Analyse der Erwartungen an die Lehrveranstaltungen Statistik I und II

DIPLOMARBEIT

zur Erlangung des Grades
einer Diplom-Volkswirtin

eingereicht an der
Wirtschaftswissenschaftlichen Fakultät
Humboldt-Universität zu Berlin

Julia Amedja
Matrikelnr. 185887

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

10. April 2008

Inhaltsverzeichnis

1 Einleitung **1**

2 Allgemeiner Überblick **4**

2.1 Erhebung der Daten 4

2.2 Erhobene Daten 5

2.3 Maryland Physics Expectations Survey 9

2.4 Fehlende Werte 11

3 Statistische Methoden **16**

3.1 Zusammenhangsmaße 16

3.1.1 Korrelation 16

3.1.2 Unsicherheitskoeffizient 17

3.1.3 Eta-Koeffizient 17

3.2 Parametervergleiche 18

3.2.1 Mann-Whitney-Test 18

3.2.2 Kruskal-Wallis-Test 19

3.3 Faktorenanalyse 20

3.3.1 Explorative Faktorenanalyse allgemein 20

3.3.2 Explorative Faktorenanalyse für ordinalskalierte Daten 25

3.3.3 Reliabilität 27

3.3.4 Konfirmatorische Faktorenanalyse 27

3.4 Clusteranalyse 30

4 Analysen **33**

4.1 Vergleich verschiedener Gruppen 33

4.2 Zusammenhänge zwischen den Items 36
4.3 Explorative Faktorenanalyse 36
4.4 Konfirmatorische Faktorenanalyse 44
4.5 Analyse der Faktorwerte 46
4.5.1 Faktor 1 46
4.5.2 Faktor 2 49
4.5.3 Faktor 4 52
4.5.4 Faktor 5 54
4.5.5 Clusteranalyse 57

5 Zusammenfassung **63**

Literaturverzeichnis **66**

Anhang A **69**

Anhang B **73**

Abbildungsverzeichnis

1.1 Evaluation Statistik I und II, aus Härdle et al. 2007 2

2.1 Balkendiagramm: Studierende nach Studiengang 6
2.2 Balkendiagramm: Studierende nach Semester 7
2.3 Balkendiagramm: Studierende nach Geschlecht 8
2.4 Boxplots: Alter weiblicher und männlicher Studierender 8
2.5 Balkendiagramm Studierende nach Schulabschluss 9
2.6 Fehlende Werte pro Item 11

4.1 Signifikante polychorische Korrelationen 37
4.2 Eigenwertediagramm 38
4.3 Fehlerbalkendiagramm: Faktorwerte 1 nach Studiengang . . . 48
4.4 Fehlerbalkendiagramm: Faktorwerte 1 nach Geschlecht 48
4.5 Fehlerbalkendiagramm Faktorwerte 1 nach Schulabschluss . . 49
4.6 Scatterplot: Faktorwerte 1 und Alter 49
4.7 Fehlerbalkendiagramm Faktorwerte 2 nach Studiengang 50
4.8 Fehlerbalkendiagramm Faktorwerte 2 nach Geschlecht 51
4.9 Fehlerbalkendiagramm Faktorwerte 2 nach Schulabschluss . . 51
4.10 Scatterplot: Faktorwerte 2 und Alter 52
4.11 Fehlerbalkendiagramm Faktorwerte 4 nach Studiengang 53
4.12 Fehlerbalkendiagramm Faktorwerte 4 nach Geschlecht 53
4.13 Fehlerbalkendiagramm Faktorwerte 4 nach Schulabschluss . . 54
4.14 Scatterplot: Faktorwerte 4 und Alter 54
4.15 Fehlerbalkendiagramm Faktorwerte 5 nach Studiengang 55
4.16 Fehlerbalkendiagramm Faktorwerte 5 nach Geschlecht 55
4.17 Fehlerbalkendiagramm Faktorwerte 5 nach Schulabschluss . . 56

4.18 Scatterplot: Faktorwerte 5 und Alter 56
4.19 Clusteranalyse: Verlauf der Fusionswerte 57
4.20 Clusteranalyse: Verlauf der letzten 10 Fusionswerte 58
4.21 Lineplot: 4 Cluster 59
4.22 Balkendiagramm: 4-Clusterlösung und Studiengang 59
4.23 Lineplot: 2 Cluster 60

Tabellenverzeichnis

4.1 Modi der Antworten von Studierenden und Experten 34
4.2 Signifikante Abweichungen von Antworten der Experten . . . 35
4.3 Eigenwerte der Korrelationsmatrix und kumulierte erklärte Varianz 37
4.4 Varimax-rotierte Ladungen >|0,4| 39
4.5 Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 1 . . 44
4.6 Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 2 . . 44
4.7 Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 4 . . 44
4.8 Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 5 . . 45
4.9 Eta-Koeffizienten der Faktorwerte und verschiedener Merkmale 46
4.10 Kruskal-Wallis-Tests für Faktorwerte des Faktors 1 47
4.11 Kruskal-Wallis-Tests für Faktorwerte des Faktors 2 50
4.12 Kruskal-Wallis-Tests für Faktorwerte des Faktors 4 52
4.13 Kruskal-Wallis-Tests für Faktorwerte des Faktors 5 56
4.14 Unsicherheitskoeffizienten (Cluster abh.) für 4-Clusterlösung . 59
4.15 Unsicherheitskoeffizienten (Cluster abh.) für 2-Clusterlösung . 61

1 Einleitung

Auf der Website der Humboldt-Universität zu Berlin heißt es unter der Rubrik „Studienangebot“:

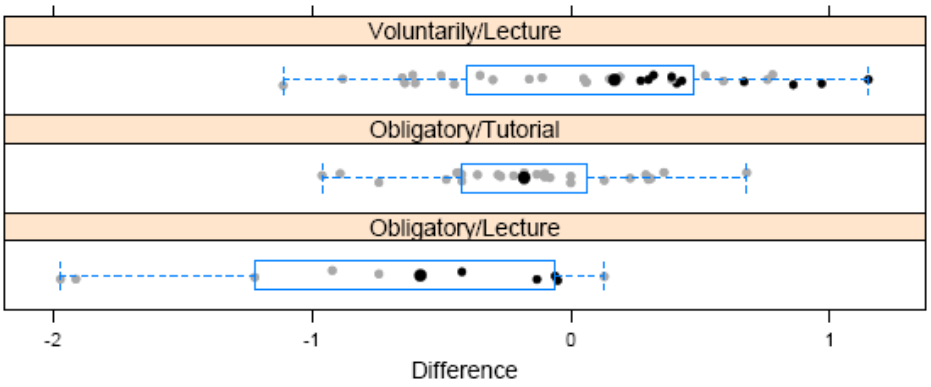
> **Warum sollten Sie BWL** [bzw. VWL, d. Verf.] **an der HU studieren?**
> Die Wirtschaftswissenschaftliche Fakultät der Humboldt-Universität hat sich [...] ein anerkanntes wissenschaftliches Profil erarbeitet, das zum überwiegenden Teil auf der Anwendung von quantitativen Methoden und Verfahren auf wirtschaftswissenschaftliche Fragestellungen beruht. Der Schwerpunkt des Studienangebots der Fakultät wird daher [...] auf dem Erlernen und Anwenden anspruchsvoller quantitativer wirtschaftswissenschaftlicher Methoden liegen. [...] Sie sollten deshalb Interesse und Begabung für quantitative Gebiete wie Statistik und Mathematik mitbringen.[1]

Ihre entschieden quantitave Orientierung ist in der Tat eines der Markenzeichen der Wirtschaftswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin. Für Studierende dürfte dies eine wichtige Rolle bei der Wahl ihrer Hochschule gespielt haben. Die Attraktivität der Fakultät spiegelt sich u.a. in den hohen NC für die Bachelorstudiengänge BWL und VWL[2] und den hohen Bewerbungszahlen wider (Vgl. Studierendenstatistik 2007/08).
Bereits in den ersten Semestern sind jedoch erfahrungsgemäß viele Studierende unzufrieden mit ihrem Studium bzw. bestimmten Lehrveranstaltungen.
Dies wird z.B. für das Statistik-Modul in den von der Evaluationsgruppe der Fakultät durchgeführten Evaluationen deutlich (Vgl. z.B. Evaluation Win-

[1] http://studium.hu-berlin.de/beratung/sgb/bwl, 09.04.2008

[2] Wintersemester 2007/08: Letzte Zulassung nach Durchschnittsnoten von 1,8 (BWL) und 2,1 (VWL).

Abb. 1.1: Evaluation Statistik I und II, aus Härdle et al. 2007



tersemester 2006/07 und Sommersemester 2007). In der aus Härdle et al. (2007, S.3) entnommenen Abbildung 1.1 ist ein Teil der Evaluationsergebnisse graphisch dargestellt. Der Durchschnitt aller Kurse der Fakultät vom Wintersemester 1999/2000 bis zum Sommersemester 2006 in den Fragen „Gesamteindruck der Veranstaltung / Erwartungen an die Veranstaltung erfüllt“ wurde gleich Null gesetzt. Im Vergleich dazu sind im unteren Teil die Bewertungen für die Pflichtveranstaltungen Statistik I und II dargestellt, in der Mitte für die freiwilligen Übungen und oben für alle übrigen freiwilligen Kurse des Instituts für Statistik und Ökonometrie der Fakultät. Die Statistik I und II werden demnach deutlich schlechter bewertet als der Durchschnitt aller Veranstaltungen.

Auch die Ergebnisse der Klausuren in Statistik I und II legen die Vermutung nahe, dass viele Studierende sich über die an sie gestellten Anforderungen nicht völlig im Klaren sind (Vgl. Brandes 2004).

In dieser Arbeit werden die Erwartungen analysiert, die Studierende am Anfang ihres Studiums an die Lehrveranstaltungen Statistik I und II stellen. In den Kapiteln 2 und 3 werden zunächst die verwendeten Daten und die zum Einsatz kommenden statistischen Methoden beschrieben.

Im Kapitel 4 wird verschiedenen Fragestellungen nachgegangen. Zum einen wird gezeigt, wie sich die Erwartungen von Studierenden und Statistik-Experten unterscheiden. Das Kernstück der Untersuchungen stellt jedoch eine Fakto-

renanalyse dar: Es wird versucht, Faktoren zu identifizieren, welche die Erwartungen der Studierenden bestimmen.
Im Kapitel 5 werden die Ergebnisse der Analysen zusammengefasst.
Die Analysen wurden unter Verwendung von Mplus, SPSS, R, XploRe und Microsoft Excel durchgeführt.

2 Allgemeiner Überblick

2.1 Erhebung der Daten

Die in dieser Arbeit verwendeten Daten wurden durch eine Befragung unter Studierenden der Wirtschaftswissenschaftlichen Fakultät der Humboldt-Universität zu Berlin erhoben. Herr Prof. Brandt stellte hierfür einige Minuten am Ende der Vorlesung Mathematik I in der sechsten Vorlesungswoche des Wintersemesters 2007/2008 zur Verfügung. Es wird angenommen, dass zu diesem Termin nahezu alle Studierenden im ersten Semester der Bachelorstudiengänge BWL, VWL und Wirtschaftspädagogik anwesend waren, da es sich hierbei um eine für das erste Semester vorgesehene Pflichtveranstaltung im Rahmen der genannten Studiengänge handelt. Da es zudem für diese Vorlesung kaum Handouts gibt und somit ein Skript nur durch eigene Mitschriften erhältlich ist, wird diese üblicherweise von den meisten Studierenden regelmäßig besucht. Die genaue Anzahl der Anwesenden am Tag der Befragung ist nicht bekannt. Laut Information des Prüfungsbüros der Wirtschaftswissenschaftlichen Fakultät waren im Wintersemester 2007/08 insgesamt 239 Studierende im ersten Semester in den Bachelorstudiengängen BWL und VWL eingeschrieben. Die Studierendenstatistik der Humboldt-Universität zu Berlin führt zudem für den Bachelorstudiengang Wirtschaftspädagogik im gleichen Zeitraum eine Anzahl von 38 Studierenden auf. Während der Befragung wurden insgesamt 229 Bögen beantwortet. Es wird davon ausgegangen, dass damit ein Großteil der relevanten Gruppe erreicht wurde.
Eine weitere Befragung wurde im Januar 2008 unter neun Teilnehmern des Privatissimums Statistik durchgeführt, welches der Vorbereitung und Anfertigung von Bachelor-, Diplom- oder Masterarbeiten am Lehrstuhl für Statistik dient.

2.2 Erhobene Daten

Der verwendete Fragebogen umfasst 38 Aussagen zum Modul Statistik, das die Studierenden planmäßig im zweiten und dritten Semester besuchen werden. Der Anhang A enthält die vollständige Liste der Items.
Die am Ende der Vorlesung Mathematik I befragten Studierenden wurden gebeten anzugeben, in welchem Maße sie den einzelnen Aussagen zustimmen bzw. nicht zustimmen. Dazu war eine der Zahlen von Eins bis Fünf mit folgenden Bedeutungen anzukreuzen:

1: Stimmt überhaupt nicht

2: Stimmt nicht

3: Neutral

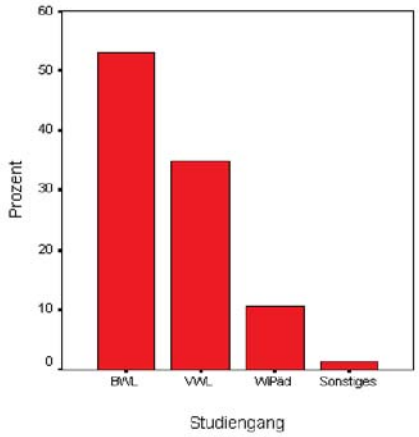
4: Stimmt

5: Stimmt vollkommen

Diese Antwortmöglichkeiten entsprechen einer klassischen, in Befragungen sehr häufig verwendeten Likert-Skala. Die Antworten auf die einzelnen Items sind somit ordinal skaliert, da zum einen eine Rangfolge gebildet werden kann, zum anderen jedoch die Abstände zwischen den einzelnen Ausprägungen nicht quantifizierbar sind (Vgl. Bleymüller 2002).
Zusätzlich wurden folgende Informationen mit den entsprechenden Antwortmöglichkeiten abgefragt:

- Studiengang: BWL, VWL, Wirtschaftspädagogik, Sonstiges
- Semester: freie Eingabe
- Geschlecht
- Alter: freie Eingabe
- Schulabschluss: Abitur in Ostdeutschland (inkl. Ost-Berlin), Abitur in Westdeutschland (inkl. West-Berlin), Abitur im Ausland, Sonstiges.

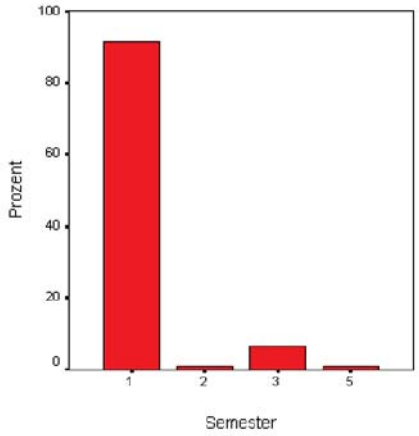
Abb. 2.1: Balkendiagramm: Studierende nach Studiengang



Die Verteilung der Studierenden nach ihren jeweiligen Studiengängen ist in der Abbildung 2.1 dargestellt. Etwas mehr als die Hälfte von ihnen studieren demnach BWL, etwas mehr als 30 Prozent VWL und weitere 10 Prozent Wirtschaftspädagogik. Diese Anteile stimmen in etwa mit den Angaben des Prüfungsbüros der Wirtschaftswissenschaftlichen Fakultät (ca. 50% BWL und 36% VWL) und der Studierendenstatistik der Humboldt-Universität (ca. 14% Wirtschaftspädagogik) überein. Bei den Befragten, die hier Sonstiges“ angekreuzt haben, könnte es sich z.B. um Austauschstudenten anderer Hochschulen handeln.

Die befragten Studierenden befanden sich erwartungsgemäß hauptsächlich im ersten Semester, wie aus der Abbildung 2.2 ersichtlich ist. Nur einzelne Anwesende studierten bereits im zweiten, dritten oder fünften Semester. Bei letzteren ist anzunehmen, dass sie in der Mathematik I bereits mindestens einmal die Prüfung nicht bestanden haben. Somit besteht die Möglichkeit, dass in dieser Gruppe eine Verzerrung vorliegt, da sie aus Studierenden besteht, welche tendenziell mehr Schwierigkeiten in Mathematik haben, was wiederum ihre Erwartungen an Statistik beeinflussen könnte. Des weiteren haben sie möglicherweise die Lehrveranstaltungen Statistik I und II bereits teilweise oder vollständig besucht, so dass ihre Antworten sich nicht mehr auf ihre Erwartungen *vor* dem Kursbesuch beziehen. Nach der Imputation der fehlenden Werte (s. Abschnitt 2.4), werden Studierende höherer Semester

Abb. 2.2: Balkendiagramm: Studierende nach Semester



daher aus den folgenden Analysen ausgeschlossen.
Die Abbildung 2.3 zeigt, dass weibliche Studierende mit einem Anteil von ca. 55% unter den Antwortenden etwas stärker vertreten sind. Ihr Anteil an den für das entsprechende Semester eingeschriebenen Studierenden in BWL und VWL beträgt hingegen lediglich 48%. Laut Sachbearbeiterin für Prüfung/Studienangelegenheiten des Instituts für Erziehungswissenschaften liegt der Frauenanteil unter den Wirtschaftspädagogik-Studierenden höher. Dennoch sind weibliche Studierende unter den Antwortenden überproportional stark vertreten. Möglicherweise besuchen diese zu einem größeren Anteil die Vorlesung oder waren eher bereit, an der Befragung teilzunehmen als ihre männlichen Kommilitonen.
Das Durchschnittsalter der Befragten Studierenden beträgt 21,77 Jahre. Wie aus Abbildung 2.4 ersichtlich ist, sind die männlichen Studierenden durchschnittlich etwas älter als ihre Kommilitoninnen (22,63 und 21,24 Jahre). Von dieser Differenz dürfte etwa ein Jahr auf den vor Studienbeginn absolvierten Wehr- oder Zivildienst männlicher Studierender zurückzuführen sein. Mehr als 90 Prozent der Befragten sind zwischen 18 und 25 Jahre alt, nur einzelne über 30.
Fast die Hälfte der befragten Studierenden hat ihr Abitur in den Neuen Bundesländern und Ost-Berlin erworben, wie aus der Abbildung 2.5 ersichtlich ist. Dies ist aufgrund der geographischen Lage Berlins nicht überraschend.

Abb. 2.3: Balkendiagramm: Studierende nach Geschlecht

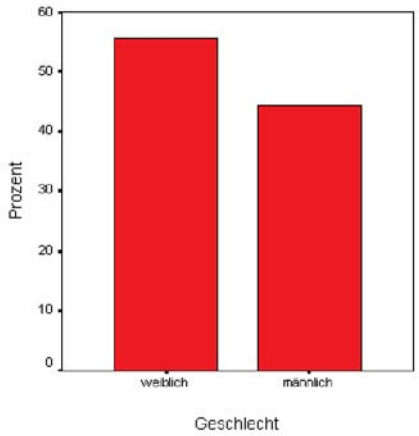


Abb. 2.4: Boxplots: Alter weiblicher und männlicher Studierender

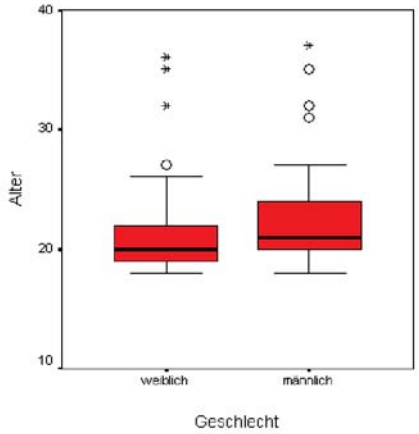
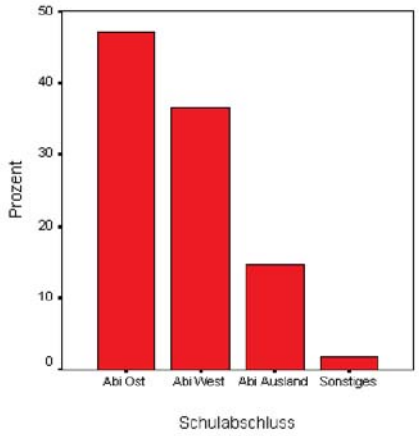


Abb. 2.5: Balkendiagramm Studierende nach Schulabschluss



Etwas mehr als ein Drittel der Antwortenden verfügt über ein Abitur aus den Alten Bundesländern oder West-Berlin und weitere ca. 14 Prozent haben ihr Abitur im Ausland abgelegt. Nur einzelne Studierende gaben als Antwort „Sonstiges“ an. Diese verfügen über einen anderen Schulabschluss als das Abitur.

Der gleiche Fragebogen, jedoch ohne persönliche Angaben, wurde auch für die Befragung im Privatissimum verwendet. Die Anwesenden, Studierende am Ende ihres Studiums und Dozenten, werden im Gegensatz zu den Studierenden im ersten Semester als *Experten* in Statistik betrachtet. Diese wurden gebeten anzugeben, wie ihrer Meinung nach Studierende antworten sollten, welche noch am Anfang ihres Studiums stehen und dieses mit Erfolg abschließen wollen bzw. eine wisschenschaftliche Karriere anstreben.

2.3 Maryland Physics Expectations Survey

Der für die Befragung verwendete Fragebogen wurde nach dem Vorbild des Maryland Physics Expectation Survey (MPEX) erstellt. Redish et al. (1998) nehmen an, dass die Erwartungen von Studierenden an eine Wissenschaft bestimmen, was sie letztendlich aus einer Lehrveranstaltung „mitnehmen“. Der Begriff *Erwartungen* der Studierenden umfasst in diesem Zusammenhang u.a. ihre Vorstellungen, Annahmen und Meinungen darüber, was sie

lernen werden, welche Fähigkeiten dafür nötig sind und welche Anstrengungen von ihnen erwartet werden. Um verschiedene Aspekte der Erwartungen von Physikstudenten zu untersuchen, wurde der MPEX Survey konzipiert. Dieser besteht aus 34 Items, die anhand der eigenen Lehrerfahrung der Autoren, Diskussionen mit anderen Experten, sowie dem Studium verwandter Literatur ausgewählt wurden. Redish et al. (1998) konzentrieren sich dabei auf die folgenden sechs Dimensionen, welche durch die in Klammern angegebenen Items aus der Liste im Anhang B gemessen werden sollen.

- **Unabhängigkeit:** Gehen Studierende davon aus, dass ihnen Physikkenntnisse „mundgerecht“ dargereicht werden oder, dass sie diese durch eigene Arbeit und Verständnis selbst erwerben müssen?
  (8, 13, 14, 17, 27)

- **Kohärenz:** Sehen Studierende Physik als eine Sammlung isolierter Teilbereiche und Fakten oder als ein kohärentes System?
  (12, 15, 16, 21, 29)

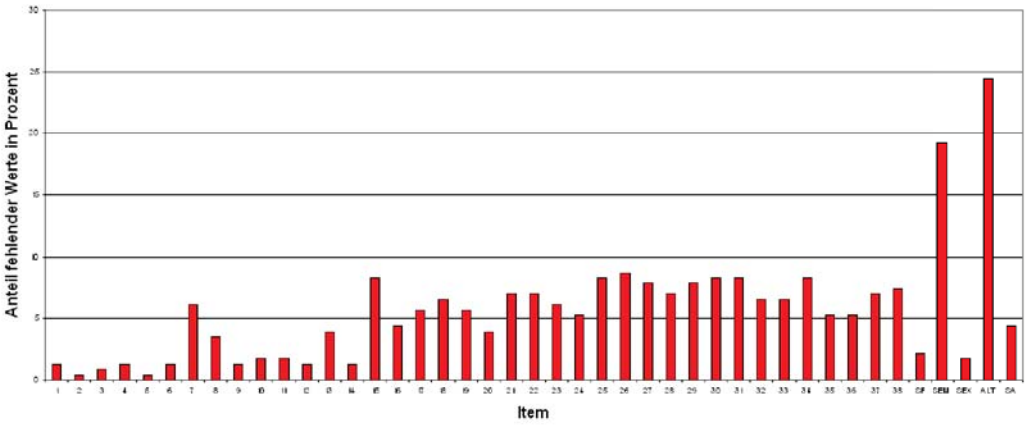
- **Konzepte:** Verstehen Studierende Physik als Formeln selbst oder als Konzepte, die diesen zugrunde liegen?
  (4, 14, 19, 23, 26, 27)

- **Realitätsbezug:** Sehen Studierende eine Verbindung zwischen den Lehrinhalten und der Außenwelt?
  (10, 18, 22, 25)

- **Mathematikbezug:** Benutzen Studierende mathematische Methoden lediglich, um Ergebnisse zu berechnen oder um physikalische Phänomene darzustellen?
  (2, 8, 15, 16, 17, 20)

- **Anstrengung:** Streben Studierende einen möglichst geringen Aufwand an oder sind sie bereit, Informationen sorgfältig zu durchdenken und zu bewerten?
  (3, 6, 7, 24, 31)

Abb. 2.6: Fehlende Werte pro Item



Um den Survey auf die in dieser Arbeit diskutierte Fragestellung anzuwenden, wurden die Items ins Deutsche übersetzt und falls erforderlich auf das Fachgebiet Statistik an Stelle von Physik übertragen. Des weiteren wurden einige der Items umformuliert und weitere hinzugefügt, um z.B. auf Besonderheiten wie die in Statistik verwendeten Formel- und Aufgabensammlungen einzugehen.

2.4 Fehlende Werte

Fehlende Werte treten in nahezu allen empirischen Datensätzen auf. Dies kann verschiedene Ursachen haben. So können z.B. die Teilnehmer einer Befragung nur selten dazu verpflichtet werden, die ihnen gestellten Fragen vollständig zu beantworten.

Bei der für diese Arbeit durchgeführten Befragung wurden die Studierenden außerdem gebeten, die Fragen, die sie nicht verstehen, unbeantwortet zu lassen. Einige unter ihnen standen zudem unter einem gewissen Zeitdruck, da im Anschluss eine andere Lehrveranstaltung stattfand, und konnten daher möglicherweise nicht alle Fragen beantworten. Die dadurch entstandenen fehlenden Werte wurden im Datensatz zunächst als „0“ codiert. Die Abbildung 2.6 zeigt den prozentualen Anteil fehlender Werte pro Item.

Den größten Anteil fehlender Werte weist mit 24,45% die Frage nach dem

Alter auf. Dies ist nicht überraschend, da diese Frage allgemein als indiskret gilt und die Studierenden durch diese Angabe möglicherweise auch ihre Anonymität gefährdet sehen. Auch die Frage nach dem Semester enthält mit 19,21% relativ viele fehlende Werte. Möglicherweise war diese Frage etwas missverständlich formuliert: Einige Studierende antworteten z.B. mit „WS 07/08“ anstatt mit der Ziffer des Semesters in dem sie derzeit studieren (also „1“, „2“ usw.). Ein weiterer Grund könnte, insbesondere für Studierende höherer Semester, wiederum die Sorge um die eigene Anonymität sein.
Des weiteren weisen die Items 15, 25, 26, 30, 31 und 34 (s. Anhang A) Anteile fehlender Werte von über 8% auf, scheinen jedoch inhaltlich in keinem deutlichen Zusammenhang zu stehen.
In der Abbildung 2.6 ist jedoch ersichtlich, dass die Anteile fehlender Werte am Anfang niedriger sind. Die Befragten mussten ihren Bogen nach der 16. Frage umblättern und waren dann möglicherweise durch den Anblick weiterer 22 Fragen demotiviert.
Bevor mit der eigentlichen Analyse des Datensatzes begonnen werden kann, ist zu entscheiden, wie mit den fehlenden Werten verfahren werden soll. Graham und Schafer (2002, S. 147) weisen eindringlich auf die Bedeutung dieser Problematik hin: „. . . ad hoc edits may do more harm than good, producing answers that are biased, inefficient (lacking in power), and unreliable.“
Von den in Frage kommenden Möglichkeiten zur Behandlung der fehlenden Werte werden an dieser Stelle einige kurz diskutiert.

Ausschluss

Eine einfache Möglichkeit zum Umgang mit fehlenden Werten wäre der Ausschluss derjenigen Items oder Fälle, in denen diese auftreten. Aus der Abbildung 2.6 ist jedoch ersichtlich, dass in diesem Fall bei allen Fragen Werte fehlen und ihr Ausschluss daher nicht in Frage kommt. Auch der Ausschluss der betroffenen Beobachtungen würde den Datensatz drastisch auf ca. die Hälfte (121 statt 229 Beobachtungen) reduzieren. Um nicht unnötig viele, auch in unvollständigen Items und Fällen vorhandene, Informationen zu verlieren, wird hier von ihrem Ausschluss abgesehen.

Beibehalten

Eine weitere Möglichkeit ist, die fehlenden Werte als solche zu definieren und im Datensatz zu belassen. Da die durchzuführenden Analysen jedoch nur gültige Werte berücksichtigen können, werden sie möglicherweise auf unterschiedlichen Fallzahlen beruhen und ihre Interpretierbarkeit und Vergleichbarkeit beeinträchtigt.
Für die Items Studiengang, Geschlecht, Alter und Schulabschluss wurde hier eine eigene Kategorie für fehlende Werte definiert. So steht eine „0“ im Datensatz für „keine Angabe“.
Für die Items 1 bis 38, die bei dieser Arbeit im Zentrum des Interesses stehen, ist diese Lösung jedoch nicht zufriedenstellend. Eine zusätzliche Kategorie „keine Angabe“ würde bedeuten, dass das Skalenniveau nicht mehr ordinal, sondern nur noch nominal ist. Dadurch stünden weniger statistische Methoden zur Analyse zur Verfügung und es ginge wiederum Information verloren.
Des weiteren werden für diese Arbeit nur die Erwartungen derjenigen Studierenden betrachtet, die noch keine Statistik-Lehrveranstaltung besucht haben. Daher wird das Item Semester benötigt, um Studierende aus höheren Semestern aus der Analyse auszuschließen. Für die klare Trennung sollten hier fehlende Werte möglichst sinnvoll ersetzt werden.

Imputation

Das Ersetzen fehlender durch gültige Werte wird allgemein als Imputation bezeichnet. Darunter fallen wiederum eine Reihe verschiedener Verfahren.

- **Imputation durch Modus**
  Ein weit verbreitetes und einfaches Verfahren zum Umgang mit fehlenden Werten bei ordinal skalierten Variablen ist es, diese durch die jeweiligen Modi zu ersetzen. Dabei ist sichergestellt, dass ein gültiger Wert eingesetzt, das ordinale Skalenniveau beibehalten und zumindest der Modus nicht verändert wird. Allerdings wird damit für alle fehlenden Werte des jeweiligen Items der gleiche Wert eingesetzt, was die Konzentration der Antworten sowie die Beziehungen zwischen den Items

(gemessen z.B. wie im Abschnitt 3.1 beschrieben) verzerrt.

- **ML-Methode**
  Die Maximum Likelihood (ML) Methode gilt aktuell laut Schafer und Graham (2002) neben der multiplen Imputation als *state-of-the-art* im Umgang mit fehlenden Werten. Eine effiziente Möglichkeit, das ML-Schätzproblem iterativ zu lösen, ist der Expectation-Maximization Algorithmus (Vgl. Schafer & Graham 2002 und Borman 2004). Dieser besteht aus den folgenden zwei Schritten:

  E-Schritt: Schätzung der fehlenden Werte für die gegebenen beobachteten Daten und derzeitgen Parameterschätzungen

  M-Schritt: Maximierung der Likelihood-Funktion unter Verwendung der im **E**-Schritt geschätzten anstelle der fehlenden Werte.

  Die Schritte werden abwechselnd wiederholt, bis die Parameterschätzungen konvergieren.

  Dieses Verfahren wurde gewählt, um die fehlenden Werte in den Items 1 bis 38 und Semester zu imputieren. Durchgeführt wurde die Imputation in der freien Software R mit einem von Herrn Dr. Klinke[3] zur Verfügung gestellten Programm, welches insbesondere das Paket `cat` verwendet.

In den Antworten der Experten traten keine fehlenden Werte auf.

[3]Institut für Statistik und Ökonometrie, Wirtschaftswissenschaftliche Fakultät und Abteilung Wirtschaftspädagogik, Institut für Erziehungswissenschaften, Philosophische Fakultät IV, Humboldt-Universität zu Berlin

3 Statistische Methoden

3.1 Zusammenhangsmaße

3.1.1 Korrelation

Ein Korrelationskoeffizient r ist allgemein als Maß für den linearen Zusammenhang zwischen zwei Variablen definiert und wie folgt zu interpretieren:

- $r = 0$: Es besteht kein linearer Zusammenhang.
- $r = \pm 1$: Es besteht ein perfekter (positiver oder negativer) linearer Zusammenhang zwischen den Variablen.

Ausgehend von der inhaltlichen Bedeutung der Variablen kann wiederum der Zusammenhang interpretiert werden. Entweder beeinflusst eine der Variablen die andere oder beide werden von derselben latenten Variable beeinflusst. Letzteres wird z.B. im Rahmen der Faktorenanalyse unterstellt (Vgl. Backhaus 2006). Je nach Skalenniveau der Daten wird der Korrelationskoeffizient auf unterschiedliche Weise berechnet. Für ordinal skalierte Variablen, die als Ausprägungen latenter stetiger Variablen angesehen werden können, ist laut Uebersax (2006) die polychorische Korrelation (s. Abschnitt 3.3.2) zu bevorzugen.
Zur Prüfung der Signifikanz der Korrelation und damit des linearen Zusammenhangs wird der Quotient aus r und dem Standardfehler gebildet und mit dem entsprechenden kritischen Wert der Standardnormalverteilung verglichen.
Im Vorfeld der explorativen Faktorenanalyse (s. Abschnitt 3.3) kann die Korrelationsmatrix bereits Hinweise darauf geben, welche (hochkorrelierten) Items einem Faktor zugeordnet werden.

3.1.2 Unsicherheitskoeffizient

Der Unsicherheitskoeffizient UC ist ein Maß für den Zusammenhang zwischen zwei nominal oder höher skalierten Variablen. Ist Y die abhängige und X die unabhängige Variable, so wird der Koeffizient berechnet als:

$$UC_{Y|X} = \frac{U_X + U_Y + U_{XY}}{U_Y}. \tag{3.1}$$

Dabei sind

- $U_X = -\sum_{j=1}^{J} f_{j+} ln f_{j+}$ die Entropie der Variable X,
- $U_Y = -\sum_{k=1}^{K} f_{+k} ln f_{+k}$ die Entropie der Variable Y und
- $U_{XY} = -\sum_{j=1}^{J} \sum_{k=1}^{K} f_{jk} ln f_{jk}$ die Entropie beider Variablen X und Y.

Der Unsicherheitskoeffizient liegt zwischen Null und Eins und gibt an, um wieviel Prozent der Vorhersagefehler der abhängigen Variable durch die Kenntnis der Ausprägung der unabhängigen Variable reduziert werden kann (Vgl. Rönz 2000).

3.1.3 Eta-Koeffizient

Der Eta-Koeffizient η misst den Zusammenhang zwischen einer metrisch skalierten Variable Y und einer nominal oder ordinal skalierten Variable X und wird wie folgt berechnet:

$$\eta = \sqrt{\frac{E_1 - E_2}{E_1}}. \tag{3.2}$$

Dabei sind

- $E_1 = \sum_{k=1}^{K} (y_k - \bar{y})^2 h_{+k}$ die Summe der quadratischen Abweichungen der K beobachteten Werte y_k vom arithmetischen Mittel $\bar{y}$ und
- $E_2 = \sum_{j=1}^{J} [\sum_{k=1}^{K} (y_{jk} - \bar{y}_j)^2 h_{jk}]$ die Summe der quadratischen Abweichungen innerhalb aller J durch die Ausprägungen x_j bestimmten Gruppen.

η kann Werte zwischen Null und Eins annehmen, wobei höhere Werte für einen stärkeren Zusammenhang sprechen. η^2 entspricht der prozentualen Reduktion des Fehlers in der Vorhersage der abhängigen Variable, welche durch die Kenntnis der Ausprägung der unabhängigen Variable erreicht wird (Vgl. Rönz 2000).

3.2 Parametervergleiche

3.2.1 Mann-Whitney-Test

Der Mann-Whitney-Test ist ein parameterfreier Test zur Überprüfung der Übereinstimmung der Verteilungen zweier Stichproben. Die Voraussetzungen für seine Anwendung sind

- unabhängige Stichproben und
- mindestens ordinal skalierte Daten.

Die Nullhypothese H_0 besagt, dass beide Stichproben aus der gleichen Grundgesamtheit stammen. In die Prüfung gehen dabei nur die Rangwerte, der Beobachtungen ein. Die Teststatistik U wird berechnet als:

$$U_i = n_1 n_2 + \frac{n_i(n_i+1)}{2} - R_i, \quad i = 1,2 \quad U = min\{U_1, U_2\}. \tag{3.3}$$

Dabei sind

- n_i der Umfang der Stichprobe i, mit $i = 1,2$, $n_1 + n_2 = N$ und
- R_i die Rangsumme der Stichprobe i.

Für $n_i > 8$ ist U approximativ normalverteilt und wird standardisiert zu:

$$z_U = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (N+1)}{12}}}. \tag{3.4}$$

Wenn Ties (Bindungen) auftreten, wird den entsprechenden Beobachtungen der Mittelwert der Ränge zugeordnet, die sie hätten wenn sie ungleich wären

und aufeinander folgen würden und die Standardabweichung σ_u (der Nenner in Gleichung 3.4) wird folgendermaßen korrigiert:

$$\sigma_u = \sqrt{\frac{n_1 n_2}{N(N-1)} \times \left(\frac{N^3 - N}{12} - \sum_{j=1}^{g} \frac{t_j^3 - t_j}{12} \right)}, \tag{3.5}$$

wobei

- g die Anzahl der Stichproben mit Ties und
- t_j die Anzahl der verbundenen Ränge in der Stichprobe j

sind.
Die Nullhypothese wird verworfen, wenn z_U größer ist, als der entsprechende kritische z-Wert der Normalverteilung (Vgl. Daly et al. 1995).

3.2.2 Kruskal-Wallis-Test

Der Kruskal-Wallis-Test ist ein nichtparametrisches Verfahren zum Vergleich der Mittelwerte mehrerer Stichproben. Er kann unter den folgenden Voraussetzungen angewendet werden:

- unabhängige Stichproben,
- stetige Verteilung und
- mindestens ordinal skalierte Daten.

Es wird die Nullhypothese H_0, dass die Mittelwerte aller C Stichproben gleich groß sind: $\mu_1 = \mu_2 = \ldots = \mu_C$, gegen die Alternativhypothese, dass sich mindestens ein Paar von Mittelwerten signifikant unterscheidet, geprüft. Anstelle der Originalwerte der Daten werden wie auch beim Mann-Whitney-Test nur ihre Rangplätze berücksichtigt. Für den Fall, dass keine Ties vorkommen, berechnet sich die Teststatistik H als:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{C} \frac{R_i^2}{n_i} - 3(N+1), \tag{3.6}$$

wobei

- C die Anzahl der Stichproben bzw. Gruppen,
- n_i die Anzahl der Beobachtungen in der i-ten Gruppe,
- N die Anzahl aller Beobachtungen $N = n_1 + n_2 + ... + n_C$ und
- R_i die Summe der Ränge der i-ten Stichprobe

sind.
Falls Ties auftreten, wird diesen Werten wiederum der Mittelwert der entsprechenden Rangplätze zugeordnet und H wird dividiert durch

$$1 - \frac{\sum_{j=1}^{g}(t^3 - t)}{N^3 - N}$$

mit j und g wie für die Formel (3.5) beschrieben.
H ist unter der Nullhypothese approximativ χ^2-verteilt mit $(C - 1)$ Freiheitsgraden. Ist H größer als der entsprechende kritische χ^2-Wert, so wird H_0 verworfen (Vgl. Kruskal & Wallis 1952).

3.3 Faktorenanalyse

3.3.1 Explorative Faktorenanalyse allgemein

Die explorative Faktorenanalyse ist ein Verfahren der Dimensionsreduktion, dessen Ziel darin besteht, die Zusammenhänge zwischen, und latenten Strukturen hinter beobachteten Variablen aufzudecken. Im Idealfall können die m beobachteten Variablen $X_j (j = 1, \ldots, m)$ durch Q gemeinsame Faktoren wie folgt erklärt werden (Vgl. Härdle & Simar 2003 und Rönz 2000):

$$X_j = \sum_{q=1}^{Q} a_{jq} F_q, \quad j = 1, ..., m \tag{3.7}$$

bzw. in Matrizenschreibweise:

$$X = FA^\top. \tag{3.8}$$

In der Praxis ist es jedoch nahezu unmöglich, die Zusammenhangsstruktur zwischen den beobachteten Variablen allein durch gemeinsame Faktoren zu erklären. Die Gleichung (3.8) wird daher für jede Variable X_j um eine Störgröße

$$\varepsilon_j = e_j \cdot U_j \tag{3.9}$$

ergänzt, wobei

- U_j der Einzelrestfaktor, der nur auf die Variable X_j wirkt, und
- e_j die zugehörige Faktorladung

sind.
So ergibt sich ein verallgemeinertes Modell der Faktorenanalyse:

$$X = FA^\top + UE \tag{3.10}$$

wobei

- $X_{(n\times m)}$ die Matrix der (häufig standardisierten) Beobachtungswerte
- $F_{(n\times Q)}$ die Faktorwertematrix der gemeinsamen Faktoren
- $A^\top_{(Q\times m)}$ die transponierte Faktorladungsmatrix der gemeinsamen Faktoren
- $U_{(n\times m)}$ die Faktorwertematrix der Einzelrestfaktoren und
- $E_{(m\times m)}$ die Diagonalmatrix mit den Faktorladungen der Einzelrestfaktoren auf der Hauptdiagonalen

sind.
Für die Faktorenanalyse müssen die folgenden Voraussetzungen erfüllt sein.

1. Die Beobachtungswerte sind unabhängig voneinander.
2. Der Stichprobenumfang n ist möglichst groß und auf jeden Fall größer als die Anzahl der Variablen m.

3. Die Variablen sind metrisch skaliert.
4. Die Variablen sind (zumindest approximativ) normalverteilt.
5. Die gemeinsamen Faktoren wirken auf mindestens zwei Variablen.
6. Die gemeinsamen Faktoren sind unkorreliert.
7. Die Einzelrestfaktoren sind unkorreliert.
8. Die Einzelrestfaktoren sind unkorreliert mit den gemeinsamen Faktoren.

Die Gleichung (3.10) ist nicht lösbar. A, F, U und E sind unbekannt und müssen schrittweise geschätzt werden.
Zunächst wird die Korrelationsmatrix R der beobachteten Variablen X_j berechnet. Nach dem Fundamentaltheorem der Faktorenanalyse gilt:

$$R = AA^\top + EE. \tag{3.11}$$

Wird R um die Einzelrestfaktoren bereinigt, so ergibt sich die reproduzierte Korrelationsmatrix

$$R_h = AA^\top \tag{3.12}$$

mit den Kommunalitäten

$$h_j^2 = \sum_{q=1}^{Q} a_{iq}^2 \tag{3.13}$$

auf der Hauptdiagonalen.
Dabei entspricht a_{jq}^2 dem Beitrag des Faktors F_q zur Erklärung der Varianz der Variablen X_j und h_j^2 somit dem Beitrag aller gemeinsamen Faktoren.
Da für die Gleichung (3.12) unendlich viele Lösungen existieren, kann sie nur geschätzt werden. Hierfür kommen z.B. folgende Extraktionsverfahren in Frage, welche zu unterschiedlichen Ergebnissen führen können:

- Hauptkomponentenmethode

- Hauptachsenmethode
- ML-Methode (Maximum Likelihood)
- ULS-Methode (Unweighted Least Squares).

Die Faktoren ergeben sich als Lösung des Eigenwertproblems der reproduzierten Korrelationsmatrix R_h. Die einzelnen Eigenwerte λ_q geben die Erklärungsbeiträge der jeweiligen Faktoren an der Varianz aller Variablen $X_1, \ldots, X_m$ an:

$$\lambda_q = \sum_{j=1}^{m} \hat{a}_{jq}^2 \tag{3.14}$$

Die Koeffizienten a_{jq} der Matrix A entsprechen den Komponenten des $q$ten Eigenvektors.
Die Anzahl der Faktoren ist vor Durchführung der Analyse nicht bekannt. Im Rahmen der Faktorenanalyse ist daher festzulegen, wieviele gemeinsame Faktoren extrahiert werden sollen. Für diese Entscheidung können z.B. folgende Kriterien herangezogen werden:

- Freiheitsgrade
  Die Anzahl der Freiheitsgrade d für ein Modell mit p Variablen und k Faktoren ergibt sich nach Härdle & Simar (2003) als:

  $$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k). \tag{3.15}$$

  Bei einer negativen Anzahl von Freiheitsgraden existieren für das Modell unendlich viele Lösungen. Die Bedingung $d \geq 0$ bestimmt daher, wieviele Faktoren maximal extrahiert werden dürfen.

- Varianzprozentanteile
  Soll ein bestimmter Anteil (häufig 90%) der Gesamtvarianz der Variablen erklärt werden, extrahiert man die ersten Q Faktoren, so dass $\frac{\sum_{q=1}^{Q} \lambda_q}{m}$ = zu erklärender Anteil der Varianz.[4]

[4] Die Gesamtvarianz standardisierter Variablen entspricht ihrer Anzahl m.

- Ellbogen-Kriterium
  Die Eigenwerte der Korrelationsmatrix werden nach ihrer Größe geordnet in einem Diagramm (Vgl. Abbildung 4.2) dargestellt. Der letzte Punkt vor dem „Knick“ in der Kurve entspricht der Anzahl der zu extrahierenden Faktoren, da der Anteil der erklärten Varianz durch die Hinzunahme weiterer Faktoren nur noch in geringem Maße ansteigt.

- Kaiser-Kriterium
  Es werden nur die Faktoren berücksichtigt, deren Eigenwerte größer als 1 sind, da nur diese mehr Varianz erklären als eine ursprüngliche standardisierte Variable.

- Horn-Kriterium
  Im Rahmen der Parallelanalyse nach Horn (Vgl. Überla 1971) vergleicht man die Eigenwerte der Korrelationsmatrix der beobachteten Variablen mit den Mittelwerten der Eigenwerte von Korrelationsmatrizen normalverteilter Zufallsvariablen. In der graphischen Darstellung dieser Eigenwerteverläufe (Vgl. Abbildung 4.2) bestimmt die Anzahl der Eigenwerte der empirischen Korrelationen oberhalb derer von Zufallskorrelationen, wieviele Faktoren zu extrahieren sind.

Da die genannten Kriterien in der Regel unterschiedliche Anzahlen von Faktoren nahelegen, ist es oftmals nötig, mehrere Modelle zu schätzen. Die Ergebnisse können dann nach ihrer Interpretierbarkeit, der Reliabilität der Faktoren (s. Abschnitt 3.3.3) sowie den Teststatistiken der Konfirmatorischen Faktorenanalyse (s. Abschnitt 3.3.4) beurteilt werden.
Die unrotierten Lösungen sind oft schwer zu interpretieren. Durch die Rotation des Koordinatensystems um seinen Ursprung können die Ladungen der Faktoren erhöht werden ohne dabei ihre Zuordnung zu den Variablen zu verändern. Das Ziel ist, eine Einfachstruktur zu erreichen, in der auf jeden Faktor jeweils einige Variablen möglichst hoch und andere möglichst niedrig laden. Ein häufig verwendetes Verfahren ist die Varimax-Rotation: Das Koordinatensystem der Faktoren wird so gedreht, dass die Varianz der quadrierten Faktorladungen pro Faktor maximiert wird.

Die Interpretation der Ergebnisse einer Faktorenanalyse wird anhand der (ggf. rotierten) Faktorladungsmatrix vorgenommen. Zwischen Variablen, die besonders hohe absolute Ladungen für einen Faktor aufweisen (wenn möglich, werden hierfür nur Ladungen $\geq |0,5|$ berücksichtigt), sind inhaltliche Zusammenhänge zu suchen, welche den gemeinsamen Faktor beschreiben.
Weitere Interpretationsmöglichkeiten bietet die Analyse der Faktorwerte. Anhand dieser können die Beziehungen zwischen den Faktoren und einzelnen oder Gruppen von Personen untersucht werden. Nachdem die Faktorladungsmatrix A aus der Gleichung (3.10) durch die Faktorenextraktion bestimmt ist, kann auch die Faktorwertematrix F geschätzt werden. In Mplus erfolgt die Schätzung im Falle ordinal skalierter Beobachtungsdaten mittels der WLSMV-Methode (Weighted Least Squares, Mean and Variance adjusted). Das Ergebnis sind standardisierte Werte mit einem Mittelwert von 0 und einer Varianz von 1, die wie folgt zu interpretieren sind:

- Ein Faktorwert f_{iq} von Null bedeutet, dass die in dem Faktor F_q zusammengefassten Items bei der Person i durchschnittlich ausgeprägt sind.

- Ein positiver (negativer) Faktorwert f_{iq} bedeutet, dass der Faktor F_q bei der Person i im Vergleich zu den anderen Personen überdurchschnittlich (unterdurchschnittlich) stark ausgeprägt ist.

3.3.2 Explorative Faktorenanalyse für ordinalskalierte Daten

In der Praxis können häufig nur ordinal skalierte Daten beobachtet bzw. erhoben werden. In diesen Fällen ist die im vorangegangenen Abschnitt beschriebene Faktorenanalyse nicht ohne weiteres anwendbar. Es existieren jedoch verschiedene Verfahren, welche die Ordinalskalierung der Variablen berücksichtigen und somit die Durchführung einer Faktorenanalyse dennoch ermöglichen. Der u.a. in Mplus implementierte Underlying Variables Ansatz nimmt an, dass jeder beobachteten ordinal skalierten Variable x eine nicht beobachtbare stetige Variable x^* zugrunde liegt. Der Zusammenhang zwi-

schen x mit m möglichen Ausprägungen bzw. Kategorien c und x^* kann wie folgt beschrieben werden (Vgl. Curran & Flora 2004):

$$x = c \Longleftrightarrow \tau_{c-1} < x^* < \tau_c, \quad c = 1, \ldots, m \tag{3.16}$$

mit den $(m-1)$ Schwellenwerten

$$\tau_{i(0)} = -\infty, \quad \tau_{i(1)} < \tau_{i(2)} < \ldots < \tau_{i(m-1)}, \quad \tau_{i(m)} = +\infty. \tag{3.17}$$

Der beobachtete Wert der ordinal skalierten Variable x ändert sich also, wenn die zugrundeliegende stetige Variable x^* einen Schwellenwert τ überschreitet. Unter der Annahme, dass die Variablen x_1^* und x_2^* standardnormalverteilt sind, können die Schwellenwerte $a_i, i = 0, \ldots, s$ und $b_j, j = 0, \ldots, r$ der beobachteten ordinal skalierten Variablen x_1 und x_2 mit Hilfe der Inversen der univariaten Standardnormalverteilungsfunktion Φ_1^{-1} und den kumulierten beobachteten Randhäufigkeiten P_{i+} und P_{+j} der Kontingenztabelle von x_1 und x_2 geschätzt werden:

$$a_i = \Phi_1^{-1}(P_{i+}) \tag{3.18}$$

und

$$b_j = \Phi_1^{-1}(P_{+j}). \tag{3.19}$$

Die Wahrscheinlichkeit $\pi_{i,j}$, dass eine Beobachtung in einer bestimmten Zelle (i, j) der Kontingenztabelle von x_1 und x_2 liegt, beträgt:

$$\pi_{i,j} = \Phi_2(a_i, b_j) - \Phi_2(a_{i-1}, b_j) - \Phi_2(a_i, b_{j-1}) + \Phi_2(a_{i-1}, b_{j-1}) \tag{3.20}$$

wobei Φ_2 die bivariate Standardnormalverteilungsfunktion ist.
Die polychorische Korrelation zwischen x_1 und x_2 wird schließlich als Schätzung derjenigen Korrelation bestimmt, die sich ergeben hätte wenn die zugrundeliegenden stetigen Variablen x_1^* und x_2^* direkt beobachtet worden wären, durch die Maximierung der log-Likelihood-Funktion

$$\ell = lnK + \sum_{i=1}^{s}\sum_{j=1}^{r} n_{i,j} ln\pi_{i,j}. \tag{3.21}$$

Dabei bezeichnet K eine Konstante und $n_{i,j}$ die beobachtete Zellhäufigkeit in der Zelle (i, j).
Mit der auf diese Weise ermittelten polychorischen Korrelationsmatrix kann eine Faktorenanalyse, wie im Abschnitt 3.2.1. beschrieben, durchgeführt werden.

3.3.3 Reliabilität

Im Rahmen der Reliabilitätsanalyse wird überprüft, inwieweit m beobachtete Items dazu geeignet sind, eine latente Variable abzubilden. Anhand der Ergebnisse können z.B. in der explorativen Faktorenanalyse ermittelte Faktoren auf ihre innere Konsistenz hin untersucht werden.
Eine Maßzahl der Reliabilität ist das Cronbach's Alpha (Vgl. Rönz 2000):

$$A = \frac{m}{m-1}\left(1 - \frac{\sum_{j=1}^{m}\sigma_j^2}{\sigma_y^2}\right). \tag{3.22}$$

σ_j^2 und σ_Y^2 stehen für die Varianz des j-ten Items bzw. der latenten Variable Y. A kann Werte zwischen Null und Eins annehmen. Je näher der Wert bei Eins liegt, desto zuverlässiger beschreiben die einbezogenen Items gemeinsam die latente Variable. Per Daumenregel sollte $A \geq 0,7$ gelten.

3.3.4 Konfirmatorische Faktorenanalyse

Die konfirmatorische Faktorenanalyse überprüft die Übereinstimmung eines theoretischen Modells mit beobachteten Daten. Das Modell ergibt sich z.B. aus den Ergebnissen einer zuvor durchgeführten explorativen Faktorenanalyse.
Jedes einbezogene beobachtete Item X_i lädt nun ausschließlich auf einen Faktor ξ_i und wird somit nur durch diesen, sowie durch den Fehlerterm δ_i, beeinflusst. Der Zusammenhang zwischen den beobachteten Variablen X_i und den Faktoren ξ_i wird zusammengefasst als:

$$X = \Lambda\xi + \delta \tag{3.23}$$

wobei

- X die beobachtete Datenmatrix
- Λ die Faktorladungsmatrix
- ξ die Faktorwertematrix und
- δ die Matrix der Fehlerterme

sind.
Das Einsetzen der Gleichung (3.23) in die empirische Kovarianzmatrix ergibt die theoretische Kovarianzmatrix $\Sigma(\theta)$

$$\Sigma(\theta) = \Lambda\Xi\Lambda^\top + \Theta \tag{3.24}$$

wobei

- Ξ die Kovarianzmatrix der Faktoren und
- Θ die Kovarianzmatrix der Fehlerterme

sind (Vgl. Schermelleh-Engel 1999).
Die Parameter θ werden in Mplus iterativ mit der WLSMV-Methode (Weighted Least Squares, Mean and Variance adjusted) geschätzt. Der Quotient aus der Parameterschätzung und dem jeweiligen Standardfehler wird zur Prüfung der Nullhypothese: $\theta = 0$ als Teststatistik benutzt. Ist der absolute Wert des Quotienten größer als der kritische Wert der Standardnormalverteilung (z.B. 1,96 für $\alpha = 0,05$), so wird die Nullhypothese verworfen und das jeweilige Item leistet einen signifikanten Beitrag zu dem Faktor.
Die Modellgüte kann u.a. anhand folgender Kriterien beurteilt werden:

- **χ^2-Anpassungstest**
  Der χ^2-Test prüft die Hypothese, dass die Differenz zwischen der theoretischen und der empirischen Kovarianzmatrix gering ist, d.h. das Modell die Datenstruktur gut beschreibt. Kleine Werte der Teststatistik

χ^2 sprechen für einen guten Modell-Fit. Der Test ist jedoch empfindlich gegenüber der Stichprobengröße und Abweichungen von der Normalverteilung.

- **CFI**
  Der Comparative Fit Index vergleicht das getestete Modell mit einem Nullmodell mit unkorrelierten Faktoren und wird nach folgender Formel berechnet:

$$CFI = 1 - \frac{max[(\chi_1^2 - df_1), 0]}{max[(\chi_0^2 - df_0), (\chi_1^2 - df_1), 0]}. \tag{3.25}$$

  Dabei sind

  - χ_1^2 und χ_0^2 die χ^2-Werte des getesteten und des Nullmodells und
  - df_1 und df_0 die Anzahl der Freiheitsgrade des getesteten und des Nullmodells.

  Der CFI ist relativ robust gegenüber der Stichprobengröße und Verletzungen der Verteilungsannahme. Der Index kann Werte zwischen 0 und 1 annehmen, wobei höhere Werte für einen besseren Modell-Fit sprechen. Ab einem CFI von 0,95 gilt das Modell als gut.

- **TLI**
  Auch der Tucker Lewis Index basiert auf dem Vergleich des getesteten Modells mit einem Nullmodell und wird wie folgt berechnet:

$$TLI = \frac{\chi_0^2/df_0 - \chi_1^2/df_1}{\chi_0^2/df_0 - 1}. \tag{3.26}$$

  Der TLI ist wiederum relativ unempfindlich gegenüber der Stichprobengröße und Verletzungen der Verteilungsannahme. Werte ab 0,95 sprechen für einen guten Modell-Fit.

- **RMSEA**
  Der Root Mean Square Error of Approximation ist ein weiteres Maß zur Beurteilung der Güte eines Modells, vergleicht dieses jedoch nicht

mit einem Nullmodell. Der RMSEA wird berechnet als:

$$RMSEA = \sqrt{\frac{\chi^2/df - 1}{N - 1}}. \tag{3.27}$$

Auch dieser Index ist relativ unempfindlich gegenüber der Stichprobengröße N. Er nimmt Werte zwischen 0 und 1 an. Ein Modell wird als gut angesehen, wenn der RMSEA nicht größer als 0,05 ist.

Ein guter Modell-Fit sollte durch möglichst viele der Kriterien belegt werden.

3.4 Clusteranalyse

Die Clusteranalyse dient dazu, Objekte (z.B. befragte Personen) in Gruppen (Cluster) einzuteilen. Die Objekte innerhalb einer Gruppe sollen sich dabei hinsichtlich mehrerer Merkmale möglichst wenig, die einzelnen Gruppen hingegen möglichst stark voneinander unterscheiden. Der Ausgangspunkt der Analyse ist eine Datenmatrix $X(n \times m)$ mit n Objekten und m Variablen. Um diese in Gruppen einzuteilen, stehen verschiedene Ansätze zur Auswahl. Sehr gebräuchlich sind hierarchisch-agglomerative Algorithmen, deren Ablauf Härdle und Simar (2003) wie folgt beschreiben:

1. Bilde die feinstmögliche Clusterung, in der jedes der n Objekte einem Cluster entspricht.

2. Berechne die Distanzmatrix $D(n \times n)$, welche die Distanzmaße für alle Paare von Objekten bzw. Clustern enthält.

3. Finde die zwei „ähnlichsten“ Cluster, d.h. diejenigen mit der geringsten Distanz.

4. Fasse diese zwei Cluster in einem neuen Cluster zusammen.

5. Berechne die Distanzmaße zwischen den neuen Clustern und erhalte die reduzierte Distanzmatrix D.

Wiederhole die Schritte 3 bis 5 solange, bis alle Cluster in nur noch einem Cluster zusammengefasst sind.

Als Maß für die Unähnlichkeit oder Distanz von Objekten in Schritt 2 wird in diesem Zusammenhang z.B. die quadrierte euklidische Distanz verwendet. Die Unähnlichkeit zwischen den Objekten i und j hinsichtlich der m berücksichtigten gleichskalierten oder standardisierten Variablen berechnet sich danach als:

$$d_{ij}^2 = \sum_{k=1}^{m} (x_{ik} - x_{jk})^2. \tag{3.28}$$

Nach dem ersten Durchlauf des Algorithmus werden nicht mehr die Abstände zwischen einzelnen Objekten, sondern allgemeiner zwischen Objekten und Gruppen bzw. zwischen verschiedenen Gruppen anhand eines erweiterten Distanzmaßes verglichen (Vgl. Wiedenbeck & Züll 2001). Dies kann je nach Agglomerationsverfahren auf verschiedene Weise erfolgen. Bei der Ward'schen Methode werden die Distanzen der Objekte zum Variablenmittelwert ihres Clusters R berechnet und aufsummiert:

$$I_R = \frac{1}{n_R} \sum_{i=1}^{n_R} d^2(x_i, \bar{x}_R), \tag{3.29}$$

um die Trägheit I_R als Maß für die Heterogenität des Clusters R zu berechnen. Wenn zwei Objekte oder Gruppen P und Q zusammengefasst werden, wird das dadurch neu gebildete Cluster ein um $\Delta(P, Q)$ größeres I_{P+Q} aufweisen. Es sind daher diejenigen Objekte P und Q zusammenzufassen, bei denen $\Delta(P, Q)$ und damit die Zunahme an Heterogenität so gering wie möglich ist (Vgl. Härdle & Simar 2003).

Die Lösung des Clusterproblems entspricht einer der Zwischenstufen auf dem Weg von einem Extrem (n Cluster) zum anderen (ein Cluster). Üblicherweise ist besonders diejenige Stufe von Interesse, nach der die Fusionswerte (Abstände zwischen den vereinigten Clusterpaaren) sprunghaft ansteigen, da damit auch die Heterogenität innerhalb der Gruppen ebenso zunimmt.

Die Anzahl der Gruppen nach der i-ten Stufe beträgt $(n - i)$.

Die ausgewählte Clusterlösung kann anhand der mittleren Ausprägung der zur Bestimmung der Ähnlichkeit einbezogenen Merkmale interpretiert werden.
Von Interesse ist weiterhin, ob einzelne Cluster durch Objekte mit bestimmten Merkmalen, wie z.B. dem Geschlecht, dominiert werden.

4 Analysen

4.1 Vergleich verschiedener Gruppen

Eine interessante Fragestellung ist, ob und in welchem Ausmaß die Erwartungen der Studierenden insgesamt bzw. bestimmter Gruppen mit den von Experten präferierten Einstellungen übereinstimmen. Der einfache Vergleich der Modi (s. Tabelle 4.1) ist in diesem Zusammenhang jedoch nur eingeschränkt aussagekräftig, da u.a. nicht ersichtlich ist, ob deren Unterschiede z.B. auf nur einem einzigen Fall oder auf sehr vielen Fällen beruhen.
Um signifikante Unterschiede aufzudecken, werden daher Mann-Whitney-Tests durchgeführt. Im Vergleich der Antworten aller Studierenden mit denen der Experten, treten bei 24 der 38 Items signifikante Unterschiede auf. Um zu untersuchen, welche Studierenden hauptsächlich für diese Unterschiede verantwortlich sind und somit in ihren Erwartungen am meisten von den Experten abweichen, werden für einzelne Gruppen weitere Tests durchgeführt:

- **Studiengang**
  Der Anteil der von den Experten abweichenden Antworten unter den BWL-Studierenden ist mit 23 Items am höchsten (Wirtschaftspädagogik: 19, VWL: 17).

- **Geschlecht**
  Weibliche Studierende stimmen in 23, männliche Studierende in 20 Items nicht mit den Experten überein.

- **Ort des Schulabschlusses**
  Der größte Anteil von Abweichungen ergibt sich mit 24 Items unter den Studierenden, die ihr Abitur in Westdeutschland abgelegt haben (Ostdeutschland: 21, Ausland: 16).

Tab. 4.1: Modi der Antworten von Studierenden und Experten

Item	Modus Studierende	Modus Experten
1	stimmt	stimmt vollkommen
2	stimmt	vollkommen stimmt
3	stimmt	stimmt vollkommen
4	stimmt vollkommen	stimmt überhaupt nicht
5	stimmt	stimmt vollkommen
6	stimmt	vollkommen stimmt
7	neutral	stimmt nicht
8	neutral	stimmt
9	stimmt	stimmt
10	stimmt	stimmt
11	neutral	stimmt überhaupt nicht
12	neutral	neutral
13	neutral	stimmt nicht
14	neutral	stimmt
15	neutral	stimmt nicht
16	neutral	stimmt nicht
17	neutral	stimmt nicht
18	neutral	stimmt nicht
19	neutral	stimmt nicht
20	neutral	stimmt nicht
21	neutral	stimmt vollkommen
22	stimmt	neutral
23	neutral	stimmt nicht
24	stimmt	stimmt
25	neutral	neutral
26	stimmt	neutral
27	neutral	stimmt nicht
28	neutral	stimmt
29	stimmt	stimmt vollkommen
30	neutral	stimmt
31	neutral	neutral
32	stimmt	stimmt nicht
33	neutral	stimmt nicht
34	stimmt	stimmt
35	stimmt	stimmt
36	stimmt	stimmt
37	neutral	stimmt nicht
38	neutral	stimmt vollkommen

Tab. 4.2: Signifikante Abweichungen von Antworten der Experten

Item	ges.	BWL	VWL	WiPäd	weibl.	männl.	Ost	West	Ausl.	≤ 21	> 21
1	×	×	×		×	×		×	×	×	
2											
3											
4	×	×	×	×	×	×	×	×	×	×	×
5	×		×			×		×	×		
6											
7	×	×	×	×	×	×	×	×	×	×	×
8	×	×			×		×	×			×
9											
10			×			×		×		×	
11	×	×	×	×	×	×	×	×	×	×	×
12				×							
13	×	×		×	×		×	×		×	
14				×							
15	×	×	×	×	×	×	×	×	×	×	×
16	×	×	×	×	×	×	×	×	×	×	×
17	×	×	×	×	×	×	×	×		×	×
18											×
19	×	×		×	×	×	×	×			×
20											
21	×	×	×	×	×	×	×	×		×	×
22	×	×	×	×	×	×	×	×	×	×	×
23	×	×		×	×	×	×		×	×	×
24											
25	×	×	×	×	×		×	×	×	×	×
26											
27	×	×	×	×	×	×	×	×	×	×	×
28	×	×			×		×	×			
29	×	×	×		×	×	×	×		×	×
30	×	×			×		×	×			×
31	×	×	×	×	×	×	×	×	×	×	×
32	×	×		×	×	×	×	×	×	×	×
33											
34	×	×			×			×	×	×	
35											
36						×					
37	×	×	×	×	×	×	×	×	×	×	×
38	×	×	×	×	×	×	×	×	×	×	×

- **Alter**
  Sowohl Studierende im Alter bis einschließlich 21 Jahre, als auch ältere antworten in jeweils 20 Items abweichend von den Experten.

Die Ergebnisse der Mann-Whitney-Tests sind in der Tabelle 4.2 zusammengefasst: Kreuze symbolisieren auf dem 5%-Niveau signifikante Unterschiede in den Antwortverteilungen.
Interessanterweise gibt es bei einigen Items (z.B. 2 und 9) keine signifikanten Unterschiede in den Antworten. Bei anderen Items hingegen weichen alle Gruppen von Studierenden von der Expertenantwort ab. Das extremste Beispiel ist das vierte Item, dessen Modi von Studierenden und Experten die völlig entgegengesetzten Antworten „stimmt vollkommen“ und „stimmt überhaupt nicht“ annehmen.
Insgesamt fällt auf, dass die Studierenden im Gegensatz zu den Experten sehr häufig mit „neutral“ geantwortet haben. Möglicherweise waren viele der

Aussagen zu spezifisch, als dass Studierende *vor* dem Besuch der Lehrveranstaltungen bereits einschätzen könnten, ob ihnen zuzustimmen sei oder nicht.

4.2 Zusammenhänge zwischen den Items

Da die Korrelationsmatrix der Items mit 38×38 Feldern sehr umfangreich und somit schwer überschaubar ist, wird diese in der Abbildung 4.1 nur schematisch dargestellt. Blau ausgefüllte Felder kennzeichnen Korrelationen, die auf dem 5% Niveau signifikant verschieden von Null sind. Es wird deutlich, dass zahlreiche Items signifikant miteinander korreliert und somit nicht unabhängig voneinander sind. Die Struktur dieser Zusammenhänge kann anhand einer Faktorenanalyse genauer analysiert werden.

4.3 Explorative Faktorenanalyse

In diesem Abschnitt wird mittels einer explorativen Faktorenanalyse untersucht, welche Faktoren den Erwartungen der Studierenden an die Lehrveranstaltungen Statistik I und II zugrundeliegen. Die Analyse wird mit Mplus unter Verwendung des Underlying Variable Ansatzes und der ULS-Schätzung durchgeführt.

Zunächst ist zu entscheiden, wieviele Faktoren aus dem Datensatz mit 38 Items (siehe Anhang A) extrahiert werden sollen. Nach der Formel (3.15) dürfen maximal 29 Faktoren extrahiert werden, da ab einer Faktorenanzahl von 30 die Anzahl der Freiheitsgrade negativ wird.[5]

Wie aus der Tabelle 4.3 ersichtlich ist, sind das Kaiser- und das 90%-Varianz-Kriterium eher ungeeignet. Bezogen auf die Anzahl der Items führen sie zur Extraktion relativ vieler und damit schwer interpretierbarer Faktoren (Kaiser: 13 Faktoren, 90% Varianz: 25 Faktoren).

Weitere Entscheidungskriterien für die Anzahl der Faktoren sind in der Abbildung 4.2 zusammengefasst. Der für das Ellbogen-Kriterium relevante „Knick“

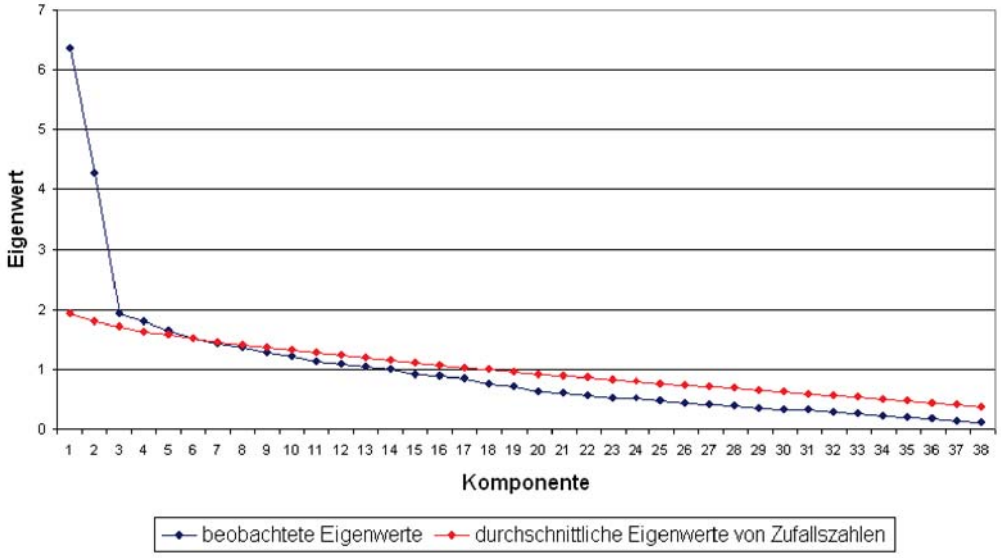
[5] Berücksichtigt man, dass beim Underlying Variables Ansatz zusätzlich auch die Schwellenwerte geschätzt werden, so liegt die Obergrenze noch etwas niedriger.

Abb. 4.1: Signifikante polychorische Korrelationen

Tab. 4.3: Eigenwerte der Korrelationsmatrix und kumulierte erklärte Varianz

Komponente	1	2	3	4	5	6	7	8	9	10
Eigenwert	6,34	4,27	1,92	1,80	1,64	1,51	1,43	1,37	1,28	1,21
erkl. Varianz	0,17	0,28	0,33	0,38	0,42	0,46	0,50	0,53	0,57	0,60
Komponente	11	12	13	14	15	16	17	18	19	20
Eigenwert	1,13	1,08	1,03	0,99	0,91	0,89	0,85	0,75	0,71	0,63
erkl. Varianz	0,63	0,66	0,68	0,71	0,73	0,76	0,78	0,80	0,82	0,84
Komponente	21	22	23	24	25	26	27	28	29	30
Eigenwert	0,62	0,56	0,53	0,51	0,47	0,44	0,40	0,40	0,35	0,32
erkl. Varianz	0,85	0,87	0,88	0,89	0,91	0,92	0,93	0,94	0,95	0,96
Komponente	31	32	33	34	35	36	37	38		
Eigenwert	0,32	0,27	0,26	0,23	0,19	0,17	0,14	0,10		
erkl. Varianz	0,96	0,97	0,98	0,98	0,99	0,99	1,00	1,00		

Abb. 4.2: Eigenwertediagramm



spricht für drei, die Parallelanalyse nach Horn, basierend auf den durchschnittlichen Eigenwerten der Korrelationsmatrizen von 20 verschiedenen mit XploRe erzeugten Zufallsdatensätzen, hingegen für fünf Faktoren.
Da die o.g. Kriterien nur Hinweise, jedoch keine verbindliche Aussage über die Anzahl der zu extrahierenden Faktoren liefern können, werden zunächst mehrere Faktormodelle mit verschiedenen Anzahlen von Faktoren berechnet. Nach dem Vergleich der Ergebnisse (insbesondere der Interpretierbarkeit und Reliabilität der jeweiligen Faktoren) und der Beratung durch Dr. Sigbert Klinke und Cornelia Wagner[6] wird ein 4-Faktoren-Modell ausgewählt, dass im folgenden genauer beschrieben wird.
Es wird zunächst ein 5-Faktoren-Modell berechnet. Die Varimax-rotierten Ladungen mit absoluten Werten größer als 0,4 sind in der Tabelle 4.4 dargestellt.
Um die einzelnen Faktoren interpretieren zu können, werden ihnen diejenigen Items zugeordnet, die diese besonders stark positiv (+) bzw. negativ (-) laden.

[6]Abteilung Wirtschaftspädagogik, Institut für Erziehungswissenschaften, Philosophische Fakultät IV, Humboldt-Universität zu Berlin

Tab. 4.4: Varimax-rotierte Ladungen >|0,4|

Item	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Faktor 5
1		0,427			
2		0,603			
3					
4					0,507
5		0,569	0,428		
6		0,600			
7					
8					
9			0,664		
10		0,413			
11	-0,650				0,479
12		0,468			
13	-0,701				
14					
15					0,529
16					0,402
17					
18					0,431
19	-0,503				
20					
21					
22					0,504
23					0,495
24					
25					
26					
27	-0,405				0,566
28	0,544				
29	0,548				
30	0,444			0,500	
31					
32					0,534
33					0,468
34				0,454	
35		0,454		0,406	
36				0,520	
37					
38				0,576	

Faktor 1: Nutzen von Statistik-Kenntnissen und Formeln

11. In Statistik erwarte ich nicht, die Formeln zu verstehen, sondern nehme sie als gegeben hin. (-)

13. Statistische Formeln haben wenig mit der realen Welt zu tun. (-)

19. Die Herleitung von Formeln in der VL/Ü oder im Skript hilft mir kaum, Aufgaben zu lösen bzw, die Prüfung zu bestehen. (-)

28. Die erworbenen Statistik-Kenntnisse werden mir in anderen Lehrveranstaltungen nützlich sein. (+)

29. Die erworbenen Statistik-Kenntnisse werden mir helfen, verschiedene Veröffentlichungen (Fachartikel, Tagespresse...) besser zu verstehen. (+)

Die fünf genannten Items beschreiben den Nutzen, den sich die Studierenden von Statistik-Kenntnissen und Formeln erhoffen: Werden letztere z.B. zum Selbstzweck verwendet (-) oder können sie auch außerhalb der Lehrveranstaltung gebraucht werden (+)? Es wird vermutet, dass dieser Faktor auch die Erwartungen der Studierenden an andere „formelintensive“ Fächer wie z.B. Mathematik oder Ökonometrie beeinflusst. Mit einem Cronbach's Alpha von 0,7279 gilt er als zuverlässig.

Faktor 2: Lernen durch Training

1. Um den Stoff zu verstehen, muss ich das Skript lesen. (+)

2. Um den Stoff zu verstehen, muss ich die Übungsaufgaben lösen. (+)

5. Um mich auf die Prüfung vorzubereiten, gehe ich meine Mitschriften aus der VL/Ü sorgfältig durch. (+)

6. Um mich auf die Prüfung vorzubereiten, löse ich die Aufgaben aus der Aufgabensammlung. (+)

10. Ich werde das Skript sorgfältig lesen und die Beispiele nachrechnen. (+)

12. Ich werde Statistik eher verstehen, wenn ich möglichst viele Aufgaben löse, als wenn ich nur wenige detailliert analysiere. (+)

35. Ich lerne aus meinen Fehlern in Übungs- und Prüfungsaufgaben. (+)

Die sieben Items beschreiben ebensoviele Aktivitäten, die Studierende sich für die Lehrveranstaltung vornehmen. Insgesamt entsteht der Eindruck, dass angestrebt wird, möglichst viel zu tun um „fitfür die Prüfung zu sein. Des weiteren spiegeln die Aussagen ein gewisses Vertrauen in bzw. eine Abhängigkeit von den zur Verfügung gestellten Materialien wie Skripten und Aufgabensammlungen wider. Der dadurch beschriebene Faktor kann somit als Lerntyp interpretiert werden, welcher wiederum auch für andere Fächer von Relevanz sein dürfte. Das Cronbach's Alpha liegt mit 0,6185 etwas niedrig, jedoch nur knapp unter den üblicherweise geforderten 0,7. Der Faktor wird daher dennoch für weitere Analysen verwendet.

Faktor 3

5. Um mich auf die Prüfung vorzubereiten, gehe ich meine Mitschriften aus der VL/Ü sorgfältig durch. (+)

9. Ich werde viel Zeit damit verbringen, die Herleitungen aus der VL/Ü oder dem Skript nachzuvollziehen. (+)

Nur zwei Items laden auf den dritten Faktor höher als |0,4|. Ein Faktor sollte jedoch *wenigstens* zwei Variablen umfassen. Mit einem Cronbach's Alpha für die Items 5 und 9 von nur 0,2768 ist dieser Faktor nicht reliabel. Er wird daher nicht interpretiert und in den weiteren Analysen nicht berücksichtigt. Das fünfte Item wird dem Faktor 2 zugeordnet, für den es die höchste Ladung aufweist.

Faktor 4: Verstehen von Konzepten

30. Bei den meisten Übungs- oder Prüfungsaufgaben, die ich löse, denke ich gezielt an die zugrundeliegenden Annahmen und Konzepte. (+)

34. Das Wichtigste, was ich in Statistik lerne, ist die Fähigkeit, statistische Zusammenhänge zu verstehen. (+)

36. Um eine Formel auf eine Aufgabe anzuwenden, muss ich mehr als nur die Bedeutung der in der Formel vorkommenden Terme kennen. (+)

38. Um Statistik zu verstehen, muss ich die Informationen aus der VL/Ü und dem Skript grundlegend durchdenken und gedanklich umstrukturieren und -organisieren. (+)

Die vier Items beschreiben wiederum Aktivitäten, die Studierende für erforderlich halten. Hierbei wird jedoch ein besonderer Wert auf das Verständnis gelegt. Der Faktor kann somit als ein anderer Lerntyp als der durch den Faktor 2 charakterisierte interpretiert werden: Während bei letzterem möglichst *viel* gerechnet und geübt wird, geschieht dies beim Faktor 4 möglichst *gründlich*. Das Cronbach's Alpha von 0,611 ist wiederum etwas kleiner als gefordert. Dennoch wird der Faktor im Weiteren beibehalten.

Faktor 5: Hinnehmen und Auswendiglernen

4. Alles, was ich durch die Herleitung einer Formel lerne, ist, dass diese Formel richtig ist und für die Lösung von Aufgaben benutzt werden kann. (+)

15. Die einzelnen Teilbereiche der Statistik haben wenig miteinander zu tun. (+)

16. Meine Note in Statistik hängt hauptsächlich davon ab, wie gut ich mir die Lehrinhalte merken kann, weniger von meinem Verständnis oder meiner Kreativität. (+)

18. Wenn beim Rechnen einer Statistik-Aufgabe mein Ergebnis deutlich von dem abweicht, was ich erwartet hätte, muss ich mich auf meine Berechnung verlassen. (+)

22. Das Wichtigste bei der Analyse einer statistischen Fragestellung ist, die richtige Formel zu finden. (+)

23. Wenn ich in der Prüfung für die Lösung einer Aufgabe eine bestimmte Formel brauche, die mir nicht mehr einfällt und nicht in der Formelsammlung enthalten ist, dann kann ich nichts (legales!) tun, um sie zu finden. (+)

27. Meine Prüfungsergebnisse helfen mir nicht mehr beim Verständnis des Stoffs. Die Veranstaltung und der damit verbundene Lernprozess sind für mich mit der Prüfung beendet. (+)

32. Lange an einer Aufgabe zu rechnen ist Zeitverschwendung. Wenn ich nicht vorankomme, frage ich lieber jemanden, der mehr davon versteht als ich. (+)

33. In Statistik ist es wichtig, alles was ich wissen muss, auswendig zu lernen. (+)

Die 9 Items betonen stark die Bedeutung des Hinnehmens und Auswendiglernens. Der Faktor 5 kann somit als ein weiterer Lerntyp beschrieben werden. Die Studierenden sehen Statistik als Pflichtfach, dass es mit möglichst geringem Aufwand (im Extremfall durch einfaches Hinnehmen und Auswendiglernen) „hinter sich zu bringen“ gilt. Das Cronbach's Alpha von 0,7452 spricht für die Reliabilität des Faktors.
Unerwarteterweise entsprechen die ermittelten Faktoren nicht den von den Autoren des MPEX Survey identifizierten Dimensionen (Vgl. Redish et al. 1998). Allein die Faktoren 1 und 2 ähneln von den zugeordneten Items her den Dimensionen „Realitätsbezug“ (gemeinsame Items: 13, 28, 29) bzw. „Anstrengung“ (gemeinsame Items: 5, 6, 10, 35). Die Items des Faktors 4 sind keiner der Dimensionen zugeordnet. Der Faktor 5 hingegen enthält Items aller Dimensionen. Eine Erklärungsmöglichkeit ist, dass die Dimensionen ein wohl reflektiertes und detailliertes *Lehr*konzept widerspiegeln. Das Faktormodell hingegen beschreibt ein gröberes und durch unreflektierte Erwartungen und Vorurteile geprägtes *Lern*konzept.

Tab. 4.5: Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 1

Item	geschätzter Parameter	Standardfehler	Teststatistik
11	1,000	0,000	0,000
13	0,743	0,062	11,898
19	0,762	0,060	12,634
28	-0,578	0,066	-8,809
29	-0,653	0,067	-9,816

Tab. 4.6: Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 2

Item	geschätzter Parameter	Standardfehler	Teststatistik
1	1,000	0,000	0,000
2	0,960	0,142	6,762
5	0,824	0,139	5,920
6	0,842	0,141	5,968
10	1,064	0,144	7,386
12	0,756	0,122	6,178
35	0,835	0,131	6,353

4.4 Konfirmatorische Faktorenanalyse

Für das im Rahmen der explorativen Faktorenanalyse definierte 4-Faktoren-Modell wird eine konfirmatorische Faktorenanalyse durchgeführt, um zu untersuchen, ob und wie gut es die Daten tatsächlich beschreibt.
Die Ergebnisse für die einzelnen Faktoren sind in den Tabellen 4.5 bis 4.8 zusammengefasst.
Für jeden Faktor wird jeweils für das Item in der ersten Zeile Eins als Start-

Tab. 4.7: Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 4

Item	geschätzter Parameter	Standardfehler	Teststatistik
30	1,000	0,000	0,000
34	0,650	0,115	5,632
36	0,725	0,105	6,927
38	0,738	0,097	7,645

Tab. 4.8: Konfirmatorische Faktorenanalyse: Ergebnisse für Faktor 5

Item	geschätzter Parameter	Standardfehler	Teststatistik
4	1,000	0,000	0,000
15	0,933	0,117	7,995
16	0,828	0,115	7,233
18	0,703	0,111	6,322
22	0,680	0,095	7,151
23	1,014	0,114	8,885
27	1,197	0,118	10,147
32	0,752	0,101	7,446
33	0,723	0,118	6,103

wert festgelegt. Sämtliche Parameterschätzungen sind auf einem Signifikanzniveau von $\alpha = 0,05$ verschieden von Null. Die einbezogenen Items leisten somit alle einen signifikanten Beitrag zum jeweiligen Faktor.

Der χ^2-Anpassungstest für das getestete Modell ergibt mit 437,024 einen relativ großen χ^2-Wert. Die Anzahl der Freiheitsgrade beträgt 84. Da der p-Wert von 0,0000 kleiner als das Signifikanzniveau von 5% ist, wird die Nullhypothese eines guten Modell-Fits abgelehnt. Für den CFI ergibt sich ein Wert von nur 0,456, was ebenfalls für eine schlechte Anpassung des Modells spricht. Der TLI liegt mit 0,508 etwas höher, jedoch noch immer deutlich niedriger als der für ein gutes Modell erforderliche Wert von 0,95. Der RMSEA von 0,141 liegt, wenn auch nicht sehr weit, über dem üblichen Cut-off Wert von 0,05 und bestätigt somit auch den schlechten Modell-Fit. Da die anderen betrachteten (hier jedoch nicht vorgestellten) Faktoren-Modelle noch schlechtere bzw. bereits in der explorativen Faktorenanalyse kaum interpretierbare Ergebnisse lieferten, wird das beschriebene Modell mit vier Faktoren trotz des schlechten Modell-Fits für die weiteren Analysen beibehalten.

Zum Vergleich wurde versucht, auch für das von den MPEX-Autoren definierte 6-Dimensionen-Modell (Vgl. Redish et al. 1998) eine konfirmatorische Faktorenanalyse unter Mplus durchzuführen. Jedoch konvergierte das Modell weder für verschiedene eindeutige Zuordnungen der Items zu jeweils nur einer Dimension, noch durch die Festlegung diverser Startwerte.

Tab. 4.9: Eta-Koeffizienten der Faktorwerte und verschiedener Merkmale

abhängig	unabhängig		
Faktorwerte	Studiengang	Geschlecht	Schulabschluss
Faktor 1	0,210	0,022	0,212
Faktor 2	0,230	0,169	0,206
Faktor 4	0,251	0,073	0,122
Faktor 5	0,256	0,048	0,178

Die Cronbach's Alphas von 0,5331 und 0,5955 sprechen zudem für eine unzureichende Reliabilität bzw. innere Konsistenz der Dimensionen „Unabhängigkeit“ und „Realitätsbezug“. Für die Dimensionen „Kohärenz“, „Konzepte“ und „Anstrengung“ kann das Alpha nicht ermittelt werden, da zwischen den Items negative Korrelationen bestehen, welche auch durch Umcodierungen nicht zu beseitigen sind. Allein die Dimension „Mathematikbezug“ kann mit einem Wert von 0,7222 als reliabel angesehen werden.

4.5 Analyse der Faktorwerte

In diesem Abschnitt wird für die vier Faktoren des Modells untersucht, wie stark diese bei verschiedenen Gruppen von Studierenden ausgeprägt sind. Dazu werden die von Mplus im Rahmen der konfirmatorischen Faktorenanalyse ausgegebenen Faktorwerte für 209 Beobachtungen[7] analysiert.

4.5.1 Faktor 1: Nutzen von Statistik-Kenntnissen und Formeln

Mit Hilfe von Kruskal-Wallis-Tests wird überprüft, ob die Faktorwerte bei verschiedenen Gruppen von Studierenden unterschiedlich verteilt sind. Die Ergebnisse sind in der Tabelle 4.10 zusammengefasst. Danach bestehen sowohl zwischen den Studiengängen als auch zwischen den Orten des Schulabschlusses signifikante Unterschiede in den mittleren Faktorwerten. Zwischen

[7] Für eine der 210 einbezogenen Beobachtungen ist die Schätzung der Faktorwerte fehlgeschlagen.

Tab. 4.10: Kruskal-Wallis-Tests für Faktorwerte des Faktors 1

Gruppierung	Teststatistik	Freiheitsgrade	Signifikanz
Studiengang	6,173	2	0,046
Geschlecht	0,103	1	0,748
Schulabschluss	9,669	2	0,008

welchen Paaren von Gruppen diese Unterschiede liegen, kann mit den Tests nicht ermittelt werden. Aus den Fehlerbalkendiagrammen in den Abbildungen 4.3 und 4.5 ist jedoch ersichtlich, dass sich insbesondere die mittleren Faktorwerte BWL- und VWL-Studierender und der Studierenden mit ostdeutschem und ausländischem Abitur stark unterscheiden. Positive Faktorwerte bedeuten hier, dass den Aussagen, welche negativ in den Faktor eingehen (s. Seite 40) eher zugestimmt und den positiv eingehenden Aussagen eher nicht zugestimmt wird. Somit erwarten sich z.B. BWL-Studierende und Studierende mit ostdeutschem Abitur im Durchschnitt einen geringeren Nutzen von Statistik-Kenntnissen und Formeln und sind u.a. eher bereit, letztere als gegeben hinzunehmen, als VWL-Studierende bzw. Studierende mit ausländischem Abitur. Zwischen männlichen und weiblichen Studierenden besteht in dieser Hinsicht fast gar kein Unterschied wie die Abbildung 4.4 und aus dem Testergebnis in der Tabelle 4.10 deutlich wird.

Diese Ergebnisse spiegeln sich auch in den Eta-Koeffizienten wider, welche in der Tabelle 4.9 für alle Faktoren dargestellt sind. Dabei werden die Faktorwerte als abhängige und die Merkmale Studiengang, Geschlecht und Ort des Schulabschlusses jeweils als unabhängige Variable betrachtet. Während durch die Kenntnis des Studiengangs oder des Schulabschlusses der Vorhersagefehler für die Ausprägung des ersten Faktors zumindest leicht (um 4,4% bzw. 4,5%) reduziert werden kann, besteht zwischen dem Faktorwert und dem Geschlecht kein Zusammenhang.

Um den Einfluss des Alters auf die Ausprägung des Faktors zu untersuchen, werden diese in der Abbildung 4.6 in einem Scatterplot gegeneinander abgetragen. Es ist kein Zusammenhang zwischen den Variablen ersichtlich. Bestätigt wird dieser Eindruck durch den nicht signifikanten Bravais-Pearson-

Abb. 4.3: Fehlerbalkendiagramm: Faktorwerte 1 nach Studiengang

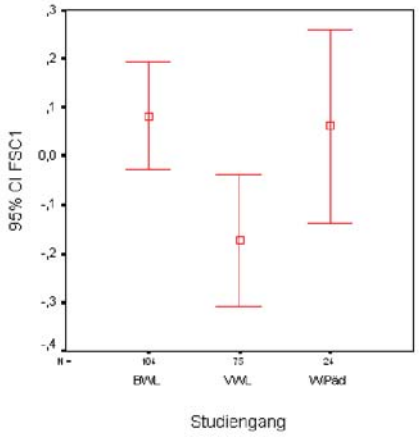


Abb. 4.4: Fehlerbalkendiagramm: Faktorwerte 1 nach Geschlecht

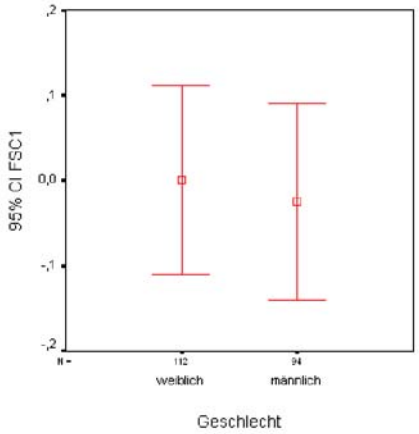


Abb. 4.5: Fehlerbalkendiagramm Faktorwerte 1 nach Schulabschluss

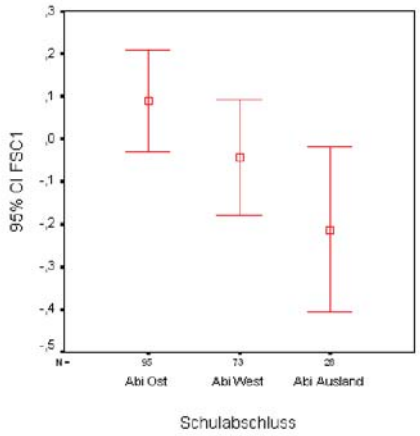
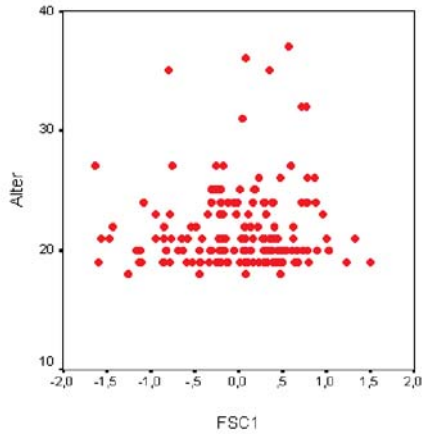


Abb. 4.6: Scatterplot: Faktorwerte 1 und Alter

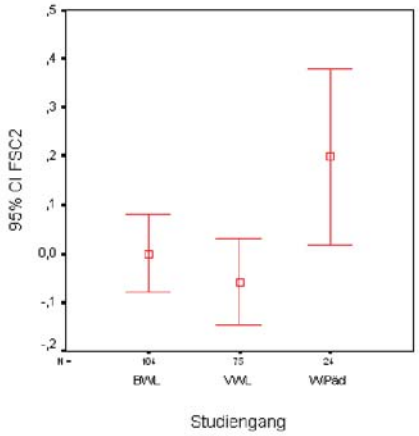


Korrelationskoeffizienten von 0,07.

4.5.2 Faktor 2: Lernen durch Training

Die Eta-Koeffizienten aus der Tabelle 4.9 deuten auf jeweils sehr leichte Zusammenhänge zwischen dem Studiengang, dem Geschlecht, dem Schulabschluss und der Ausprägung des zweiten Faktors hin, welche hier wiederum etwas genauer betrachtet werden.
Laut den Ergebnissen der Kruskal-Wallis-Tests in der Tabelle 4.11 bestehen nach allen Gruppierungen signifikante Unterschiede zwischen den mittleren

Abb. 4.7: Fehlerbalkendiagramm Faktorwerte 2 nach Studiengang



Tab. 4.11: Kruskal-Wallis-Tests für Faktorwerte des Faktors 2

Gruppierung	Teststatistik	Freiheitsgrade	Signifikanz
Studiengang	7,276	2	0,026
Geschlecht	5,117	1	0,024
Schulabschluss	10,468	2	0,005

Ausprägungen des Faktors. Welche Paare von Gruppen sich am stärksten unterscheiden, ist wiederum aus den Fehlerbalkendiagrammen in den Abbildungen 4.7 bis 4.9 abzulesen.

Die durchschnittlichen Faktorwerte von Studierenden der BWL und VWL liegen etwa auf gleicher Höhe, der Mittelwert für Wirtschaftspädagogik-Studierende hingegen deutlich höher. Die durch den Faktor beschriebene Lernstrategie ist also unter letzteren überdurchschnittlich stark ausgeprägt. Der Vergleich der mittleren Faktorwerte nach dem Geschlecht der Studierenden, zeigt eine überdurchschnittliche Ausprägung des zweiten Faktors bei weiblichen Studierenden. Ebenso haben Studierende, die über ein ostdeutsches Abitur verfügen, einen stärkeren Einfluss auf den Faktor als diejenigen mit einem westdeutschen oder ausländischen Abitur.

Das Alter der Studierenden hat wiederum keinen Einfluss auf die Ausprägung des Faktors, wie in der Abbildung 4.10 und durch den nicht signifikanten Korrelationskoeffizienten von -0,042 deutlich wird.

Abb. 4.8: Fehlerbalkendiagramm Faktorwerte 2 nach Geschlecht

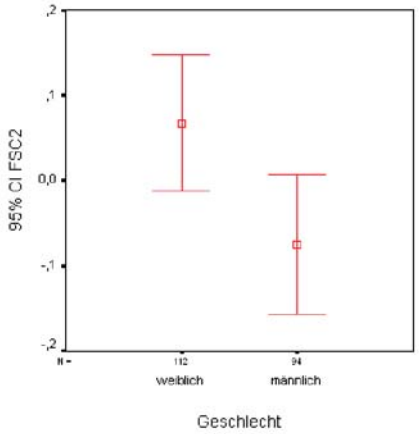


Abb. 4.9: Fehlerbalkendiagramm Faktorwerte 2 nach Schulabschluss

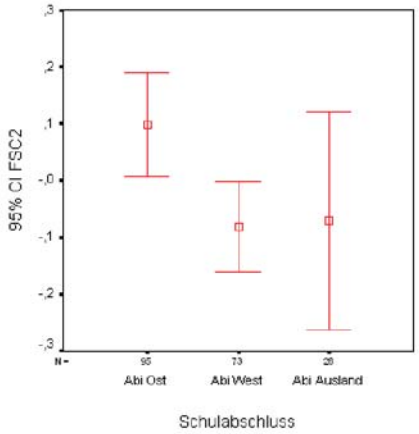
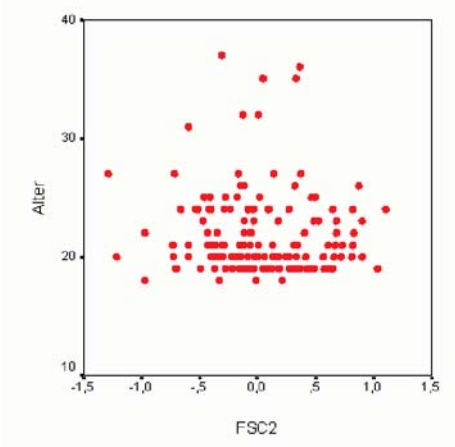


Abb. 4.10: Scatterplot: Faktorwerte 2 und Alter



Tab. 4.12: Kruskal-Wallis-Tests für Faktorwerte des Faktors 4

Gruppierung	Teststatistik	Freiheitsgrade	Signifikanz
Studiengang	10,235	2	0,006
Geschlecht	0,857	1	0,355
Schulabschluss	3,105	2	0,212

4.5.3 Faktor 4: Verstehen von Konzepten

Laut Eta-Koeffizient besteht zwischen der Ausprägung des vierten Faktors und dem Studiengang ein leichter Zusammenhang. Der mittlere Faktorwert ist, wie in Abbildung 4.11 zu sehen ist, bei den Studierenden der VWL und der Wirtschaftspädagogik in etwa gleich hoch. BWL-Studierende hingegen weisen einen niedrigeren Mittelwert auf und haben somit einen unterdurchschnittlichen Einfluss auf den Faktor. Zwischen den Geschlechtern und den Schulabschlüssen (s. Abbildungen 4.12 und 4.13 und Tabelle 4.12) bestehen hinsichtlich der durchschnittlichen Ausprägung des vierten Faktors keine signifikanten Unterschiede. Auch die Eta-Koeffizienten bestätigen, dass zwischen den Ausprägungen dieser beiden Merkmale und dem Faktor so gut wie kein Zusammenhang besteht.

Auch das Alter hat keinen Einfluss auf die Ausprägung des vierten Faktors, wie sowohl in der Abbildung 4.14, als auch durch die nicht signifikante Kor-

Abb. 4.11: Fehlerbalkendiagramm Faktorwerte 4 nach Studiengang

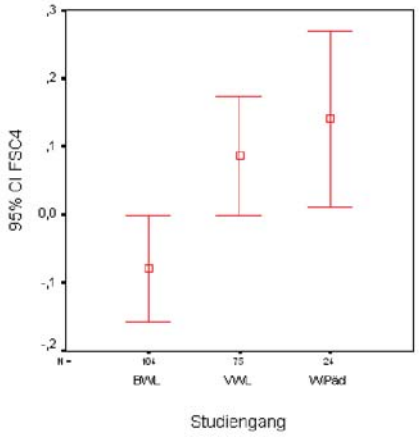


Abb. 4.12: Fehlerbalkendiagramm Faktorwerte 4 nach Geschlecht

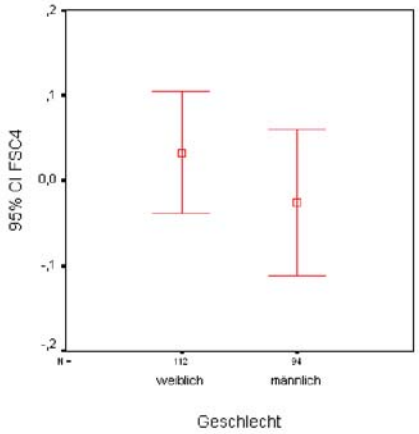


Abb. 4.13: Fehlerbalkendiagramm Faktorwerte 4 nach Schulabschluss

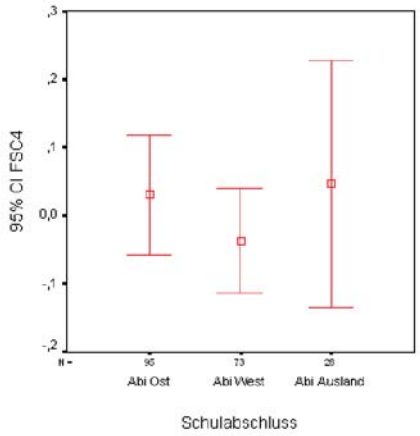
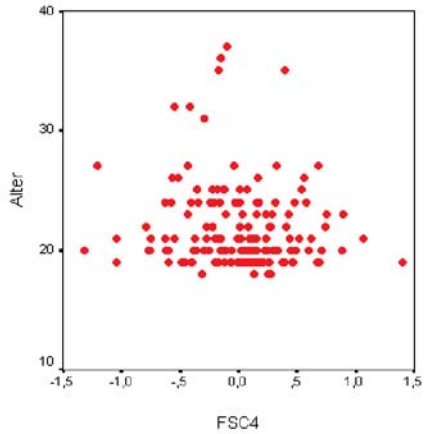


Abb. 4.14: Scatterplot: Faktorwerte 4 und Alter



relation beider Merkmale von -0,114 deutlich wird.

4.5.4 Faktor 5: Hinnehmen und Auswendiglernen

Der Fehler bei der Vorhersage der Ausprägung des fünften Faktors kann durch die Kenntnis des Studiengangs leicht (um 6,5%) reduziert werden. Das Ergebnis des Kruskal-Wallis-Tests in der Tabelle 4.13 spricht wiederum für signifikante Unterschiede in den Faktorwerten verschiedener Studiengänge. Wie aus der Abbildung 4.15 ersichtlich, ist der fünfte Faktor bei BWL-Studierenden deutlich stärker ausgeprägt als bei VWL-Studierenden,

Abb. 4.15: Fehlerbalkendiagramm Faktorwerte 5 nach Studiengang

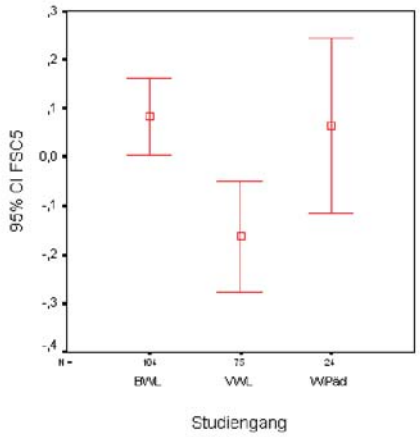
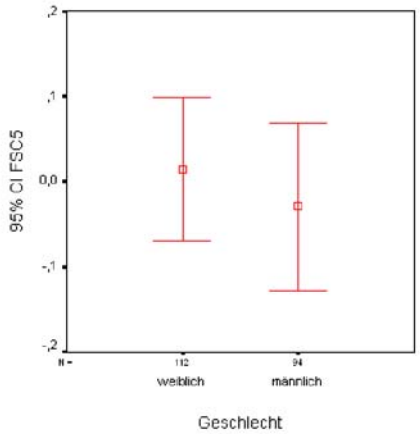


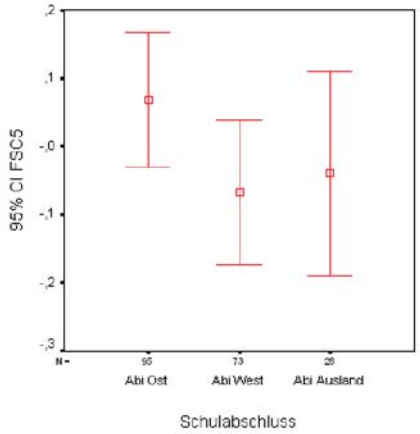
Abb. 4.16: Fehlerbalkendiagramm Faktorwerte 5 nach Geschlecht



bei Wirtschaftspädagogik-Studierenden hingegen durchschnittlich. Zwischen weiblichen und männlichen Studierenden, sowie zwischen Studierenden mit Schulabschlüssen verschiedener Regionen (s. Abbildungen 4.16 und 4.17 und Tabelle 4.13) bestehen keine signifikanten Unterschiede in den mittleren Faktorwerten. Die Eta-Koeffizienten bestätigen, dass das Geschlecht nicht und der Ort des Schulabschlusses nur sehr gering mit dem Faktorwert zusammenhängen.

Ausgehend von der Abbildung 4.18 könnte zunächst ein positiver Zusammenhang zwischen dem Alter und der Ausprägung des Faktors vermutet werden. Der Korrelationskoeffizient beträgt jedoch lediglich 0,029.

Abb. 4.17: Fehlerbalkendiagramm Faktorwerte 5 nach Schulabschluss



Tab. 4.13: Kruskal-Wallis-Tests für Faktorwerte des Faktors 5

Gruppierung	Teststatistik	Freiheitsgrade	Signifikanz
Studiengang	8,045	2	0,018
Geschlecht	0,377	1	0,539
Schulabschluss	4,682	2	0,096

Abb. 4.18: Scatterplot: Faktorwerte 5 und Alter

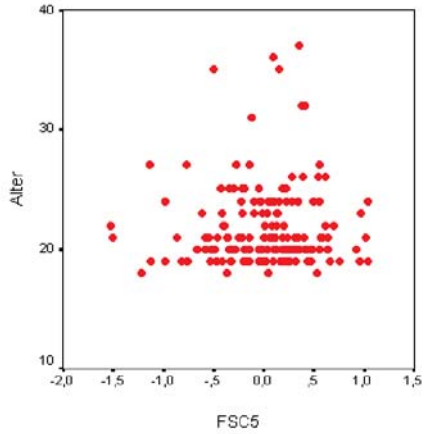
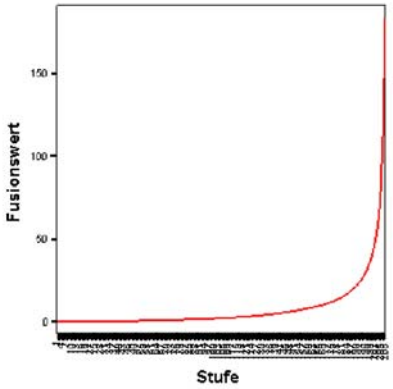


Abb. 4.19: Clusteranalyse: Verlauf der Fusionswerte



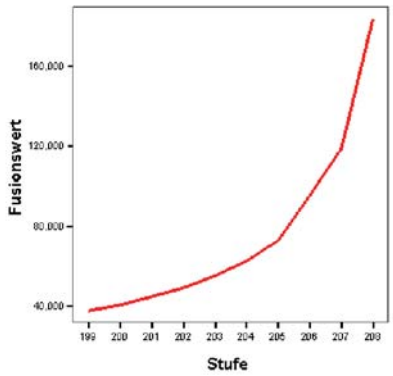
4.5.5 Clusteranalyse

Mit den ermittelten Faktorwerten wird in diesem Abschnitt eine Clusteranalyse durchgeführt. Ziel ist es, möglichst homogene Gruppen von Studierenden hinsichtlich der Ausprägungen der vier einbezogenen Faktoren zu bilden.
Üblicherweise wird der Verlauf der Clusterung in einem Dendogramm zusammengefasst. Da dieses jedoch für eine Anzahl von 209 Objekten bereits sehr unübersichtlich und somit nur sehr begrenzt aussagekräftig ist, wird hier auf seine Darstellung verzichtet.
In der Abbildung 4.19 ist stattdessen der Verlauf der Fusionswerte dargestellt, welcher ebenfalls einen Überblick über die Clusterung liefert. Ein sprunghafter Anstieg ist darin aufgrund der großen Anzahl von 208 Stufen kaum sichtbar. Daher werden in der Abbildung 4.20 nochmals die Fusionswerte der letzten zehn Stufen als ein vergrößerter Ausschnitt dargestellt. Darin ist jeweils ein Knick nach der 205. und nach der 207. Stufe zu erkennen. Im folgenden werden daher die 4- und die 2-Clusterlösung etwas näher betrachtet.

4-Clusterlösung

Die einzelnen Cluster können anhand der mittleren Faktorwerte inhaltlich beschrieben werden. Der Line-Plot in Abbildung 4.21 stellt diese für die 4-Clusterlösung dar. Die mittleren Faktorwerte der ersten zwei Cluster lie-

Abb. 4.20: Clusteranalyse: Verlauf der letzten 10 Fusionswerte

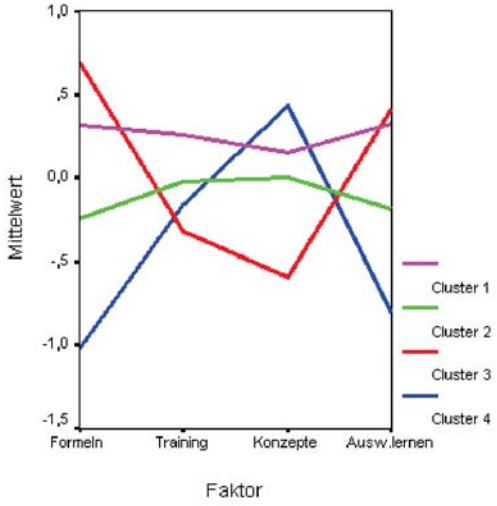


gen jeweils relativ nahe bei Null, haben also einen eher durchschnittlichen Einfluss auf alle vier Faktoren. Interessanter sind die Kurven der übrigen Cluster. So wird im dritten Cluster ein überdurchschnittlich hoher Wert auf das Hinnehmen und Auswendiglernen gelegt, während die Bedeutung des Verstehens von Konzepten unterdurchschnittlich ausgeprägt ist. Der hohe Wert des ersten Faktors bedeutet, dass der Nutzen von Statistik-Kenntnissen und Formeln geringer als im Durchschnitt aller Studierender eingeschätzt wird. Für das vierte Cluster hingegen sind die jeweiligen mittleren Faktorwerte und die dadurch beschriebenen Prioritäten fast exakt umgekehrt verteilt.

Von Interesse ist nun, ob die Cluster bezüglich der Zusammensetzung der enthaltenen Objekte Besonderheiten aufweisen. Anhand des Unsicherheitskoeffizienten wird daher überprüft, ob zwischen den Verteilungen der Studierenden nach Studiengängen, Geschlecht und Ort des Schulabschlusses und ihrer Clusterzugehörigkeit Zusammenhänge bestehen. Die Ergebnisse sind in der Tabelle 4.14 zusammengefasst. Der Koeffizient deutet darauf hin, dass zwischen dem Studiengang der Befragten und ihrer Clusterzugehörigkeit ein signifikanter Zusammenhang besteht.

Die Abbildung 4.22 stellt die Verteilung der Studiengänge nach der Clusterzugehörigkeit in einem gruppierten Balkendiagramm dar. Es ist deutlich zu erkennen, dass VWL-Studierende im vierten Cluster stark überproportional vertreten sind (Vgl. Abbildung 2.1).

Abb. 4.21: Lineplot: 4 Cluster



Tab. 4.14: Unsicherheitskoeffizienten (Cluster abh.) für 4-Clusterlösung

	Uns.koeff.	Sig.
Studiengang	0,047	0,003
Geschlecht	0,002	0,738
Schulabschluss	0,027	0,137

Abb. 4.22: Balkendiagramm: 4-Clusterlösung und Studiengang

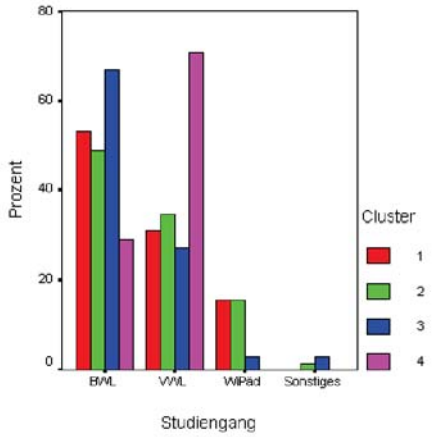
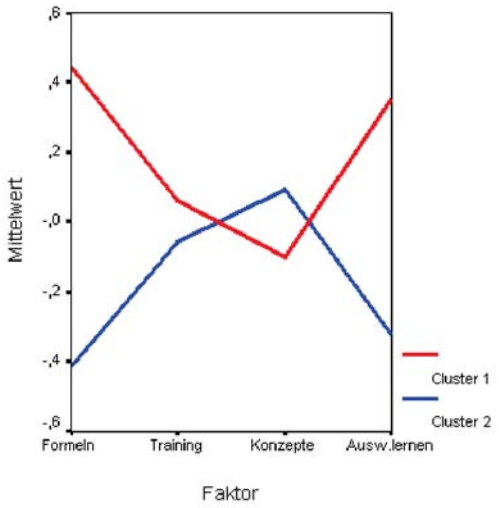


Abb. 4.23: Lineplot: 2 Cluster



Des weiteren besteht zwischen der Clusterzugehörigkeit (abhängig) der Studierenden und ihrem Alter (unabhängig) nur ein sehr geringer Zusammenhang: Der entsprechende Eta-Koeffizient beträgt 0,271.

2-Clusterlösung

Die Abbildung 4.23 enthält den Line-Plot für die 2-Clusterlösung. Die mittleren Faktorwerte der zwei Cluster ähneln im Verlauf stark dem dritten und vierten Cluster aus der 4-Clusterlösung (Vgl. Abbildung 4.21). Allerdings liegen die Linien jetzt näher bei Null und somit beieinander. Die Cluster spiegeln also die gleichen Eigenschaften wie die im vorangegangenen Abschnitt beschriebenen wieder, durch die Vereinigung mit den zwei mittleren Clustern jedoch nur noch in abgeschwächter Form. Aus der Tabelle 4.15 ist wiederum ersichtlich, dass die Merkmale Studiengang, Geschlecht und Ort des Schulabschlusses sich in den zwei Clustern auf einem Signifikanzniveau von 5% nicht signifikant unterscheiden.
Der Eta-Koeffizient von 0,359 hingegen deutet auf eine gewisse Abhängigkeit der Clusterzugehörigkeit der Studierenden von ihrem Alter hin. Die Studierenden im zweiten Cluster sind durchschnittlich ein halbes Jahr jünger.
Allgemein lässt sich also feststellen, dass die Merkmale Studiengang, Ge-

Tab. 4.15: Unsicherheitskoeffizienten (Cluster abh.) für 2-Clusterlösung

	Uns.koeff.	Sig.
Studiengang	0,014	0,253
Geschlecht	0,002	0,446
Schulabschluss	0,025	0,070

schlecht, Ort des Schulabschlusses und Alter der Studierenden in den betrachteten Clustern beider Lösungen nicht exakt gleich verteilt sind. Mit Ausnahme des überproportionalen Vorkommens von VWL-Studierenden im vierten Cluster der 4-Clusterlösung sind die Abweichungen in den Verteilungen jedoch statistisch nicht signifikant. So umfassen die Cluster zwar verschiedene Typen von Studierenden hinsichtlich ihres durch die Faktoren beschriebenen Lernverhaltens. Anhand der erhobenen Merkmale kann für diese jedoch kein Profil bzw. ein typischer Vertreter erstellt werden.

Anmerkung

Die konfirmatorische Faktorenanalyse wurde mit allen 38 Items durchgeführt. Für diejenigen Items, die im Modell nicht benutzt werden, erscheint im Mplus Output eine Warnung. Schließt man die nicht benutzten Items aus, so verbessern sich die Teststatistiken der konfirmatorischen Faktorenanalyse etwas, erreichen jedoch auch nicht die für einen guten Modell-Fit geforderten Werte. Die Parameterschätzungen sind identisch und die Faktorwerte korrelieren perfekt mit denjenigen, die unter Einbeziehung aller Items ausgegeben werden. Die in diesem Kapitel dargestellten Ergebnisse würden daher durch den Ausschluss der nicht verwendeten Items nicht oder nur sehr geringfügig verändert.

5 Zusammenfassung

Mit Hilfe verschiedener statistischer Verfahren wurden in dieser Arbeit hauptsächlich zwei Fragestellungen untersucht: Was bestimmt die Erwartungen von Studierenden an die Lehrveranstaltungen Statistik I und II und welche Unterschiede bestehen dabei zwischen verschiedenen Gruppen von Studierenden? Misst man die Unterschiede zwischen den Antwortverteilungen verschiedener Gruppen durch Mann-Whitney-Tests, so zeigen sich die meisten Übereinstimmungen (im Vergleich zu den jeweils anderen Gruppen) mit den Expertenmeinungen unter

- VWL-Studierenden
- männlichen Studierenden und
- Studierenden mit ausländischem Abitur.

Um latente Strukturen hinter den Antworten der Studierenden aufzudecken, wurde eine explorative Faktorenanalyse durchgeführt, welche zur Identifikation der folgenden Faktoren führte:

- Nutzen von Statistik-Kenntnissen und Formeln
- Lernen durch Training
- Verstehen von Konzepten und
- Hinnehmen und Auswendiglernen.

Der Modell-Fit erwies sich in der Prüfung durch die konfirmatorische Faktorenanalyse als nicht zufriedenstellend. Da jedoch durch die Variation der

Faktorenzahl keine besseren (im Sinne des Modell-Fits) und gleichzeitig interpretierbaren Ergebnisse erzielt werden konnten, wurde das 4-Faktoren-Modell dennoch für die weiteren Untersuchungen beibehalten.
Die Analyse der Faktorwerte ergab, dass die Faktoren teilweise bei verschiedenen Gruppen von Studierenden signifikant unterschiedlich stark ausgeprägt sind:

- VWL-Studierende und Studierende, die ihr Abitur im Ausland abgelegt haben, versprechen sich einen höheren Nutzen von Statistik-Kenntnissen und Formeln als der Durchschnitt aller Studierender.
- Der Lerntyp „Lernen durch Training“ ist überdurchschnittlich stark ausgeprägt bei Wirtschaftspädagogik-Studierenden, weiblichen Studierenden und Studierenden mit ostdeutschem Abitur.
- Studierende der Studiengänge VWL und Wirtschaftspädagogik legen einen größeren Wert auf das Verstehen von Konzepten als BWL- Studierende.
- BWL-Studierende sind eher zum Hinnehmen und Auswendiglernen bereit als VWL-Studierende.

Die Clusteranalyse der Faktorwerte identifizierte u.a. eine Gruppe von Studierenden, bei denen der „Nutzen von Statistik-Kenntnissen und Formeln“ überdurchschnittlich hoch (negativer Faktorwert) eingeschätzt wird und die Faktoren „Hinnehmen und Auswendiglernen“ stark unterdurchschnittlich bzw. „Verstehen von Konzepten“ stark überdurchschnittlich ausgeprägt sind, welche mit überproportional vielen VWL-Studierenden besetzt ist. Ein weiteres Cluster weist genau entgegengesetzte Faktorausprägungen auf, weicht jedoch wie auch die übrigen Cluster in seiner Zusammensetzung kaum von der Gesamtheit der befragten Studierenden ab.
Insgesamt fällt in den Analysen auf, dass es wenige oder nur schwach ausgeprägte Unterschiede zwischen den Antwortverteilungen verschiedener Gruppen von Studierenden gibt. Dies könnte zum einen daran liegen, dass sie nahezu alle erst seit einigen Wochen studieren. Sie haben somit einen sehr

ähnlichen Hintergrund, welcher noch nicht durch u.a. unterschiedliche besuchte Lehrveranstaltungen geprägt ist, so dass beispielsweise die Ähnlichkeit zwischen BWL- und VWL-Studierenden nicht überrascht.
Des weiteren wurde von den Studierenden sehr häufig die Antwort „neutral“ gewählt. Möglicherweise sind viele der Aussagen zu spezifisch, als dass Studierende zu ihnen bereits *vor* dem Besuch der Lehrveranstaltung Stellung beziehen könnten. Interessant wäre es daher, die Befragung am Ende der Lehrveranstaltung Statistik II zu wiederholen und die Ergebnisse zu vergleichen. In der MPEX-Studie zeigte sich z.B., dass die Antworten der Studierenden *nach* dem Besuch der Lehrveranstaltung *stärker* von denen der Experten abwichen als *vorher* (Vgl. Redish et al. 1998). In solch einem Fall wäre es eventuell angebracht, das Vorgehen in einer Lehrveranstaltung oder gar ihre Ziele neu zu definieren. Durch das Auslassen der Antwortkategorie „neutral“ wären die Studierenden zwar eher zu einer Entscheidung angehalten. Dies könnte jedoch wiederum zu einem höheren Anteil fehlender Werte im Datensatz führen.
Unerwarteterweise wurde kein Modell gefunden, das die von den MPEX-Autoren definierten Dimensionen widerspiegelt. Dies könnte dadurch bedingt sein, dass es sich bei letzterem um ein *Lehr*konzept handelt, während die Faktoren eher einem *Lern*konzept entsprechen. Trotzdem stellt sich die Frage, ob die Befragung tatsächlich *das* misst, was sie beabsichtigt. Abweichungen könnten z.B. durch die Übertragung des Fragebogens vom Fachgebiet Physik auf Statistik, vom US-amerikanischen auf das deutsche Bildungssystem, von der englischen in die deutsche Sprache oder auch in der Konzeption des MPEX-Fragebogens selbst begründet sein.

Literaturverzeichnis

[1] *Studienordnung für den Bachelorstudiengang Betriebswirtschaftslehre.* 2004. – Humboldt-Universität zu Berlin

[2] *Studienordnung für den Bachelorstudiengang Volkswirtschaftslehre.* 2004. – Humboldt-Universität zu Berlin

[3] *Evaluation Sommersemester 2007.* 2007. – Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin

[4] *Evaluation Wintersemester 2006/07.* 2007. – Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin

[5] *Studien- und Prüfungsordnung für das Bachelorstudium Wirtschaftspädagogik.* 2007. – Humboldt-Universität zu Berlin

[6] *Studierendenstatistik.* 2007/08. – Humboldt-Universität zu Berlin

[7] Backhaus, K. ; Erichson, B. ; Plinke, W. ; Weiber, R.: *Multivariate Analysemethoden.* 11. Berlin : Springer, 2006

[8] Bleymüller, J. ; Gehlert, G. ; Gülicher, H.: *Statistik für Wirtschaftswissenschaftler.* 13. München : Verlag Franz Vahlen, 2002

[9] Borman, S.: *The Expectation Maximization Algorithm - A short tutorial.* http://www.seanborman.com/publications, 2006. – 09.04.2008

[10] Bortz, J.: *Statistik für Human- und Sozialwissenschaftler.* 6. Berlin : Springer, 2005

[11] Brandes, U.: *Statistische Bewertung und Analyse der Klausurergebnisse Statistik (Grundstudium)*, Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin, Diplomarbeit, 2004

[12] Curran, P. J. ; Flora, D. B.: An Empirical Evaluation of Alternative Methods of Estimation for Confirmatory Factor Analysis With Ordinal Data. In: *Psychological Methods* 9 (2004), Nr. 4, S. 466–491

[13] Daly, F. ; Hand, D.J. ; Jones, M.C. ; Lunn, A.D. ; McConway, K.J.: *Elements of Statistics.* Addison-Wesley Publishing Company, 1995

[14] Garson, G.D.: *Statnotes: Topics in Multivariate Analysis.* – http://www2.chass.ncsu.edu/garson/pa765/statnote.htm, 07.04.2007

[15] Härdle, W. ; Klinke, S. ; Ziegenhagen, U.: On the Utility of E-Learning in Statistics / Sonderforschungsbereich 649, Humboldt Universität zu Berlin. 2007 (2007-050). – SFB 649 Discussion Paper. – http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2007-050.pdf

[16] Härdle, W. ; Simar, L.: *Applied Multivariate Statistical Analysis.* 1. Berlin : Springer Verlag, 2003

[17] Jöreskog, K.G. ; Moustaki, I.: Factor Analysis of Ordinal Variables: A Comparison of Three Approaches. In: *Multivariate Behavioral Research* 3 (2001), Nr. 36, S. 347–387

[18] Klinke, S. ; Wagner, C.: Visualizing exploratory factor analysis models / Sonderforschungsbereich 649, Humboldt Universität zu Berlin. 2008 (2008-012). – SFB 649 Discussion Paper. – http://sfb649.wiwi.hu-berlin.de/papers/pdf/SFB649DP2008-012.pdf

[19] Kruskal, W.H. ; Wallis, W.A.: Use of Ranks in One-Criterion Variance Analysis. In: *Journal of the American Statistical Association* 47 (1952), 12, Nr. 260, S. 583–621

[20] Muthén, B.O. ; Muthén, L.K.: *Mplus User's Guide.* 5. Los Angeles, CA : Muthén & Muthén, 2007

[21] Pohlenz, P.: *Konfirmatorische Faktorenanalyse.* – Universität Potsdam

[22] Redish, E. F. ; Saul, J. M. ; Steinberg, R. N.: Student Expectations in Introductory Physics. In: *The American Journal of Physics* 66 (1998), S. 212–224

[23] Rönz, B.: Skript Computergestützte Statistik II. (2000). – Institut für Statistik und Ökonometrie, Wirtschaftswissenschaftliche Fakultät, Humboldt-Universität zu Berlin

[24] Schafer, J.L. ; Graham, J.W.: Missing Data: Our View of the State of the Art. In: *Psychological Methods* 7, Nr.2 (2002), S. 147–177

[25] Schermelleh-Engel, K.: Konfirmatorische Faktorenanalyse mit LISREL: Ein Übungsbeispiel mit allen Ablaufschritten. (1999), Nr. 3. – Institut für Psychologie, Johann Wolfgang Goethe-Universität Frankfurt am Main

[26] Ueberla, K.: *Faktorenanalyse.* 2. Berlin Heidelberg New York : Springer-Verlag, 1971

[27] Uebersax, J. S.: *The tetrachoric and polychoric correlation coefficients. Statistical Methods for Rater Agreement web site.* http://ourworld.compuserve.com/homepages/jsuebersax/tetra.htm, 2006. – 09.04.2008

[28] Wiedenbeck, M. ; Züll, C.: Klassifikation mit Clusteranalyse: Grundlegende Techniken hierarchischer und K-means-Verfahren. In: *ZUMA How-to-Reihe* (2001), Nr. 10. – Zentrum für Umfragen, Methoden und Analysen, Mannheim

Anhang A

Die Items der in der Vorlesung Mathematik I im Wintersemester 2007/2008 durchgeführten Befragung:

1. Um den Stoff zu verstehen, muss ich das Skript lesen.
2. Um den Stoff zu verstehen, muss ich die Übungsaufgaben lösen.
3. Um den Stoff zu verstehen muss ich während der VL/Ü aufpassen.
4. Alles, was ich durch die Herleitung einer Formel lerne, ist, dass diese
5. Formel richtig ist und für die Lösung von Aufgaben benutzt werden kann.
6. Um mich auf die Prüfung vorzubereiten, gehe ich meine Mitschriften aus der VL/Ü sorgfältig durch.
7. Um mich auf die Prüfung vorzubereiten, löse ich die Aufgaben aus der Aufgabensammlung.
8. Um eine statistische Fragestellung zu beantworten, sind passende Formeln für den entsprechenden Sachverhalt zu suchen und Werte einzusetzen, um eine Zahl zu erhalten.
9. Die erworbenen Statistikkenntnisse werden meine Sicht auf einige alltägliche Zusammenhänge verändern.
10. Ich werde viel Zeit damit verbringen, die Herleitungen aus der VL/Ü oder dem Skript nachzuvollziehen.

11. Ich werde das Skript sorgfältig lesen und die Beispiele nachrechnen. In Statistik erwarte ich nicht, die Formeln zu verstehen, sondern werde sie als gegeben hinnehmen.

12. Ich werde Statistik eher verstehen, wenn ich möglichst viele Aufgaben löse, als wenn ich nur wenige detailliert analysiere.

13. Statistische Formeln haben wenig mit der realen Welt zu tun.

14. Ein gutes Verständnis für Statistik ist wichtig für meine berufliche Karriere. Eine gute Note allein reicht nicht aus.

15. Die einzelnen Teilbereiche der Statistik haben wenig miteinander zu tun.

16. Meine Note in Statistik hängt hauptsächlich davon ab, wie gut ich mir die Lehrinhalte merken kann, weniger von meinem Verständnis oder meiner Kreativität.

17. Statistikkenntnisse zu erwerben bedeutet vor allem, Formeln und Konzepte in der VL/Ü und aus dem Skript kennenzulernen.

18. Wenn beim Rechnen einer Statistik-Aufgabe mein Ergebnis deutlich von dem abweicht, was ich erwartet hätte, muss ich mich auf meine Berechnung verlassen.

19. Die Herleitung von Formeln in der VL/Ü oder im Skript hilft mir kaum, Aufgaben zu lösen bzw. die Klausur zu bestehen.

20. Nur einige wenige sehr qualifizierte Leute sind in der Lage, Statistik wirklich zu verstehen.

21. Um Statistik zu verstehen, beziehe ich mein Wissen aus anderen Bereichen ein und verbinde dies mit der Problemstellung.

22. Das Wichtigste bei der Analyse einer statistischen Fragestellung ist, die richtige Formel zu finden.

23. Wenn ich in der Klausur eine bestimmte Formel für die Lösung einer Aufgabe brauche, die mir nicht mehr einfällt und nicht in der Formelsammlung enthalten ist, dann kann ich nichts (legales!) tun, um sie zu finden.

24. Wenn zwei verschiedene Lösungsansätze für eine Aufgabe unterschiedliche Ergebnisse liefern, wähle ich einfach das Ergebnis aus, das mir sinnvoller scheint.

25. Statistik hat mit der realen Welt zu tun. Für das Verständnis kann es hilfreich sein, an diese Verbindung zu denken. Um die Anforderungen in der VL/Ü zu erfüllen, hilft diese Verbindung nur selten.

26. Das Wichtigste, was ich in dieser Veranstaltung lerne, ist statistische Fragestellungen zu lösen.

27. Meine Klausurergebnisse helfen mir nicht mehr beim Verständnis des Stoffs. Die Veranstaltung und der damit verbundene Lernprozess sind für mich mit der Klausur beendet.

28. Meine Statistik-Kenntnisse werden mir in anderen Lehrveranstaltungen nützlich sein.

29. Meine Statistik-Kenntnisse werden mir helfen, verschiedene Veröffentlichungen (Fachartikel, Tagespresse...) besser zu verstehen.

30. Bei den meisten Übungs- oder Klausuraufgaben, die ich löse, denke ich gezielt an die zugrundeliegenden Annahmen und Konzepte.

31. Statistik "verstehen"heisst im Grunde, sich an Dinge, die einem gezeigt wurden oder die man gelesen hat, zu erinnern.

32. Lange an einer Aufgabe zu rechnen ist Zeitverschwendung. Wenn ich nicht vorankomme, frage ich lieber jemanden, der mehr davon versteht als ich.

33. In Statistik ist es wichtig, alles was ich wissen muss, auswendig zu lernen.

34. Das Wichtigste, was ich in Statistik lerne, ist die Fähigkeit, statistische Zusammenhänge zu verstehen.

35. Ich lerne aus meinen Fehlern in Übungs- und Klausuraufgaben.

36. Um eine Formel auf eine Aufgabe anzuwenden, muss ich mehr als nur die Bedeutung der in der Formel vorkommenden Terme kennen.

37. Es ist möglich, zu bestehen ohne Statistik sehr gut zu verstehen.

38. Um Statistik zu verstehen, muss ich die Informationen aus der VL/Ü und dem Skript grundlegend durchdenken und gedanklich umstrukturieren und -organisieren.

Anhang B

Die Items des MPEX Survey:

1. All I need to do to understand most of the basic ideas in this course is just read the text, work most of the problems, and/or pay close attention in class.

2. All I learn from a derivation or proof of a formula is that the formula obtained is valid and that it is OK to use it in problems.

3. I go over my class notes carefully to prepare for tests in this course.

4. Problem solvingïn physics basically means matching problems with facts or equations and then substituting values to get a number.

5. Learning physics made me change some of my ideas about how the physical world works.

6. I spend a lot of time figuring out and understanding at least some of the derivations or proofs given either in class or in the text.

7. I read the text in detail and work through many of the examples given there.

8. In this course, I do not expect to understand equations in an intuitive sense; they must just be taken as givens.

9. The best way for me to learn physics is by solving many problems rather than by carefully analyzing a few in detail.

10. Physical laws have little relation to what I experience in the real world.
11. A good understanding of physics is necessary for me to achieve my career goals. A good grade in this course is not enough.
12. Knowledge in physics consists of many pieces of information each of which applies primarily to a specific situation.
13. My grade in this course is primarily determined by how familiar I am with the material. Insight or creativity has little to do with it.
14. Learning physics is a matter of acquiring knowledge that is specifically located in the laws, principles, and equations given in class and/or in the textbook.
15. In doing a physics problem, if my calculation gives a result that differs significantly from what I expect, I'd have to trust the calculation.
16. The derivations or proofs of equations in class or in the text has little to do with solving problems or with the skills I need to succeed in this course.
17. Only very few specially qualified people are capable of really understanding physics.
18. To understand physics, I sometimes think about my personal experiences and relate them to the topic being analyzed.
19. The most crucial thing in solving a physics problem is finding the right equation to use.
20. If I don't remember a particular equation needed for a problem in an exam there's nothing much I can do (legally!) to come up with it.
21. If I came up with two different approaches to a problem and they gave different answers, I would not worry about it; I would just choose the answer that seemed most reasonable. (Assume the answer is not in the back of the book.)

22. Physics is related to the real world and it sometimes helps to think about the connection, but it is rarely essential for what I have to do in this course.

23. The main skill I get out of this course is learning how to solve physics problems.

24. The results of an exam don't give me any useful guidance to improve my understanding of the course material. All the learning associated with an exam is in the studying I do before it takes place.

25. Learning physics helps me understand situations in my everyday life.

26. When I solve most exam or homework problems, I explicitly think about the concepts that underlie the problem.

27. Ünderstandingphysics basically means being able to recall something you've read or been shown.

28. Spending a lot of time (half an hour or more) working on a problem is a waste of time. If I don't make progress quickly, I'd be better off asking someone who knows more than I do.

29. A significant problem in this course is being able to memorize all the information I need to know.

30. The main skill I get out of this course is to learn how to reason logically about the physical world.

31. I use the mistakes I make on homework and on exam problems as clues to what I need to do to understand the material better.

32. To be able to use an equation in a problem (particularly in a problem that I haven't seen before), I need to know more than what each term in the equation represents.

33. It is possible to pass this course (get a Cör better) without understanding physics very well.

34. Learning physics requires that I substantially rethink, restructure, and reorganize the information that I am given in class and/or in the text.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Julia Amedja

Berlin, 10.04.2008