahlstich, Ulrich-Neitzert (2002)]

In der Hochschulforschung ist Evaluation ein Sammelbegriff für unterschiedliche Bewertungs- und Analyseverfahren, die von der Hörer- oder Absolventenbefragung über die Durchführung von Einzelstudien bis hin zu dem Modell der so genannten internen und externen Evaluation reichen. Erst zu Beginn der 90er Jahre gewannen Maßnahmen zur Sicherung der Qualität der Lehre an den Hochschulen in Deutschland an Bedeutung. In dieser Arbeit geht es um das Verfahren der Veranstaltungsbewertung durch Studierende. 1994 wurde an der wirtschaftswissenschaftlichen Fakultät der Humboldt Universität von Studenten die Evaluationstelle gegründet und noch heute von Studenten geführt. Diese gibt in jedem Semester in der 12. und 13. Vorlesungswoche einen standardisierten Fragebogen an alle Lehrveranstaltungen aus. Der daraus resultierende Evaluationsbericht gibt eine Einschätzung von Lehrveranstaltungen aus Sicht der Studierenden. Dabei werden zentrale Dimensionen wie Aufbau der Veranstaltung, Angemessenheit von Stoffmenge und Schwierigkeitsgrad, Didaktik, Skripte und Literatur und Lernerfolgseinschätzung erhoben.

In einer ersten Voruntersuchung wird im zweiten Kapitel die Struktur der Evaluationsdaten des Sommersemesters 2005 und 2006 betrachtet. Hierzu gehört auch, die Auswahl der Variablen für die weiteren Untersuchungen und die Behandlung der fehlenden Werte mit geeigneten Verfahren. Am Ende des ersten Kapitels liegen vier vollständige Datensätze vor (Vorlesungen SS05, Vorlesungen SS06, Übungen SS05 und Übungen SS06).

Das dritte Kapitel beschäftigt sich mit den statistischen Methoden, die für die Untersuchung der Daten verwendet werden. Hierzu gehören u.a. die explorative und konfirmatorische Faktorenanalyse angewendet auf ordinale Daten. Die explorative Faktorenanalyse liefert als hypothesengenerierendes Verfahren ein Konstrukt aus latenten und manifesten Variablen, welches mit Hilfe der konfirmatorischen Faktorenanalyse in Bezug auf die vorliegenden Daten getestet wird. Die konfirmatorische Faktorenanalyse liefert auch die Faktorwerte, die zur weiteren Untersuchung der Daten verwendet werden.

Im vierten und fünften Kapitel wird dann mit Hilfe der statistischen Verfahren und der Daten versucht zu klären, welche Aussagen getroffen werden können. Zum einen ist interessant, in wie fern sich die Ergebnisse über die Evaluationsdaten der Sommersemester 2002 und 2003 von Yilan Zhou (2004) im Vergleich zu den Ergebnissen in dieser Arbeit über die Evaluationsdaten der Sommersemester 2005 und 2006 verändern bzw. bestätigen lassen. Des Weiteren stellt sich die Frage, ob die Evaluationsdaten eine Bewertung von guten und schlechten Lehrveranstaltungen liefern und wenn ja, ob diese Bewertung Aussagen darüber zulässt, welche Merkmale bzw. Eigenschaften eine gute

bzw. schlechte Lehrveranstaltung ausmachen.

Als Software zur Umsetzung der verschiedenen Verfahren wurde XploRe, Mplus und SPSS verwendet.

2 Übersicht über die Daten

2.1 Verwendete Daten

Die verwendeten Datensätze resultieren aus den Evaluationsbögen der wirtschaftswissenschaftlichen Fakultät der Humboldt Universität zu Berlin der Sommersemester 2005 und 2006. Insgesamt wurden 186 Kurse mit 8667 Beobachtungen ausgewertet. Von der Evaluationsstelle werden für Vorlesungen, Übungen und Seminare verschiedene Fragebögen ausgegeben. Da die Struktur des Seminarfragebogens anders ist als bei den Fragebögen für Vorlesungen und Übungen, wird dieser aus der Betrachtung herausgenommen. Somit ergeben sich vier Datensätze - Vorlesung SS06, Übung SS06, Vorlesung SS05 und Übung SS05 -, die sich auf der beigefügten CD befinden. Die Fragebögen für Vorlesung und Übung gliedern sich in 7 Blöcke:

1. Allgemeine Fragen

- Studiengang (12 Ausprägungen)
- Geschlecht (m/w)
- Wie oft Vorlesung versäumt? (nie/1x/2x/3x/öfter)
- Warum Vorlesung versäumt? (4 Ausprägungen und Textfeld)

- Gesamteindruck (5 Ausprägungen 1 - 5)
2. Dozent/Übungsleitung
3. Konzept/Kommunikation
- jeweils 5 Ausprägungen (sehr gut - schlecht)
 - ersten 3 Fragen zum Dozent/Übungsleitung
 - 2 Fragen zur Bereitschaft und Qualität Beantwortung Zwischenfragen
 - 2 Fragen zu Schwerpunkte, Gliederung des Stoffes
 - auf dem Fragebogen für Übungen befindet sich eine zusätzliche Frage: Abstimmung Vorlesung und Übung
4. Lehrmaterial
- jeweils 5 Ausprägungen (sehr gut - schlecht)
 - 3 Fragen zu Skripten und Literatur
 - Qualität
 - Verfügbarkeit
 - Internetpräsenz
 - auf dem Fragebogen für Vorlesungen befindet sich eine zusätzliche Frage:
Veranschaulichung theoretischer Inhalte anhand praktischer Beispiele
5. Leistungsanforderung
- 3 Fragen mit 5 Ausprägungen (zu hoch - zu niedrig)
 - Frage Vor- und Nachbereitungszeit mit 5 Ausprägungen

- 1 = 0
- 2 = bis 30min
- 3 = bis 60min
- 4 = bis 90min
- 5 = mehr als 90min

6. Selbsteinschätzung

- jeweils 5 Ausprägungen (hoch - niedrig)
- 2 Fragen zu Aufmerksamkeit, Erkenntnis- bzw. Lernzuwachs
- auf dem Fragebogen für Vorlesungen befindet sich eine zusätzliche Frage:
Interesse am Fach

7. Atmosphäre

- jeweils 5 Ausprägungen
- stressfrei - stressig
- interessant - langweilig
- diszipliniert - chaotisch
- motivierend - geistig blockierend

Den gesamten Fragebogen für Vorlesungen und Übungen befindet sich im Anhang und auf der beigelegten CD. Alle Variablen sind ordinalskaliert. In der Abbildung 2.1 wird die prozentuale Verteilung von männlichen und weiblichen Studenten im Sommersemester 2005, die an der Evaluation teilgenommen haben im Vergleich zu allen

immatrikulierten Studenten an der Fakultät abgebildet. Die Verteilung liegt jeweils bei ca. 50%.

In der Abbildung 2.2 wird die Verteilung der Studenten nach ihrem Studiengang

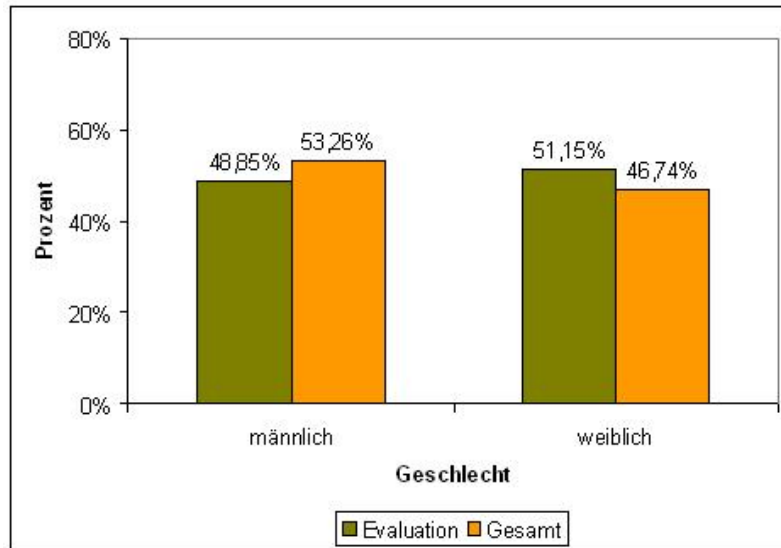


Abbildung 2.1: Auswertung der Studenten nach Geschlecht

im Sommersemester 2005 dargestellt. Hierbei wurden die Studiengänge BWL Diplom, Bachelor und NZ/FZ zu BWL, die Studiengänge VWL Diplom, Bachelor und NZ/FZ zu VWL und die anderen Studiengänge in Andere zusammengefasst. Im Vergleich der Verteilungen bei den evaluierten Studenten und den Gesamtzahlen ist Verteilung ungefähr gleich.

Vergleicht man wie in Abbildung 2.3 die Anzahl der evaluierten Kurse im Grund- und Hauptstudium mit den abgegebenen Bögen (Beobachtungen) im Sommersemester 2005, wird deutlich, dass im Grundstudium prozentual weniger Kurse angeboten

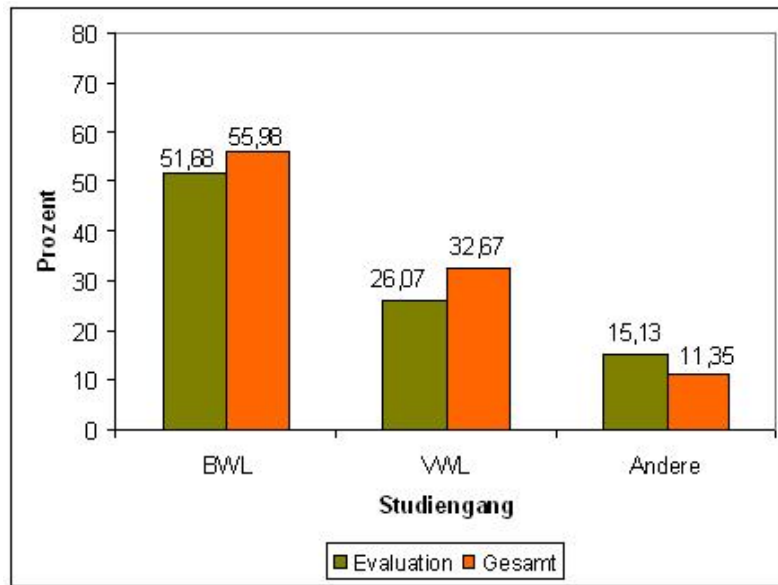


Abbildung 2.2: Auswertung nach Studiengang

werden und die Anzahl der Studenten im Grund- und Hauptstudium in etwa gleich ist. Daraus lässt sich schließen, dass in den Kursen im Grundstudium mehr Studenten sitzen als in den Kursen im Hauptstudium. Ob dies einen Einfluss auf die Bewertung der Kurse hat, wird zu einem späteren Zeitpunkt untersucht werden.

In der Abbildung 2.4 sieht man die prozentuale Verteilung von männlichen und weiblichen Lehrkräften in den Sommersemestern 2005 und 2006, die eine Lehrveranstaltung gehalten haben im Vergleich zu allen Lehrkräften an der Humboldt Universität zu Berlin.

Im Gegensatz zu der Geschlechterverteilung bei den Studenten, sind über 80% der Dozenten männlich, sowohl in den Semestern 2005 und 2006, als auch an der gesamten Universität.

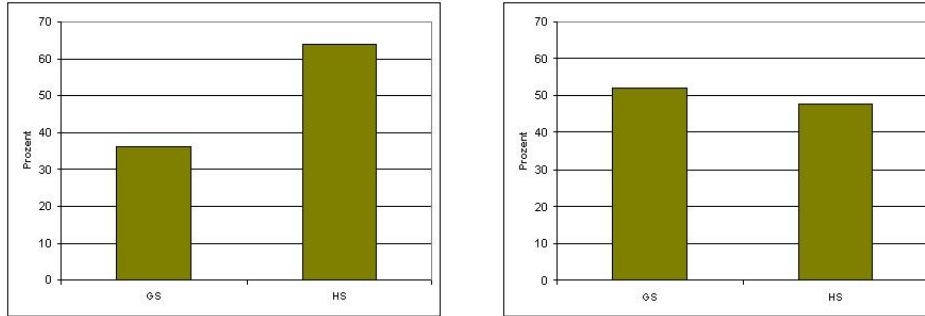


Abbildung 2.3: Links: Prozentuale Verteilung der Kurse, Rechts: Prozentuale Verteilung der ausgewerteten Bögen, HS=Hauptstudium, GS=Grundstudium

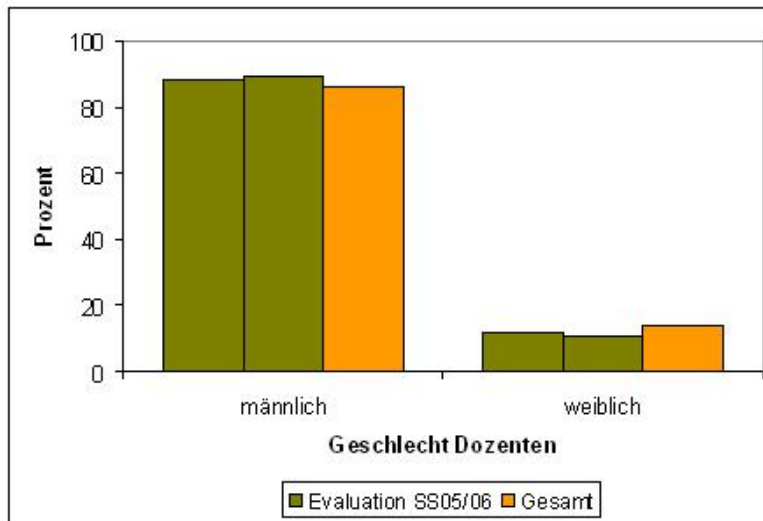


Abbildung 2.4: Auswertung der Lehrkräfte nach Geschlecht

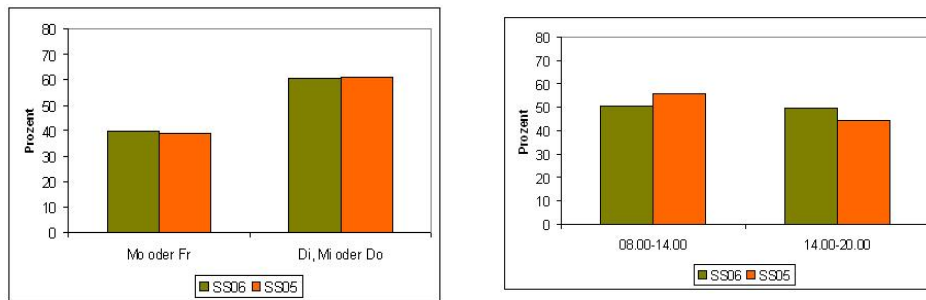


Abbildung 2.5: Links: Prozentuale Verteilung der Lehrveranstaltungen an den Wochentagen, Rechts: Prozentuale Verteilung der Lehrveranstaltung zu bestimmter Tageszeit

Die meisten Lehrveranstaltungen finden zwischen Montag und Freitag in der Zeit von 08.00 Uhr bis 20.00 Uhr statt. In Abbildung 2.5 sieht man, dass zum einen weniger Kurse an den Randtagen stattfinden, d.h., Montag und Freitag und dass zum anderen die Verteilung der Lehrveranstaltungen auf den Vor- oder Nachmittag ungefähr gleich ist.

2.2 Die fehlenden Daten

Wie in den meisten Daten aus der Realität fehlen auch bei dieser Untersuchung Kurse, Beobachtungen oder einzelne Werte. Das Fehlen kompletter Kurse kann verschiedene Ursachen haben:

- Es wurde ein falscher Bogen ausgefüllt, z.B. in einer Vorlesung wurde ein Übungsbogen ausgefüllt.
- Es wurden keine Bögen an die Evaluationsstelle zurückgegeben.
- Die Veranstaltungsgröße ist zu gering. Aus datenschutzrechtlichen Gründen dür-

fen Veranstaltungen, die weniger als 10 Teilnehmer haben, nicht ausgewertet werden.

In den vier Datensätzen sind nicht alle Studenten (Beobachtungen) erfasst, da

- der Bogen am Ende der Lehrveranstaltung nicht abgegeben wurde, oder
- der Student am Tag der Evaluation nicht anwesend war.

In den vier Datensätzen fehlen insgesamt 5,66% der Daten. Innerhalb der vier Datensätze liegt der Anteil der fehlenden Werte pro Beobachtung meist $\leq 5\%$. Als Veranschaulichung werden in Abbildung 2.6 die Beobachtungen des Datensatzes Vorlesung SS05 den prozentual fehlenden Werten gegenübergestellt. Die horizontale Linie markiert die 5% Marke. Abbildung 2.7 zeigt die fehlenden Werte in Prozent für jede Variable. Bei den meisten Variablen liegt der Anteil der fehlenden Werte unter 5%. Bei den Variablen 2 (Geschlecht), 3 (Gesamteindruck) und 24 (Vor- und Nachbereitungszeit) fehlen jeweils mehr als 10%. Bei der Variable 5 (Warum hast du gefehlt?) fehlen in allen Datensätzen mehr als 40%. Dies liegt unter anderem daran, dass man hier auf ein strukturelles Problem des Fragebogens trifft.

In Variable 4 wird gefragt, wie oft man die Vorlesung versäumt hat. Wenn dies nie der Fall war, wird die befragte Person, bei der Frage warum sie gefehlt hat, nichts angeben. Dies sieht man auch in der Kreuztabelle 2.1, bei der die beiden Variablen für den Datensatz Vorlesung SS05 gegenübergestellt wurden. Zum einen geben mehr als 10% der Studenten keinen Grund für ihr Fehlen an und zum anderen geben 14 Studenten aus diesem Datensatz einen Grund an, obwohl sie vorher angegeben haben,

Vorlesung 05

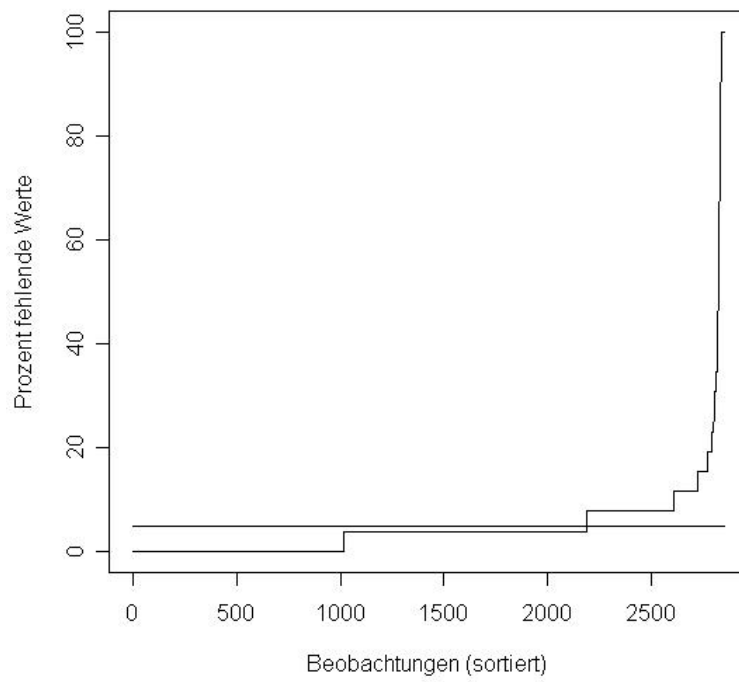


Abbildung 2.6: Anteil der fehlenden Werte pro Beobachtung

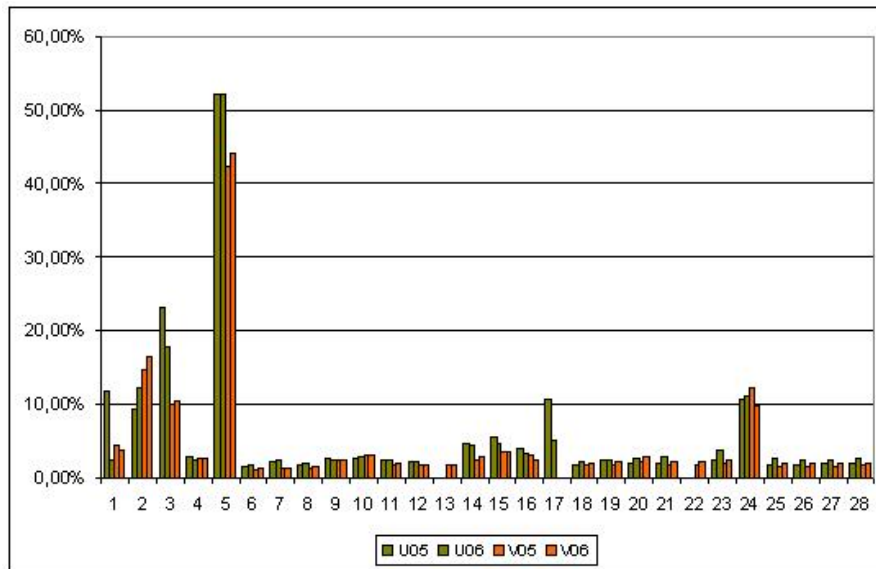


Abbildung 2.7: Fehlende Werte pro Variable für 4 Datensätze

dass sie gar nicht gefehlt haben.

Um die fehlenden Daten näher zu untersuchen wird folgende Bezeichnung eingeführt.

Z bezeichnet die Datenmatrix, wobei die Spalten von Z die Variablen enthalten und die Zeilen von Z die Beobachtungen. Der beobachtete Teil der Datenmatrix Z wird mit Z_{obs} bezeichnet und der fehlende Teil mit Z_{mis}

$$Z = (Z_{obs}, Z_{mis}).$$

Für die Einordnung der fehlenden Daten gibt es verschiedene Mechanismen.

1. Ein Fehlend-Muster gibt an, welche Werte der Datenmatrix beobachtet und welche fehlend sind. Man unterscheidet in monoton und zufällig oder nichtmonotone Fehlend-Muster. Ein Fehlend-Muster in der Datenmatrix Z heißt monoton, wenn die Zeilen und Spalten von Z so neu geordnet werden können, dass die folgende

Wie oft?/Warum?	zu früh/zu spät	Überschneidung	gibt mir nichts	Belastung	Sonstiges	k.A.
nie	3	1	3	1	6	780
1x	59	98	16	270	94	188 (26.04%)
2x	61	85	35	216	110	92 (15.36%)
3x	42	45	31	78	76	46 (14.47%)
öfter	29	71	66	77	57	40 (11.77%)

Tabelle 2.1: Prozentuale Verteilung bei den Variablen 4 und 5

Bedingung zutrifft.

Für $a < b$ gilt: Aus der Beobachtung einer bestimmten Variablen x_{ib} für einen Studenten $i, i = 1, \dots, n$ folgt, dass auch die Variable x_{ia} für diesen Fall beobachtet ist [Morton, Lipsitz (2001)].

Um dies zu untersuchen führt man eine Indikatormatrix $R = (r_{ij})$ ein, deren Elemente $r_{ij} = 1$ sind, wenn z_{ij} beobachtet ist und $r_{ij} = 0$ sind, wenn z_{ij} fehlt. Liegt eine monotonen Fehlend-Muster vor, können einfache Imputationsmethoden angewendet werden. In der Praxis ist das Fehlend-Muster aber häufig völlig zufällig.

2. Eine weitere Einordnung gibt drei Möglichkeiten für das Fehlen der Daten an. Nach Schäfer [Schäfer (1997)] werden Fehlend-Mechanismen formalisiert, indem die Indikatoren für fehlende Werte wie Zufallsvariablen behandelt werden. Die Fehlend-Mechanismen werden formal als Wahrscheinlichkeitsmodelle für das Fehlen von Daten definiert.

Missing Completely At Random (MCAR) Die Wahrscheinlichkeit, einen bestimmten Wert zu beobachten, hängt weder von den beobachteten Werten noch von den fehlenden Werten ab. Das bedeutet, die Verteilungen in Z haben keinen Einfluss auf die Verteilung von R .

$$P(R|Z) = P(R) \forall Z$$

Missing At Random (MAR) Die Wahrscheinlichkeit, einen bestimmten Wert zu beobachten, ist abhängig von den fehlenden Werten, aber unabhängig von den beobachteten Werten. Das bedeutet, dass nur die beobachteten Daten Z_{obs} in Z Einfluss auf die bedingte Verteilung von P gegeben Z haben.

$$P(R|Z) = P(R|Z_{obs}) \forall Z_{mis}$$

Not Missing At Random (NMAR) Die Wahrscheinlichkeit, einen bestimmten Wert zu beobachten, hängt von den fehlenden Werten ab. Das bedeutet, die bedingte Wahrscheinlichkeit $P(R|Z)$ kann nicht vereinfacht werden.

$$P(R|Z) = P(R|Z_{obs}, Z_{mis}) = P(R|Z) \forall Z$$

Der Umgang mit fehlenden Werten kann auf verschiedene Weise geschehen.

- Casewise Deletion

Bei dieser Methode werden alle Fälle ausgeschlossen, in denen mindestens eine Ausprägung fehlt. Der Vorteil an dieser Methode ist, dass sie sehr einfach umzusetzen ist und dass auf die vollständigen Daten Standardanalysen angewendet

werden können. Besonders nachteilig wirkt sich die Reduzierung der Fälle aber auf die Ergebnisse aus.

- Die berechneten Schätzer sind verzerrt, besonders dann, wenn der Fehlend-Mechanismus nicht MCAR ist.
- Schätzungen verlieren an Präzision.
- Es ergeben sich größere Standardfehler und breitere Konfidenzintervalle.

- Analyse mit Fehlend-Kategorie

Bei dieser Methode erhalten alle unvollständig beobachteten Variablen eine zusätzliche Kategorie für die fehlenden Werte. Die zusätzliche Kategorie wird wie eine weitere Merkmalsausprägung der Variablen behandelt. Auch hier werden die Ergebnisse stark verzerrt.

- Imputation

Das Ziel der Imputation ist die Ersetzung von fehlenden Werten durch „passende“ Werte. Es gibt eine Vielzahl von Verfahren, die unterschieden werden in single und multiple Imputation. Bei der single Imputation wird für jeden fehlenden Wert ein Wert generiert, während bei der multiple Imputation für jeden fehlenden Wert mehrere Werte generiert werden. Die fehlenden Daten in dieser Arbeit wurden mit der single Imputation ersetzt.

Umsetzung in R

Mit Hilfe der Software R werden die imputierten Datensätze erzeugt. Der Expectation-Maximization-(EM)-Algorithmus wird durchgeführt, um den Maximum-Likelihood-

Schätzer unter dem saturierten Modell zu finden. Das saturierte Modell enthält alle möglichen Interaktionen bzw. Wechselwirkungen zwischen den Variablen. Die zugrunde liegende Idee ist, dass die fehlenden Werte Z_{mis} basierend auf einem Startwert für den Parameter Θ aufgefüllt werden. Dann wird Θ basierend auf den beobachteten Daten Z_{obs} und den aufgefüllten Werten für Z_{mis} neu geschätzt. Dieser Prozess wird iterativ so lange durchgeführt, bis die Schätzungen für den Parameter konvergieren. Die hauptsächlich verwendeten Funktionen aus dem Paket `cat` sind:

prelim.cat Diese Funktion hilft bei der Erstellung der Fehlend-Muster. Hierzu wird der kategorielle Datensatz sortiert, eventuell gruppiert und eine Liste erzeugt, die zur Eingabe für die Funktion `em.cat`, `da.cat` und `imp.cat` benötigt wird.

da.cat Bei dieser Funktion wird die Methode der Data augmentation zum Auffüllen der fehlenden Daten angewendet.

em.cat Diese Funktion berechnet den Maximum-Likelihood-Schätzer der Zellwahrscheinlichkeiten einer Kontingenztafel unter dem saturierten Modell.

imp.cat Diese Funktion imputiert die fehlenden Werte in einem kategoriellen Datensatz mit Hilfe eines benutzerdefinierten Wertes für die zugrunde liegenden Zellwahrscheinlichkeiten.

Weitere Ausführungen zur Theorie der single und multiple Imputation und der Umsetzung in R findet man unter anderem bei Schäfer [Schäfer (1997)].

3 Statistische Methoden

In dem folgenden Kapitel werden die statistischen Verfahren beschrieben, die im weiteren Verlauf verwendet werden.

3.1 Faktorenanalyse

Die Faktorenanalyse ist ein Verfahren der Datenreduktion. Den Ausgangspunkt bildet eine große Anzahl an Variablen. Vor der Analyse ist nicht bekannt, ob und wie die einzelnen Variablen zusammenhängen. Mittels explorativer und konfirmatorischer Faktorenanalyse wird untersucht, ob Zusammenhänge bestehen, inwiefern die Variablen latente Strukturen bilden und wie diese Konstrukte interpretiert werden können. Da die Art und Anzahl der Faktoren zunächst nicht bekannt ist, wird zunächst eine explorative Faktorenanalyse durchgeführt. Innerhalb dieser Analyse wird untersucht, ob zwischen den Variablen Zusammenhänge existieren und ob die Variablen Konstrukte bilden. Diese latenten Konstrukte werden in einem Modell formuliert. Hypothesen bezüglich der zugrunde liegenden Faktorenstruktur werden innerhalb der konfirmatorischen Faktorenanalyse geprüft, um eine Aussage über die Struktur der Daten zu machen.

3.1.1 Explorative Faktorenanalyse

Ziel der explorativen Faktorenanalyse ist es, nicht messbare Strukturen aufzudecken, die hinter p beobachtbaren Variablen in der Datenmatrix X vermutet werden. Im Idealfall können alle Informationen, die in X stecken durch eine kleinere Anzahl von Faktoren reproduziert werden. Dieser Fall tritt ein, wenn alle $x = (x_1, \dots, x_p)^\top$ wie folgt erklärt werden können [Härdle (2003)]:

$$x_j = \sum_{l=1}^k q_{jl} f_l + \mu_j \quad j = 1, \dots, p.$$

Es wird angenommen, dass es k Faktoren f_l ($l = 1, \dots, k$) gibt, wobei $k < p$ und die Anzahl k der Faktoren unbekannt ist. In Matrixschreibweise sieht das Modell wie folgt aus:

$$X = QF + \mu$$

In der Faktorenanalyse ist es üblich, den Einfluss der Faktoren in gemeinsame und Einzelrestfaktoren zu unterteilen. Es ergibt sich das allgemeine Modell:

$$X = QF + U + \mu$$

wobei

- $X_{(p \times 1)}$ die Matrix der standardisierten Beobachtungswerte z_{ij} mit $j = 1, \dots, m$ und $i = 1, \dots, n$,
- $Q_{(p \times k)}$ die q_{jl} die Ladungen der j -ten Variablen auf den l -ten Faktor,

- $F_{(k \times 1)}$ die gemeinsamen Faktoren,
- $U_{(p \times 1)}$ die Einzelrestfaktoren und
- $\mu_{(p \times 1)}$ die Mittelwert der Variablen j sind.

Annahmen für das Modell der Faktorenanalyse sind, dass die Faktorvariablen F unkorreliert sind, dass die Einzelrestfaktoren unkorreliert sind und dass die Kovarianz zwischen den Einzelrestfaktoren und den gemeinsamen Faktoren 0 ist. In Matrixschreibweise ausgedrückt:

$$EF = 0$$

$$\text{Var}(F) = I_k$$

$$EU = 0$$

$$\text{Cov}(U_i, U_j) = 0$$

$$\text{Cov}(F, U) = 0.$$

Da die Matrizen bis auf X alle unbekannt sind, ist das Modell in dieser Form nicht lösbar. Schritt für Schritt werden die einzelnen Matrizen auf Grundlage der Korrelationsmatrix nacheinander geschätzt.

Bei metrisch skalierten Variablen wird hierfür die Bravais-Pearson-Korrelation verwendet. Gemeinsame Faktoren existieren nur für Variablen, die stark miteinander korreliert sind. Variablen, die geringe Korrelationen mit anderen Variablen aufweisen, können eventuell unberücksichtigt bleiben. Die Faktorladungen können nur iterativ geschätzt werden. Es existieren unterschiedliche Extraktionsmethoden, die zu verschiedenen Fak-

torlösungen gelangen können:

- Maximum-Likelihood-Methode
- Hauptkomponentenanalyse
- Hauptachsenmethode

Bei kategoriellen Daten gibt es drei Ansätze, um eine Faktorenanalyse durchzuführen:

1. Man fasst die kategoriellen Daten in binäre Daten zusammen und wendet auf diese eine Faktorenanalyse für binäre Daten an.
2. Man verwendet ein multinomiales Logit-Modell. Dieser Ansatz wird als item response function bezeichnet.
3. Man verwendet den Ansatz der *underlying variable*, der auch in der hier verwendeten Software Mplus implementiert ist.

Beim *underlying variable* Ansatz nimmt man an, dass für jede kategorielle Variable x_i eine kontinuierliche *underlying variable* x_i^* existiert. Die *underlying variable* ist normalverteilt mit dem Mittelwert μ und der Varianz σ_i^2 . Die Verbindung zwischen x_i mit m Kategorien und x_i^* ist

$$x_i = s \Leftrightarrow \tau_{i(s-1)} < x_i^* < \tau_{i(s)} \quad s = 1, 2, \dots, m_i$$

die Parameter

$$-\infty = \tau_{i(0)} < \tau_{i(1)} < \tau_{i(2)} < \dots < \tau_{i(m-1)} < \tau_{i(m)} = +\infty$$

werden als Schwellenwerte bezeichnet. Wenn die Daten in m Kategorien vorliegen, so existieren $(m - 1)$ Schwellenwerte.

Analog zum Faktorenmodell für metrische Daten wird das Faktorenmodell mit der *underlying variable* formuliert:

$$x_i^* = \alpha_{i1}^* y_1 + \alpha_{i2}^* y_2 + \dots + \alpha_{iq}^* y_q + e_i$$

mit α_{ij}^* als Faktorladungen und y_i als latente Variablen. Die Annahmen des Modells sind:

- Die latenten Variablen y_i sind unabhängig und standardnormalverteilt mit dem Mittelwert 0 und der Varianz 1.
- Die Residuen e_i sind ebenfalls unabhängig und standardnormalverteilt.
- Die *underlying variable* x_i^* ist normalverteilt.

Anstatt der Bravais-Pearson-Korrelation verwendet man bei kategoriellen Variablen die polychorische Korrelation. Für die Variable x_i^* kann man im Prinzip jede stetige Verteilung wählen. Günstigerweise wählt man die Standardnormalverteilung. Dann lässt sich die Wahrscheinlichkeit, dass x_i^* unterhalb des Schwellenwertes τ_i liegt mit Hilfe der Normalverteilungsfunktion an der Stelle τ_i berechnen

$$P(x_i^* \leq \tau_i) = \int_{-\infty}^{\tau_i} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} x_i^{*2}\right) dx_i^* = \Phi(\tau_i)$$

Die polychorische Korrelation ist die Korrelation in der zweidimensionalen Normalverteilung der *underlying variable* x_i^* und x_j^* . Die Schwellenwerte werden mit der Inversen

der Standardnormalverteilung Φ^{-1} geschätzt.

$$\hat{\tau}_i = \Phi^{-1}(p_1 + p_2 + \dots + p_i) \quad i = 1, \dots, m - 1.$$

Nachdem die Schwellenwerte geschätzt wurden und Korrelationen zwischen den Variablen bestimmt wurden, werden dann die Faktorladungen des Modells aus den Korrelationen durch die gewichteten kleinsten Quadrate geschätzt.

Durch Drehung des Koordinatensystems der Faktoren können die Faktorladungen erhöht werden. Die Zuordnungen der Variablen ändert sich dadurch nicht. Mit Hilfe der Rotationsmethoden sucht man eines der Koordinatensysteme in dem gemeinsamen Faktorenraum. Die Faktorladungsmatrix beschreibt den Zusammenhang zwischen Variablen und Faktoren. Durch die Transformation der Faktoren sind die Ergebnisse besser interpretierbar. Eine orthogonale Transformation auf orthonormierte Faktoren angewendet ergibt wieder orthonormierte Faktoren. Bei einer nichtorthogonalen (obliquen Transformation) werden auch abhängige Faktorvariablen zugelassen.

1. Varimax-Methode

Bei der Varimax-Methode versucht man Faktoren herzustellen, die einige Variablen hoch, in den anderen aber sehr niedrig geladen sind. Der Name kommt daher, dass bei dieser Methode die Varianz der Ladungsquadrate maximiert werden soll. Die Maximierung geschieht iterativ. Die Methode liefert gute Ergebnisse, wenn die Daten Gruppenfaktoren zugrunde liegen und keine Generalfaktoren vorliegen. Gruppenfaktoren sind Faktoren, die nur einen Teil der Variablen erklären.

Generalfaktoren bestimmen jede Variable mit.

$$\mathbf{L} = \begin{pmatrix} x & & x & x \\ x & & x & \\ x & & x & x \\ & x & x & x \\ & x & x & x \\ & x & x & x \\ & & x & x & x \\ & & x & x & x \end{pmatrix}$$

Die ersten drei Spalten der Matrix L zeigen einen Gruppenfaktor und die letzten beiden Spalten einen Generalfaktor.

2. Promax-Methode

Die Promax-Rotation verbessert das Ergebnis einer orthogonalen Varimax-Rotation durch eine oblique Transformation. Bei dieser Rotation sollen die Ladungsquadrate noch näher an 1 bzw. 0 gebracht werden.

3. Oblimin-Methode

Die Methode ist eine weitere nichtorthogonale Methode. Unter dem Oberbegriff Oblimin werden verschieden Kriterien zusammengefasst, die eine Minimierung der 4. Momente von Faktorenladungen zum Ziel haben.

3.1.2 Konfirmatorische Faktorenanalyse

Im Gegensatz zur explorativen Faktorenanalyse findet bei der konfirmatorischen Faktorenanalyse (CFA) keine Datenreduktion statt, sondern man untersucht die Übereinstimmung eines theoretischen Modells mit den empirischen Daten. Hierbei laden die Indikatoren nur auf einen Faktor und nicht wie bei der explorativen Faktorenanalyse

auf mehrere Faktoren gleichzeitig. Die konfirmatorische Faktorenanalyse ist ein Spezialfall von Strukturgleichungsmodellen (SEM). Ein Strukturgleichungsmodell besteht aus zwei Teilen: zum einen besteht es aus einem multivariaten Regressionsmodell, welches den Zusammenhang zwischen den abhängigen beobachteten Variablen und den latenten Variablen darstellt und zum anderen aus einem strukturellen Modell, welches die latenten Variablen miteinander verbindet. Zur graphischen Darstellung der CFA verwendet man ein Pfaddiagramm (siehe Abbildung 3.1).

In den großen Ovalen stehen die latenten Variablen und in den Quadraten stehen die

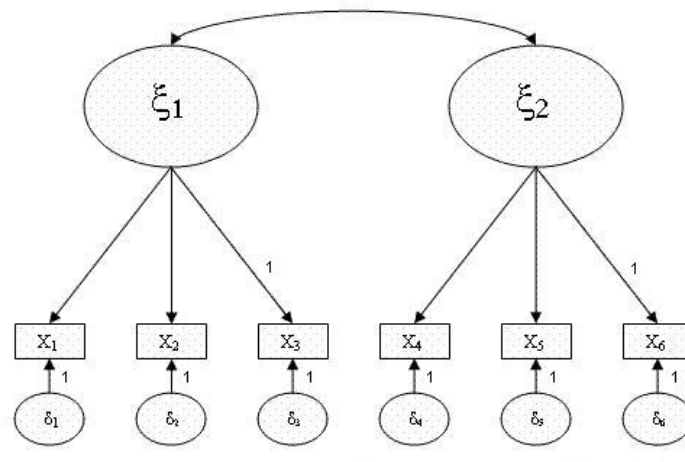


Abbildung 3.1: Beispiel Pfaddiagramm

beobachteten Variablen, die aufgrund der explorativen Analyse den latenten Variablen zugeordnet wurden. Unter den Quadraten befinden sich die Kreise für die Fehlerterme δ_i . Die Fehlerterme sind untereinander unabhängig. Zwischen den latenten Variablen besteht des Weiteren noch eine Korrelation, welche durch den Doppelpfeil symbolisiert

wird. Jeweils ein Pfad zwischen der beobachteten Variablen und der latenten Variablen wird mit einer 1 fixiert. Dies bedeutet, dass eine Veränderung um 1 in der beobachteten Variablen mit einer Veränderung um 1 in der latenten Variablen einhergeht. In einer Formel ausgedrückt ergibt sich für die CFA:

$$X = \Lambda\xi + \delta$$

mit

- X als Vektor der beobachteten Variablen,
- Λ der Matrix der Ladungen,
- ξ dem Vektor der gemeinsamen Faktoren und
- δ dem Vektor der Fehlerterme.

Es wird angenommen, dass $E[\delta] = 0$ und dass die Faktoren und die Fehlerterme unkorreliert sind $E[\xi\delta] = 0$. Die allgemeine Formel kann für das in der Abbildung 3.1 dargestellte Beispiel wie folgt umgeschrieben werden

$$\begin{aligned} x_1 &= \lambda_{11}\xi_1 + \delta_1 & x_2 &= \lambda_{21}\xi_1 + \delta_2 & x_3 &= \lambda_{31}\xi_1 + \delta_3 \\ x_4 &= \lambda_{42}\xi_2 + \delta_4 & x_5 &= \lambda_{52}\xi_2 + \delta_5 & x_6 &= \lambda_{62}\xi_2 + \delta_6 \end{aligned}$$

An dieser Stelle sieht man auch deutlich die Gemeinsamkeit mit der Regressionsanalyse, denn jedes x_i wird durch eine lineare Funktion der Faktoren und eines Fehlerterms dargestellt. Der Hauptunterschied zwischen der Regressionsanalyse und der CFA ist, dass die ξ_i unbekannt sind. Die vorher beschriebenen Beschränkungen, dass eine Faktorladung mit 1 festgelegt wird, ist notwendig, da sonst das Modell nicht eindeutig

bestimmt ist.

Die Schätzungen der Parameter bei kategoriellen Variablen erfolgt unter Mplus durch die Methode WLSMV (weighted least squares mean and variance adjusted). Diese Methode wird einer Maximum-Likelihood-Schätzung vorgezogen, da Muthén [Muthén (1984)] gezeigt hat, dass die berechneten Schätzer konsistent und asymptotisch normalverteilt sind. Bei dieser Methode wird zur Schätzung der Parameter eine Gewichtungsmatrix W verwendet, welche die geschätzte Varianz der Parameter des Modells ohne Beschränkungen enthält. Mit dieser wird dann die asymptotische Kovarianzmatrix geschätzt.

3.1.3 Faktorwerte

Die unter Mplus im Zuge der konfirmatorischen Faktorenanalyse ausgegebenen Faktorwerte geben an, wie stark die in einem Faktor zusammengefassten Merkmale in einer Variablen ausgeprägt sind. Sie ergeben sich aus der gewichteten Linearkombination der manifesten Variablen. Die zugehörigen Gewichte der Items w_i werden als „factor score coefficients“ bezeichnet. Für eine gegebene Beobachtung und Faktor werden die Faktorwerte wie folgt berechnet:

1. Für jede Variable einer Beobachtung wird der standardisierte Wert genommen.
2. Der standardisierte Wert wird mit den korrespondierenden Faktorladungen der Variablen für den gegebenen Faktor multipliziert.
3. Abschließend werden die Produkte aufsummiert.

$$\hat{F}_{ik} = \sum_{j=1}^J w_{jk} * Z_{X_{ij}}$$

Darin ist

- \hat{F}_{ik} der Faktorwert der Beobachtung i auf den Faktor k ,
- w_{jk} sind die Faktorgewichte der Variablen j auf den Faktor k und
- $Z_{X_{ij}}$ sind die z-standardisierten Werte der Bewertung der Beobachtung i auf die Variable j

Die Bestimmung der Faktorwerte hängt von der verwendeten Methode bei der Faktorenanalyse ab. Gegebenenfalls wird die Formel noch um die Matrix der Residuen/Fehlerterme der Faktorwerte erweitert. Da in die Berechnung alle Faktorladungen der Faktorladungsmatrix einfließen, haben auch kleine Ladungen Einfluss auf die Größe der Faktorwerte. Die Werte können positiv oder negativ ausfallen bzw. (näherungsweise) bei Null liegen.

- Positive Faktorwerte bedeuten, dass ein Objekt in Bezug auf diesen Faktor im Vergleich zu allen anderen Objekten überdurchschnittlich ausgeprägt ist.
- Faktorwerte gleich 0 bedeuten, dass ein Objekt in Bezug auf diesen Faktor durchschnittlich ausgeprägt ist.
- Negative Faktorwerte bedeuten, dass ein Objekt in Bezug auf diesen Faktor im Vergleich zu allen anderen Objekten unterdurchschnittlich ausgeprägt ist.

Die Auswertung von Faktorwerten kann aber problematisch sein, da u.a.

- konkurrierende Methoden zur Bestimmung von Faktorwerten existieren,
- verschiedene Rotationsmethoden zu unterschiedlichen Faktorwerten führen und
- die Faktorwerte durch Fehler beeinflusst werden.

Daher sollten Faktorwerte nicht als vollkommenes Maß für Faktoren betrachtet werden. Nimmt man die Faktorwerte als Variablen, können diese für nachfolgende Modellierungen genutzt werden.

3.2 Analyse der Faktorwerte

Um die durch die Faktorwerte neu gebildeten Variablen weiter zu untersuchen werden im weiteren Verlauf u.a. folgenden Verfahren verwendet.

1. Kruskal-Wallis-Test

Der Kruskal-Wallis-Test ist ein nichtparametrisches Verfahren, welches zum Vergleich von Mittelwerten verwendet wird und eine Verallgemeinerung des Wilcoxon Rangsummentests ist. Voraussetzungen für die Anwendung des Testes sind:

- Es liegen unabhängige Stichproben vor.
- Die Daten folgen einer stetigen Verteilung.
- Die Daten sind mindestens ordinalskaliert.

Die Nullhypothese besagt, dass alle Mittelwerte gleich sind und die Alternativhypothese, dass die Gleichheit der Mittelwerte abgelehnt wird. Die Testgröße

lautet:

$$K = \frac{12}{N(N+1)} \sum_{j=1}^k \frac{R_j}{n_j} - 3(N+1)$$

mit

- n_j als dem Umfang der j -ten Stichprobe (Gruppe),
- N der Anzahl der Beobachtungen $N = n_1 + n_2 + \dots + n_k$,
- k die Anzahl der Stichproben und
- R_j die Summe der Ränge der j -ten Stichprobe.

Die Testgröße K ist unter H_0 approximativ χ^2 -verteilt mit $(k-1)$ Freiheitsgraden. Treten sogenannte Bindungen (Ties), d.h., gleiche Werte und damit gleiche Ränge über die Gruppen hinweg auf, so wird allen betreffenden Werten das arithmetische Mittel der betreffenden Rangplätze zugewiesen.

2. Kerndichteschätzung

Der im folgenden beschriebene Kerndichteschätzer ist eine Verallgemeinerung des Histogramms und ermöglicht eine stetige Schätzung einer unbekanntem Verteilung. Der Kerndichteschätzer einer Wahrscheinlichkeitsfunktion f , basierend auf einer Stichprobe X_1, X_2, \dots, X_n aus f , hat die Form

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - X_i)$$

mit

$$K_h(\bullet) = \frac{1}{h} K(\bullet/h).$$

$K(\bullet)$ ist eine beliebiger Kern z.B. Gausskern, Epanechnikov-Kern und h ist die

Bandbreite. Um ein optimales h zu finden kann man unter anderem die Regel von Silvermann verwenden, die unter der Annahme, dass man einen Gausskern verwendet und die Referenzfunktion die Normalverteilung ist, folgende Schätzung angibt

$$\hat{h} = 1.06\hat{\sigma}n^{-1/5}.$$

3. Jarque-Bera-Test

Der Jarque-Bera-Test wird genutzt, um zu überprüfen, ob eine Zufallsvariable normalverteilt ist. Die Teststatistik lautet:

$$JB = \frac{n}{6} \left(\hat{\gamma}_1^2 - \frac{\hat{\gamma}_2 - 3}{4} \right)$$

mit

- n der Anzahl der Beobachtungen,
- $\hat{\gamma}_1$ der Schiefe $\hat{\gamma}_1 = \frac{\mu_3}{(\sigma^2)^{3/2}}$ mit μ_3 als drittes zentrales Moment und σ^2 als Varianz und
- $\hat{\gamma}_2$ die Wölbung $\hat{\gamma}_2 = \frac{\mu_4}{(\sigma^2)^2}$ mit μ_4 als viertes zentrales Moment und σ^2 als Varianz.

Die Teststatistik ist asymptotisch χ^2 -verteilt mit zwei Freiheitsgraden.

3.3 Clusteranalyse

Die Clusteranalyse ist ein Verfahren zur Einteilung einer Anzahl von Objekten in homogenen Gruppen. Die Beobachtungen oder Variablen einer Stichprobe werden so in

die Gruppen eingeteilt, dass das Assoziationsmaß zwischen den Mitgliedern der gemeinsamen hoch und zwischen den Mitgliedern aus verschiedenen Gruppen niedrig ist. Das Maß, mit dem die Ähnlichkeit oder Unähnlichkeit zwischen den Objekten numerisch ausgedrückt wird, muss vor der Analyse festgelegt werden. Als Maß für Unähnlichkeiten werden Metriken in endlich-dimensionalen Räumen oder davon abgeleitete Größen wie die Euklidische Distanz oder deren quadrierten Werte verwendet. Auch zur Bildung der Gruppen existieren verschiedene Algorithmen. Zu den meistverwendeten Verfahren gehören die hierarchischen Verfahren, welche noch einmal in agglomerative und diversive unterschieden werden. Zu den hierarchisch-agglomerativen Verfahren gehören u.a. single linkage und complete linkage, average linkage between und within groups, Median, Ward und Centroid. Des Weiteren gibt es noch partitionierende Verfahren u.a. Austauschverfahren und Optimierungsverfahren. Da das Clusterverfahren später zur Auswertung der Faktorwerte verwendet wird, werde ich an dieser Stelle nur das hierarchisch-agglomerative Verfahren näher erläutern.

Die Aggregation beginnt mit den kleinstmöglichen Gruppen, d.h., jede Gruppe besteht aus einem Objekt. Durch zusammenfassen der zwei, im Sinne des Abstandsmaßes, ähnlichsten Objekten wird eine erste zweielementige Gruppe gebildet. Daraufhin wird die ursprüngliche Definition des Abstandes zwischen zwei Objekten erweitert zu einer Definition von Abständen zwischen verschiedenen Gruppen. Im nächsten Schritt werden Paare von Gruppen und/oder Objekten mit dem jeweils kleinsten Abstand zu neuen Gruppen zusammengefasst. Nach jedem Aggregationsschritt werden die Abstände neu berechnet. In jedem weiteren Schritt wird eine neue Gruppe durch Vereinigung zwei-

er bereits konstruierter Gruppen gebildet. Die Gruppierung besteht nach dem i -ten Schritt aus $(n - i)$ Gruppen, wobei n die Anzahl aller Objekte ist. Das Aggregationschema besteht aus folgenden Iterationen:

1. Berechnung von Distanzen zwischen der Gruppe der i -ten Stufe.
2. Vereinigung der Gruppen mit den entsprechenden Abständen der i -ten Stufe führt zur Gruppierung der $(i+1)$ -ten Stufe. Diese enthält eine Gruppe weniger und stimmt in $(n - i - 2)$ Gruppen mit der i -ten Stufe überein.

Der Prozess wird bis zur Aggregation aller n Objekte in einer einzigen Gruppe fortgesetzt, d.h., nach $(n-1)$ Schritten. Da die Gruppen möglichst homogen sein sollen, wenn sie als Cluster gelten sollen, wird man als Lösung des Clusterproblems die Gruppen auf der Stufe unmittelbar bevor es zu einem sprunghaften Anstieg der Fusionswerte kommt, denn damit steigt auch die Heterogenität sprunghaft.

Wie oben beschrieben, werden die hierarchisch-agglomerativen Verfahren auch nach der Wahl des Abstandsmaßes unterschieden.

single linkage („nearest neighbor“): Der minimale Abstand wird betrachtet

complete linkage („furthest neighbor“): Der maximale Abstand wird betrachtet

average linkage („between groups linkage“): Der Mittelwert der Abstände zwischen allen Paaren von Objekten wird betrachtet

Diese Verfahren besitzen Monotonieeigenschaften, d.h., die Heterogenität wächst im Verlauf des Verfahrens. Anders verhält es sich bei den beiden folgenden Verfahren:

Centroid Clustering: Dieses Abstandsmaß ist ein Mittelweg zwischen dem single und complete linkage. Die Monotonieeigenschaft wird hier nicht allgemein garantiert.

Ward: Bei dieser Metrik wird die Binnenvarianz der Gruppierungen minimiert. Die Binnenvarianz berücksichtigt die Heterogenität. Es werden konvexe Gruppen gebildet und die gleichmäßige Besetzung der Gruppen wird begünstigt.

Im Kapitel 5 werden die hier beschriebenen Verfahren auf die Mittelwerte der Faktorewerte angewendet.

4 Faktorenanalyse mit den Daten der Evaluation

In diesem Kapitel wird mit Hilfe der Faktorenanalyse untersucht, welche Eigenschaften der Kurse entscheidend sind.

4.1 Explorative Faktorenanalyse

4.1.1 Anzahl der Faktoren

In diesem Abschnitt soll untersucht werden, wie viele gemeinsame Faktoren durch die Datensätze repräsentiert werden. Die Faktorenanalyse basiert auf dem *underlying variable* Ansatz und wurde mit der Software Mplus durchgeführt. Es werden die Variablen aus Block 2 bis 7 verwendet.

Zur Bestimmung der Faktorwerte der gemeinsamen Faktoren werden die Eigenwerte λ_q ($q = 1, \dots, Q$) der Korrelationsmatrix berechnet. Der Eigenwert eines Faktors gibt an, wie groß der Erklärungsbeitrag des jeweiligen Faktors an der Varianz aller Variablen ist. Verschiedene Kriterien helfen, die Anzahl der zu extrahierenden Faktoren festzulegen. Dazu zählt das Kaiser-Kriterium, das Faktoren mit Eigenwerten größer als 1 auswählt. Man betrachtet weiterhin die Varianzprozentanteile, die angeben, wie-

viel Prozent der Gesamtvarianz durch die extrahierten Faktoren erklärt werden. Als weitere Orientierung kann man den Screeplot heranziehen. In dieser Grafik werden die nach der Größe sortierten Eigenwerte dargestellt. Ein Knick im Verlauf trennt die zu extrahierenden Faktoren.

Mit Hilfe des Kaiser-Kriteriums (Eigenwerte > 1) werden zwischen 4 und 5 Faktoren

Datensatz	erklärte Varianz 4 Faktoren	erklärte Varianz 5 Faktoren
Vorlesung SS05	61.69%	66.28%
Vorlesung SS06	61.81%	
Übung SS05	60.89%	
Übung SS06	62.57%	67.54%

Tabelle 4.1: Erklärte Varianz für vier bzw. fünf Faktoren

extrahiert, die mehr als 60% der Varianz aller Variablen erklärt (siehe Tabelle 4.1).

Die Screeplots in Tabelle 4.2 zeigen bei den Datensätzen Vorlesung SS06 und Übung SS05 einen Knick bei 4 Faktoren und bei den Datensätzen Vorlesung SS05 und Übung SS06 einen Knick bei 5 Faktoren.

4.1.2 Teststatistiken

Mit Hilfe der in Mplus ausgegebenen Teststatistiken wird geprüft, wie gut das gefundene Modell die Daten beschreibt.

1. Chi-Quadrat-Test

Der Chi-Quadrat-Test testet die Nullhypothese, dass eine vorgegebene Anzahl von Faktoren die Daten vollständig erklären. Das Maß entspricht einer Likelihood-Ratio-Teststatistik, die in der Nullhypothese testet, dass die geschätzte Kovarianz-

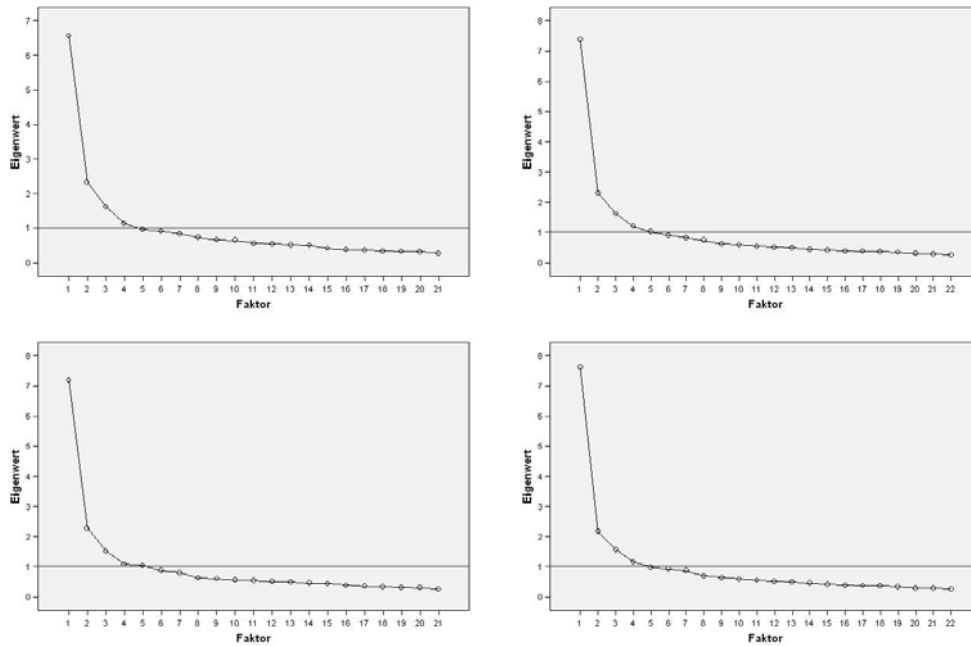


Tabelle 4.2: Screeplots für Übung SS05 (oben links), Vorlesung SS05 (oben rechts), Übung SS06 (unten links) und Vorlesung SS06 (unten rechts)

Matrix Σ der empirischen Kovarianz-Matrix S entspricht. Es berechnet sich als

$$\chi^2 = (n - 1)F(S, \Sigma)$$

mit F als Funktion des Minimums. Die p -Werte für die vier Datensätze sind in der Tabelle 4.3 dargestellt.

Die Nullhypothese wird für alle vier Datensätze abgelehnt. Es sei an dieser Stelle aber darauf hingewiesen, dass der Test empfindlich gegenüber großen Stichproben ist, d.h., umso größer die Stichprobe (> 200), um so größer ist die Wahrscheinlichkeit die Nullhypothese abzulehnen.

2. RMSEA (Root Mean Square Error of Approximation)

Dieser Index testet ebenfalls, ob das Modell die Daten hinreichend gut approximiert und ist robust gegenüber großen Stichproben. Er berechnet sich durch

$$RMSEA = \sqrt{\frac{\chi^2/df - 1}{N - 1}}$$

mit

- N der Stichprobengröße,
- χ^2 dem Chi-Quadrat-Wert des zu testenden Modells und
- df der Anzahl der Freiheitsgrade des zu testenden Modells.

Wenn der Wert < 0.06 , ist das Modell zufriedenstellend spezifiziert und wenn der Wert < 0.08 , ist das Modell gerade noch ausreichend spezifiziert. Für die Datensätze Vorlesung SS06 und Übung SS05 ergibt sich eine inakzeptable Modellanpassung (siehe Tabelle 4.3) und für die anderen beiden Datensätze eine gerade noch akzeptable Anpassung.

3. RMR (Root Mean Square Residual)

Dieser Index ist ein Maß für die Abweichung der empirischen Varianz-/Kovarianzgrößen von den aus dem Modell geschätzten Größen. Er berechnet sich durch

$$RMR = \sqrt{2 \sum_{i=1}^n \sum_{j=1}^i \frac{(s_{ij} - \sigma_{ij})}{n(n+1)}}.$$

Hierbei ist n die Anzahl der Variablen. Kleine Werte von RMR sprechen für eine gute Modellanpassung. Werte < 0.05 geben an, dass das Modell vollständig bestätigt ist. Liegt der Wert unter 0.10, sagt man, das Modell ist tendentiell

bestätigt und liegt der Wert über 0.10, wird das Modell abgelehnt. Wie in Tabelle 4.3 zu sehen, wird das Modell für alle Datensätze vollständig bestätigt.

Datensatz	p-Wert des χ^2 -Tests	RMSEA	RMR
Vorlesung SS06	0.000	0.094	0.037
Vorlesung SS05	0.000	0.077	0.028
Uebung SS06	0.000	0.067	0.024
Uebung SS05	0.000	0.089	0.036

Tabelle 4.3: Teststatistiken

4.1.3 Interpretation der Faktorladungen

Für die Interpretation der Faktoren schaut man sich die Ladungen der einzelnen Variablen an. Besonders von Interesse sind Ladungen q_{jl} mit $|q_{jl}| > 0.5$. Mit Hilfe der Ladungen möchte man herausfinden, welche Variablen mit welchem Faktor verwandt sind und welche Variablen gemeinsam von einem Faktor bestimmt werden. Die Ergebnisse aus Tabelle 4.1 legen zunächst ein Fünf-Faktoren-Modell nahe. Da man aber in den Datensätzen Vorlesung SS05 und Übung SS06 für den fünften Faktor keine Ladungen findet, die > 0.5 sind, beschränke ich mich in der weiteren Vorgehensweise auf das Vier-Faktoren-Modell. Es wird die varimaxrotierte Ladungsmatrix betrachtet. Da die Varimax-Rotation eine orthogonale Rotation ist, bleibt die Unkorreliertheit der Faktoren erhalten und das Modell ist leicht zu interpretieren, da die Ladungen als bivariate Korrelation zwischen den Variablen und den Faktoren betrachtet werden können. Tabelle 4.4 enthält die Ladungen aus Mplus.

Eine Interpretation der Ladungsmatrix aufgrund ihrer Ladungen, die > 0.5 , zeigt die folgende Variablengruppierungen auf.

	Dozent				Konzept							Anford.			Selbststeinsch.				Atmosphäre					
Vorlesung06	Erklärungsvermögen	B1	B3	B4	B6	B62	C1	C2	C4	C5	C6	C7	C72	D1	D2	D3	E1	E2	E3	E4	F1	F2	F3	F4
	Qualität Folien	.6	.3	.5	.7	.8	.4	.4	.3	.2	.0	.1	.0	.0	.0	.0	.1	.1	.3	.0	.2	.3	.3	.3
	Didakt.Kompetenz	.3	.4	.3	.1	.1	.4	.4	.3	.6	.8	.7	.0	.0	.0	.0	.1	.1	.2	.0	.1	.1	.1	.2
	Bereit. Zwischenfragen	.4	.2	.4	.2	.2	.3	.3	.4	.2	.0	.1	.0	.1	.1	.1	.7	.7	.6	-1	.2	.7	.4	.7
Qualität Zwischenfragen	.1	.0	.0	.0	.0	.1	.1	.2	.0	.0	.0	.0	.6	.6	.8	.0	-1	.0	.3	.3	.0	.0	.1	
Vorlesung05	Schwerpunkte	.4	.4	.4	.2	.0	.1	.4	.4	.2	.0	.1	.0	.0	.0	.0	.1	.1	.3	.0	.2	.4	.3	.4
	Gliederung	.3	.4	.3	.6	.7	.6	.3	.4	.3	.6	.7	.6	.0	.0	.0	.1	.1	.2	.0	.1	.1	.1	.1
	Veranschaulichung	.1	.1	.1	.0	.0	.0	.1	.1	.2	.1	.0	.0	.6	.7	.8	.0	-1	.0	.3	.3	.0	.0	.1
	Skripte	.4	.2	.4	.1	.2	.0	.0	.3	.2	.3	.2	.0	.0	.1	.0	.0	.6	.7	.8	.0	.1	.0	.1
Verfügbarkeit	.3	.2	.3	.2	.0	.0	.0	.3	.2	.3	.2	.0	.0	.0	.1	.0	.6	.7	.8	.0	.1	.0	.1	
Internet	.1	.1	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Abstimmung	.3	.3	.3	.2	.0	.1	.1	.2	.0	.0	.0	.0	.6	.7	.8	.0	.6	.7	.8	.0	.1	.0	.1	
Übung06	Geschwindigkeit	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.1	.1	.3	.0	.2	.4	.3	.4
	Formalisierung	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.1	.1	.2	.0	.1	.1	.1	.1
	Schwierigkeitsgrad	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.6	.7	.8	.0	-1	.0	.3	.3	.0	.0	.1
	Interesse	.0	.1	.0	.0	.0	.0	.0	.6	.7	.6	-1	.2	.7	.4	.6	.6	.7	.6	-1	.2	.7	.4	.6
Übung05	Aufmerksamkeit	.1	.1	.2	.0	.0	.0	.1	.2	.0	.0	.0	.0	.0	.0	.0	.1	.1	.2	.0	.0	.1	.1	.1
	Lernzuwachs	.0	.1	.2	.0	.0	.0	.1	.2	.0	.0	.0	.0	.0	.0	.0	.1	.2	.0	.0	.1	.1	.1	.1
	Vorbereitung	.0	.1	.2	.0	.0	.0	.1	.2	.0	.0	.0	.0	.0	.0	.0	.1	.2	.0	.0	.1	.1	.1	.1
	Stressfrei	.2	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Interessant	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Diszipliniert	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Motivierend	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Vorlesung06	Erklärungsvermögen	.7	.5	.7	.6	.7	.4	.4	.2	.1	.2	.3	.0	.0	.0	.0	.1	.2	.0	.0	.2	.3	.3	.3
	Qualität Folien	.2	.3	.3	.2	.2	.4	.4	.6	.7	.7	.4	.0	.0	.0	.0	.1	.2	.0	.0	.0	.1	.1	.1
	Didakt.Kompetenz	.3	.2	.3	.2	.2	.3	.3	.2	.1	.0	.2	.0	.0	.0	.0	.0	.5	.6	.1	.4	.7	.5	.7
	Bereit. Zwischenfragen	.0	.0	.0	.0	.0	.1	.1	.0	.0	.0	.0	.0	.6	.7	.8	-2	.0	.2	.2	.0	.0	.1	.1
Übung05	Schwerpunkte	.4	.4	.4	.2	.1	.2	.3	.0	.0	.0	.0	.0	.0	.0	.0	.1	.2	.0	.0	.2	.3	.3	.3
	Gliederung	.4	.4	.4	.2	.1	.2	.3	.0	.0	.0	.0	.0	.0	.0	.0	.1	.2	.0	.0	.0	.1	.1	.1
	Veranschaulichung	.1	.1	.1	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
	Skripte	.3	.3	.3	.2	.1	.0	.2	.0	.0	.0	.1	.0	.0	.0	.0	.0	.6	.7	.8	.0	.1	.0	.1
Verfügbarkeit	.3	.3	.3	.2	.1	.0	.2	.0	.0	.0	.1	.0	.0	.0	.0	.0	.6	.7	.8	.0	.1	.0	.1	
Internet	.1	.1	.1	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Abstimmung	.3	.3	.3	.2	.0	.0	.1	.1	.2	.0	.1	.0	.0	.0	.0	.0	.6	.7	.8	.0	.1	.0	.1	
Vorlesung06	Geschwindigkeit	.7	.4	.7	.7	.8	.4	.4	.2	.1	.2	.3	.0	.0	.0	.0	.1	.2	.0	.0	.2	.3	.3	.3
	Formalisierung	.0	.2	.1	.1	.1	.4	.4	.5	.7	.7	.3	.0	.0	.0	.0	.0	.1	.0	.0	.1	.1	.1	.1
	Schwierigkeitsgrad	.1	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
	Interesse	.3	.2	.3	.1	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.6	.7	.8	.0	.1	.0	.1
Übung05	Aufmerksamkeit	.4	.4	.4	.2	.1	.2	.3	.0	.0	.0	.0	.0	.0	.0	.0	.1	.2	.0	.0	.0	.1	.1	.1
	Lernzuwachs	.1	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
	Vorbereitung	.0	.1	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
	Stressfrei	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0
Interessant	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Diszipliniert	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	
Motivierend	.2	.0	.2	.0	.2	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	

Tabelle 4.4: Faktorladungen, rot: Ladungen $\in [0.5, 0.7]$, blau: Ladungen $\in [0.7, 1]$

1. Faktor: Lehrbefähigung

Bei diesem Faktor laden alle Variablen positiv und umfassen Aussagen zu den Fähigkeiten des Dozenten.

- Qualität der Beantwortung fachlicher Zwischenfragen (+)
- Bereitschaft zur Beantwortung fachlicher Zwischenfragen (+)
- Erklärungsvermögen (+)
- Didaktische Kompetenz/ Gestaltung des Lernangebotes (+)

2. Faktor: Skripte

Dieser Faktor umfasst die Fragen, die Auskunft über die Qualität und Bereitstellung von Skripten innerhalb der Vorlesung bzw. Übung geben.

- Verfügbarkeit der Literaturlauswahl und Skripte(+)
- Internetpräsenz (+)
- Qualität der Literaturlauswahl und Skripte (+)

3. Faktor: Selbsteinschätzung

Bei diesem Faktor schätzen die Studenten ihr Interesse an der Vorlesung bzw. Übung und ihre Aufmerksamkeit und den Lernzuwachs selbst ein.

- interessant/langweilig (+)
- Interesse am Fach/ Thema (+)
- Aufmerksamkeit in der Veranstaltung (+)
- Erkenntnis- bzw. Lernzuwachs (+)

4. Faktor: Anforderung

Auffällig bei diesem Faktor ist, dass alle Ladungen ein negatives Vorzeichen haben, d.h., niedrige Ausprägungen sind negativ zu bewerten. Die Anforderungen sind also zu hoch. Dieser Faktor umfasst die Variablen

- Schwierigkeitsgrad (-)
- Geschwindigkeit des Vorgehens (-)
- Formalisierung/ Mathematisierung (-)

Vergleicht man die Ergebnisse mit den Ergebnissen aus der Arbeit von Frau Zhou, so werden einige Unterschiede sichtbar. Durch die Veränderung des Fragebogens kam es zu einer Reduktion der Faktoren von fünf auf vier Faktoren. Die Faktoren „Communication Skills“ und „Question Answering“ aus Frau Zhou's Arbeit wurden zu einem Faktor „Lehrbefähigung“ in meiner Arbeit zusammengefasst. Die weiteren drei Faktoren ähneln sich in den Variablen, die auf sie laden.

4.2 Konfirmatorische Faktorenanalyse

Mit Hilfe der Ergebnisse aus der explorativen Faktorenanalyse werden latente Konstrukte für die konfirmatorische Faktorenanalyse gebildet. Die Durchführung in Mplus erfolgt mit 4 Konstrukten.

- Dozent (Variablen b1, b4, b6 und b62)
- Skripte (Variablen c5, c6 und c7)
- Selbsteinschätzung (Variablen e2, e3, f2 und f4)

- Anforderung (Variablen d1, d2 und d3)

4.2.1 Teststatistiken

1. Chi-Quadrat Test

Dieser Test ist ein erstes Maß für die Gesamtanpassung des Modells. Es wird die Annahme geprüft, ob das Modell die Daten gut beschreibt. Die Likelihood-Quotienten-Teststatistik ist bei Gültigkeit der Nullhypothese asymptotisch χ^2 -verteilt. Je größer der Wert der Teststatistik ist, umso eher ist die Nullhypothese abzulehnen. Die Nullhypothese wird verworfen, wenn der p -Wert kleiner als 0.05 ist. Der Test ist zum einen empfindlich gegenüber großen Stichproben. Mit wachsendem Stichprobenumfang wächst die Wahrscheinlichkeit H_0 zu verwerfen. Ein weiterer kritischer Punkt im Zusammenhang mit der χ^2 -Teststatistik ist die Nichteinhaltung der Normalverteilungsannahme. Bei nichtnormalverteilten Daten, z.B. mit erheblicher Schiefe, ist die χ^2 -Verteilung im Allgemeinen eine schlechte Approximation für die Likelihood-Quotienten-Teststatistik.

Wie in Tabelle 4.5 zu sehen, wird die Nullhypothese für alle vier Datensätze abgelehnt.

2. TLI (Tucker-Lewis-Index)

Der TLI gehört zu den inkrementellen oder komperativen Fit-Indizes. Diese testen die proportionale Verbesserung der Anpassung des zu testenden Modells gegenüber einem restriktiven Modell (Nullmodell) Im Nullmodell bestehen zwischen den Variablen keinerlei Abhängigkeiten. Dieser Index wurde u.a. von Le-

dyard R. Tucker (1910-2004), Professor der Psychologie (1960-79), entwickelt, der als Pionier im Bereich der Psychometrie, d.h. der Theorie und Methode des psychologischen Messens, gilt. Dieser Index ist weniger sensitiv gegenüber einfachen Modellspezifikation und der Verletzung der Verteilungsannahme. Der TLI wird durch folgende Formel berechnet

$$TLI = \frac{\chi_0^2/df_0 - \chi_1^2/df_1}{\chi_0^2/df_0 - 1}$$

mit

- χ_0^2 als dem Chi-Quadrat-Wert des Nullmodells,
- df_0 der Anzahl der Freiheitsgrade des Nullmodells,
- χ_1^2 dem Chi-Quadrat-Wert des zu testenden Modells und
- df_1 der Anzahl der Freiheitsgrade des zu testenden Modells.

Bei einem guten Modell liegt der Index über 0.95. Für die vier verwendeten Datensätze liegt der Wert immer nahe bzw. über 0.95 (siehe Tabelle 4.5).

3. CFI (Comparative-Fit-Index)

Auch der CFI gehört zu den relativen Fit-Indizes. Er vergleicht das gegebene Modell mit einem Modell, in dem die Variablen unkorreliert sind. Er ist unabhängig von der Stichprobengröße und relativ robust gegen die Verletzung der Verteilungsannahme. Aber der Index ist abhängig von der Modellkomplexität, d.h., sparsam besetzte Modelle werden schlechter bewertet. Der CFI ist auf einem Intervall von 0 (keine Übereinstimmung mit den Daten) und 1 (volle Überein-

stimmung mit den Daten) normiert. Er wird durch folgende Formel berechnet:

$$CFI = 1 - \frac{\max[(\chi_1^2 - df_1), 0]}{\max[(\chi_0^2 - df_0), (\chi_1^2 - df_1), 0]}$$

mit den Bezeichnungen wie beim TLI. Um von einer hinreichenden Übereinstimmung von Modell und Daten zu sprechen, sollte der CFI mindestens 0.90 betragen; ein CFI, der kleiner als 0.90 ist, zeigt eine ungenügende Übereinstimmung von Modell und Daten an, ab einer Größe von circa 0.95 wird von einer guten Übereinstimmung gesprochen. Wie in Tabelle 4.5 zu sehen ist, liegt der CFI für alle Datensätze über 0.9 und für die Übungsdatensätze sogar über 0.95, d.h., es liegt eine gute Modellanpassung vor.

4. RMSEA (Root Mean Square Error Of Approximation)

Der RMSEA gehört zu den absoluten Fit-Indizes. Diese beurteilen unmittelbar, wie gut ein a priori Modell durch die Stichprobendaten reproduziert wird, d.h., das geschätzte Modell wird mit einem saturierten Modell verglichen. Dieser Index hängt ebenfalls von der Modellkomplexität ab, d.h., je komplexer das Modell ist, um so größer der RMSEA. Berechnet wird er durch

$$RMSEA = \sqrt{\frac{\chi^2/df - 1}{N - 1}}$$

mit

- N der Stichprobengröße,
- χ^2 dem Chi-Quadrat-Wert des zu testenden Modells und
- df der Anzahl der Freiheitsgrade des zu testenden Modells.

Werte, die ≤ 0.05 sind, zeigen eine gute Modellanpassung an und Werte ≤ 0.08 eine akzeptable Modellanpassung. Bei allen vier Datensätzen liegen Werte größer als 0.08 vor.

Datensatz	p-Wert	TLI	CFI	RMSEA
Vorlesung SS06	0.000	0.964	0.936	0.119
Vorlesung SS05	0.000	0.965	0.931	0.115
Uebung SS06	0.000	0.976	0.952	0.092
Uebung SS05	0.000	0.966	0.941	0.105

Tabelle 4.5: Teststatistiken

4.2.2 Ergebnisse

Nachfolgend werden die Ergebnisse von Mplus der einzelnen Konstrukte ausgewertet. In der ersten Spalte sieht man die geschätzten Koeffizienten für jede Variable. In der zweiten Spalte stehen die Standardfehler und in der dritten Spalte steht das Ergebnis der Division von geschätzter Parameter durch dazugehörigen Standardfehler. Die letzten beiden Spalten sind standardisierte Koeffizienten für jeden geschätzten Parameter im Modell. Diese verwendet man, wenn die Variablen unterschiedlich skaliert sind. Beispielhaft werden hier die Ergebnisse des Datensatzes für die Vorlesung SS06 dargestellt.

1. Konstrukt: Lehrbefähigung

Wie man sieht, sind die Werte der geschätzten Parameter alle recht hoch. Der Parameter für die Variable B1 war mit 1 im Programmaufruf gesetzt worden. Die geschätzten Koeffizienten geteilt durch ihre Standardfehler testen die Nullhypothese, dass der geschätzte Koeffizient gleich Null ist. Da alle Werte größer

Variablen	Estimates	S.E.	Est./S.E.	Std	StdYX
B1	1.000	0.000	0.000	0.846	0.846
B4	1.009	0.014	71.761	0.853	0.853
B6	0.914	0.015	60.125	0.773	0.773
B62	0.996	0.014	72.672	0.843	0.843

Tabelle 4.6: Ergebnisse für das Konstrukt Lehrbefähigung

als 1,96 (kritischer Wert $Z_{1-0.05/2}$ der standardisierten Normalverteilung) sind, wird die Nullhypothese auf dem Niveau 0.05 für alle Variablen verworfen. Dies bedeutet, dass die verwendeten Variablen einen signifikanten Beitrag zum Konstrukt Lehrbefähigung leisten.

2. Konstrukt: Lehrmaterial

Auch bei diesem Konstrukt sind die geschätzten Koeffizienten alle relativ hoch.

Variablen	Estimates	S.E.	Est./S.E.	Std	StdYX
C5	1.000	0.000	0.000	0.826	0.826
C6	0.882	0.026	33.393	0.729	0.729
C7	0.880	0.027	33.103	0.727	0.727

Tabelle 4.7: Ergebnisse für das Konstrukt Skripte

Die Hypothese, dass der Koeffizient 0 ist, wird für alle Variablen abgelehnt. Damit leisten die verwendeten Variablen einen signifikanten Beitrag zum Konstrukt Lehrmaterial.

3. Konstrukt: Selbsteinschätzung

Der Parameter für die Variable E2 wurde wieder mit 1 gesetzt. Die geschätzten Parameter für die Variablen E3, F2 und F4 sind hoch und haben einen statistisch signifikanten Einfluss auf das Konstrukt Selbsteinschätzung.

Variablen	Estimates	S.E.	Est./S.E.	Std	StdYX
E2	1.000	0.000	0.000	0.671	0.671
E3	1.189	0.027	43.673	0.798	0.798
F2	1.279	0.026	48.379	0.858	0.858
F4	1.319	0.028	46.886	0.885	0.885

Tabelle 4.8: Ergebnisse für das Konstrukt Selbsteinschätzung

4. Konstrukt: Leistungsanforderung

In der ersten Zeile steht bei dem Wert für den geschätzten Parameter wieder eine

Variablen	Estimates	S.E.	Est./S.E.	Std	StdYX
D1	1.000	0.000	0.000	0.614	0.614
D2	1.065	0.028	37.406	0.654	0.654
D3	1.541	0.063	24.434	0.947	0.947

Tabelle 4.9: Ergebnisse für das Konstrukt Leistungsanforderung

1, da dieser Wert vorgegeben ist. Die Parameter der Variablen D2 und D3 haben einen statistisch signifikanten Einfluss auf das Konstrukt Leistungsanforderung.

Neben den Ergebnissen für die einzelnen Konstrukte werden in Mplus die Korrelation zwischen den Konstrukten untersucht.

SKRIPT	WITH				
DOZENT	0.391	0.014	26.978	0.559	0.559
SELBST	WITH				
DOZENT	0.451	0.013	35.815	0.795	0.795
SKRIPT	0.291	0.013	21.949	0.524	0.524
FORDERUN	WITH				
DOZENT	-0.079	0.012	-6.419	-0.152	-0.152

SKRIPT	-0.043	0.013	-3.248	-0.085	-0.085
SELBST	-0.068	0.009	-7.455	-0.166	-0.166

Die Korrelation zwischen den vier Konstrukten ist signifikant verschieden von 0. Des Weiteren werden Thresholds (Schwellenwerte) ausgegeben. Die Anzahl der Schwellenwerte ist die Anzahl der Kategorien minus Eins. Die Thresholds sind Schwellenwerte, die überschritten werden müssen, damit eine bestimmte Kategorie angenommen wird. Mit Hilfe der Thresholds und der Residualvarianz können Wahrscheinlichkeiten berechnet werden, dass bestimmte Werte angenommen werden z.B. wie hoch der Erwartungswert ist, dass ein Übergang von dem Wert 0 zu dem Wert 1.00 bei der kategoriellen Ausgangsvariablen stattfindet.

4.3 Ein-Faktor-Modell

In der Tabelle 4.10 sind die Ladungen der einzelnen Variablen bei einem Ein-Faktor-Modell dargestellt. Man sieht, dass alle Variablen in dem Bereich Anforderungen, sowie die Variablen Vorbereitungszeit und ob die Atmosphäre stressig oder stressfrei ist, nur gering mit dem ersten Faktor korrelieren. Einen besonders wichtigen Beitrag zum ersten Faktor, an den hohen Ladungen erkennbar, liefern die Variablen aus dem Bereich Dozent, die Variablen Schwerpunkte und Gliederung der Vorlesung und ob die Vorlesung/Übung interessant und motivierend ist. Dieser eine Faktor repräsentiert also die generelle Fähigkeit eines Dozenten eine gute Vorlesung/Übung zu gestalten.

	Dozent					Konzept							Anford.			Selbststeinsch.				Atmosphäre			
	B1	B3	B4	B6	B62	C1	C2	C4	C5	C6	C7	C72	D1D2D3	D1D2D3	D1D2D3	E1	E2	E3	E4	F1	F2	F3	F4
Vorlesung06	.8	.5	.8	.6	.7	.7	.6	.6	.5	.4	.4		.1	.2	.2	-.6	-.6	.7	.0	-.4	.8	.5	.8
Vorlesung05	.8	.5	.8	.6	.7	.7	.6	.6	.5	.4	.4		.1	.2	.2	.5	.6	.7	.0	.4	.8	.5	.8
Übung06	-.8	-.6	-.8	-.7	-.7	-.7	-.7	-.6	-.6	-.5	-.5		-.1	-.3	-.3	-.4	-.6	-.6	.0	-.4	-.7	-.4	-.7
Übung05	.8	.5	.7	.7	.7	.7	.7	.5	.4	.4	.4	-.3	-.3	-.3		.4	.6	.0		.4	.7	.5	.7

Tabelle 4.10: Faktorladungen, rot: Ladungen $\in [0.5, 0.7]$, blau: Ladungen $\in [0.7, 1]$

5 Bewertung der Lehrveranstaltungen mit Hilfe der Faktorwerte

5.1 Faktorwerte der Ein-Faktor-Modells

Von den in Mplus im Zuge der konfirmatorischen Faktorenanalyse ausgegebenen Faktorwerten wird der Mittelwert innerhalb der Kurse berechnet. Der Mittelwert gibt Auskunft über das Ranking des Kurses, d.h., je größer der Faktorwert, desto niedriger ist der Rang. Um die Auswertung und speziell die Graphiken leichter interpretierbar zu machen, wurden die Faktorwerte umkodiert. Daher gilt im weiteren Verlauf der Arbeit, dass je größer der Faktorwert ist, desto höher ist der Rang der Lehrveranstaltung. Angelehnt an die Arbeit von Frau Zhou stellt sich die Frage, ob Merkmale einer Lehrveranstaltung existieren, die nicht direkt vom Dozenten beeinflusst werden und die einen gut bzw. schlecht bewerteten Kurs ausmachen z.B. Tag, Uhrzeit, Level, Fachbereich, Größe des Kurses.

In der Tabelle 5.1 sieht man im oberen Teil die fünf best bewerteten Vorlesungen und im unteren Teil die fünf best bewerteten Übungen des SS06. Die Anzahl gibt die ausgewerteten Bögen innerhalb der Veranstaltung an. Im Gegensatz zu den Aussagen

KAPITEL 5. BEWERTUNG DER LEHRVERANSTALTUNGEN MIT HILFE DER
FAKTORWERTE

von Frau Zhou ist nicht klar erkennbar, dass überwiegend Kurse mit einer geringen Teilnehmerzahl positiv bewertet wurden. Die Spalte Fach gibt an, aus welchem Bereich die Lehrveranstaltung kommt, BWL, VWL oder Methodisch. Bei den Übungen überwiegen eindeutig die Veranstaltungen aus dem methodischen Bereich. Das Level gibt an, ob die Lehrveranstaltung von den Studenten im Haupt- oder Grundstudium besucht wird. Die gut bewerteten Vorlesungen wurden alle im Hauptstudium belegt, während die Übungen hauptsächlich im Grundstudium belegt wurden. Dies könnte u.a. daran liegen, dass im Hauptstudium meist nur eine Übung zu einer Vorlesung angeboten wird, während im Grundstudium teilweise mehrere Übungen zu einer Vorlesung angeboten werden. In den letzten beiden Spalten sind der Wochentag und die Uhrzeit angegeben, an welchen der Kurs stattgefunden hat. Es fällt auf, dass fast alle Veranstaltungen in der Mitte der Woche und in den Nachmittagstunden stattfanden.

In der Tabelle 5.2 sind die Faktorwerte der Lehrveranstaltungen angegeben, die von

Kurs	FS	Anzahl	Fach	Level	Tag	Zeit
V1	0.906	11	MTH	HS	Di bis Do	14-20
V2	0.718	10	VWL	HS	Di bis Do	14-20
V3	0.68	26	VWL	HS	Di bis Do	14-20
V4	0.662	53	BWL	HS	Di bis Do	8-14
V5	0.654	49	VWL	HS	Mo oder Fr	8-14
U1	0.942	31	MTH	GS	Di bis Do	8-14
U2	0.799	18	MTH	GS	Di bis Do	14-20
U3	0.686	16	MTH	GS	Di bis Do	14-20
U4	0.666	67	MTH	GS	Di bis Do	8-14
U5	0.637	15	VWL	HS	Di bis Do	14-20

Tabelle 5.1: Faktorwerte des Ein-Faktor-Modells der best bewerteten Lehrveranstaltungen im SS06

den Studenten schlechter bewertet wurden. Bei der Anzahl der ausgewerteten Bögen

findet man sowohl größere als auch kleinere Kurse. Bei den Fächern überwiegt keiner der drei Bereiche. Wie auch bei den besser bewerteten Kursen überwiegen auch hier die Veranstaltungen aus dem Hauptstudium. Dies liegt sicherlich auch u.a. daran, dass im gesamten Datensatz die Lehrveranstaltungen aus dem Hauptstudium überwiegen (siehe Kapitel 2). Ebenfalls sieht man bei den schlechter bewerteten Kursen, dass wie bei den besser bewerteten Kursen, die Kurse, die in der Mitte der Woche und in den Nachmittagsstunden stattfinden, überwiegen.

Im weiteren Verlauf der Arbeit, wenn nichts anderes angegeben ist, wie auch in diesem

Kurs	FS	Anzahl	Fach	Level	Tag	Zeit
V48	-0.628	66	MTH	GS	Mo oder Fr	8-14
V47	-0.619	10	VWL	HS	Di bis Do	8-14
V46	-0.594	16	VWL	HS	Mo oder Fr	14-20
V45	-0.587	57	BWL	HS	Mo oder Fr	8-14
V44	-0.562	21	VWL	HS	Di bis Do	8-14
U37	-1.374	20	VWL	HS	Di bis Do	14-20
U36	-0.649	52	MTH	GS	Di bis Do	14-20
U35	-0.573	16	VWL	HS	Di bis Do	14-20
U34	-0.563	67	BWL	HS	Di bis Do	14-20
U33	-0.538	11	BWL	HS	Di bis Do	8-14

Tabelle 5.2: Faktorwerte des Ein-Faktor-Modells der schlechter bewerteten Lehrveranstaltungen im SS06

Abschnitt, beschränke ich mich auf die Auswertung der Daten aus dem Sommersemester 2006. Die anderen Ergebnisse finden sich auf der beigefügten CD. Im nächsten Abschnitt werden die hier erwähnten Tendenzen auf ihre Signifikanz hin überprüft.

5.1.1 Auswertung der Größe, Level, Tag und Uhrzeit der LV

Anhand des Boxplots in Abbildung 5.1 bestätigt sich die Tendenz, dass Kurse im Hauptstudium besser bewertet werden als im Grundstudium. Führt man einen Kruskal-

Wallis-Test durch, so ergibt sich ein asymptotischer Signifikanzwert von 0.222. Da der Wert größer als 0.05 ist, sind die Mittelwerte auf dem 5% Niveau nicht signifikant verschieden. Das gleiche Ergebnis ergibt sich bei den Übungen im SS06 sowie bei den Vorlesungen und Übungen im SS05.

Vergleicht man die Größe des Kurses, d.h., die Anzahl der ausgewerteten Fragebögen

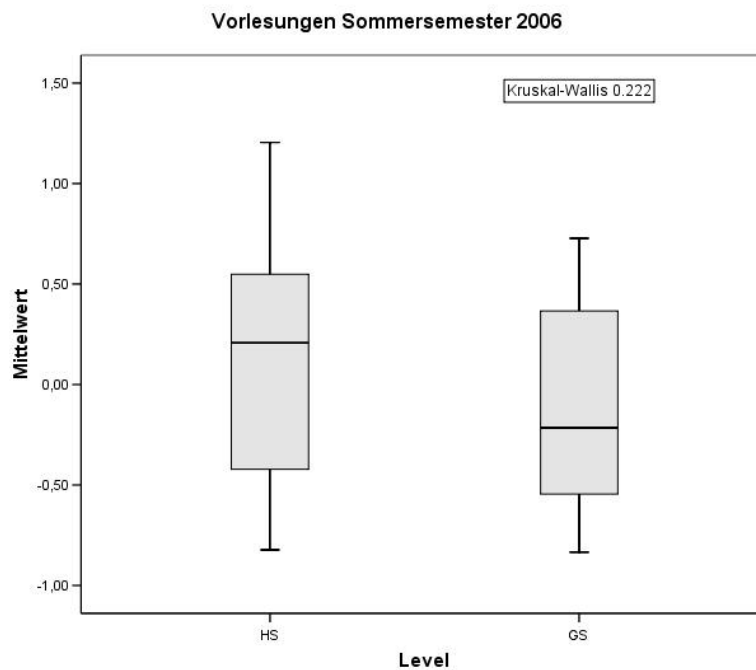


Abbildung 5.1: Auswertung der Vorlesungen nach Grund-und Hauptstudium

mit dem mittleren Faktorwert des Kurses, ergibt sich der Scatterplot in Abbildung 5.2. Würde ein Zusammenhang zwischen diesen beiden Größen bestehen, d.h., je größer der Kurs, desto schlechter die Bewertung, müssten die Werte annähernd auf einer Geraden liegen. Zwar sind die größten Kurse unterdurchschnittlich bewertet worden, aber ein

eindeutiger Zusammenhang ist nicht sichtbar.

Des Weiteren wurde vermutet das Kurse, die in der Mitte der Woche gehalten wer-

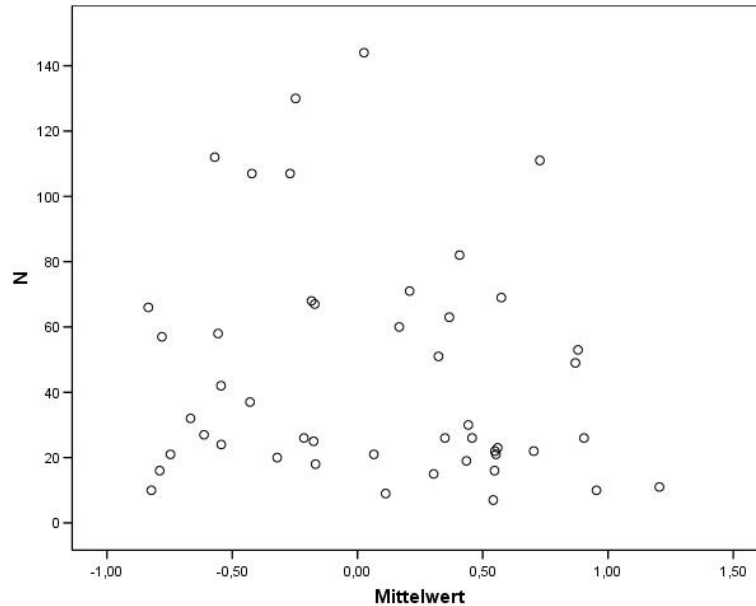


Abbildung 5.2: Auswertung der Vorlesungen nach der Größe

den, d.h., dienstags, mittwochs oder donnerstags, tendenziell besser bewertet werden. Schaut man sich den Boxplot in Abbildung 5.3 an, so scheint sich dies zunächst zu bestätigen. Schaut man aber auf den Signifikanzwert des Kruskal-Wallis-Tests, der bei 0.17 liegt, so ergibt sich, dass die Mittelwerte nicht signifikant verschieden sind. Die gleichen Ergebnisse erhält man für die anderen Datensätze.

Abschließend wurde noch ein Zusammenhang zwischen der Bewertung der Lehrveranstaltung und der Tageszeit, an der sie stattfand, vermutet. In dem Boxplot in Abbildung 5.4 wird dies für die Vorlesungen des SS06 bestätigt. Auch der asymptotische

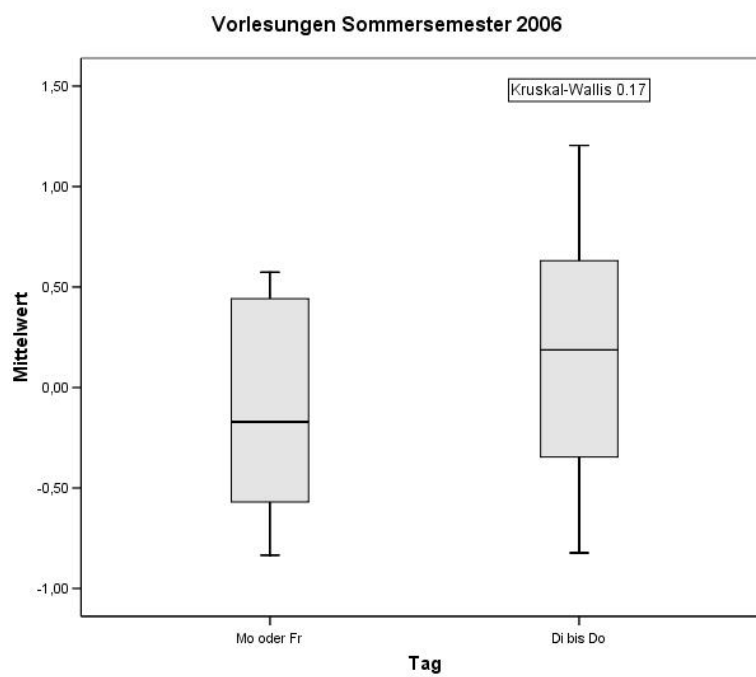


Abbildung 5.3: Auswertung der Vorlesungen nach dem Wochentag

Signifikanzwert der Kruskal-Wallis-Teststatistik 0.032 bestätigt, dass die Mittelwerte zum 5% Niveau signifikant verschieden sind.

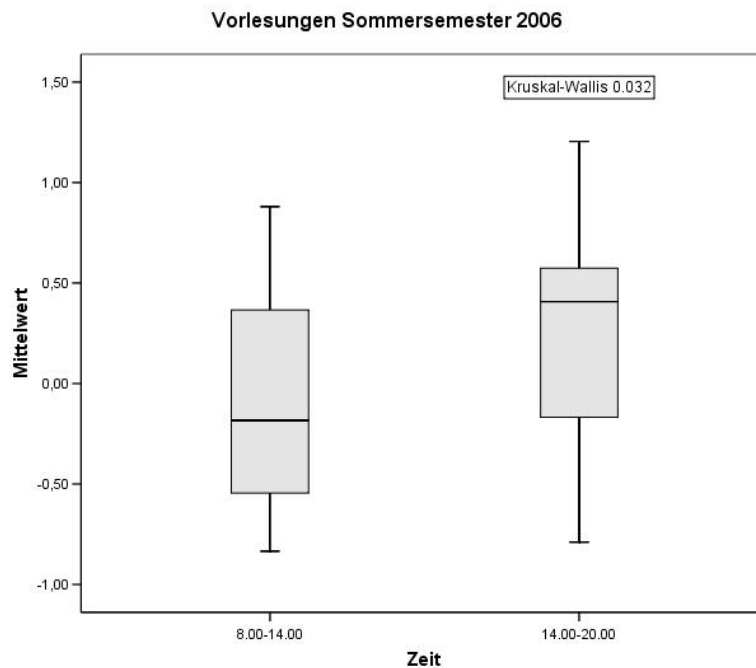


Abbildung 5.4: Auswertung der Vorlesungen nach der Uhrzeit

5.1.2 Abhängigkeit vom Dozenten

Im letzten Kapitel hatte die Faktorenanalyse gezeigt, dass die Lehrbefähigung des Dozenten ein wesentlicher Faktor bei der Bewertung von Lehrveranstaltungen ist. In Abbildung 5.5 und 5.6 wurde untersucht, inwiefern die Änderung des Dozenten eine Veränderung in der Bewertung nach sich zieht.

KAPITEL 5. BEWERTUNG DER LEHRVERANSTALTUNGEN MIT HILFE DER FAKTORWERTE

Auf der Abszissenachse wurden die Faktorwerte der Lehrveranstaltungen des SS05

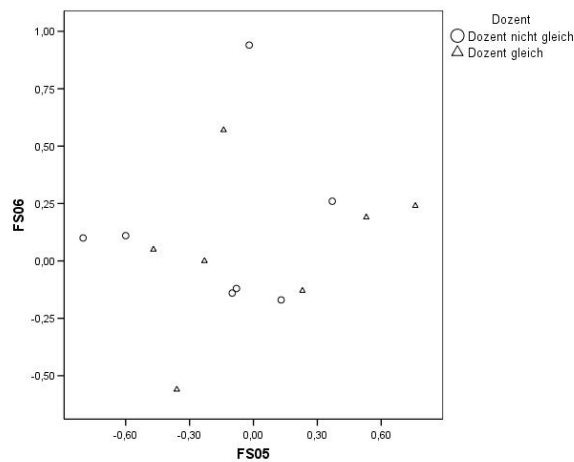


Abbildung 5.5: Übungen

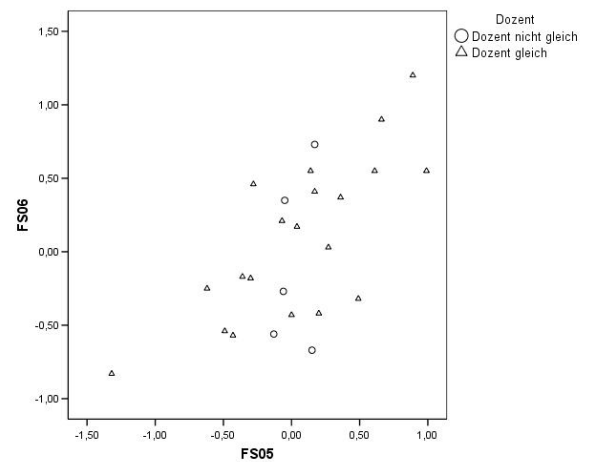


Abbildung 5.6: Vorlesungen

abgetragen und auf der Ordinatenachse die Faktorwerte des SS06. Die geringe Anzahl an Datenpunkten erklärt sich dadurch, dass nur Lehrveranstaltungen betrachtet wurden, die in beiden Semestern angeboten wurden. Zu erwarten war, dass sich die Kreise, welche für die Lehrveranstaltungen stehen, bei denen sich der Dozent geändert hat, im oberen linken und unteren rechten Bereich befinden, während sich die Dreiecke, welche für die Lehrveranstaltungen stehen, bei denen der Dozent gleich geblieben ist, auf der Winkelhalbierenden befinden. Bei den Übungen in Abbildung 5.5 ist keine Tendenz zu erkennen. Bei den Vorlesungen überwiegen Vorlesungen, bei denen der Dozent gleich geblieben ist. Diese befinden sich auch ungefähr auf der Winkelhalbierenden. Die Vorlesungen bei denen der Dozent gewechselt hat, sind zu gering vertreten um eine allgemeine Aussage treffen zu können.

5.1.3 Verteilung der Faktorwerte

In den Abbildungen 5.7 und 5.8 sind die geschätzten Verteilungen und die dazugehörigen Konfidenzintervalle für die Vorlesungen und Übungen im SS06 dargestellt. Die Dichte wurde mit Hilfe der Kerndichtschätzung bestimmt. Der p -Wert des Jarque-Bera-Tests gibt an, ob die Annahme der Normalverteilung angenommen oder abgelehnt wird. In drei der vier Datensätze wird der Test auf Annahme der Normalverteilung auf dem 5% Niveau abgelehnt.

Entsprechend dem Ranking in den Tabellen 5.1 und 5.2 wurde die Vorlesung V4 mit

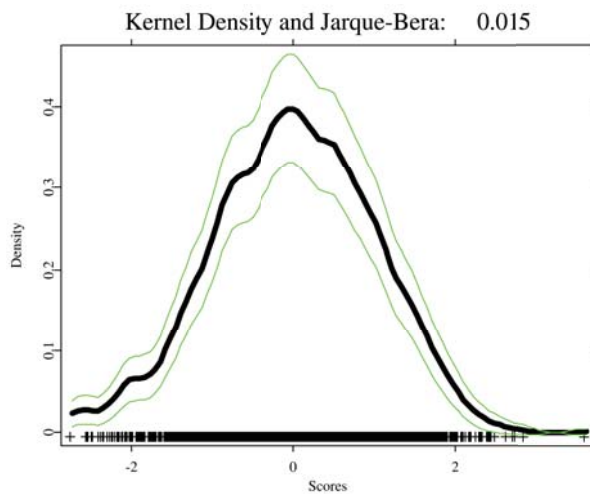


Abbildung 5.7: Vorlesungen SS06

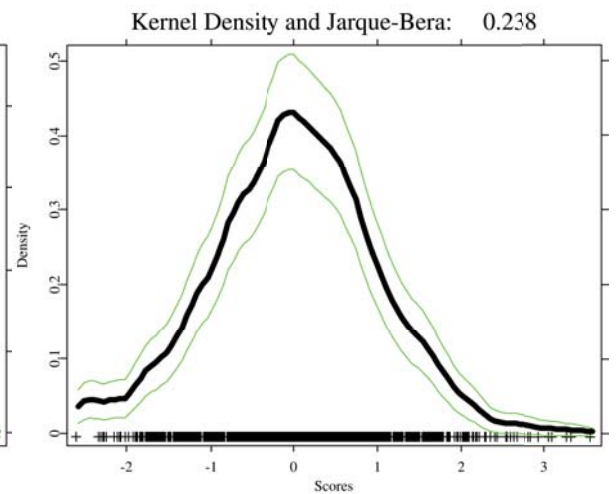


Abbildung 5.8: Uebungen SS06

einem relativ hohen Mittelwert und die Vorlesung V47 mit einem relativ niedrigem Mittelwert ausgewählt und die Verteilung in Abbildung 5.9 betrachtet. Beide Kurse weisen in etwa die gleiche Anzahl von Beobachtungen auf. Man sieht, dass der schlechtere Kurs eine größere Varianz als der gute Kurs aufweist. Des Weiteren konzentriert

sich die Verteilung des guten Kurses stärker auf die positiven Faktorwerte.

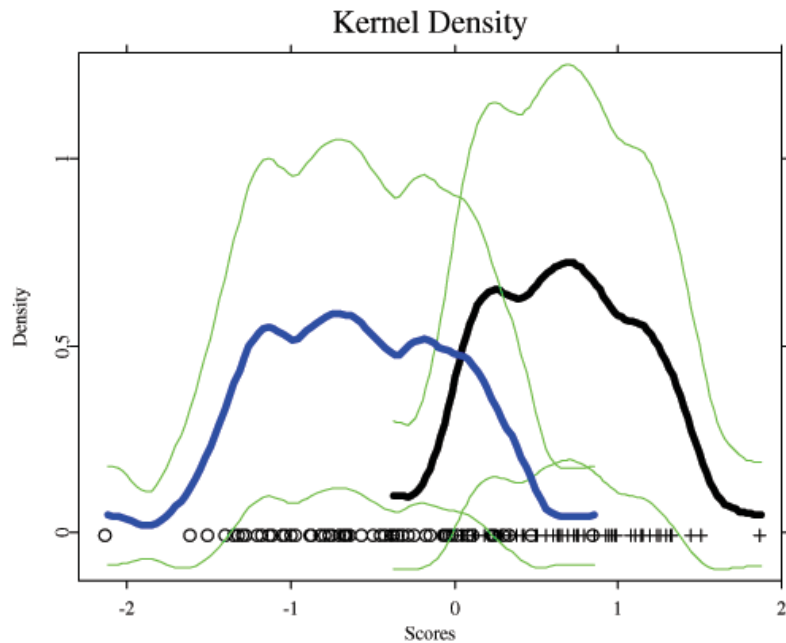


Abbildung 5.9: Dichteschätzung der Faktorwerte, blau: schlechter Kurs, schwarz: guter Kurs

5.1.4 Clusteranalyse

Die Clusteranalyse wurde dazu verwendet, die Vorlesungen in homogene Gruppen einzuteilen. Als Maß für die Unähnlichkeit zwischen den Objekten wurde die Euklidische Distanz verwendet und als Clusterverfahren das hierarchisch-agglomerative Ward-Verfahren.

Die Knoten des Dendrogramms werden von den Gruppen des hierarchischen Systems gebildet. Durch die Kanten werden die Inklusionsrelationen zwischen einer Gruppe

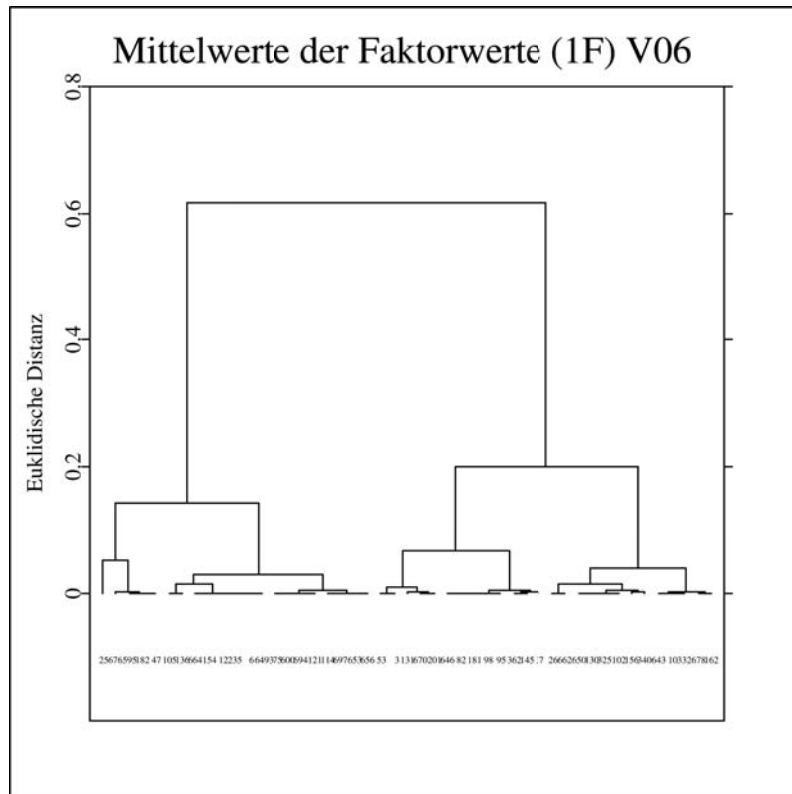


Abbildung 5.10: Dendrogramm der Mittelwerte der Faktorwerte der Vorlesungen SS06

der i -ten Stufe und ihren Subgruppen der $(i - 1)$ -ten Stufe dargestellt. In dem Dendrogramm werden des Weiteren die Abstandsrelationen zwischen den Gruppen und damit die Reihenfolge der Vereinigung dargestellt. In Abbildung 5.10 sieht man, dass vier Gruppen gebildet werden, die in sich sehr homogen sind. Hierbei werden die fünf best bewerteten Kurse und die fünf schlechtesten Kurse in jeweils einer Gruppe zusammengefasst. Der mittlere Teil wird noch einmal in zwei Gruppen, zum einen die Kurse mit einer positiven Tendenz und zum anderen die Kurse mit einer negativen Tendenz unterteilt.

5.2 Faktorwerte des Vier-Faktor-Modells

Wie auch beim Ein-Faktor-Modell wurden die Faktorwerte für das Vier-Faktor-Modell im Zuge der konfirmatorischen Faktorenanalyse in Mplus ausgegeben. Die Faktorwerte des Ein-Faktor-Modells wurden u.a. dazu genutzt, ein Ranking für die Lehrveranstaltung aufzustellen. Da in den vier Faktoren mehr Informationen über den Gesamtdatensatz enthalten sind, stellt sich die Frage, ob die Faktorwerte des Vier-Faktor-Modells die Aussagen des Ein-Faktor-Modells unterstützen.

5.2.1 Clusteranalyse

Wie auch bei einem Faktor, wurden die Vorlesungen des SS06 mit Hilfe der Faktorwerte geclustert. Das Dendrogramm in Abbildung 5.11 ähnelt dem Dendrogramm in Abbildung 5.10. In der ersten Gruppe sind die elf best bewerteten Vorlesungen zusammengefasst. Die mittleren Kurse wurden zusammengefasst, sowie die drei am schlechtesten bewerteten Kurse.

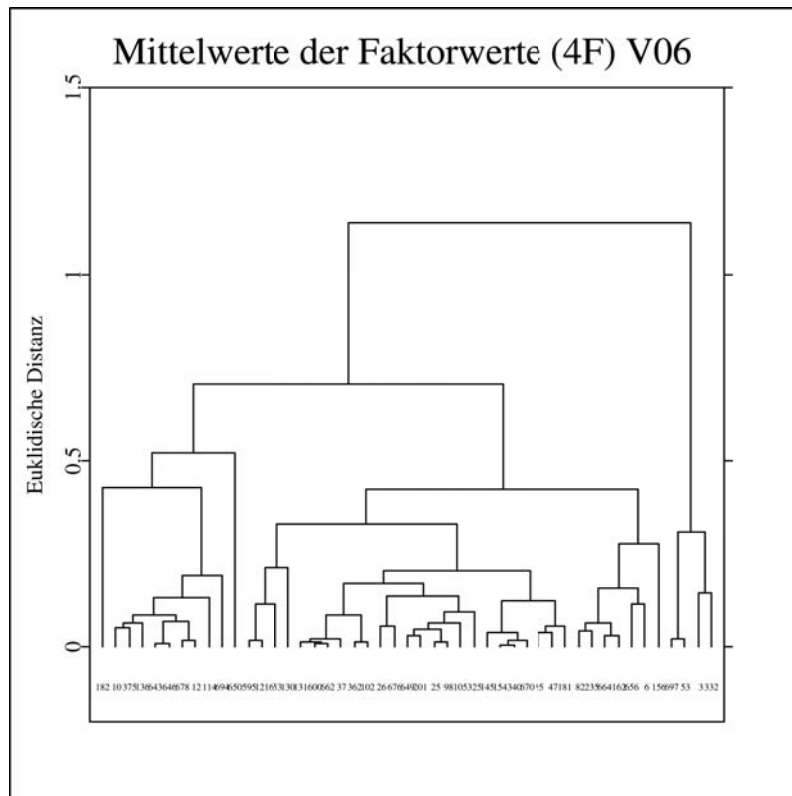


Abbildung 5.11: Dendrogramm der Mittelwerte der Faktorwerte der Vorlesungen SS06
(4F)

5.2.2 Auswertung der Faktoren

Für die Übungen des SS06 wurden die mittleren Faktorwerte für jede Veranstaltung in Scatterplots gegenübergestellt. Die Zuordnung der in den Abbildungen vergebene Nummer für jede Veranstaltung kann in der beigefügten CD nachgeschlagen werden.

Faktor Lehrbefähigung und Skripte

In der Abbildung 5.12 sieht man, dass die Übungen mit der best bewerteten Lehrbefähigung und mit den best bewerteten Skripten in etwa den best bewerteten Kursen aus der Tabelle 5.1 des Ein-Faktor-Modells entsprechen. Fünf Übungen sind aus dem Bereich Methodisch/Quantitativ, zwei Übungen aus dem Bereich VWL und eine Übung aus dem Bereich BWL. Auch die schlecht bewerteten Übungen in Bezug auf die Lehrbefähigung und die Skripte im unteren linken Bereich entsprechen den Kursen in Tabelle 5.2. Des Weiteren sieht man an Hand der hohen positiven Korrelation von 0.894, dass ein großer Zusammenhang zwischen diesen beiden Faktoren besteht.

Faktor Lehrbefähigung und Selbsteinschätzung

In der Abbildung 5.13 sieht man, dass die Übungen mit der best bewerteten Lehrbefähigung und mit der höchsten positiven Selbsteinschätzung in etwa den best bewerteten Kursen in der Abbildung 5.12 entsprechen. Auch die schlecht bewerteten Übungen in Bezug auf die Lehrbefähigung und die Selbsteinschätzung im unteren linken Bereich entsprechen den Kursen in Abbildung 5.12. Des Weiteren sieht man an Hand der hohen positiven Korrelation von 0.941, dass wiederum ein großer Zusammenhang zwischen diesen beiden Faktoren besteht.

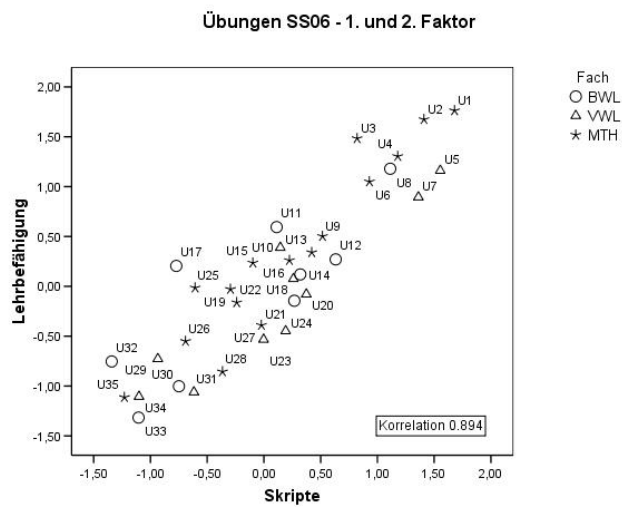


Abbildung 5.12: Scatterplot der Mittelwerte der Faktorwerte des 1. und 2. Faktors für die Übungen SS06

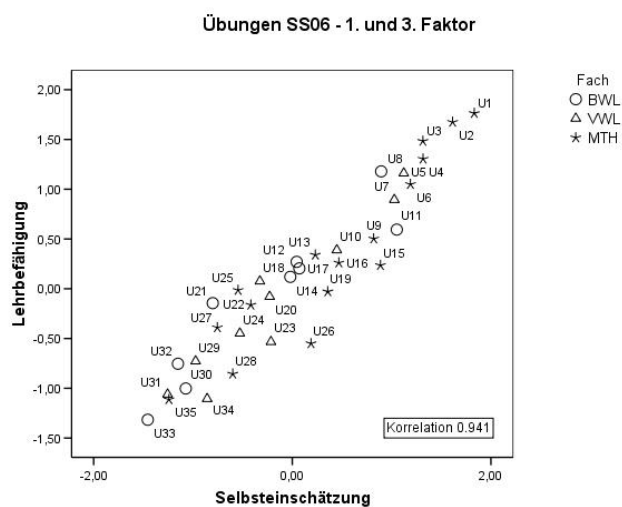


Abbildung 5.13: Scatterplot der Mittelwerte der Faktorwerte des 1. und 3. Faktors für die Übungen SS06

Faktor Lehrbefähigung und Anforderung

Die Interpretation bei dem ersten und vierten Faktor muss etwas anders geführt werden. Übungen, die sich im linken Teil der Abbildung 5.14 finden, wurden in ihrer Anforderung als zu niedrig bewertet. Bei den Kursen in der Mitte ist die Anforderung genau richtig und bei den Kursen im rechten Teil sind die Anforderungen zu hoch. Anhand der Verteilung und der berechneten Korrelation von -0.142 sieht man, dass kein direkter Zusammenhang zwischen der Lehrbefähigung und der Anforderung besteht.

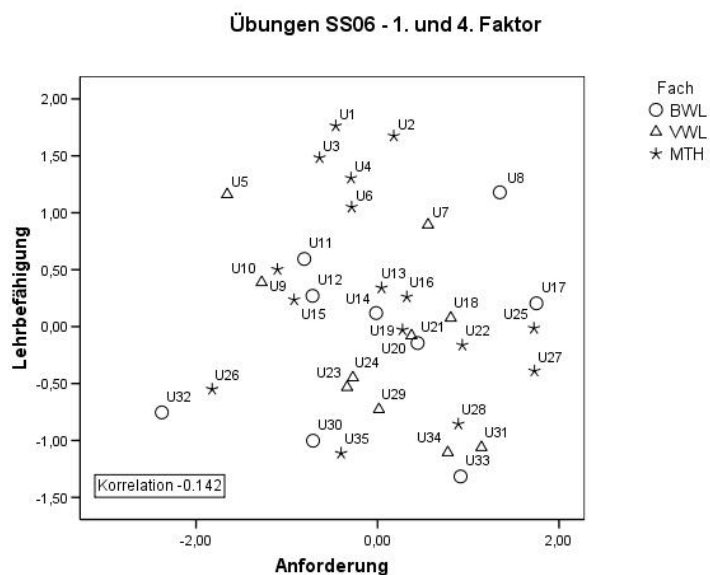


Abbildung 5.14: Scatterplot der Mittelwerte der Faktorwerte des 1. und 4. Faktors für die Übungen SS06

Aufgrund der Ergebnisse aus den vorangegangenen Abbildungen lässt sich vermuten, dass hinter den Faktoren, besonders zwischen den ersten drei, ein Generalfaktor

steht. Diese Vermutung wird bestärkt, schaut man sich die Lineplots für Übungen und Vorlesungen des SS06 an. In den Abbildungen 5.15 und 5.16 wurden nur die Faktorwerte der ersten drei Faktoren für jeweils alle Übungen und Vorlesungen eines Semesters betrachtet, da mit dem vierten Faktor kaum eine Korrelation existiert. Bei den Vorlesungen ergeben sich besonders für die Faktoren Lehrbefähigung und Selbsteinschätzung parallel zu Abzissenachse verlaufende Linien. Auch bei den Übungen ergeben sich überwiegend Linien, die parallel zur Abzissenachse verlaufen. Um die durch Gra-

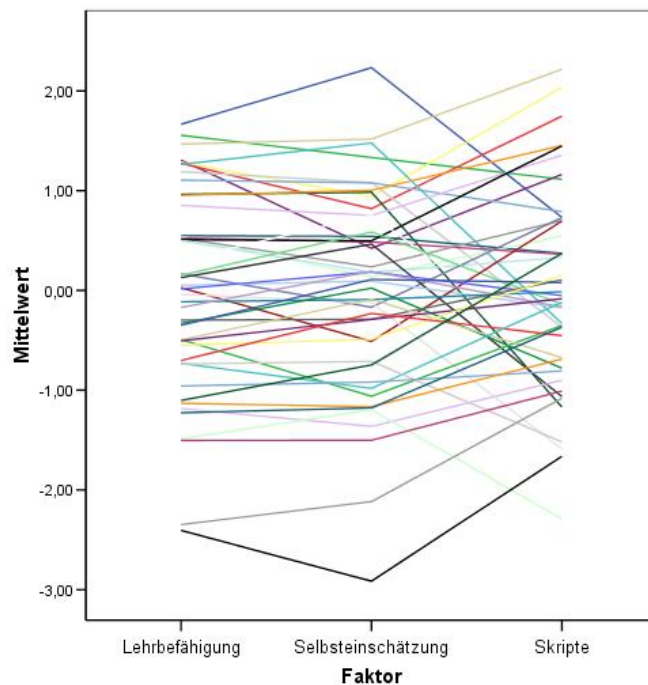


Abbildung 5.15: Lineplot der Faktorwerte (Ordinate) der ersten 3 Faktoren (Abszisse)
- Vorlesungen SS06

phiken gestützte Vermutung zu untersuchen, wird eine Faktormodell zweiter Ordnung

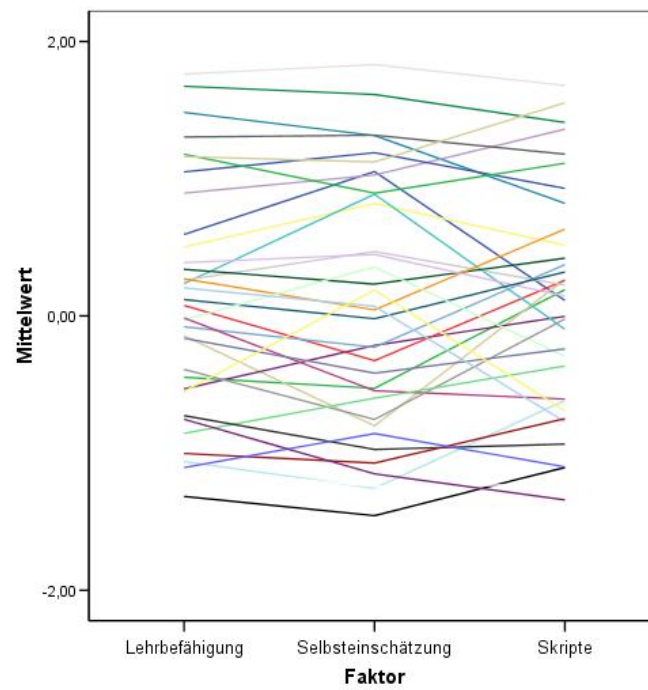


Abbildung 5.16: Lineplot der Faktorwerte (Ordinate) der ersten 3 Faktoren (Abszisse)
- Übungen SS06

aufgestellt. Bei dem Faktor zweiter Ordnung handelt es sich nicht um eine eigenständig gemessene latente Variable, sondern er stellt eine Zerlegung der Kovarianzstruktur der abhängigen Faktoren dar. Diese Theorie findet besondere Verwendung in der Intelligenzforschung z.B. im Intelligenzmodell von Cattell. Die vier Faktoren erster Ordnung haben sich mit Hilfe der explorativen Faktorenanalyse ergeben. Da drei der Faktoren stark miteinander korrelieren, erlauben sie es, weiter faktorisiert zu werden. Dadurch erhält man Faktoren zweiter Ordnung. Das zu schätzende Modell zweiter Ordnung ist in Abbildung 5.17 dargestellt. In Mplus ergeben sich ähnlich wie bei der konfir-

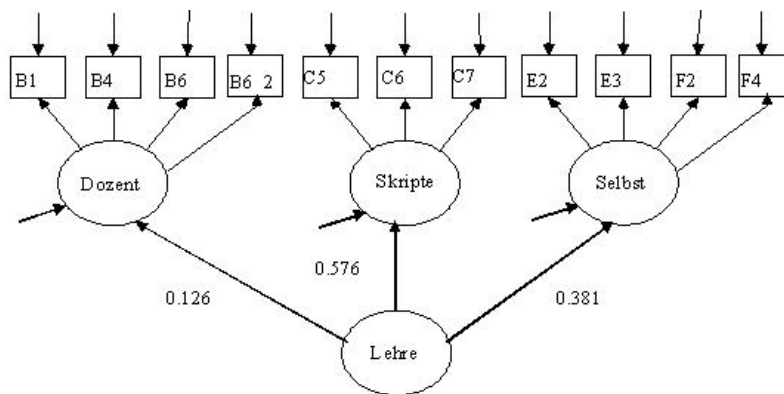


Abbildung 5.17: Konfirmatorisches Faktorenmodell zweiter Ordnung Übungen SS06

matorischen Faktorenanalyse verschiedene Teststatistiken und Größen, die das Modell bestätigen. Besonders interessant ist die Varianz der drei Konstrukte Lehrbefähigung, Skripte und Selbsteinschätzung, die durch den Generalfaktor „Wie finde ich die Veranstaltung“ erklärt wird. Die Residualvarianz ist in den Abbildungen 5.17 und 5.18 an die Pfeile zwischen dem Konstrukt Lehre und den drei ursprünglichen Konstrukten

geschrieben.

Wie durch das Ein-Faktor-Modell und die Auswertung der Faktorwerte des Vier-

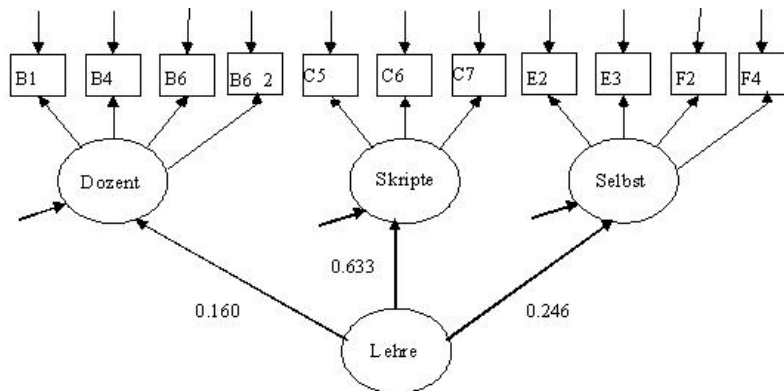


Abbildung 5.18: Konfirmatorisches Faktorenmodell zweiter Ordnung Vorlesungen SS06

Faktor-Modells vermutet, wird durch den Faktor zweiter Ordnung 'Lehre':

- zwischen 75% und 91% der Varianz des Konstrukts Lehrbefähigung in allen vier Datensätzen
- durchschnittlich 30% der Varianz des Konstrukts Skripte in allen vier Datensätzen
- durchschnittlich 70% der Varianz des Konstrukts Skripte in allen vier Datensätzen

erklärt. Die Korrelation zwischen dem Konstrukt Lehre und dem Konstrukt Anforderung, welches nicht in das Konstrukt Lehre mit eingeflossen ist, entspricht in etwa der Korrelation zwischen Lehrbefähigung und Anforderung (siehe Abbildung 5.19).

KAPITEL 5. BEWERTUNG DER LEHRVERANSTALTUNGEN MIT HILFE DER FAKTORWERTE

Durch die Betrachtung der Faktorwerte des Ein- und Vier- Modells ist es möglich, ein

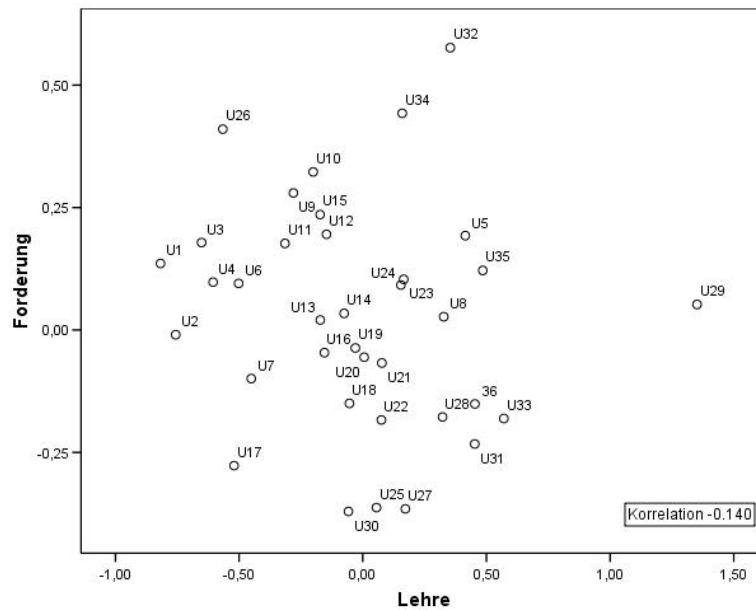


Abbildung 5.19: Zusammenhang zwischen Faktor zweiter Ordnung Lehre und Faktor Anforderung der Übungen im SS06

Ranking für die Lehrveranstaltungen aufzustellen. Die Betrachtung der Scatter- und Lineplots hat aber gezeigt, dass es nicht möglich ist, zu sagen, welche Aspekte einen guten bzw. schlechten Kurs ausmachen, da hinter den vier extrahierten Faktoren ein Generalfaktor steht. Von den Studenten wird hauptsächlich bewertet, ob ihnen der Kurs gefallen hat und wie hoch die Anforderung sind. Eine differenzierte Bewertung der einzelnen Aspekte findet im Großen und Ganzen nicht statt.

6 Zusammenfassung

In dieser Arbeit wurden verschiedene statistische Verfahren herangezogen, um die Evaluationsdaten des Sommersemester 2005 und 2006 auszuwerten. Ziel war es herauszufinden, inwieweit sich die Ergebnisse mit den Ergebnissen von Frau Zhou ähneln bzw. unterscheiden und des Weiteren herauszufinden, ob es eine Möglichkeit gibt, Lehrveranstaltungen zu bewerten und Merkmale zu extrahieren, welche eine gute bzw. schlechte Lehrveranstaltung kennzeichnen.

Die explorative Faktorenanalyse lieferte jeweils vier Faktoren, die auch in der konfirmatorischen Faktorenanalyse bestätigt wurden. Durch das Entfernen von Fragen aus den Fragebögen nach der Masterarbeit von Frau Zhou, zeigt sich in dieser Arbeit eine Verbesserung in der Struktur des Fragebogens, da die meisten Variablen den vier Konstrukten

- Lehrbefähigung
- Skripte
- Selbsteinschätzung
- Anforderung

eindeutig in allen Datensätzen zugeordnet werden konnten. Im Ein-Faktor-Modell wurden die Variablen der ersten drei Faktoren zusammengefasst.

In den weiteren Schritten wurden die Faktorwerte des Ein- und Vier-Faktor-Modells untersucht, speziell auch ihren Zusammenhang mit externen Faktoren. Hierbei zeigten sich die gleichen Tendenzen, wie bei Frau Zhou, die sich aber als statistisch nicht signifikant erwiesen. Hierzu gehört der Zusammenhang mit den Merkmalen

- Level des Kurse (Grund- oder Hauptstudium)
- Anzahl der Teilnehmer
- Uhrzeit und Tag des Kurses
- Dozentenwechsel.

Betrachtet man die Faktorwerte des Vier-Faktor-Modells, so zeigte sich eine sehr hohe Korrelation zwischen den ersten drei Faktoren Lehrbefähigung, Skripte und Selbsteinschätzung. Dies führte zu der Annahme, dass hinter diesen drei Faktoren ein Generalfaktor steht. Im Faktormodell zweiter Ordnung wurde dies bestätigt.

Zusammenfassend kommt man zu der Schlussfolgerung, dass mit den Evaluationsdaten zwar eine Bewertung des Kurses stattfindet, d.h., man ist in der Lage zu sagen, welcher Kurs gut bzw. schlecht von den Studenten bewertet wurde, aber man nicht sagen kann, welche Bereiche diese Bewertung hervorrufen. Um den Dozenten einen Anhaltspunkt zu geben, welche Anforderungen geändert werden müssen, um die Bewertung eines Kurses zu verbessern, sind die Evaluationsdaten nicht geeignet, d.h.,

die Aussagekraft beziehungsweise die Validität der Ergebnisse ist mit einer Reihe von Erhebungs- und Deutungsproblemen verbunden. Dies liegt zum Einen an dem Fragebogen, bei dem es sicherlich sinnvoll wäre, diesem unter sozialwissenschaftlicher Anleitung zu überarbeiten, um z.B. die Aussagekraft für den Dozenten zu erhöhen. Zum Anderen sind die befragten Studierenden so genannte Alltags-Evaluatoren. Prof. Dr. Helmut Kromrey beschreibt in seiner Arbeit dieses Problem in Bezug auf die Evaluation wie folgt [Kromrey (2001)]: „Jeder einzelne von ihnen bewertet irgend etwas (was er mit dem in der Frage angesprochenen Sachverhalt ad hoc assoziiert) irgendwie ('alles in allem' oder 'aus aktueller Erfahrung' oder 'mit Blick auf das Wesentliche' oder...) unter irgendwelchen Gesichtspunkten (Nutzen für sein Studium oder vermuteter Nutzen für den angestrebten Beruf oder aktuelles persönliches Interesse oder abstrakt-verallgemeinertes Interesse der Studierenden oder...)“. Darüber hinaus stellt zum Beispiel die Auswahl der zu Befragenden ein Problem dar. Werden wie in diesem Fall die Studenten in der 12. und 13. Vorlesungswoche befragt, bedeutet dies, dass durch die bis zu diesem Zeitpunkt erfolgte Selbstselektion die Bewertung positiv beeinflusst sein wird, denn diejenigen Studierenden, die der Veranstaltung und/oder dem Lehrenden gegenüber eine eher negative Einschätzung zukommen lassen, nehmen zu dem Datum an der Lehrveranstaltung bereits nicht mehr teil. Befragt man Studierende, die eine Pflichtveranstaltung ohne Wahlmöglichkeit zwischen alternativen Angeboten besuchen, fällt deren Urteil systematisch negativer aus. Die Gültigkeit der Ergebnisse wird ferner durch die Abgabe der Fragebögen mitbestimmt. Da die Abgabe freiwillig ist, setzt die Befragung ein aktives Beteiligungsverhalten seitens der Studenten voraus.

Im Grundstudium, gerade in den ersten Semestern fühlen sich vielleicht die Studenten zur Teilnahme verpflichtet. Nach vielen Semestern der Evaluation in jeder belegten Lehrveranstaltung setzt eine gewissen Ermüdung ein. Meist sehen die Studenten auch nicht mehr den Sinn oder Zweck der Befragung. Also kann es passieren, dass aufgrund fehlender Motivation die aktive Beteiligung an der Umfrage eindeutig sinkt.

Literaturverzeichnis

- [1] Ahlstick, K., Ulrich-Neitzert, T.(2004), *Grundlagen der Evaluation*, veröff. In: Manfred Hennen (Hg.): *Evaluation- Erfahrungen und Perspektiven*, Mainzer Beiträge zur Hochschulentwicklung, Bd. 4, S.1-21.
- [2] Enzmann, Dr.D.: *Die Technik und Logik von linearen Strukturgleichungsmodellen*.
- [3] Fahrmeier, L.,Hamerle, A., Tutz, G. (199), *Multivariate statistische Verfahren*, Berlin: de Gruyter.
- [4] Flora, D.B., Curran, P.J. (2004), *An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data*, in *Psychological Methods*.
- [5] Härdle, W., Simar, L. (2003), *Applied Multivariate Statistical Analysis*, Heidelberg: Springer Verlag.
- [6] Härdle, W., Müller, M., Sperlich, S., Werwatz, A. (2004), *Nonparametric and Semiparametric Models*, Heidelberg: Springer Verlag.
- [7] Kromrey, Helmut (2005), *Evaluation- Ein Überblick*, veröff. In: Heidrun Schöch

- (Hg.): Was ist Qualität. Die Entzauberung eines Mythos, Berlin: Wissenschaftl. Verlag (Schriftenreihe Wandel und Kontinuität in Organisationen, Band 6), S. 31-85.
- [8] Kühnlein, A. (2005), *Multiple Imputation von fehlenden Werten in epidemiologischen Beobachtungsstudien*.
- [9] Langer, Dr. W. (2002), *Einführung in die konfirmatorische Faktoren- und Pfadanalyse mit LISREL*, in Methoden V: Konfirmatorische Faktorenanalyse
- [10] Ludwig-Mayerhofer, W. (2004), *ILMES*.
- [11] Muthén, Linda K./Muthén, Bengt O., *Mplus User's Guide*.
- [12] Peng, S., *Comparison of Proc Impute and Schafer's Multiple Imputation Software*, Working Paper Series National Center for Education Statistics.
- [13] Pohlentz, Philipp, *Konfirmatorische Faktorenanalyse*.
- [14] Rönz, B. (2000), *Skript Computergestützte Statistik II*.
- [15] Rubin, D.B., Little, R.J.A. (2002), *Statistical analysis with missing data*, Wiley & Sons Inc.
- [16] Schafer, J.(1997), *Analysis of incomplete Multivariate Data*, Chapman and Hall/-CRC Press, London.
- [17] Überla, K. (1971), *Faktorenanalyse*, Berlin: Springer Verlag.

- [18] Wiedenbeck, M., Züll, C. (2001), *Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means Verfahren*, ZUMA How-to-Reihe Nr. 10.
- [19] *Lehre und Studium an der Humboldt-Universität zu Berlin- Evaluation von Studiengängen 2003- 2005*, Humboldt Universität zu Berlin (2005).
- [20] *Studierenden- Statistik- Akademisches Jahr 2004/05*, Humboldt Universität zu Berlin (2006).
- [21] *Evaluation Sommersemester 2006*, Humboldt Universität zu Berlin (2006).

A Appendix

Die beigefügte CD enthält die Datensätze, XploRe-, Mplus und R- Programme sowie die Ergebnisse der Auswertungen. Im Anhang finden sich Tabellen und Graphiken, auf welche im Text Bezug genommen wurde.

- In der Tabelle A.1 und A.2 werden die verwendeten Daten näher beschrieben.
- Tabelle A.3 zeigt den prozentualen Anteil an fehlenden Werten für jede Variable der vier Datensätze.
- In der Tabelle A.4 sieht man, welche Variable den Items des Übungs- und Vorlesungsfagebogens zugeordnet wird.
- In den Tabellen A.5 bis A.8 sind die Mittelwerte der Faktorwerte für jede Lehrveranstaltung angegeben. Die Zuordnung der Veranstaltungsnummer zu der entsprechenden Veranstaltung findet man auf der beigefügten CD. Die Anzahl gibt die ausgewerteten Bögen innerhalb der Veranstaltung an. Die Spalte Fach gibt an, aus welchen Bereich die Lehrveranstaltung kommt, BWL, VWL oder Methodisch. Das Level gibt an, ob die Lehrveranstaltung von den Studenten im

Haupt- oder Grundstudium besucht wird. In den letzten beiden Spalten sind der Wochentag und die Uhrzeit angegeben, an welchen der Kurs stattgefunden hat.

- Die Tabellen A.9 bis A.12 geben die Korrelation zwischen den vier Faktoren für jeden Datensatz an.
- Abbildung A.1 zeigt den Fragebogen für Vorlesungen und Abbildung A.2 zeigt den Fragebogen für Übungen.

	Anzahl	Studiengang			Geschlecht	
		BWL	VWL	Andere	Männlich	Weiblich
Vorlesung SS05	2851	54%	25.2%	16.4%	41.2%	44.1%
Übung SS05	1671	47.8%	27.8%	14%	45.1%	45.7%
Vorlesung SS06	2127	53.2%	27.2%	15.8%	42.1	41.6
Übung SS06	1635	47%	34.2%	16.6%	40.8%	47.1%

Tabelle A.1: Prozentuale Verteilung des Studienganges und Geschlechts

	08.00-14.00	14.00-20.00	Mo oder Fr	Di, Mi oder Do
SS06	50,35%	49,65%	39,72%	60,28%
SS05	55,37%	44,63%	38,84%	61,16%

Tabelle A.2: Prozentuale Verteilung der Uhrzeiten und Tage

	Vorlesungen		Übungen	
	2005	2006	2005	2006
Studiengang	4.4%	3.8%	10.4%	2.2%
Geschlecht	14.6%	16.4%	9.2%	12.1%
Gesamteindruck	9.9%	10.4%	21.9%	17.6
Wie oft gefehlt	2.6%	2.6%	2.8%	2.5%
Warum gefehlt	42.5%	44.2%	52.2%	52.3%
Erklärungsvermögen	1.0%	1.4%	1.4%	1.6%
Qualität Folien/Tafelbilder	1.2%	2.2%	2.2%	2.5%
Didaktische Kompetenz	1.4%	1.5%	1.7%	2.0%
Inhaltliche Schwerpunktsetzung	1.8%	2.3%	2.3%	2.2%
Klarheit der Gliederung	1.6%	1.7%	2.0%	2.2%
Bereitschaft Beantwortung Zwischenfragen	2.2%	2.4%	2.6%	2.3%
Qualität Beantwortung Zwischenfragen	3.0%	3.0%	2.6%	2.8%
Veranschaulichung	1.7%	1.6%		
Qualität Literatur/Skripte	2.5%	2.9%	4.5%	4.3%
Verfügbarkeit Literatur/Skripte	3.4%	3.4%	5.6%	4.6%
Internetpräsenz	3.0%	2.4%	3.9%	3.3%
Abstimmung Übung/Vorlesung			10.5%	5.1%
Geschwindigkeit Vorgehen	1.6%	2.0%	1.6%	2.0%
Formalisierung/Mathematisierung	1.8%	2.1%	2.2%	2.3%
Schwierigkeitsgrad	2.0%	2.9%	1.8%	2.7%
Interesse am Fach/Thema	1.6%	2.1%		
Aufmerksamkeit Veranstaltung	1.8%	2.1%	2.0%	2.8%
Erkenntnis- bzw. Lernzuwachs	2.0%	2.5%	2.3%	3.6%
Vor- und Nachbereitungszeit	12.2%	9.6%	10.5%	11.2%
stressfrei/stressig	1.5%	2.0%	1.7%	2.6%
interessant/langweilig	1.6%	1.8%	1.6%	2.5%
diszipliniert/chaotisch	1.5%	1.8%	1.9%	2.5%
motivierend/geistig blockierend	1.7%	1.9%	1.9%	2.7%

Tabelle A.3: Anteil der fehlenden Werte pro Variable

Variablenname	Kodierung Vorlesungen	Kodierung Übungen
Lehrbefähigung		
Erklärungsvermögen	b1	b1
Qualität Folien/Tafelbilder	b3	b3
Didaktische Kompetenz	b4	b4
Inhaltliche Schwerpunktsetzung	c1	c1
Klarheit der Gliederung	c2	c2
Bereitschaft Beantwortung Zwischenfragen	b6	b6
Qualität Beantwortung Zwischenfragen	b62	b62
Lehrmaterialien		
Veranschaulichung	c4	
Qualität Literatur/Skripte	c5	c5
Verfügbarkeit Literatur/Skripte	c6	c6
Internetpräsenz	c7	c7%
Abstimmung Übung/Vorlesung		c72
Leistungsanforderung		
Geschwindigkeit Vorgehen	d1	d1
Formalisierung/Mathematisierung	d2	d2
Schwierigkeitsgrad	d3	d3
Selbsteinschätzung		
Interesse am Fach/Thema	e1	
Aufmerksamkeit Veranstaltung	e2	e2
Erkenntnis- bzw. Lernzuwachs	e3	e3
Vor- und Nachbereitungszeit	e4	e4
Atmosphäre		
stressfrei/stressig	f1	f1
interessant/langweilig	f2	f2
diszipliniert/chaotisch	f3	f3
motivierend/geistig blockierend	f4	f4

Tabelle A.4: Zuordnung der Variablen

Kurs	FS	Anzahl	Fach	Level	Tag	Uhrzeit
V1	0.91	11	MTH	HS	Di bis Do	14-20
V2	0.72	10	VWL	HS	Di bis Do	14-20
V3	0.68	26	VWL	HS	Di bis Do	14-20
V4	0.66	53	BWL	HS	Di bis Do	8-14
V5	0.65	49	VWL	HS	Di bis Do	8-14
V6	0.55	111	VWL	GS	Di bis Do	14-20
V7	0.53	22	MTH	HS	Di bis Do	14-20
V8	0.43	69	BWL	HS	Mo oder Fr	14-20
V9	0.42	23	MTH	HS	Di bis Do	14-20
V10	0.42	21	BWL	HS	Mo oder Fr	8-14
V11	0.41	22	BWL	HS	Mo oder Fr	8-14
V12	0.41	16	VWL	HS	Di bis Do	8-14
V13	0.41	7	BWL	HS	Di bis Do	14-20
V14	0.34	26	BWL	HS	Mo oder Fr	14-20
V15	0.33	30	BWL	HS	Mo oder Fr	8-14
V16	0.33	19	MTH	HS	Mo oder Fr	14-20
V17	0.31	82	VWL	GS	Di bis Do	14-20
V18	0.28	63	MTH	GS	Di bis Do	8-14
V19	0.26	26	MTH	HS	Mo oder Fr	14-20
V20	0.24	51	BWL	HS	Di bis Do	14-20
V21	0.23	15	BWL	HS	.	.
V22	0.16	71	BWL	HS	Di bis Do	14-20
V23	0.13	60	VWL	HS	Di bis Do	14-20
V24	0.08	9	MTH	HS	Di bis Do	8-14
V25	0.05	21	VWL	HS	Mo oder Fr	8-14
V26	0.02	144	VWL	GS	Di bis Do	8-14
V27	-0.13	18	BWL	HS	Mo oder Fr	14-20
V28	-0.13	67	BWL	HS	Di bis Do	8-14
V29	-0.13	25	MTH	HS	Mo oder Fr	8-14
V30	-0.14	68	VWL	GS	Di bis Do	8-14
V31	-0.16	26	VWL	HS	Mo oder Fr	8-14
V32	-0.19	130	BWL	GS	Di bis Do	8-14
V33	-0.20	107	BWL	GS	Di bis Do	8-14
V34	-0.24	20	VWL	HS	Mo oder Fr	14-20
V35	-0.32	107	BWL	HS	Di bis Do	8-14
V36	-0.32	37	BWL	HS	Di bis Do	8-14
V37	-0.41	24	BWL	HS	Di bis Do	14-20
V38	-0.41	42	MTH	GS	Di bis Do	8-14
V39	-0.42	58	BWL	HS	Mo oder Fr	8-14
V40	-0.43	112	BWL	GS	Mo oder Fr	14-20
V41	-0.46	27	VWL	HS	Mo oder Fr	14-20
V42	-0.50	32	BWL	HS	Di bis Do	8-14
V43	-0.56	21	VWL	HS	Di bis Do	8-14
V44	-0.59	57	BWL	HS	Mo oder Fr	8-14
V45	-0.59	16	VWL	HS	Mo oder Fr	14-20
V46	-0.62	10	VWL	HS	Di bis Do	8-14
V47	-0.63	66	MTH ₉₁	GS	Mo oder Fr	8-14

Tabelle A.5: Faktorwerte der Vorlesungen SS06

Kurs	FS	Anzahl	Fach	Level	Tag	Uhrzeit
V051	0.99	11	BWL	HS	Mo oder Fr	8-14
V052	0.89	15	MTH	HS	Di bis Do	14-20
V053	0.82	42	MTH	HS	Di bis Do	8-14
V054	0.66	12	VWL	HS	Di bis Do	14-20
V055	0.61	80	VWL	HS	Di bis Do	8-14
V056	0.58	41	BWL	HS	Mo oder Fr	8-14
V057	0.55	48	BWL	HS	Mo oder Fr	8-14
V058	0.49	33	VWL	HS	Mo oder Fr	14-20
V059	0.36	112	MTH	GS	Di bis Do	8-14
V0510	0.31	25	VWL	HS	Di bis Do	8-14
V0511	0.27	202	VWL	GS	Di bis Do	8-14
V0512	0.22	225	VWL	HS	Mo oder Fr	8-14
V0513	0.21	51	BWL	HS	Di bis Do	8-14
V0514	0.20	42	BWL	HS	Mo oder Fr	14-20
V0515	0.20	53	BWL	HS	Mo oder Fr	14-20
V0516	0.20	65	VWL	HS	Di bis Do	14-20
V0517	0.17	154	VWL	GS	Mo oder Fr	8-14
V0518	0.17	105	VWL	GS	Mo oder Fr	8-14
V0519	0.15	121	BWL	HS	Di bis Do	8-14
V0520	0.14	32	BWL	HS	Mo oder Fr	14-20
V0521	0.13	14	MTH	HS	Di bis Do	8-14
V0522	0.13	20	BWL	HS	Mo oder Fr	14-20
V0523	0.08	32	BWL	HS	Di bis Do	8-14
V0524	0.05	16	MTH	HS	Di bis Do	8-14
V0525	0.04	72	VWL	HS	Di bis Do	14-20
V0526	-0.00	38	BWL	HS	Di bis Do	8-14
V0527	-0.02	25	BWL	HS	Di bis Do	8-14
V0528	-0.03	13	BWL	HS	Di bis Do	14-20
V0529	-0.05	10	MTH	HS	Di bis Do	8-14
V0530	-0.05	22	BWL	HS	Mo oder Fr	8-14
V0531	-0.06	47	BWL	GS	Di bis Do	8-14
V0532	-0.07	15	BWL	HS	Mo oder Fr	8-14
V0533	-0.07	58	BWL	HS	Mo oder Fr	14-20
V0534	-0.13	58	BWL	HS	Di bis Do	8-14
V0535	-0.20	13	MTH	HS	Mo oder Fr	8-14
V0536	-0.20	92	BWL	HS	Di bis Do	8-14
V0537	-0.24	18	BWL	HS	Di bis Do	14-20
V0538	-0.26	51	BWL	HS	Mo oder Fr	14-20
V0539	-0.28	60	BWL	HS	Di bis Do	8-14
V0540	-0.30	44	MTH	HS	Mo oder Fr	8-14
V0541	-0.36	99	BWL	HS	Mo oder Fr	14-20
V0542	-0.43	146	BWL	GS	Mo oder Fr	14-20
V0543	-0.45	63	VWL	HS	Mo oder Fr	8-14
V0544	-0.49	28	MTH	GS	Mo oder Fr	14-20
V0545	-0.60	61	BWL	HS	Mo oder Fr	8-14
V0546	-0.61	161	BWL	GS	Mo oder Fr	8-14
V0547	-0.63	38	BWL ₉₂	HS	Mo oder Fr	14-20
V0548	-0.80	15	VWL	HS	Di bis Do	8-14
V0549	-1.32	52	MTH	GS	Mo oder Fr	8-14

Tabelle A.6: Faktorwerte der Vorlesungen SS05

Kurs	FS	Anzahl	Fach	Level	Tag	Uhrzeit
U1	0.94	31	MTH	GS	Di bis Do	8-14
U2	0.80	18	MTH	GS	Di bis Do	14-20
U3	0.69	16	MTH	GS	Di bis Do	14-20
U4	0.67	67	MTH	GS	Di bis Do	8-14
U5	0.64	15	MTH	HS	Mo oder Fr	8-14
U6	0.58	70	MTH	GS	Di bis Do	8-14
U7	0.57	26	VWL	HS	Di bis Do	14-20
U8	0.45	23	BWL	HS	Di bis Do	14-20
U9	0.30	12	MTH	HS	Mo oder Fr	8-14
U10	0.26	33	VWL	HS	Mo oder Fr	8-14
U11	0.24	31	BWL	GS	Di bis Do	8-14
U12	0.21	51	BWL	HS	Di bis Do	14-20
U13	0.19	119	MTH	GS	Di bis Do	8-14
U14	0.12	118	BWL	GS	Mo oder Fr	8-14
U15	0.12	28	MTH	GS	Mo oder Fr	8-14
U16	0.10	25	MTH	GS	Di bis Do	14-20
U17	0.07	24	BWL	HS	Di bis Do	8-14
U18	0.05	47	VWL	HS	Di bis Do	8-14
U19	0.04	18	MTH	GS	Mo oder Fr	8-14
U20	0.00	88	VWL	HS	Di bis Do	8-14
U21	-0.01	21	BWL	HS	Di bis Do	14-20
U22	-0.11	49	MTH	GS	Di bis Do	14-20
U23	-0.12	159	VWL	GS	Mo oder Fr	8-14
U24	-0.13	119	VWL	GS	Mo oder Fr	14-20
U25	-0.14	55	MTH	GS	Mo oder Fr	14-20
U26	-0.16	18	VWL	HS	Di bis Do	14-20
U27	-0.17	40	MTH	GS	Di bis Do	8-14
U28	-0.34	32	MTH	GS	Mo oder Fr	8-14
U29	-0.38	14	BWL	HS	Mo oder Fr	14-20
U30	-0.49	15	VWL	HS	Mo oder Fr	14-20
U31	-0.49	119	VWL	HS	Mo oder Fr	14-20
U32	-0.54	11	BWL	HS	Di bis Do	8-14
U33	-0.56	67	BWL	HS	Di bis Do	14-20
U34	-0.57	20		HS	Di bis Do	14-20
U35	-0.65	52	MTH	GS	Di bis Do	14-20
U36	-1.37	20	VWL	HS	Di bis Do	14-20

Tabelle A.7: Faktorwerte der Übungen SS06

Kurs	FS	Anzahl	Fach	Level	Tag	Uhrzeit
U051	0,82	62	BWL	HS	Di bis Do	8-14
052	0,78	32	BWL	GS	Di bis Do	8-14
U053	0,76	16	BWL	GS	Di bis Do	8-14
U054	0,75	25	MTH	GS	Di bis Do	14-20
U055	0,53	150	MTH	GS	Di bis Do	8-14
U056	0,51	18	BWL	HS	Di bis Do	14-20
U057	0,37	73	VWL	HS	Di bis Do	8-14
U058	0,23	101	VWL	GS	Mo oder Fr	14-20
U059	0,18	49	BWL	GS	Mo oder Fr	8-14
U0510	0,15	30	MTH	GS	Di bis Do	14-20
U0511	0,15	18	MTH	HS	Mo oder Fr	8-14
U0512	0,13	59	MTH	GS	Di bis Do	8-14
U0513	0,11	33	VWL	GS	Di bis Do	14-20
U0514	0,06	59	VWL	HS	Di bis Do	8-14
U0515	-0,02	12	MTH	GS	Di bis Do	8-14
U0516	-0,04	68	BWL	GS	Mo oder Fr	8-14
U0517	-0,05	48	BWL	GS	Di bis Do	14-20
U0518	-0,08	120	VWL	GS	Di bis Do	8-14
U0519	-0,10	12	MTH	GS	Di bis Do	8-14
U0520	-0,14	39	VWL	HS	Di bis Do	14-20
U0521	-0,19	16	MTH	GS	Mo oder Fr	8-14
U0522	-0,23	23	BWL	HS	Mo oder Fr	14-20
U0523	-0,23	136	VWL	HS	Di bis Do	14-20
U0524	-0,26	37	MTH	GS	Di bis Do	8-14
U0525	-0,33	69	BWL	GS	Di bis Do	8-14
U0526	-0,36	124	BWL	HS	Mo oder Fr	8-14
U0527	-0,38	40	BWL	GS	Di bis Do	8-14
U0528	-0,47	78	VWL	HS	Di bis Do	8-14
U0529	-0,60	30	.	GS	.	.
U0530	-0,74	68	BWL	GS	Di bis Do	8-14
U0531	-0,80	26	MTH	GS	Di bis Do	14-20

Tabelle A.8: Faktorwerte der Übungen SS05

	Lehrbefähigung	Skripte	Selbsteinschätzung	Anforderung
Lehrbefähigung	1	0.716	0.951	-0.403
Skripte		1	0.716	-0.184
Selbsteinschätzung			1	-0.484
Anforderung				1

Tabelle A.9: Korrelation zwischen den Faktoren für Übungen SS05

	Lehrbefähigung	Skripte	Selbsteinschätzung	Anforderung
Lehrbefähigung	1	0.894	0.941	-0.142
Skripte		1	0.833	-0.100
Selbsteinschätzung			1	-0.266
Anforderung				1

Tabelle A.10: Korrelation zwischen den Faktoren für Übungen SS06

	Lehrbefähigung	Skripte	Selbsteinschätzung	Anforderung
Lehrbefähigung	1	0.653	0.943	-0.287
Skripte		1	0.556	-0.280
Selbsteinschätzung			1	-0.327
Anforderung				1

Tabelle A.11: Korrelation zwischen den Faktoren für Vorlesungen SS05

	Lehrbefähigung	Skripte	Selbsteinschätzung	Anforderung
Lehrbefähigung	1	0.698	0.949	-0.095
Skripte		1	0.599	0.119
Selbsteinschätzung			1	-0.223
Anforderung				1

Tabelle A.12: Korrelation zwischen den Faktoren für Vorlesungen SS06

Seite 2

Bitte ausmalen: ● *NICHT so:* ⊗

41803

Dein Studiengang

BWL Dipl MEMS

BWL BA Wipäd Dipl

BWL NF/ZF LA WiWi

VWL Dipl WInf Master

VWL BA Stat. BA

VWL NF/ZW Stat. Master

Dein Geschlecht

weiblich männlich

Wie oft hast Du bisher in diesem Semester diese Vorlesung versäumt?

nie 1x 2x 3x öfter

Warum hast Du gefehlt? (Bitte nur ein Feld ausfüllen!)

Vorlesung zu früh / zu spät

Überschneidung mit anderen Veranstaltungen

diese Vorlesung gibt mir nichts


zu starke Belastung durch andere Fächer

Sonstige, und zwar

Bitte deutlich schreiben

nicht auszufüllen!

Nr. Intern



Gesamteindruck der Veranstaltung

	1	2	3	4	5
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Lehrbefähigung des Dozenten/ der Dozentin

Erklärungsvermögen	sehr gut	▶	sehr schlecht
Qualität der Folien und / oder Tafelbilder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Didaktische Kompetenz/ Gestaltung des Lernangebots	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Inhaltliche Schwerpunktsetzung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Klarheit der Gliederung des Stoffes ("roter Faden")	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Kommunikation mit den Studierenden

Bereitschaft zur Beantwortung fachlicher Zwischenfragen	sehr gut	▶	sehr schlecht
Qualität der Beantwortung fachlicher Zwischenfragen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Lehrmaterialien

Veranschaulichung der theoretischen Inhalte anhand von praktischen Beispielen	sehr gut	▶	sehr schlecht
Qualität der Literatúrauswahl und Skripte	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Verfügbarkeit der Literatúrauswahl und Skripte (Bibliothek, Copyshop, Internet)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Internetpräsenz (Skripte, Aktualität)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Leistungsanforderung

Geschwindigkeit des Vorgehens	zu hoch	▶	genau richtig	▶	zu niedrig
Formalisierung / Mathematisierung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Schwierigkeitsgrad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vor- und Nachbereitung auf eine einzelne Veranstaltung (in Minuten) 0 bis 30 bis 60 bis 90 mehr als 90

Selbsteinschätzung

Interesse am Fach / Thema	hoch	▶	niedrig
Aufmerksamkeit in der Veranstaltung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Erkenntnis- bzw. Lernzuwachs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie empfindest Du die Atmosphäre in der Veranstaltung?

stressfrei	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	stressig
interessant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	langweilig
diszipliniert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	chaotisch
motivierend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	geistig blockierend

Unsere Internetseiten findet Ihr unter: www.wiwi-evaluation.de

EMail: eva@wiwi.hu-berlin.de

Abbildung A.1: Fragebogen für Vorlesungen



Bitte ausmalen: ● NICHT so: ⊗

nicht auszufüllen

--	--	--	--

 Nr. intern

Dein Studiengang

- BWL Dipl MEMS
- BWL BA Wipäd Dipl
- BWL NF/ZF LA WiWi
- VWL Dipl Winf Master
- VWL BA Stat. BA
- VWL NF/ZF Stat. Master

Wie oft hast Du bisher in diesem Semester diese Übung versäumt?

- nie 1x 2x 3x öfter

Warum hast Du gefehlt?
 (Bitte nur ein Feld ausfüllen!)

- Übung zu früh / zu spät
- Überschneidung mit anderen Veranstaltungen
- diese Übung gibt mir nichts
- zu starke Belastung durch andere Fächer
- Sonstige, und zwar

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

Bitte deutlich schreiben



Dein Geschlecht

- weiblich männlich

Übungsleitung

- Erklärungsvermögen
- Didaktische Kompetenz/ Gestaltung des Lernangebots
- Qualität der Folien und / oder Tafelbilder
- Bereitschaft zur Beantwortung fachlicher Zwischenfragen
- Qualität der Beantwortung der Zwischenfragen

Gesamteindruck der Veranstaltung:

	1	2	3	4	5
sehr gut	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
sehr schlecht	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Konzept der Veranstaltung

- Inhaltliche Schwerpunktsetzung
- Klarheit der Gliederung des Stoffes ("roter Faden")
- Inhaltliche Abstimmung von Vorlesung und Übung

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Lehrmaterialien

- Qualität der Skripte und Literatur
- Verfügbarkeit der Literaturlauswahl und Skripte (Bibliothek, Copyshop, Internet)
- Internetpräsenz (Übungsblätter / Skripte, Aktualität)

<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Leistungsanforderung

- Geschwindigkeit des Vorgehens
- Formalisierung / Mathematisierung
- Schwierigkeitsgrad

zu hoch	genau richtig	zu niedrig
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Vor- und Nachbereitung auf eine einzelne Veranstaltung (in Minuten) 0 bis 30 bis 60 bis 90 mehr als 90

Selbsteinschätzung

- Aufmerksamkeit in der Veranstaltung
- Erkenntnis- bzw. Lernzuwachs

hoch	niedrig
<input type="radio"/>	<input type="radio"/>
<input type="radio"/>	<input type="radio"/>

Wie empfindest Du die Atmosphäre in der Veranstaltung?

stressfrei	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	stressig
interessant	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	langweilig
diszipliniert	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	chaotisch
motivierend	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	geistig blockierend

Wenn Du Lust hast, Dich mit uns für die Verbesserung der Lehre zu engagieren, uns bei der Evaluation zu helfen, oder eigene Ideen einbringen möchtest, dann melde dich doch kurz per EMail an eva@wiwi.hu-berlin.de

Abbildung A.2: Fragebogen für Übungen