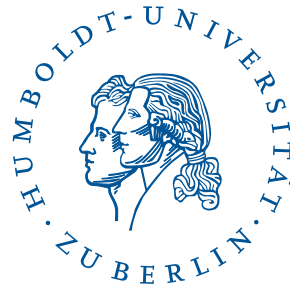


Masterarbeit
Untersuchung der Ergebnisse der Faktoranalyse
bei Anwendung auf ordinale Daten



zur Erlangung des Grades
Master of Science in Statistic

eingereicht bei:

Erstgutachter: Prof. Dr. Bernhard Rönz

Zweitgutachter: PD Dr. Marlene Müller

Betreuer: Dr. Sigbert Klinke

Student: Elke Pari Schatz(166685)

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

30. Oktober 2005

Erklärung

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

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

Elke Pari Schatz

Berlin, den

Inhaltsverzeichnis

1	Einleitung	6
2	Daten	8
2.1	Datenherkunft	8
2.2	Datenstruktur	9
2.2.1	Personencodierung und nominale Variablen	9
2.2.2	Ordinale Variablen	9
2.3	Datenaufbereitung	13
2.4	Fehlende Werte	14
2.4.1	Analyse fehlender Werte	14
2.4.2	Missing Data Algorithmus	22
3	Statistische Modelle	27
3.1	Explorative Faktoranalyse(EFA)	27
3.2	Underlying Variable Analyse(UVA)	32
4	Simulationen	35
4.1	Simulation unter der bivariaten Normalverteilung	37
4.2	Simulation unter der bivariaten Chiquadratverteilung	43
5	Modellanwendung	46
5.1	EFA und UVA	46
5.2	Vorgehen	51
5.3	Erklärungsverlust der Daten	55
6	Faktorladungen und Faktorwerte	57
6.1	Fragebereich 1	58
6.2	Fragebereich 2	60
6.3	Fragebereich 3	68
7	Fazit	73
8	Quellen	75
A	Appendix	77

Tabellenverzeichnis

1	Teilnahme der Lehrer an der Befragung je Schule	8
2	Aussagestruktur: „...zum gegenwärtigen Stand“	10
3	Aussagestruktur: „...wahrgenommene Veränderung“	10
4	5 Fragebereiche	11
5	Anzahl der tatsächlich angekreuzten Variablen f0201-f0212	13
6	Anzahl der tatsächlich angekreuzten Variablen f0403 und f0404	14
7	Spezifikation fehlender Werte	17
8	Erweiterte Kategorienanzahl der Aussagestruktur „...wahrgenommene Veränderung“	18
9	Anzahl der fehlenden Variablen pro Beobachtung	21
10	Bravais-Pearson Korrelationskoeffizienten	36
11	Polychorische Korrelationskoeffizienten	36
12	Bravais-Pearson Korrelationskoeffizienten, angewandt auf kategorisierte Daten	38
13	Prozentualer Anteil der signifikanten Variablenkombinationen (auf dem 5%-Niveau) zwischen verschiedenen Blockkombination	54
14	Verlorene Erklärungsanteile berechnet über Eigenwerte und Zusammenhangsmaße	56
15	Eigenwerte > 1, im Block 1	58
16	Faktorladungen des 1.Faktors des Variablenblocks 1, Bereich 1	58
17	Faktorladungen des 2. Faktors des Variablenblocks 1, Bereich 1	59
18	Eigenwerte > 1, im Block 5 und 7	60
19	Faktorladungen des 1. Faktors der Variablenblöcke 5 und 7, Bereich 2	61
20	Faktorladungen des 2. Faktors der Variablenblöcke 5 und 7, Bereich 2	63
21	Faktorladungen des 3.Faktors der Variablenblöcke 5 und 7, Bereich 2	64
22	Faktorladungen des 4.Faktors der Variablenblöcke 5 und 7, Bereich 2	65
23	Faktorladungen des 5.Faktors der Variablenblöcke 5 und 7, Bereich 2	66
24	Faktorladungen des 6.Faktors der Variablenblöcke 5 und 7, Bereich 2	67
25	Faktorladungen des (7.Faktors) der Variablenblöcke 5 und 7, Bereich 2	67
26	Korrelationen der jeweiligen Faktoren aus EFA und UVA	71
27	Kategorienaufteilung(symmetrisch)	78
28	Korrelation der simulierten kategorisierten Daten, n=400, 10.000 Durchläufe	79
29	Korrelation der simulierten kategorisierten Daten, n=400, 10.000 Durchläufe (Fortsetzung)	80
30	Korrelation der simulierten asymmetrisch klassifizierten Daten in 2 Kategorien, n=400, 10.000 Durchläufe	81
31	Eigenwerte der jeweils gemeinsam untersuchten Blöcke mit Werten > 1	82
32	Faktorladungen des 1.Faktors der Variablenblöcke 6 und 8, Bereich 2	83
33	Faktorladungen des 2.Faktors der Variablenblöcke 6 und 8, Bereich 2	83
34	Faktorladungen des 3.Faktors der Variablenblöcke 6 und 8, Bereich 2	83
35	Faktorladungen des 4.Faktors der Variablenblöcke 6 und 8, Bereich 2	84
36	Faktorladungen des 5.Faktors der Variablenblöcke 6 und 8, Bereich 2	84
37	Faktorladungen des 1.Faktors der Variablenblöcke 9 und 11, Bereich 3	84
38	Faktorladungen des 2.Faktors der Variablenblöcke 9 und 11, Bereich 3	85
39	Faktorladungen des 3.Faktors der Variablenblöcke 9 und 11, Bereich 3	85
40	Faktorladungen des 4.Faktors der Variablenblöcke 9 und 11, Bereich 3	85
41	Faktorladungen des 5.Faktors der Variablenblöcke 9 und 11, Bereich 3	85
42	Faktorladungen des 7.Faktors der Variablenblöcke 9 und 11, Bereich 3	86

43	Faktorladungen des 8.Faktors der Variablenblöcke 9 und 11, Bereich 3 . . .	86
44	Faktorladungen des 9.Faktors der Variablenblöcke 9 und 11, Bereich 3 . . .	87
45	Faktorladungen des 1.Faktors der Variablenblöcke 10 und 12, Bereich 3 . .	88
46	Faktorladungen des 2.Faktors der Variablenblöcke 10 und 12, Bereich 3 . .	89
47	Faktorladungen des 3.Faktors der Variablenblöcke 10 und 12, Bereich 3 . .	89
48	Faktorladungen des 4.Faktors der Variablenblöcke 10 und 12, Bereich 3 . .	90
49	Faktorladungen des 5.Faktors der Variablenblöcke 10 und 12, Bereich 3 . .	90

Abbildungsverzeichnis

1	Fehlende Werte der Blöcke 1 - 12 in %	15
2	Fehlende Werte der Blöcke 13 - 16 in %	16
3	Fehlende Werte der Blöcke 6 und 8 in %, nach Korrektur	18
4	Fehlende Werte der Blöcke 9 und 10 in %, nach Korrektur	19
5	Simulation, gleich große Klassen, 2-10 Kategorien	39
6	Simulation, versch. große Klassen, 2 Kategorien	41
7	rel. Anteil(linke Achse) an wahrer Korrelation(rechte Achse), verschieden große Klassen, 3 und 4 Kategorien, 400 Beobachtungen, 10.000 Durchläufe .	42
8	rel. Anteil(linke Achse) an wahrer Korrelation(rechte Achse), gleich große Klassen, 400 Beobachtungen, 10.000 Durchläufe, biv. χ^2 -Vtlg., mit FG(4,4)	44
9	rel. Anteil(linke Achse) an wahrer Korrelation(rechte Achse), gleich große Klassen, 400 Beobachtungen, 10.000 Durchläufe, biv. χ^2 -Vtlg. mit FG(4,100)	45
10	P-Values des Tests auf bivariate Normalverteilung	49
11	P-Values des Tests auf bivariate Normalverteilung(rechte und linke Achse: Variablenblöcke); ohne die Hauptdiagonale; PVALindep.R	53
12	Plot der Werte des 1. Faktors des Variablenblocks 1, Bereich 1, aus UVA(rechte Achse) und EFA(linke Achse)	59
13	Plot der Werte des 1. Faktors der Variablenblöcke 5 und 7, Bereich 2, aus UVA(rechte Achse) und EFA(linke Achse)	62
14	Plot des Faktors 2 in Block 5+7 aus UVA(rechte Achse) und EFA(linke Achse)	63
15	Plot der Werte des 3. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linke Achse)	64
16	Plot der Werte des 4. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linke Achse)	65
17	Plot des 5. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linke Achse)	66
18	Plots des Faktorwerte der Variablenblöcke 6+8	91
19	Plots des Faktorwerte der Variablenblöcke 9+11	92
20	Plots des Faktorwerte der Variablenblöcke 10+12	93
21	Fragebogen, Seite 1	94
22	Fragebogen, Seite 2	95
23	Fragebogen, Seite 3	96
24	Fragebogen, Seite 4	97
25	Fragebogen, Seite 5	98
26	Fragebogen, Seite 6	99
27	Fragebogen, Seite 7	100
28	Fragebogen, Seite 8	101
	.70.9 .30.9	

1 Einleitung

Der zugrundeliegende empirische Datensatz aus 443 Beobachtungen und 219 Variablen, der vom Institut für Erziehungswissenschaften erhoben wurde, wird zunächst auf fehlende Daten untersucht und gegebenenfalls aufbereitet.

Basismodelle, mit denen der aufbereitete Datensatz im nachfolgenden ausführlich untersucht wird, sind die Explorative Faktoranalyse(EFA) und die Underlying Variable Analyse(UVA). Beide Modelle gehören der übergeordneten Gruppe der Faktoranalyse an. Aus technischen Gründen, die später näher erläutert werden, ist ebenfalls die Verwendung der Confirmativen Faktoranalyse(CFA) notwendig.

Die Explorative Faktoranalyse als auch die Underlying Variable Analyse verfügen über die gleichen Ziele: Das Aufdecken von latenten Strukturen, die sich hinter einer Menge von Variablen verbergen, sowie die Dimensionsreduktion von Variablengruppen durch die Extraktion von Faktoren.

Der entscheidende Unterschied zwischen den Modellen der EFA und der UVA ist jedoch die Modellausrichtung: So ist die Explorative Faktoranalyse nach den Modellannahmen nur für metrisch skalierte Daten geeignet und die Underlying Variable Analyse wurde speziell für ordinale Daten konzipiert.

Daraus stellt sich die Frage, inwieweit sinnvolle Ergebnisse bei Anwendung der EFA auf ordinale Daten gegeben sind und inwieweit dann diese Ergebnisse mit denen der UVA vergleichbar sind, bei Anwendung auf den gleichen Datensatz, mit ordinalem Skalenniveau.

In der folgenden Analyse wird daher sowohl das Modell der Underlying Variable Analyse als auch das Modell der Explorativen Faktoranalyse auf den vorliegenden ordinalen Datensatz angewendet. Die Ergebnisse beider Studien werden anschließend miteinander verglichen.

Die Idee und das Vorhaben, auf ordinale Daten dafür ungeeignete statistische Methoden unzulässigerweise anzuwenden und zu untersuchen, ist nicht

neu: So liegen bereits zahlreiche Simulationsstudien zu diesem Gebiet vor (siehe Bollen und Barb(1981), Green und Rao(1970)). Auch im Rahmen dieser Arbeit werden noch einmal eigene angefertigte Simulationsstudien dieser Art angefertigt, um die Auswirkungen der Anwendung von „ungeeigneten Daten“ auf bestimmte statistische Methoden aufzuzeigen. Für die Analyse der Daten wurden folgende Software-Programme verwendet: Mplus 2.12, R 2.1.0, SAS 9.1 und SPSS 12.0.

2 Daten

2.1 Datenherkunft

Die der nachfolgenden Analyse zugrundeliegenden Daten wurden vom Institut für Erziehungswissenschaften der Humboldt-Universität zu Berlin erhoben. Das Institut hat einen Fragebogen konzipiert, der sich mit Aussagen über die Schulentwicklung an Berliner berufsbildenden Schulen befasst. Dieser Fragebogen wurde im Jahr 2004 an die Lehrer von 6 berufsbildenden Schulen in Berlin gerichtet.

Der Hintergrund dieser Befragung liegt auf der Erfassung der Einschätzung und Einstellung seitens der Lehrkräfte zu ihrer pädagogischen Arbeit als auch zu ihrer Schule zum Zeitpunkt der Befragung, da den Berliner Schulen durch eine Gesetzesänderung mehr Autonomie verliehen wurde. Die Teilnahme an dieser Befragung war freiwillig. Die Anteile der befragten Lehrer je Schule, sowie die, die an der Befragung auch tatsächlich teilgenommen haben, sind in folgender Tabelle 1 wiedergegeben:

Schule	Anzahl der befragten Lehrer Gesamt	Anzahl der befragten Lehrer Teilgenommen	Teilnahme in %
1	72	72	100.00
2	161	124	77.02
3	191	43	25.51
4	65	43	66.15
5	160	100	62.50
6	91	61	67.03

Tabelle 1: Teilnahme der Lehrer an der Befragung je Schule

2.2 Datenstruktur

Der Datensatz enthält in seiner ursprünglichen Form 443 Beobachtungen, die sich aus den 6 Schulen zusammensetzen und 218 Variablen. Die Daten weisen 3 Typen von Variablen auf:

- Variablen für die eindeutige Personencodierung(2)
- nominalskalierte Variablen(19)
- ordinalskalierte Variablen(197)

2.2.1 Personencodierung und nominale Variablen

Die Variablen der Personencodierung erfordern die Angabe des Geburtsdatums der Mutter ohne Jahresangabe, wie auch die ersten drei Buchstaben des Vornamens der Mutter.

14 der 19 nominalen Variablen sind binär, also dichotom. So lässt sich bei 14 Variablen für den Fall, dass die jeweilige Aussage zutrifft, ein Kreuz setzen. In jedem anderen Fall wird kein Kreuz gesetzt. Inhaltlich befassen sich die dichotomen Variablen überwiegend mit Aussagen zu Verbesserungsvorschlägen in der Tätigkeit als Lehrer. Andere inhaltliche Aussagen der nominalen Variablen beziehen sich auf allgemeine Angaben zur Person und Tätigkeit der Lehrer.

2.2.2 Ordinale Variablen

Die ordinalen Variablen machen mit 90% den größten Anteil von Variablen im gesamten Datensatz aus.

Es existieren zwei Aussagestrukturen in den ordinalen Variablen:

Endet eine Variable auf „...zum gegenwärtigen Stand“, besitzt die Variable 6 Kategorien, die in Tabelle 2 dargestellt sind:

Kategorie	1	2	3	4	5	6
Aussage	trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft über- wiegend zu	trifft völlig zu

Tabelle 2: Aussagestruktur: „... zum gegenwärtigen Stand“

In der weiteren möglichen Aussagestruktur endet die Variable auf „...wahrgenommene Veränderung“ und besitzt dann 4 Kategorien, die in Tabelle 3 dargestellt sind:

Kategorie	1	2	3	4
Aussage	negativer	eher negativer	eher positiver	positiver

Tabelle 3: Aussagestruktur: „... wahrgenommene Veränderung“

Das bedeutet, dass abhängig von der Endung der Aussage, eine 4 kategoriale oder eine 6 kategoriale Variable entsteht. Um eine eindeutigere Vorstellung dieser Variablenentstehung zu erhalten und um alle Variablen inhaltlich nachvollziehen zu können, sei im Appendix der gesamte Fragebogen angehängen.

Da die Anzahl der ordinalen Variablen einen großen Umfang besitzt, werden im folgenden bestimmte Gruppen von Variablen mit bestimmten Blocknummern bezeichnet. Zusätzlich lassen sich alle ordinalen Variablen übergeordneten Fragebereichen zuordnen. Die nachfolgende Tabelle 4 stellt den Zusammenhang zwischen den Variablen, den zugehörigen Blöcken und den Fragebereichen dar.

Die Inhalte der Fragebereiche sind:

1. Aussagen über die Arbeitsbedingungen in der Abteilung
2. Aussagen über die Schule als Institution
3. Aussagen über die Arbeitstätigkeit

Frage- bereich	6 Kategorien Name (Block)	Variablen- anzahl	4 Kategorien Name (Block)	Variablen- anzahl
1			f0101-f0111(1)	11
2	f0501-f0514(5)	14	f0601-f0614(6)	14
2	f0701-f0716(7)	16	f0801-f0816(8)	16
3	f0901-f0917(9)	16	f1001-f1017(10)	16
3	f1101-f1117(11)	17	f1201-f1217(12)	17
4	f1301-f1316(13)	16	f1401-f1416(14)	16
5	f1501-f1514(15)	14	f1601-f1614(16)	14

Tabelle 4: 5 Fragebereiche

4. Aussagen über die Schulleitung

5. Aussagen über das Schulprogramm.

Im nachfolgenden werden noch einmal alle ordinalen Variablen des Fragebogens mit ihren zugehörigen Fragebereichen dargestellt:

1. Fragebereich: Aussagen über die Arbeitsbedingungen in der Abteilung

Variablenanzahl: 11

Antwortmöglichkeiten: 4 Kategorien

Variablenkennziffern: f0101-f0111

Block: 1

2. Fragebereich: Aussagen über die Schule als Institution

Variablenanzahl: 30(6 Kategorien)/30(4 Kategorien)

Antwortmöglichkeiten:

(a) Bewertung des gegenwärtigen Zustandes: 6 Kategorien

Variablenkennziffern: f0501-f0514 und f0701-f0716

Block: 5 und 7

(b) Bewertung der wahrgenommenen Veränderung: 4 Kategorien

Variablenkennziffern: f0601-f0614 und f0801-f0816

Block: 6 und 8

3. Fragebereich: Aussagen über die Arbeitstätigkeit

Variablenanzahl: 33(6 Kategorien)/33(4 Kategorien)

Antwortmöglichkeiten:

(a) Bewertung des gegenwärtigen Zustandes: 6 Kategorien

Variablenkennziffern: f0901-f0917 [f0909 fehlt]

und f1101-f1117

Block: 9 und 11

(b) Bewertung der wahrgenommenen Veränderung: 4 Kategorien

Variablenkennziffern: f1001-f1017 [f1009 fehlt]

und f1201-f1217

Block: 10 und 12

4. Fragebereich: Aussagen über die Schulleitung

Variablenanzahl: 16(6 Kategorien)/16(4 Kategorien)

Antwortmöglichkeiten:

(a) Bewertung des gegenwärtigen Zustandes: 6 Kategorien

Variablenkennziffern: f1301-f1316

Block: 13

(b) Bewertung der wahrgenommenen Veränderung: 4 Kategorien

Variablenkennziffern: f1401-f1416

Block: 14

5. Fragebereich: Aussagen über das Schulprogramm

Variablenanzahl: 14(6 Kategorien)/14(4 Kategorien)

Antwortmöglichkeiten:

(a) Bewertung des gegenwärtigen Zustandes: 6 Kategorien

Variablenkennziffern: f1501-f1514

Block: 15

(b) Bewertung der wahrgenommenen Veränderung: 4 Kategorien

Variablenkennziffern: f1601-f1614

Block: 16

Im Rahmen der spezifischen Analyse dieser Arbeit werden allerdings nicht alle vorgestellten Variablen verwendet.

2.3 Datenaufbereitung

Es wird eine Datenaufbereitung im Hinblick auf die Variablen f0201-f0212 vorgenommen, wie auch der Variablen f0403 und f0404. Der Grund dafür ist, dass die Variablen in ihrer ursprünglichen Form binär kodiert sind.

Zusätzlich wird im Bereich der Variablen f0201-f0212 gefordert, dass nur 2 der 12 möglichen Variablen angekreuzt werden.

Um eventuell Strukturen in der Beantwortung der Fragen zu erkennen, werden die auftretenden Ankreuzmuster der Personen gebildet. Neben Personen, die innerhalb dieser zwölf Variablen gar kein Kreuz gesetzt haben, gibt es auch Personen, die nur ein Kreuz gesetzt haben und solche, die mehr als 2 Kreuze gesetzt haben.

Anzahl angekreuzter Variablen	0	1	2	3	4	6	7	8	9	10	11
Anzahl der Fälle	13	33	375	10	4	1	3	1	1	1	1

Tabelle 5: Anzahl der tatsächlich angekreuzten Variablen f0201-f0212

An der Tabelle 5 ist abzulesen, dass 46 Personen nicht die erforderlichen zwei Variablen aus 12 Variablen gewählt haben. Diese Personen haben entweder gar keine Variable oder nur eine Variable angekreuzt.

Diese 46 Beobachtungen, auf die 10.3% aller Daten entfallen, werden als fehlende Beobachtungen behandelt. Zusätzlich haben 22 Personen mehr als zwei Variablen gewählt. Um die Informationen dieser 22 Beobachtungen nicht

zu verlieren, wurde eine Umlegung über die Ankreuzmuster nach Häufigkeit vorgenommen.

Nach einer Analyse der Ankreuzmuster hat sich ergeben, dass 52 (verschiedene) gewählte Ankreuzmuster existieren. Entsprechend wurden die ursprünglich binär kodierten Variablen zu einer nominalen Variablen mit 52 Ausprägungen über die verschiedenen Ankreuzmuster umkodiert. Diese nun neu generierte Variable sei im folgenden mit dem Namen f0200 benannt.

Ein ähnliches Vorgehen wird für die Variablen f0403 und f0404 vorgenommen. Über die Fragestellung ist es hierbei erlaubt eine der zwei Variablen oder beide binären Variablen zu wählen. Entsprechend existieren 3 verschiedene Ankreuzmuster. 48 Personen haben auf keine der beiden Variablen geantwortet, also 10.8%. Diese 48 Personen werden als fehlende Daten behandelt. Auch aus dieser binären Variable wurde eine nominale Variable entwickelt

Anzahl angekreuzter Variablen	0	1	2
Fälle	48	158	237

Tabelle 6: Anzahl der tatsächlich angekreuzten Variablen f0403 und f0404

mit 3 Ausprägungen, die im folgenden mit f0400 bezeichnet sei.

(Alle Ankreuzmuster der neugenerierten Variablen f0200 und f0400 mit entsprechenden Häufigkeiten sind auf der beiliegenden CD verfügbar.)

2.4 Fehlende Werte

2.4.1 Analyse fehlender Werte

Allison(2001) bemerkte einst, „...dass die beste Lösung für das Problem fehlender Werte sei, keine fehlenden Werte zu haben“.

Arbeitet man mit empirischen Datensätzen, finden sich in der Regel immer fehlende Werte in dem zu untersuchenden Datensatz. So enthält auch der vorliegende Datensatz fehlende Werte, teilweise in erheblichem Ausmaß,

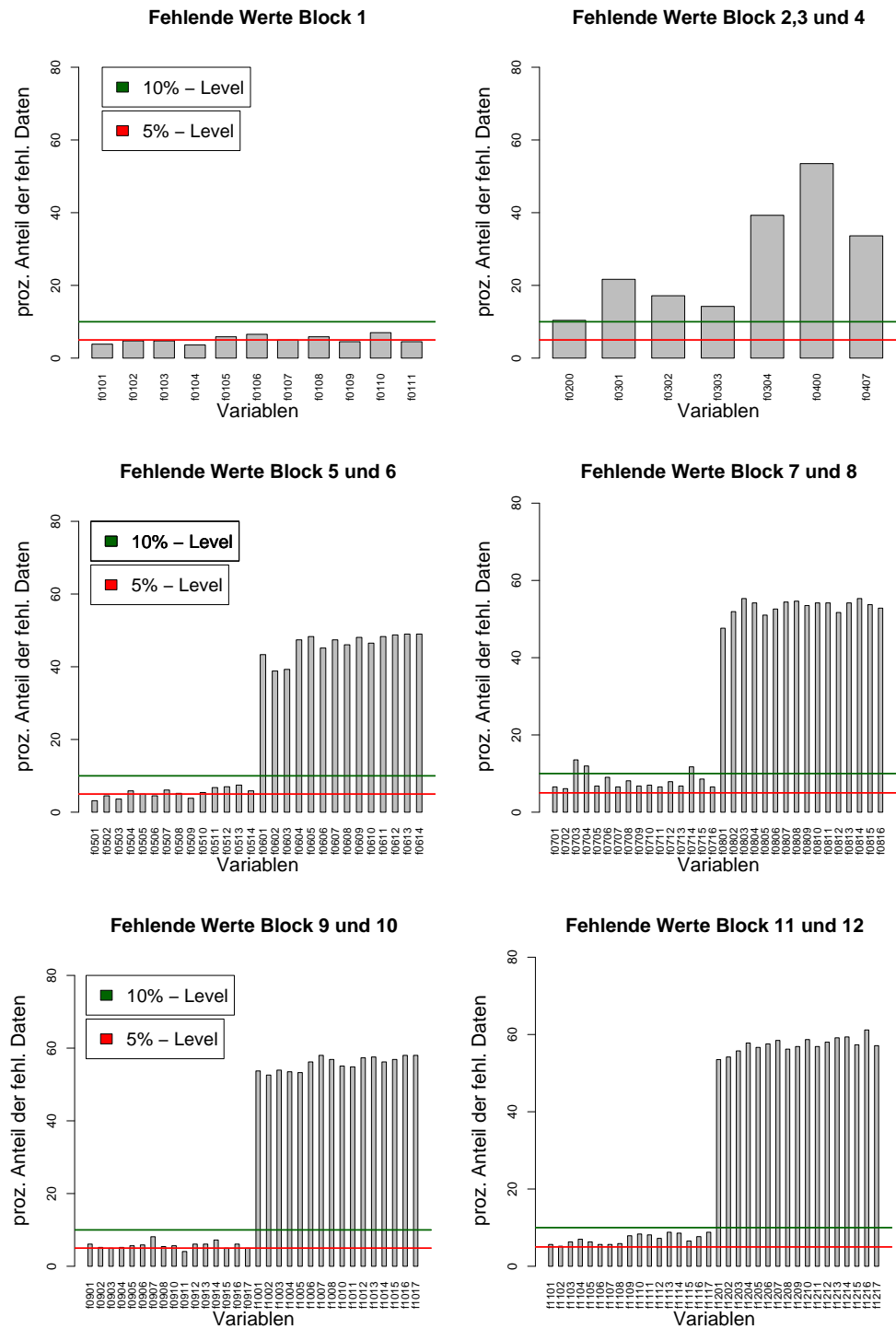


Abbildung 1: Fehlende Werte der Blöcke 1 - 12 in % (rechte Achse: Variablen, linke Achse: proz. Anteil)

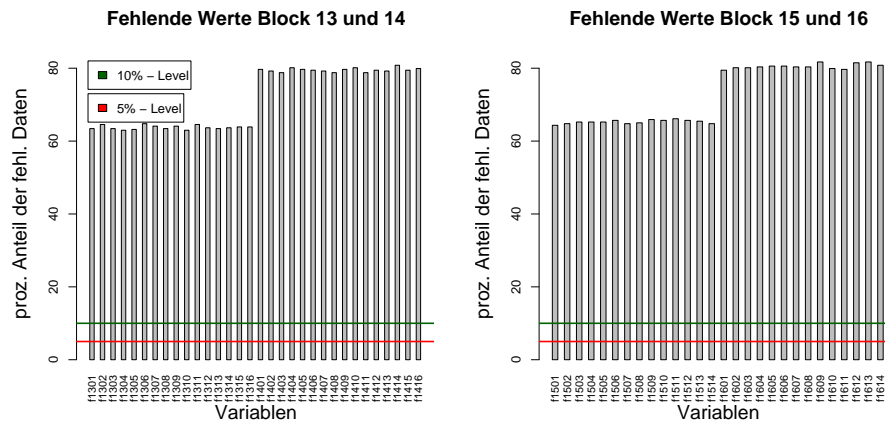


Abbildung 2: Fehlende Werte der Blöcke 13 - 16 in % (rechte Achse: Variablen, linke Achse: proz. Anteil)

wie aus den Abbildungen 1 und 2 ersichtlich ist. Hinzu kommt, dass der vorliegende Datensatz bereits a priori von einem außeruniversitären Institut vorkodiert wurde. Die ausgefüllten Fragebögen sind in diesem Institut verblieben. Es besteht daher kein Zugang mehr zu diesen Fragebögen. Dieses Institut hat drei Arten von fehlenden Daten spezifiziert: Zum einen wurde die Variable einer Beobachtung, also einer Person, mit dem Wert „0“ versehen, wenn diese keine der möglichen Antwortkategorien gewählt hat. Wenn sich eine Person entgegen der Vorgabe für mehrere Antwortkategorien entschieden hat, wurde die Variable bei der jeweiligen Beobachtung mit einer „9“ versehen.

Die „9“ wird daher ebenfalls als fehlende Variable einer Beobachtung angesehen, weil keine exakte Zuordnung zu einer der möglichen Kategorien möglich ist. Da leider kein Zugriff mehr auf das Material der Originaldaten besteht, ist eine Aufschlüsselung der als „9“-kodierte Variablen pro Beobachtung nicht mehr möglich. Daneben hat das Institut bei Schulen, die vorsätzlich und geschlossen die Aussage zu bestimmten Variablen verweigern, die entsprechenden Variablen dieser Fälle mit keinerlei Wertung versehen,

sondern nur mit einem Punkt.

Fehlend, weil	keine Wahl	zu häufige Wahl	Aussageverweigerung
Werte	„0“	„9“	„ . “

Tabelle 7: Spezifikation fehlender Werte

Im Rahmen der genaueren Analyse der Daten ist zu erkennen, dass bei einigen Beobachtungen und Variablen vermehrt fehlende Werte auftreten. Aus der Abbildung 2 ist abzulesen, dass Variablen mit mehr als 80% fehlender Daten im Datensatz vorhanden sind.

Bevor weitere statistische Analysen mit den Daten vorgenommen werden können, muss zunächst eine Datenkorrektur vorgenommen werden. Dabei besteht ein Großteil dieser fehlenden Daten aus dem Wert „0“ und „ . “. Der Block 1 weist innerhalb seiner 11 Variablen nur Variablen mit weniger als 10% fehlender Daten auf, wie anhand der Abbildung 1 zu erkennen ist. Dieser Anteil an fehlenden Daten ist noch akzeptabel. Da die Variablen aus Block 3 und 4 zwischen 14.22% und 53.50% fehlende Daten aufweisen, werden diese von nachfolgenden Analysen ausgeschlossen. Es ist jedoch ersichtlich, dass noch weitere Variablen hohe Anteile fehlender Daten aufweisen.

In Abbildung 1 ist eine gewisse Systematik zwischen den Variablenblöcken 5 und 6, 7 und 8, 9 und 10, sowie 11 und 12 zu erkennen. Es besteht durchgängig ein großes Gefälle zwischen den jeweiligen Blöcken, im Hinblick auf die fehlenden Daten. Während sich bei den Blöcken 5, 7, 9 und 11 eine prozentuale Spanne zwischen 3.16% und 13.54% an fehlenden Werten beobachten lässt, weisen die geradzahligen Blöcke 6, 8, 10 und 12 eine Spanne zwischen 38.82% und 61.17% an fehlenden Daten auf.

Wie bereits oben erklärt, verfügen die geradzahligen Variablenblöcke ab Block 6 über 4 mögliche Antwortkategorien. Die ungeraden Variablenblöcke ab Block 5 besitzen 6 Kategorien. Bei genauerer Analyse und der Berücksich-

tigung der Aussagen der Antwortmöglichkeiten liegt folgende Vermutung nahe: In den 4 Antwortmöglichkeiten ist es nur möglich über „wahrgenommene Richtungsänderungen“ eine Aussage zu machen. Wenn nun aber der Befragte keinerlei Richtungsänderung wahrnehmen kann, wird er folglich keine der Antworten ankreuzen. Über diese Interpretation wäre der hohe Anteil fehlender Daten in den Blöcken 6, 8, 10 und 12 zu erklären. Damit diese Variablen nicht von der späteren Analyse ausgeschlossen werden müssen, besteht die Idee, die zusätzliche Kategorie „keine wahrgenommene Veränderung“ einzuführen, siehe Tabelle 8.

Kategorie	1	2	3	4	5
wahrgenommene Veränderung	negativer	eher negativer	keine Veränderung	eher positiver	positiver

Tabelle 8: Erweiterte Kategorienanzahl der Aussagestruktur „... wahrgenommene Veränderung“

Nach der Erweiterung der Kategorien in den Blöcken 6, 8, 10 und 12 ergibt sich nachfolgende Struktur der fehlenden Daten, siehe Abbildung 3 und 4.

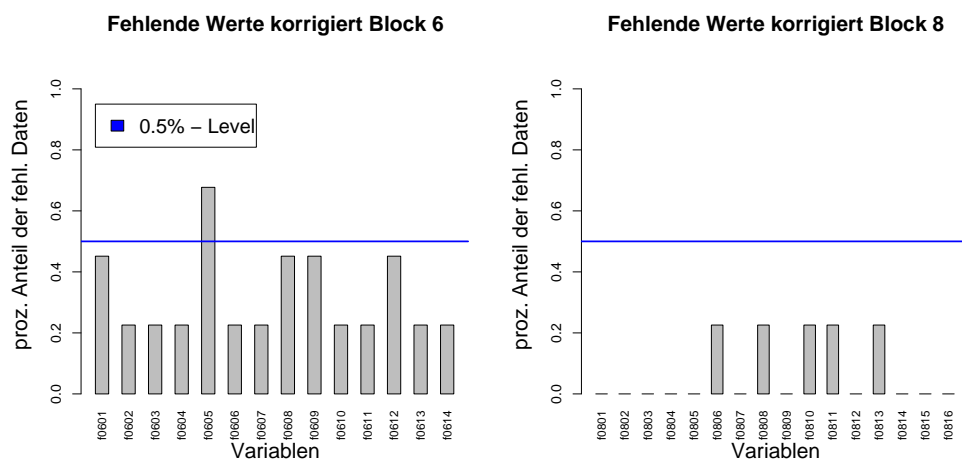


Abbildung 3: Fehlende Werte der Blöcke 6 und 8 in %, nach Korrektur

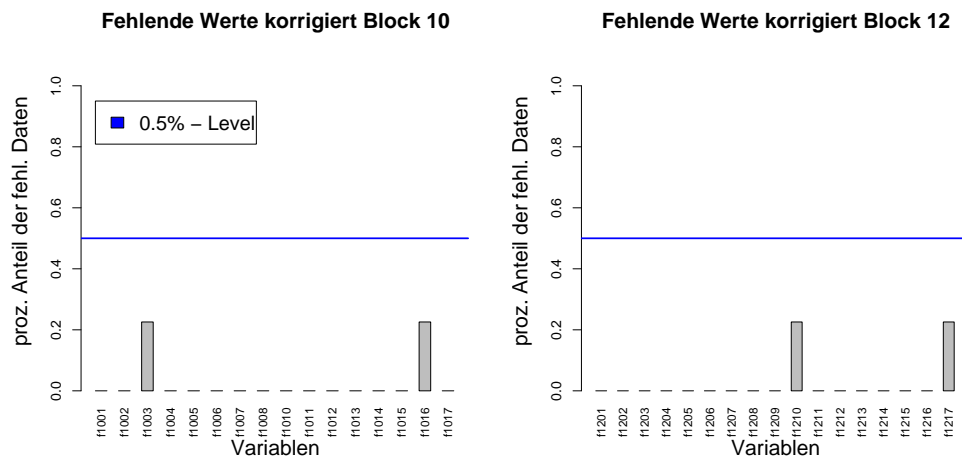


Abbildung 4: Fehlende Werte der Blöcke 9 und 10 in %, nach Korrektur

Es ist zu beobachten, dass kaum noch fehlende Daten auftreten. Aber mit der Datenkorrektur wurde auch der Anteil fehlender Daten mit dieser neuen Kategorie „keine wahrgenommene Veränderung“ gleichgesetzt, der tatsächlich „wahre fehlende Daten“ darstellt. Damit sei der Anteil von Personen benannt, die mit dem Nichtankreuzen keine Aussage tätigen wollten. Nach der Korrektur der Blöcke 6, 8, 10 und 12 liegen die Anteile fehlender Werte zwischen 0% und 0.68%, wie anhand der Abbildungen 3 und 4 nachzuvollziehen ist. Dies erscheint im Vergleich zu den Ergebnissen der fehlenden Daten in den Blöcken 5, 7, 9 und 11 jedoch unplausibel. Es besteht kein ersichtbarer Grund dafür, dass die Personen hier keine „wahren fehlenden Daten“ hinterlassen sollten. Da der Anteil der Beobachtungen, die mit dem Nichtankreuzen keine Aussage tätigen wollten, nicht messbar ist, sei an dieser Stelle unbedingt darauf hingewiesen, dass eine Verfälschung der Aussagen durch das beschriebene Vorgehen gegeben sein kann. In späteren Analysen wird mit dem veränderten Datenmaterial dennoch gearbeitet, um nicht weitere Informationen zu verlieren.

Damit sind jedoch noch nicht alle Charakteristika der fehlenden Daten er-

klärt. Die Variablenblöcke 13, 14, 15 und 16 weisen durchgängig fehlende Daten in einer Spannweite von 62.98% bis 81.72% auf. Bei genauerer Analyse der Beobachtungen lässt sich allerdings feststellen, dass 3 der 6 befragten Schulen mit 267 Personen, vorsätzlich keine Aussage zu den Variablenblöcke 13, 14, 15, 16 gemacht haben. Dies entspricht einem Anteil von etwa 60% der Lehrer insgesamt.

Die Variablen der Beobachtungen der betroffenen Schulen sind an dieser Stelle mit einem „.“ kodiert. Aufgrund der hohen Anteile fehlender Daten, die aus der beabsichtigten Vernachlässigung dieser Variablenblöcke durch 3 Schulen entstanden, werden die Variablenblöcke 13 - 16 von der nachfolgenden Analyse ausgeschlossen.

Im Rahmen der Analyse fehlender Daten müssen ebenfalls die fehlenden Daten pro Beobachtung untersucht werden, um hier gegebenenfalls Beobachtungen mit sehr hohen Anteilen fehlender Daten auszuschließen. Es erfolgt eine Betrachtung der fehlenden Daten pro Beobachtung, ohne die Variablenblöcke 3, 4 und 13-16.

Über die Entwicklung in Tabelle 9 wird entschieden, alle Beobachtungen mit mehr als 22 „Missings“ pro Beobachtung von der Analyse auszuschließen. Dieser Sachverhalt entspricht 27 Beobachtungen. Denn ab 22 „Missings“ pro Beobachtung wird die Anzahl der betroffenen Beobachtungen immer weniger und die Häufigkeit der „Missings“ pro Beobachtung steigt danach überproportional an.

Da die generierte Variable f0200 mit 52 Ausprägungen nominal, nur schwer mit den anderen verbleibenden Variablen vergleichbar ist und für nachfolgende Darstellungen keine Verwendung mehr findet, wird auch diese Variable von der folgenden Analyse ausgeschlossen.

Für alle weiteren Analysen in dieser Arbeit verbleiben somit 416 Beobachtungen und 137 Variablen.

Fehlende Werte pro Beob.	0	1	2	3	4	5	6	7	8	9
Anzahl der Fälle	220	81	46	18	11	5	1	2	4	3
Fehlende Werte pro Beob.	10	11	12	13	14	15	16	17	18	19
Anzahl der Fälle	1	4	2	2	1	2	2	1	2	2
Fehlende Werte pro Beob.	20	21	22	23	25	26	29	31	33	34
Anzahl der Fälle	1	3	2	1	1	3	1	1	1	3
Fehlende Werte pro Beob.	42	43	48	49	54	64	65	71	73	74
Anzahl der Fälle	1	1	1	1	1	1	1	1	1	7

Tabelle 9: Anzahl der fehlenden Variablen pro Beobachtung

2.4.2 Missing Data Algorithmus

Nach der oben beschriebenen Analyse und Bereinigung des Datensatzes um Variablen und Beobachtungen mit hohen Anteilen fehlender Daten verbleiben dennoch in einem gewissen Ausmaß fehlende Daten. Um alle verbleibenden Beobachtungen und Variablen trotzdem vollständig berücksichtigen zu können, soll daher ein Missing Data Algorithmus verwendet werden: Multiple Random Imputation(MRI). Dafür wird angenommen, dass die Daten MAR sind, also „Missing at Random“.

Die Annahme „Missing at Random“ bedeutet, dass die fehlenden Daten unabhängig von den Werten sind, die fehlen. Diese Annahme ist prinzipiell nicht überprüfbar und dient nur der formalen Vollständigkeit, siehe Schafer (1997).

Auf der Basis von IVEware unter SAS wird ein Imputationsalgorithmus über das Multinomiale Logit Modell verwendet. IVEware wurde von den Wissenschaftlern Raghunathan, Lepkowski, van Hoewyk und Solenberger(2001) entwickelt und im Jahr 2002 für die Verwendung bereitgestellt.

Die Imputation Der Imputationsprozess erfolgt nach folgendem Schema:

Es sei n die Größe des Stichprobenumfanges. $X(n \times p)$ sei eine Submatrix des gesamten Datensatzes G . Im vorliegenden Datensatz besitzt X die Größe (416×64) . Diese Matrix X enthält dabei ausschließlich die Variablen des gesamten Datensatzes $G(416 \times 137)$, die keine fehlenden Werte aufweisen.

Daneben existiert die Submatrix Y aus G , die ein monotonen Pattern darstellt. Y habe in bezug auf den zugrundeliegenden Datensatz die Größe (416×73) . Die Variablen Y_1, Y_2, \dots, Y_{73} seien aufsteigend nach der Anzahl ihrer fehlenden Werte geordnet.

Es folgt eine modellbasierte Imputation. Dabei kann die gemeinsame bedingte

Dichte von Y_1, Y_2, \dots, Y_{73} , gegeben X , wie folgt faktorisiert werden:

$$\begin{aligned} f(Y_1, Y_2, \dots, Y_{73}|X, \delta_1, \delta_2, \dots, \delta_{73}) = \\ f_1(Y_1|X, \delta_1) \cdot f_2(Y_2|X, Y_1, \delta_2) \cdot \dots \\ f_{73}(Y_{73}|X, Y_1, Y_2, \dots, Y_{72}, \delta_{73}) \end{aligned} \quad (1)$$

dabei seien $f_t, t = 1, 2, \dots, 73$ die bedingten Dichtefunktionen und δ_t ein Vektor von Parametern in der bedingten Dichtefunktion, wie zum Beispiel den Regressionskoeffizienten. Damit ist es möglich, jede beliebige bedingte Dichtefunktion durch ein passendes Regressionsmodell mit den unbekannten Parametern δ_t anhand von Ziehungen aus der zugehörigen „predictive distribution“ der fehlenden Daten und den beobachtbaren Daten zu modellieren. Es wird dafür angenommen, dass für die „prior distribution“ der Parameter gilt $\delta = (\delta_1, \delta_2, \dots, \delta_{73}) = \pi(\delta) \propto 1$ sei. Da die Daten ordinal sind, verwendet IVEware das Multinomiale Logit Modell.

Jede Imputation beinhaltet mehrere Runden. Durch jeden Imputations-Durchlauf wird eine Y_t -Variable in die Matrix X aufgenommen. In der ersten Runde wird zunächst Y_1 , also die Variable, die am wenigsten fehlende Daten aufweist, auf X regressiert. X sei dann als Matrix U bezeichnet, also die zuletzt erweiterte X Matrix. Die aufgenommene Variable Y_1 wird anschließend vervollständigt. Über das Multinomiale Logit Modell werden die fehlenden Werte in der Variable Y_1 ersetzt.

Der exakte Ablauf gliedert sich wie folgt: Es existiert die kategoriale Variable Y_t mit K Ausprägungen, $k = 1, \dots, K$. Es sei

$$\pi_k = P(Y_t = k|U). \quad (2)$$

Es wird angenommen, dass die abhängige Variable Y_t sich aus dem logarithmierten Verhältnis aus π_k , also der Wahrscheinlichkeit des Auftretens der k -ten Kategorie, zu π_K , der Wahrscheinlichkeit für das Auftreten der K -ten

Kategorie und hier die Referenzkategorie, ergibt, in Abhängigkeit aller erklärenden Variablen U :

$$\log\left(\frac{\pi_k}{\pi_K}\right) = U\beta_k \quad (3)$$

für $k = 1, \dots, K - 1$.

Die Referenzkategorie π_K sei unter der Annahme, dass $\sum_k^K \pi_k = 1$ gilt, in folgender Weise definiert:

$$\pi_k = [1 + \sum_{k=1}^{K-1} \exp\{U\beta_k\}]^{-1} \quad (4)$$

B sei die Schätzung des Maximum-Likelihood Schätzers der Regressionskoeffizienten $(\beta_1^\top, \beta_2^\top, \dots, \beta_{K-1}^\top)$ und T die Cholesky Zerlegung,

$$TT^\top = (U^\top U)^{-1}. \quad (5)$$

Für die Imputation sind dann nach Raghunathan, Lepkowski, van Hoewyk und Solenberger(2001) detailliert folgende 3 Schritte erforderlich:

1. Es sei $\beta_* = B + Tz$, mit z als Vektor von Standardabweichungen einer normalverteilten Zufallsstichprobe.
2. U_{mis} seien diejenigen Zeilen von U , die fehlende Werte enthalten durch die Aufnahme von Y_t . Die bedingte Wahrscheinlichkeit für das Auftreten der i -ten Kategorie der Responsevariablen $Y_{mis,t}$ für die 1 bis $K-1$ Kombinationen der erklärenden Variablen U_{mis} sei:

$$P_i^* = \exp\{U_{mis}\beta_{*i}\} / \{1 + \sum_i \exp(U_{mis}\beta_{*i})\}, \quad (6)$$

wobei β_{*i} der Regressionskoeffizient der jeweiligen Kategorie i ist mit $i = 1, \dots, K - 1$ ist und $P_K^* = 1 - \sum_i P_i^*$.

3. $R_0 = 0, R_k = \sum_i^k P_i^*$ und $R_K = 1$ seien die kumulativen Summen der Wahrscheinlichkeiten. Um die fehlenden Werte zu ersetzen, werden gleichverteilte Zufallszahlen u erzeugt. Es wird die Kategorie k gewählt, als der zu imputierende Wert, falls gilt $R_{k-1} \leq u \leq R_k$.

Nach diesem 3-stufigen Ablauf besteht nun eine um eine Variable erweiterte Matrix X , also U . Danach wird die Variable Y_2 , mit der nächst größeren Anzahl von fehlenden Variablen, auf X regressiert und somit auch Y_2 nun vervollständigt. Dieser Ablauf erfolgt solange, bis alle Variablen des gesamten Datensatzes keine fehlenden Werte mehr enthalten. Je nach weiterer Handhabung, ist es natürlich möglich, sich mehrere imputierte Datensätze zu erzeugen. Diese sind nach dem oben beschriebenen Ablauf natürlich nicht identisch.

So wurden mit IVEware auf beschriebene Weise 10 imputierte Datensätze aus dem zugrundeliegenden (unvollständigen) Datensatz heraus generiert.

Data Augmentation Die nachfolgende Behandlung dieser imputierten Datensätze soll über Data Augmentation erfolgen: Es sei θ ein Parameter der Grundgesamtheit. Dabei sei $\hat{\theta}$ ein Schätzer für θ . Dieser Schätzer wäre für den Fall vollständiger Daten verwendet worden. Die Schätzfunktion $\hat{\theta}$ muss wenigstens approximativ normalverteilt sein.

$$\frac{(\hat{\theta} - \theta)}{\sqrt{Var(\hat{\theta})}} \stackrel{\text{as.}}{\sim} N(0, 1) \quad (7)$$

Aus D ergänzten Datensätzen können dann D Schätzer $\hat{\theta}_d$ und ihre Varianz $Var(\hat{\theta}_d)$ berechnet werden, für $d = 1, \dots, D$. Der Multiple Random Imputation(MRI) Schätzer $\hat{\theta}_{MRI}$ ergibt sich aus dem einfachen ungewichteten Mittel aller Schätzergebnisse.

Entsprechend sollen die vorgestellten Methoden nun auf den zugrundeliegenden 10-fach imputierten Datensatz angewandt werden.

Es besteht allerdings ein technisches Problem bei der Anwendung des MRI-Schätzers im Rahmen der verwendeten Software: Die zentralen MRI-Schätzer stellen in den nachfolgenden Analysen die Matrizen der Zusammen-

hangsmaße der Pearson Korrelation und der polychorischen Korrelation¹ dar (für jeweils 2 Variablen x_j und x_h mit $j, h=1, \dots, 137$). Da in späteren Analysen mit verschiedenen Software-Systemen gearbeitet wird, die nicht die konkrete Eingabe des MRI-Schätzers $\hat{\rho}_{x_j x_h}^{MRI}$ erlauben, muss ein einfachimputierter Datensatz, Single Random Imputation(SRI), verwendet werden. Es ist zu erwarten, dass der MRI-Schätzer $\hat{\rho}_{x_j x_h}^{MRI}$ bessere Ergebnisse liefert, als ein SRI-Schätzer $\hat{\rho}_{x_j x_h}^{SRI}$. Um zu überprüfen, inwieweit Abweichungen bestehen, wird auf dem 5%-Niveau ein Test auf Gleichheit der Korrelation durchgeführt.

Diese Überprüfung wird mit Hilfe des Tests auf Gleichheit der Korrelationen über die Fisher-Transformation durchgeführt, nach Härdle und Simar(2003, S.88):

$$H_0 : \rho_{x_j x_h}^{MRI} = \rho_{x_j x_h}^{SRI} \quad (8)$$

$$H_1 : \rho_{x_j x_h}^{MRI} \neq \rho_{x_j x_h}^{SRI}$$

Die Teststatistik lautet:

$$W = \frac{1}{2} \log \left(\frac{1 + r_{x_j x_h}^{MRI}}{1 - r_{x_j x_h}^{MRI}} \right) \quad (9)$$

Der approximative Erwartungswert und Varianz von W seien:

$$E(W) \approx \frac{1}{2} \log \left(\frac{1 + r_{x_j x_h}^{SRI}}{1 - r_{x_j x_h}^{SRI}} \right) \quad (10)$$

$$Var(W) \approx \frac{1}{(n-3)} \quad (11)$$

$$Z = \frac{W - E(W)}{\sqrt{Var(W)}} \xrightarrow{\mathcal{L}} N(0, 1) \quad (12)$$

Der Test wird zunächst mit den Schätzergebnissen der Bravais-Pearson Korrelation aus MRI und SRI durchgeführt. Dabei sei $\rho_{x_j x_h}$ die wahre Korrelation zwischen zwei Variablen x_j und x_h und $r_{x_j x_h}$ der Bravais-Pearson

¹ diese Korrelation sei später ausführlich definiert

Korrelationskoeffizient mit der Formel:

$$r_{x_j, x_h} = \frac{Cov(x_j, x_h)}{Var(x_j)Var(x_h)}, \quad (13)$$

mit $-1 \leq r_{x_j, x_h} \leq +1$.

Der kleinste p-Value in %, den man aus dem beschriebenen Test erhält, für alle möglichen Variablenkombinationen, ist 26.77%. Der kleinste p-Value aus dem Test, bezogen auf die polychorische Korrelation, sei 25.57%. Das Testergebnis ist, dass sich die verschiedenen Korrelationsmatrizen aus SRI und MRI nicht wesentlich voneinander unterscheiden (siehe p-Values im CD-Anhang). Daher wird in der folgenden statistischen Analyse mit dem einfachimputierten Datensatz aus SRI gearbeitet.

3 Statistische Modelle

Wie bereits anfänglich erläutert, werden im Verlauf der Analyse dieser Arbeit sowohl das Underlying Variable Modell, dass für ordinale Daten konzipiert wurde, als auch das Modell der Explorativen Faktoranalyse auf die aufbereiteten ordinalen Daten angewandt. Vollständigkeitshalber werden beide Modelle jedoch vorerst noch einmal vorgestellt und die fundamentalen Unterschiede beider Modelle herausgearbeitet.

3.1 Explorative Faktoranalyse(EFA)

Die Ausgangsgleichung des Modells der Explorativen Faktoranalyse sei:

$$Z_{(n \times m)} = F_{(n \times Q)} A_{(Q \times m)}^T + U_{(n \times m)} E_{(m \times m)} \quad (14)$$

mit Beobachtungen $i=1, \dots, n$, Variablen $j=1, \dots, m$ und Faktoren $q=1, \dots, Q$.

Die einzelnen Variablen des Modells seien wie folgt definiert(nach Rönz (2000), S.187ff):

- Z: Matrix der standardisierten Variablen z_j (vor der Standardisierung mit x_j benannt).
- F: Matrix der Faktorwerte f_{iq} jedes Faktors F_q für jeden Fall i bei jeder Variablen Z_j . Dabei sind die Werte der Faktoren und die Anzahl Q der Faktoren unbekannt und müssen geschätzt werden. Darüber hinaus wird gefordert, dass jeder Faktor mindestens auf zwei Variablen wirken soll. Daher werden diese Faktoren auch als gemeinsame Faktoren bezeichnet.
- A: Matrix der Faktorladungen. Die einzelnen Faktorladungen a_{jq} repräsentieren Gewichte, mit denen die jeweiligen Faktoren in die Variablen Z_j eingehen. Die Werte der Matrix A müssen ebenfalls geschätzt werden.
- U und E: Analog zur Idee im linearen Regressionsmodell wird eine Störgröße ε_j je Variable aufgenommen, die sich wie folgt zusammensetzt:

$$\varepsilon_j = e_j \cdot U_j \quad (15)$$

Mit U_j als sogenannter Einzelrestfaktor und e_j als Ladung des Einzelrestfaktors U_j .

Die Voraussetzungen für Daten, auf die EFA angewandt werden, sind :

1. metrisch skalierte Variablen
2. unabhängige Beobachtungen
3. (approximativ)normalverteilte Variablen
4. großer Stichprobenumfang

Die Annahmen, die hinsichtlich der Faktoren im EFA-Modell getroffen werden sind:

1. Wirkung der gemeinsamen Faktoren auf mindestens zwei Variablen
2. Unkorreliertheit der gemeinsamen Faktoren untereinander

3. Unkorreliertheit der Einzelrestfaktoren untereinander
4. Unkorreliertheit der Einzelrestfaktoren mit den gemeinsamen Faktoren

Da alle Variablen auf der rechten Seite der Gleichung 14 unbekannt sind, erfolgt die Herleitung der unbekannten Variablen über die einzig bekannte Variable der Gleichung 14, Z . Es wird eine Korrelationsmatrix gebildet, aus dem Bravais-Pearson Korrelationskoeffizienten zwischen den standardisierten Variablen:

$$r_{jh} = \frac{1}{n-1} \sum_{i=1}^n z_{ij} z_{ih}; j, h = 1, \dots, m \quad (16)$$

bzw.

$$R_{(m \times m)} = Z_{(m \times n)}^\top Z_{(n \times m)} / (n-1) \quad (17)$$

Über Gleichung 14 und R aus Gleichung 17 wird anschließend die Fundamentalgleichung der Explorativen Faktoranalyse entwickelt,

$$R = (FA^\top + UE)^\top (FA^\top + UE) / (n-1) \quad (18)$$

die sich bei geltenden Modellannahmen an die Faktoren wie folgt ergibt:

$$R = AA^\top + EE \quad (19)$$

bzw. für jeweils zwei Variablen auf der Hauptdiagonalen:

$$r_{jj} = \sum_{q=1}^Q a_{jq}^2 + e_j^2 = 1, \quad \text{mit } j = 1, \dots, m. \quad (20)$$

und zwei Variablen außerhalb der Hauptdiagonalen:

$$r_{jh} = a_{j1}a_{h1} + \dots + a_{jq}a_{hq} + \dots + a_{jQ}a_{hQ} + \nu_{jh}e_je_h$$

mit $\nu_{jh}=0$, falls $j \neq h$; $j, h=1, \dots, m$ und $q=1, \dots, Q$.

Da die Beobachtungen standardisiert sind,

ist die Varianz jeder Variablen 1,

$$r_{jj} = \text{Cov}(Z_j, Z_j) = \text{Var}(Z_j) = 1 \quad (21)$$

Die Bravais-Pearson Korrelation zwischen jeweils zwei Variablen ergibt sich somit stets durch die Gleichung:

$$r_{jh} = \frac{Cov(Z_j, Z_h)}{Var(Z_j)Var(Z_h)} = Cov(Z_j, Z_h) \quad (22)$$

Nach der Gültigkeit von Gleichung 20 stellt a_{jq}^2 den Erklärungsbeitrag des Faktors F_q an der Varianz $Var(Z_j)$ dar.

Die Korrelationsmatrix ist damit der zentrale Mittelpunkt der Faktoranalyse. Abhängig von dieser Matrix wird über die Güte der geschätzten Werte F und A bestimmt. Somit rückt diese Matrix in späteren Simulationen in den Focus der Untersuchung.

Stellt man die Formel 20 um, erhält man

$$h_j^2 = \sum_{q=1}^Q a_{jq}^2 = 1 - e_j^2, j = 1, \dots, m. \quad (23)$$

Dabei sei h_j^2 die Summe der Varianzbeiträge der gemeinsamen Faktoren. Diese werden als Kommunalitäten bezeichnet. Im nächsten Schritt der Faktoranalyse wird die Reproduzierte Korrelationsmatrix R_h gebildet:

$$R_h = R - EE = AA^\top. \quad (24)$$

Dabei befinden sich auf der Diagonalen der Matrix R_h die geschätzten Kommunalitäten \hat{h}_j^2 und außerhalb der Diagonalen die bekannten Bravais-Pearson Korrelationskoeffizienten.

Da es für die Gleichung 24 unendlich viele Lösungsmöglichkeiten gibt, muss die Gleichung in einem iterativen Prozess geschätzt werden. In dem iterativen Prozess wird mit vorgegebenen Anfangswerten begonnen. Dabei werden diese im Verlauf der Faktorextraktion verbessert, bis ein Genauigkeitskriterium erfüllt worden ist. In dieser der Arbeit zugrundeliegenden Analyse sollte zunächst die Hauptachsenmethode verwendet werden. In der Hauptachsenmethode wird davon ausgegangen, dass ein Einzelrestfaktor existiert. Tatsächlich wurde in der Analyse die Hauptkomponentenanalyse gewählt,

da sich gezeigt hat, dass die Ladungshöhe der Variablen bei den jeweiligen Faktoren eindeutiger sind, als im Fall der Hauptachsenmethode.

Die Hauptkomponentenanalyse geht davon aus, dass kein Einzelrestfaktor besteht. Daraus folgt, dass die Anfangswerte im iterativen Prozess, also die Ausgangskommunalitäten, stets 1 sind. Am Ende des iterativen Prozesses werden die geschätzten Kommunalitäten kleiner als 1 sein, aufgrund des Fehlens von gemeinsamen Faktoren.

Die Bestimmung der Anzahl der Faktoren wird in der Analyse dieser Arbeit über das Kaiser-Kriterium entschieden. Das Kriterium besagt, dass die Anzahl der Faktoren durch die Anzahl von Eigenwerten der Matrix R_h bestimmt sei, die größer als 1 sind. Dies würde aufgrund der standardisierten Variablen bedeuten, dass der extrahierte Faktor in jedem Fall mehr Varianz erklärt, als eine einzelne Variable.

Für eine verbesserte Möglichkeit der Interpretation der Faktoren und der Faktorladungen wird zusätzlich die Varimaxrotation verwendet. In der Varimaxrotation wird das Achsensystem der Faktoren so lange gedreht, bis die Varianz der quadrierten Faktorladungen der gemeinsamen Faktoren ein Maximum erreicht.

Die letztliche Schätzung der Faktoren erfolgt über die nachfolgende Gleichung, die aus der Grundgleichung 14 abgeleitet ist:

$$ZA(A^\top A)^{-1} = FA^\top A(A^\top A)^{-1} \quad (25)$$

nach Umstellung der Gleichung 25 ergibt sich:

$$F = ZA(A^\top A)^{-1}$$

Die Faktorwerte haben folgende Eigenschaft: Sie besitzen einen Mittelwert von Null und eine Varianz vom Wert eins.

Anmerkung: Die Berechnung der Explorativen Faktorenanalyse für metrische Daten erfolgt in dieser Arbeit unter SPSS.

3.2 Underlying Variable Analyse(UVA)

Dieses Modell ist das äquivalente Modell der Faktorenanalyse für den Fall ordinaler Daten. Dabei werden hier in ähnlicher Weise latente Strukturen aufgedeckt. Im Rahmen dieses Modells wird angenommen, dass für jede Variable x_j eine unvollständig beobachtbare Variable x_j^* existiert, die jeweilige Underlying Variable. Eine Modellannahme ist, dass die Underlying Variable normalverteilt ist.

Das Modell der UVA stellt sich wie folgt dar:

$$X_{(n \times m)}^* = F_{(n \times Q)}^* A_{(Q \times m)}^{\top *} + U_{(n \times m)}^* E_{(m \times m)}^* \quad (26)$$

Mit

- X^* : Matrix der Underlying Variables
- F^* : Matrix der Faktorwerte der gemeinsamen Faktoren
- A^* : Matrix der Faktorladungen der gemeinsamen Faktoren
- U^* : Matrix der Faktorwerte der Einzelrestfaktoren
- E^* : Matrix mit den Faktorladungen der Einzelrestfaktoren.

Der Modellaufbau ist prinzipiell ein Analogon zum EFA Modell. Daher soll im folgenden nur auf die wesentlichen Eigenschaften des UVA Modells und die Unterschiede zum EFA Modell eingegangen werden.

Die Verbindung zwischen der vollständig beobachtbaren Variablen x_j und der unvollständig beobachtbaren Variablen x_j^* ist über folgende Treppenfunk-

tion gegeben(nach Reinecke ²):

$$x_j = \begin{cases} 1, & \text{wenn } x_j^* \leq \tau_1 \\ 2, & \text{wenn } \tau_1 < x_j^* \leq \tau_2 \\ \vdots & \vdots \\ K-1, & \text{wenn } \tau_{K-2} < x_j^* \leq \tau_{K-1} \\ K, & \text{wenn } \tau_{K-1} < x_j^* \end{cases} \quad (27)$$

mit k als Anzahl der Kategorien der Variable x_j , $k = 1, \dots, K$ und τ_k als Schwellenwert.

Äquivalent zu EFA ist hier die einzig „bekannte“ Variable nun X^* . Die Herleitung der interessierenden Variablen erfolgt nur hier nicht über die Bravais-Pearson Korrelationsmatrix, sondern über die polychorische Korrelationsmatrix als das entsprechend geeignete Zusammenhangsmaß für ordinale Daten. Jöreskog und Sörbom(1993) zeigen in einer Simulationsstudie für ordinale Variablen, dass die polychorische Korrelation als bestes Zusammenhangsmaß für dieses Meßniveau anzusehen ist. Die polychorische Korrelation berechnet sich wie folgt:

Da die Skalierung der x_j willkürlich ist, wird angenommen, dass diese standardnormalverteilt sind. Damit ist die Wahrscheinlichkeit P , dass ein Wert von x_j^* unterhalb des Schwellenwertes τ_k liegt, entsprechend dem Wert der Standardnormalverteilungsfunktion an der Stelle τ_k :

$$P(x_j^* \leq \tau_k) = \int_{-\infty}^{\tau_k} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x_j^*\right\} = \Phi(\tau_k) \quad (28)$$

mit Φ als Funktion der Standardnormalverteilung. Die Schwellenwerte können daher über die Inverse der Standardnormalverteilung ermittelt werden:

$$\tau_k = \Phi^{-1}\left(\sum_{k=1}^c \frac{N_k}{N}\right) \quad (29)$$

²siehe Quellenverzeichnis

, mit N_k als kumulierte Häufigkeiten bis zur Kategorie c .

Die Wahrscheinlichkeiten jeder Zelle der Kontingenztabelle der Variablen x_j und x_h , $j, h=1, \dots, m$, berechnen sich dann wie folgt:

$$\pi_{vk} = \Phi_2(\tau_k, \tau_v) - \Phi_2(\tau_{k-1}, \tau_v) - \Phi_2(\tau_k, \tau_{v-1}) + \Phi_2(\tau_{k-1}, \tau_{v-1}) \quad (30)$$

mit $\Phi_2(\dots)$ als bivariate Standardnormalverteilungsfunktion.

Dabei sei k eine Kategorienausprägung der Variable x_j , mit $k=1, \dots, K$ und v eine Kategorienausprägung der Variable x_h , mit $v=1, \dots, V$. Die polychorische Korrelation ergibt sich letztlich als Maximum Likelihood Schätzung. Dafür wird die Loglikelihood Funktion gebildet:

$$\ln L(r_{x_h^* x_j^*}^*) = \sum_{v=1}^V \sum_{k=1}^K N_{vk} \cdot \ln(\pi_{vk}) \quad (31)$$

Dabei sei N_{vk} die Häufigkeit für das Auftreten der Kategorien v und k und π_{vk} die Wahrscheinlichkeit aus Gleichung 30. Damit ist hier der Maximum Likelihood Schätzer diejenige Korrelation, die die Wahrscheinlichkeit maximiert, dass die Daten der Kontingenztabelle der Variablen x_j und x_h einer bivariaten Standardnormalverteilung folgen.

Es sei allerdings darauf hinzuweisen, dass es sich bei der beschriebenen polychorischen Korrelation um die „two-step“-Methode handelt der polychorischen Korrelation. Das bedeutet, dass die Schwellenwerte und die polychorische Korrelation nicht gleichzeitig geschätzt werden, sondern nacheinander.

Dabei sei in der Analyse der Daten ebenfalls die „two-step“-Methode gewählt, aufgrund einer viel geringeren Rechnerzeit gegenüber der gewöhnlichen ML-Methode, in der Schwellenwerte und Korrelation simultan geschätzt werden. Simulationen zeigen jedoch, dass die Ergebnisse aus der „two-step“-Methode und der Methode des gleichzeitigen Schätzens der Schwellenwerte und der polychorischen Korrelation, keinen bemerkenswerten Unterschied aufweisen(siehe Olsson, Drasgow und Dorans(1982)).

Der Ablauf der UVA gliedert sich somit wie folgt:

1. Schätzung der Schwellenwerte
2. Berechnung der polychorischen Korrelation
3. Schätzung des Faktormodells aus Gleichung 26 über die polychorische Korrelation, aber sonst analog zu Gleichung 14 in der Durchführung

Als Modellspezifikation werden auch hier ebenso das Kaiser-Kriterium für die Bestimmung der Anzahl der zu extrahierenden Faktoren gewählt, als auch die Varimaxrotation für die Verbesserung der Interpretation der Ladungen und der Faktorwerte.

Die Voraussetzungen an die Faktoren des UVA Modells sind ebenso analog zu denen aus dem EFA Modell. Daher sei an dieser Stelle nur auf diese auf Seite 29 verwiesen.

Anmerkung: Die Berechnung der Underlying Variable Analyse erfolgt in dieser Arbeit unter Mplus.

Problem: Es besteht ein technisches Problem in der praktischen Herleitung der Faktorwerte unter Mplus. So ist es in der Version Mplus 2.12 noch nicht möglich, sich im Rahmen der UVA auch die entsprechenden Faktorwerte ausgeben zu lassen. Daher wurde die Confirmatorische Faktoranalyse (CFA) herangezogen. CFA basiert auf dem a priori Wissen, dass sich hinter bekannten Gruppen von Variablen latente Strukturen verbergen. Dieses Wissen wird genutzt, um sich Faktorwerte zu generieren. Da CFA unter Mplus 2.12 verfügbar ist, wurden die Faktorwerte über CFA gebildet.

4 Simulationen

Es ist deutlich zu erkennen, dass die Korrelationsmatrizen in beiden Modellen eine entscheidende Rolle spielen. Wie bereits oben erwähnt, ist von den

jeweiligen Zusammenhangsmaßen abhängig, wie „gut“ die Schätzungen der unbekannten Variablen des Modells der UVA und der EFA sind.

Bartholomew, Steele, Moustaki und Galbraith(2002) betonen die Relevanz der UVA und weisen darauf hin, dass für den Fall, dass die Bravais-Pearson Korrelation auf ordinale Daten angewandt wird, generell eine Unterschätzung der wahren Korrelation vorliegt. Um ihre Aussagen zu unterstützen, beziehen sie sich auf ordinalskaliertes Datenmaterial der “Eurobarometer Umfrage 1992, Bereich: wissenschaftliche und technologische Daten”, mit 7 Variablen, die jeweils 4 Ausprägungen haben und 392 Beobachtungen, Bartholomew u.a(2002, S.222). Zunächst wenden sie die Bravais-Pearson Korrelation auf 4 der 7 Variablen des Datensatzes an und anschließend die polychorische Korrelation. Das benannte Beispiel sei im folgenden kurz angeführt in den Tabellen 10 und 11

Pearson Korrelation/Variable	Var1	Var2	Var3	Var4
Var1	1.00			
Var2	0.15	1.00		
Var3	0.28	0.40	1.00	
Var4	0.33	0.17	0.31	1.00

Tabelle 10: Bravais-Pearson Korrelationskoeffizienten, aus Bartholmew u.a.(2002, S.222)

Polychorische Korrelation/Variable	Var1	Var2	Var3	Var4
Var1	1.00			
Var2	0.20	1.00		
Var3	0.35	0.48	1.00	
Var4	0.41	0.21	0.38	1.00

Tabelle 11: Polychorische Korrelationskoeffizienten, aus Bartholmew u.a.(2002, S.222)

Auf der anderen Seite bemerken Bartholomew u.a(2002, S.209), dass wenn ordinale Daten 7 Kategorien aufweisen, sich keine nachteiligen Ergebnisse

mehr durch die Anwendung von Methoden für metrisch skalierte Daten gegeben werden.

Um diese Behauptungen näher zu untersuchen, werden in den folgenden Abschnitten verschiedene Simulationen vorgenommen und ausgewertet. Für die eigentliche Ergebnispräsentation seien im nachfolgenden häufig Abbildungen dargestellt. [Viele Originalergebnisse seien im Appendix angehängen oder auf der beigefügten CD verfügbar].

Anmerkung: Die Simulationen wurden in der Programmiersprache R durchgeführt.

4.1 Simulation unter der bivariaten Normalverteilung

Gleiche Klassengrößen Es werden Simulationen durchgeführt, die auf der Standardnormalverteilung basieren. Dabei werden in Anlehnung an den tatsächlich vorhandenen Datensatz stets 400 bivariate standardnormalverteilte Beobachtungen erzeugt. Über die Varianzmatrix der simulierten Verteilung wird die wahre Korrelationen festgelegt, die zwischen den marginalen Verteilungen besteht. Anschließend werden die stetigen Daten umkodiert, indem eine Klassenbildung erfolgt. Die Klassengrenzen stellen dabei die Werte der Quantile der Standardnormalverteilung dar. Die Klassen werden zunächst gleich groß und symmetrisch um Null gelegt (Idee nach Bollen und Barb(1981)).

So wird z.B. um 5 Kategorien aus den metrischen Daten zu erhalten, das 20%-, 40%-, 60%- und das 80%-Quantil aus der Standardnormalverteilung bestimmt. Und anschließend die 400 simulierten Beobachtungen in die Klassen umgelegt. Dann erfolgt die Umkodierung, so dass z.B. die Werte die in die Quantilsgrenzen [0%-20%] fallen, fortan der Wert 1 zugewiesen wird. Nach gleichem Vorgehen werden alle weiteren Klassen gebildet.

Nachdem die Werte umkodiert sind, wird die Bravais-Pearson Korrelati-

on auf die künstlich kategorisierten Daten angewandt. Nach diesem Vorgehen erfolgen 10.000 Durchläufe. Anschließend wird das einfache Mittel über die Durchläufe und die Standardabweichung der Durchläufe berechnet. Die nachfolgende Tabelle 12 zeigt die gemittelten Werte aus 10.000 Durchläufen, erstellt für 2 bis 10 Kategorien und verschiedenen wahren Korrelationen der zugrundeliegenden marginalen stetigen Standardnormalverteilungen.

Anzahl der Kategorien	2	3	4	5	6	7	8	9	10
Wahre Korrelation Pearson Korr.									
0.2	0.13	0.16	0.17	0.18	0.18	0.18	0.19	0.19	0.19
0.4	0.26	0.32	0.35	0.36	0.37	0.37	0.37	0.38	0.38
0.5	0.33	0.41	0.44	0.45	0.46	0.47	0.47	0.47	0.47
0.6	0.41	0.50	0.53	0.55	0.56	0.56	0.57	0.57	0.57
0.8	0.59	0.69	0.73	0.75	0.76	0.77	0.77	0.77	0.78
0.9	0.71	0.80	0.84	0.85	0.86	0.87	0.88	0.88	0.88

Tabelle 12: Bravais-Pearson Korrelationskoeffizienten, angewandt auf kategorisierte Daten

[Die Werte der Klassengrenzen und die absoluten Werte der Korrelationskoeffizienten mit den jeweiligen Standardabweichungen seien im Appendix in den Tabellen 27, 28 und 29 angehängen.]

Für eine besseren Übersicht der Entwicklung nach Kategorien sei auf die Abbildung 5 verwiesen.

Es ist in der Abbildung 5 und Tabelle 12 abzulesen, dass mit zunehmender wahrer Korrelationsstärke zwischen den marginalen Verteilungen, eine deutliche Annäherung in den Ergebnissen der kategorisierten Daten an die der metrischen Daten zu erkennen ist. Entscheidend für die Güte der Annäherung ist allerdings die Anzahl der Kategorien der kategorisierten Daten. Es ist zu erkennen, dass bereits ab 5 Kategorien eine Übereinstimmung von 88% besteht und ab 6 Kategorien eine Übereinstimmung von mindestens 91%, bei einer wahren Korrelation von 0.2 zwischen den marginalen Verteilungen.

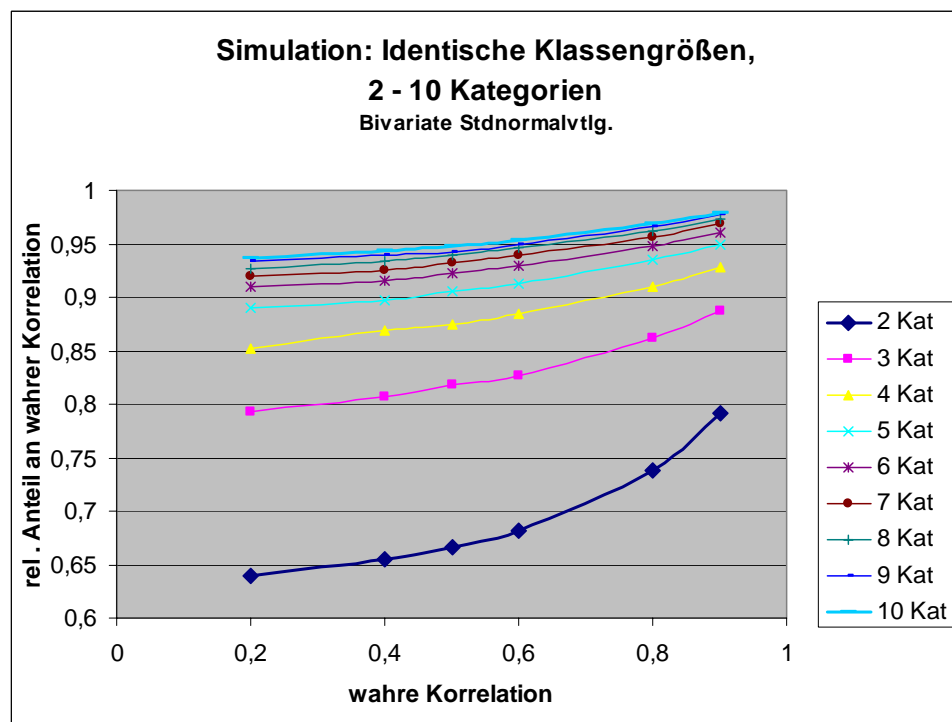


Abbildung 5: rel. Anteil (linke Achse) an wahrer Korrelation (rechte Achse), gleich große Klassen, 400 Beobachtungen, 10.000 Durchläufe

Bartholomew u.a.(2002) ist in der Hinsicht zuzustimmen, dass die ordinalen Daten die wahre Korrelation stets unterschätzen, und dass mit zunehmender Kategorienanzahl sich die Ergebnisse der ordinalen Daten an die der metrischen Daten annähern. Anzumerken sei jedoch, dass dieses Ergebnis nur dann zutrifft, wenn die Klassen jeweils gleich groß gewählt sind und die Verteilung symmetrisch ist.

Verschiedene Klassengrößen Um andere mögliche Umstände zu untersuchen, sei im folgenden nach gleichen Vorgaben wieder eine Simulationsstudie angefertigt. Nun jedoch mit verschiedenen großen Klassenaufteilungen zur Kategorienbildung bei einer bestimmten Kategorienanzahl. Es wird mit der Simulation im Rahmen von 2 Kategorien begonnen.

Die Abbildung 6 zeigt somit die Ergebnisse der Simulationsstudie, in der die Bravais-Pearson Korrelation auf 2 künstlich kodierte Variablen, denen jeweils die Standardnormalverteilung zugrundeliegt, mit 2 Kategorien angewandt wurde. In der Legende der Abbildung ist angegeben welche Quartilsgrenzen der Standardnormalverteilung gewählt wurden, um die Klassengrößen zu bilden.

Wie intuitiv zu erwarten war, lässt sich anhand der Abbildung 6 ablesen, dass die Unterschätzung der wahren Korrelation dann abnimmt, wenn die Aufteilung der Klassen symmetrischer wird.[Die absoluten Werte der Simulationsergebnisse seien im Appendix in Tabelle 30 zu finden.]

Als weitere Darstellung dieses Verhältnisses seien im folgenden noch die Simulationen von 3 und 4 Kategorien in asymmetrischer Klassenaufteilung in Abbildung 7 dargestellt. Diese Darstellung unterstützt die Aussage der Abbildung 6, dass mit zunehmend symmetrischen Klassenverhältnis die relativen Anteile an der wahren Korrelation der marginalen Verteilungen zunehmen. Und andersherum, die Klassenaufteilungen, die in einem stark asymmetrischen Verhältnis zueinander stehen, die schlechtesten Ergebnisse erzeugen.

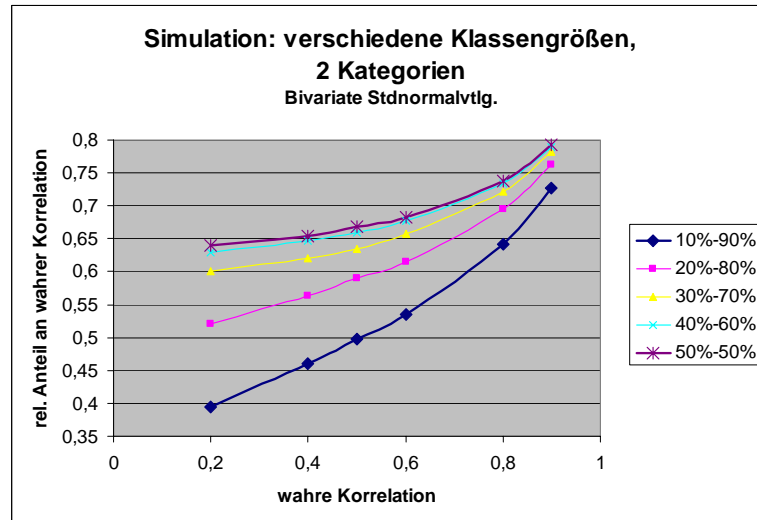


Abbildung 6: rel. Anteil (linke Achse) an wahrer Korrelation (rechte Achse), verschieden große Klassen, 2 Kategorien, 400 Beobachtungen, 10.000 Durchläufe

Nachdem nun einige Simulationen durchgeführt wurden, besteht nun folgende Vermutung: Die gewöhnliche Faktoranalyse für metrische Daten kann eventuell trotz Anwendung auf ordinale Daten sinnvolle Ergebnisse erzeugen, wenn eine günstige Beschaffenheit der Daten vorliegt.

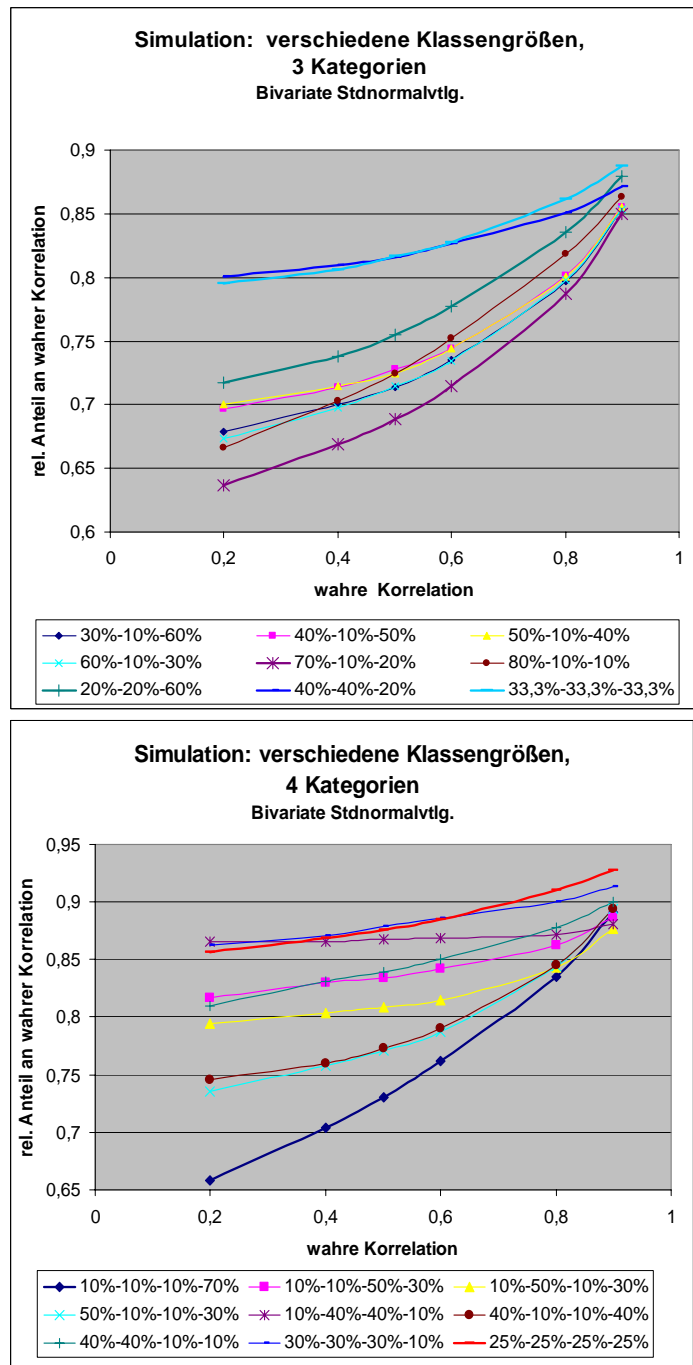


Abbildung 7: rel. Anteil(linke Achse)an wahrer Korrelation(rechte Achse), verschieden große Klassen, 3 und 4 Kategorien, 400 Beobachtungen, 10.000 Durchläufe

4.2 Simulation unter der bivariaten Chiquadratverteilung

Eine höchst relevante Modellannahme, die vom EFA-Modell als von dem UVA-Modell gefordert wird, ist die Normalverteilungsannahme. So fordert die polychorische Korrelation, dass die Underlying Variables der zugrundeliegenden Verteilung standardnormalverteilt sind.

Daher wird nun in der folgenden Simulationsstudie untersucht, wie gut die polychorische Korrelation den wahren Zusammenhang zwischen zwei kategorisierten Variablen widerspiegelt, wenn diese ursprünglich aus einer bivariaten χ^2 -Verteilung stammen. Entgegen der Anforderung wird die Underlying Variable nun eine schiefe Verteilung besitzen.

Der Aufbau der Simulation ist analog zu dem Aufbau der Simulation unter 4.1, Abschnitt: Gleiche Klassengröße. Mit dem Unterschied der zugrundeliegenden bivariaten χ^2 -Verteilung.

Die Bildung der bivariaten χ^2 -Verteilung erfolgte nach Kotz, Balakrishnan und Johnson(2000): Es sei

$$C_1 = Z_1^2 + \dots + Z_f^2 \quad (32)$$

$$C_2 = W_1^2 + \dots + W_f^2 + U_1^2 + \dots + U_{b-f}^2$$

,wobei U_t standardnormalverteilt ist, $U_t \sim N(0, 1)$, $t=1, \dots, b-f$

und $(Z_j, W_j) \sim N((0, 0), \Sigma)$ mit $\Sigma = (1, \rho, \rho, 1)$, da $\rho_j = \rho \quad \forall \quad j = 1, \dots, f$.

C_1 sei eine marginale χ^2 -Verteilung mit f Freiheitsgraden, wie C_2 mit b Freiheitsgraden, $f < b$.

Das Zusammenhangsmaß sei hier:

$$Korrelation(C_1, C_2) = \sqrt{\frac{f}{b}} \rho^2 \quad (33)$$

Der Grund für die Wahl der beschriebenen χ^2 -Verteilung ist, dass diese die Voraussetzungen erfüllt, die hinsichtlich der Simulation erforderlich sind:

- die beiden marginalen Verteilungen sind ebenfalls χ^2 -verteilt

- es existiert ein Korrelationsmaß, um die Zusammenhänge zwischen den marginalen Verteilungen zu bestimmen.

Zunächst wird der Fall einer bivariaten χ^2 -Verteilung mit jeweils vier Freiheitsgraden der marginalen Verteilungen simuliert. Die Grenzen der Klassen werden auch hier wieder über die Quantile der Verteilung bestimmt. Es werden 2-10 Kategorien gebildet.

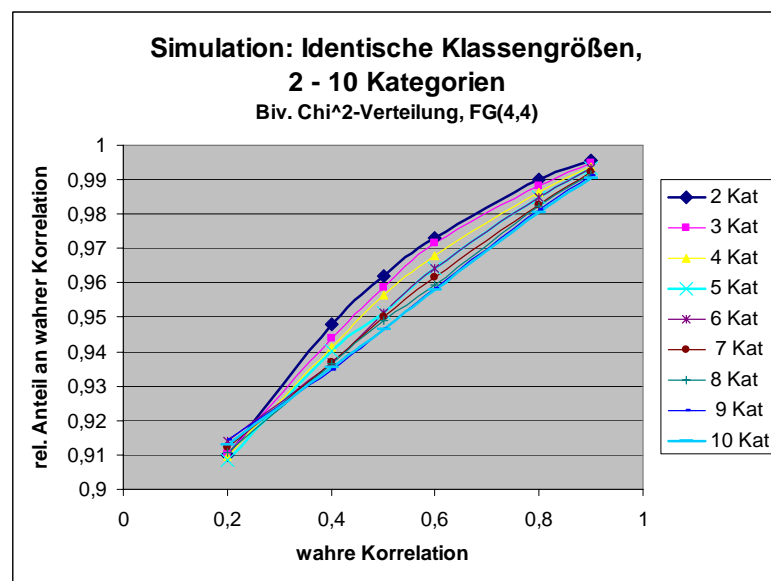


Abbildung 8: rel. Anteil(linke Achse) an wahrer Korrelation(rechte Achse), gleich große Klassen, 400 Beobachtungen, 10.000 Durchläufe, biv. χ^2 -Vtlg., mit FG(4,4)

Im Ergebnis der Simulation, in Abbildung 8, zeigt sich, dass die polychorische Korrelation mit relativen Anteilen zwischen 90% und 99% gute Ergebnisse aufzeigt, trotz der schiefen zugrundeliegenden Verteilung. Das bedeutet, dass die polychorische Korrelation, als Zusammenhangsmaß für ordinale Daten, relativ robust bei Abweichungen von der Normalverteilung der Underlying Variables reagiert, wenn die Verteilung eine bivariate χ^2 -Verteilung ist, mit jeweils vier Freiheitsgraden der marginalen Verteilungen. Dennoch

lässt sich eine Tendenz aus der Abbildung 8 ablesen: Mit zunehmender Anzahl der Kategorien wird für die polychorische Korrelation die Abweichung von der Normalverteilung immer sichtbarer. Mit zunehmender Anzahl von Kategorien nimmt damit der relative Anteil an der wahren Korrelation ab.

Ein weiterer simulierter Fall, sei die kategorisierte bivariate χ^2 -Verteilung mit jeweils 4 und 100 Freiheitsgraden, der marginalen Verteilungen. In dieser Darstellung unterscheiden sich beide Verteilungen massiv voneinander, so nähert sich die marginale χ^2 -Verteilung mit 100 Freiheitsgraden in der Gestalt an eine Normalverteilung an. Die Ergebnisse der Anwendung der polychorischen Korrelation seien in Abbildung 9 zu finden, dabei wurde die Verteilung jeweils in 3-5 Kategorien umgelegt.

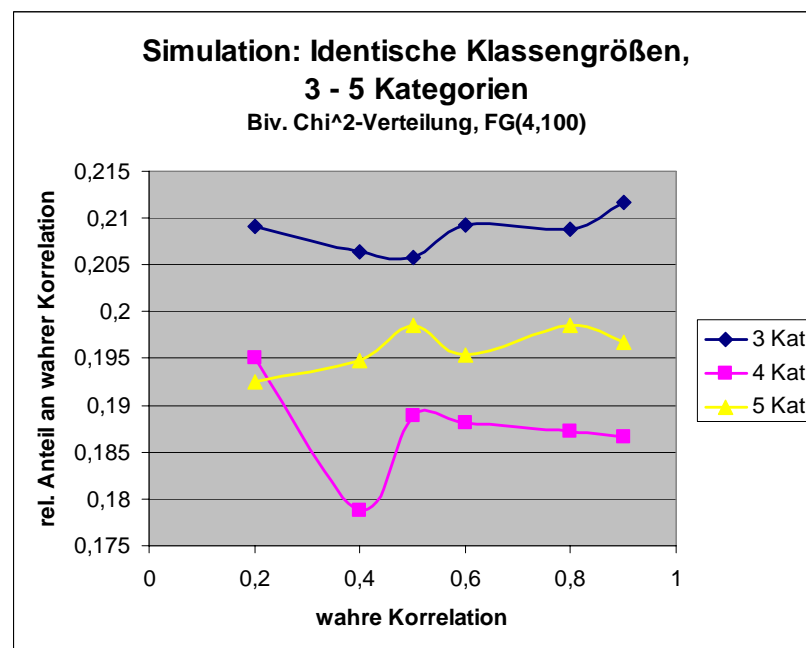


Abbildung 9: rel. Anteil(linke Achse) an wahrer Korrelation(rechte Achse), gleich große Klassen, 400 Beobachtungen, 10.000 Durchläufe, biv. χ^2 -Vtlg. mit FG(4,100)

Das Ergebnis dieser Simulation weicht stark von dem vorherigen Ergebnis aus Abbildung 8 ab. Die relativen Anteile an der wahren Korrelation be-

wegen sich im Rahmen dieser 3-5 Kategorien zwischen 18.72% und 21.17%. Hinzukommt, dass die Abbildung 8 keine Tendenzen durch eine erweiterte Kategorienanzahl erkennen lässt. Falls eine UVA auf Daten ähnlicher Struktur, wie in der letztgenannten Simulation, angewandt werden würde, wäre es nicht möglich, etwas über die Qualität der resultierenden Ergebnisse auszusagen.

5 Modellanwendung

Bevor die Modelle auf die ausführlich beschriebenen ordinalskalierten Daten aus Kapitel 2 angewendet werden, muss überprüft werden, inwieweit die Daten die Annahmen der beiden Modelle erfüllen:

5.1 EFA und UVA

Die a priori Voraussetzungen an die Daten, um eine für metrische Daten konzipierte Explorative Faktorenanalyse anzuwenden, sind wie folgt erfüllt:

1. Unabhängige Beobachtungen: ⚡

Diese Voraussetzung ist nicht erfüllt. So hatten zum einen die Lehrer prinzipiell die Möglichkeit sich untereinander abzusprechen.

Hinzu kommt, dass die Lehrer, die einer Schule angehören schon deswegen nicht voneinander unabhängig sind, weil sie die gleichen Erfahrungen an der Schule gesammelt haben. Es hat sich somit eine Art „Gruppenmeinung“ bilden können.

Ferner war die Teilnahme freiwillig. So dass nicht klar ist, ob eventuell nur die Personen an der Umfrage teilgenommen haben, die eine extreme Einstellung besitzen.

2. Großer Stichprobenumfang: ✓

Der Datensatz enthält nach der Datenaufbereitung noch 416 Beobachtungen und kann damit eindeutig als „groß“ eingestuft werden.

3. Metrisch skalierte Variablen: $\frac{1}{2}$

Diese Voraussetzung ist als einzige der Voraussetzungen absichtlich nicht gegeben.

4. (approximativ) normalverteilte Variablen:

In der EFA ist diese Voraussetzung natürlich darauf aufgebaut, dass die Daten metrisch skaliert sind. Da dieser Fall aber nicht vorliegt, wird daher ein χ^2 -Test auf bivariate Normalverteilung für die 137 Variablen durchgeführt. Mit diesem Test wird überprüft, ob die hinter den ordinalen Daten liegenden metrischen Daten einer bivariaten Standardnormalverteilung folgen. Die zu erfüllenden Approximationsbedingungen des Tests sind:

- die erwartete Häufigkeit $\hat{e}_{kv,SN}$ jeder Zelle der Kontingenztabelle muss größer als 1 sein
- höchstens 20% der Zellen der Kontingenztabelle dürfen erwartete Häufigkeiten $\hat{e}_{kv,SN}$ kleiner als 5 aufweisen.

mit k als Kategorie der Variable x_j , $k = 1, \dots, K$

und v als Kategorie der Variable x_h , $h = 1, \dots, V$, wobei $j, h = 1, \dots, m$.

Die Daten wurden entsprechend den Approximationsbedingungen so aufbereitet, dass diese die Anforderungen erfüllen. So mussten die Randkategorien der einzelnen Variablen häufig zusammengelegt werden. (Der die genannten Approximationsbedingungen erfüllende Datensatz ist in der beigefügten CD zu finden, zusammen mit dem R-Programm, dass die Approximationsbedingungen im Datensatz überprüft.)

Zusätzlich wird gefordert, dass

- die Aufteilung der Variablenausprägung vollständig ist ✓

- die Beobachtungen voneinander unabhängig sind \nmid

Auch hier sei die letztere Bedingung aus den bereits angeführten Gründen nicht erfüllt.

$$H0 : (x_j, x_h) \stackrel{\text{as.}}{\sim} N(0, \Sigma); \quad \text{mit} \quad \Sigma = (1, \rho_{jh}, \rho_{jh}, 1) \quad (34)$$

$$H1 : (x_j, x_h) \stackrel{\text{as.}}{\not\sim} N(0, \Sigma); \quad \text{mit} \quad \Sigma = (1, \rho_{jh}, \rho_{jh}, 1)$$

Teststatistik:

$$V = \sum_{k=1}^K \sum_{v=1}^V \frac{(H_{kv} - \hat{e}_{kv,SN})^2}{\hat{e}_{kv,SN}} \stackrel{\text{as.}}{\sim} \chi_{(K-1)(V-1)}^2 \quad (35)$$

mit $\hat{e}_{kv,SN}$ als erwartete Häufigkeiten, die sich ergeben, wenn (x_j, x_h) aus der bivariaten Standardnormalverteilung stammen.

Die Ergebnisse seien in nachfolgender Abbildung 10 in Form von p-Values wiedergegeben. Dabei gibt die Farbe der p-Values deren Wertebereich wieder.

Liegt der Wert eines p-Values unterhalb von 5% so ist der p-Value in der Abbildung rot markiert. Die entsprechende Interpretation wäre dann, dass der χ^2 -Test auf bivariate Normalverteilung auf dem 5%-Niveau signifikant ist. Also die Nullhypothese, dass die Daten der Kontingenztabelle einer bivariaten Normalverteilung folgen, abgelehnt würde.

Besitzt der p-Value eine grüne Farbe, liegen die prozentualen Werte des p-Values zwischen 5% und 10%. Dann würde das Ergebnis des Tests noch auf dem 10%-Niveau signifikant sein.

Alle blauen p-Values deuten hingegen darauf hin, dass der Test auf bivariate Normalverteilung nicht abgelehnt werden konnte, da die p-Values oberhalb von 10% liegen.

chi²-Test auf biv. Normalvtlg.; p-values

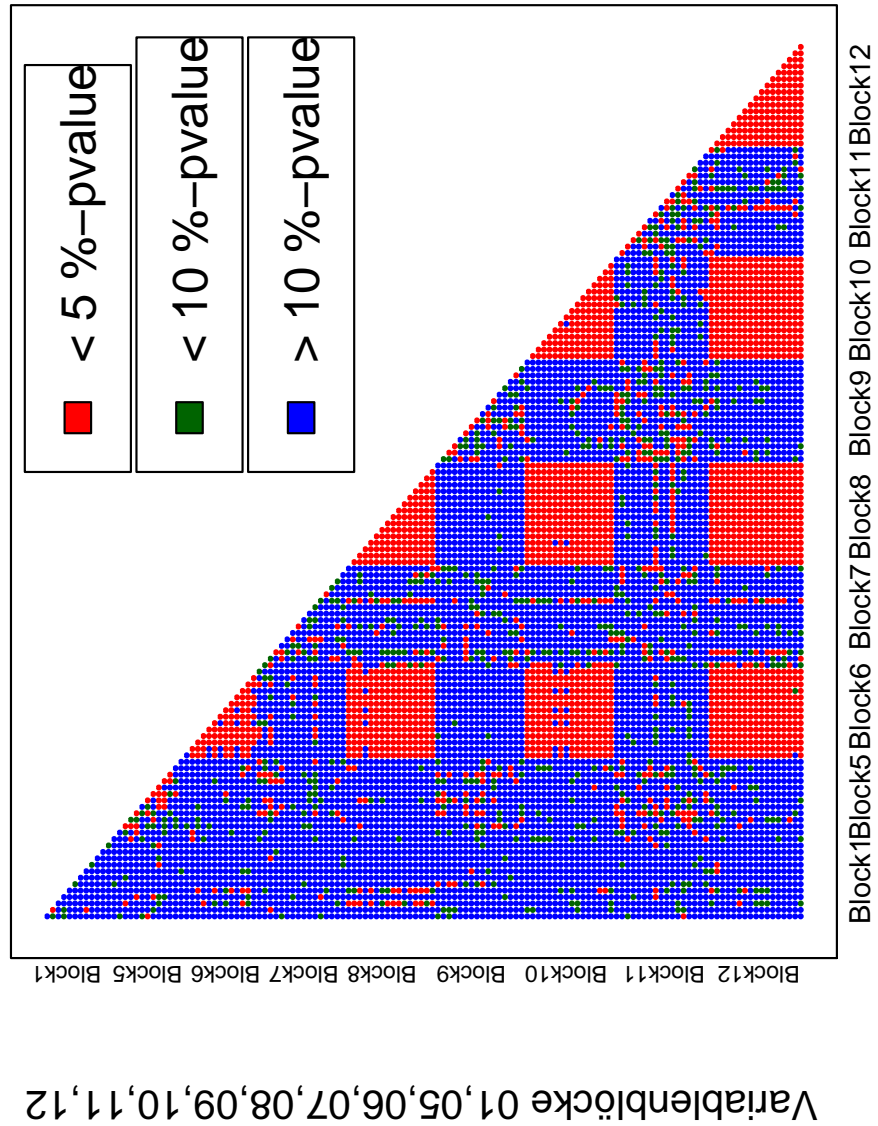


Abbildung 10: P-Values des Tests auf bivariate Normalverteilung zwischen jeweils zwei verschiedenen Variablen, (rechte und linke Achse: Variablenblöcke); ohne die Hauptdiagonale; PVALnorm.R

Es ist deutlich zu erkennen, dass zwischen den geradzahligen Variablenblöcken 6, 8, 10 und 12 der Test auf bivariate Normalverteilung durch den approximativen χ^2 -Test nahezu durchgängig abgelehnt wird. Hierbei sei allerdings nochmals anzumerken, dass gerade die Variablen meist den Test ablehnen, die durch eine zusätzliche Kategorie künstlich erweitert wurden, (siehe Abschnitt 2.4). Daher kann dieses schlechte Testergebnis der geradzahligen Variablenblöcke darauf zurückzuführen sein. Im weiteren ist nur vereinzelt zu erkennen, dass zwischen einigen Variablen der Test auf dem 5%- bzw. 10%-Niveau abgelehnt wurde. Weitere Systematiken sind nicht erkennbar.

Die Voraussetzung auf bivariate Normalverteilung ist damit nicht bei allen Variablen erfüllt.

Die einzige Voraussetzung der UVA an die Daten ist die Normalverteilung der Underlying Variables: Da diese Voraussetzung bereits oben geprüft wurde, sei auf das oben genannte Ergebnis verwiesen. So wird für die UVA die Voraussetzung an die Daten nicht vollständig erfüllt.

Empfehlung: EFA Im Rahmen der Überprüfung der Modellannahmen hat sich gezeigt, dass nur eine Voraussetzung der EFA vollständig erfüllt werden kann.

Die ursprüngliche Idee der Analyse dieser Arbeit, nur die Voraussetzung „metrisch skalierte Variablen“ der EFA unbeachtet zu lassen und die Konsequenzen dieses Sachverhalts durch die Anwendung auf ordinale Daten zu untersuchen, ist nun nicht mehr möglich.

Es sei daher darauf hingewiesen, dass die Ergebnisse aus der Analyse nicht nur auf Basis der fehlenden Voraussetzung „metrisch skalierte Variablen“ verzerrt sein könnten. Daher sollte in jedem Fall die Interpretation unter Vorbehalt stattfinden.

Empfehlung: UVA In der Überprüfung der Modellannahmen hat sich gezeigt, dass auch die Voraussetzung der UVA an die vorliegenden Daten nicht vollständig erfüllt wurde. Das bedeutet für die spätere Analyse, dass die Ergebnisse aus der UVA nicht mehr ohne Vorbehalt als Referenz der wahren Ergebnisse einer Faktorenanalyse für ordinale Daten verwendet werden können.

5.2 Vorgehen

Da wir davon ausgehen, dass latente Strukturen hinter den ordinalen Variablen bestehen, folgt daraus, dass die Variablen nicht unabhängig voneinander sein können. Aufgrund der gemeinsamen inhaltlichen Fragestellung innerhalb ganzer Variablenblöcke wird davon ausgegangen, dass sich besonders hinter den Variablen, die Fragebereichen zugeordnet sind, latente Strukturen verbergen. Daher werden in der folgenden Durchführung der Modelle nicht alle 137 Variablen gemeinsam untersucht, sondern lediglich die Variablen gemeinsam in einem Modell untersucht, die auch den gleichen Fragebereichen angehören und die gleiche Aussagestruktur besitzen (siehe Abschnitt 2.2.2). Um dieses Vorgehen statistisch abzusichern, wird daher der χ^2 -Unabhängigkeitstest durchgeführt.

Die Voraussetzungen für die Daten des χ^2 -Unabhängigkeitstest sind die gleichen, wie die des χ^2 -Tests auf bivariate Normalverteilung. Daher müssen diese hier nicht noch einmal aufgeführt werden. Die Hypothesen des Tests lauten:

H0: die Zufallsvariablen x_j und x_h sind
stochastisch voneinander unabhängig

H1: die Zufallsvariablen x_j und x_h sind nicht
stochastisch voneinander unabhängig

Die Teststatistik lautet:

$$V = \sum_{k=1}^K \sum_{v=1}^V \frac{(H_{kv} - \hat{e}_{kv})^2}{\hat{e}_{kv}} \stackrel{\text{as.}}{\sim} \chi^2_{(K-1)(V-1)} \quad (36)$$

Das Testergebnis ist in Abbildung 11 beigefügt. Auch hier sei die Erklärung der p-Values exakt die gleiche wie die des χ^2 -Tests auf bivariate Normalverteilung.

χ^2 -Test auf Unabhängigkeit; p-values

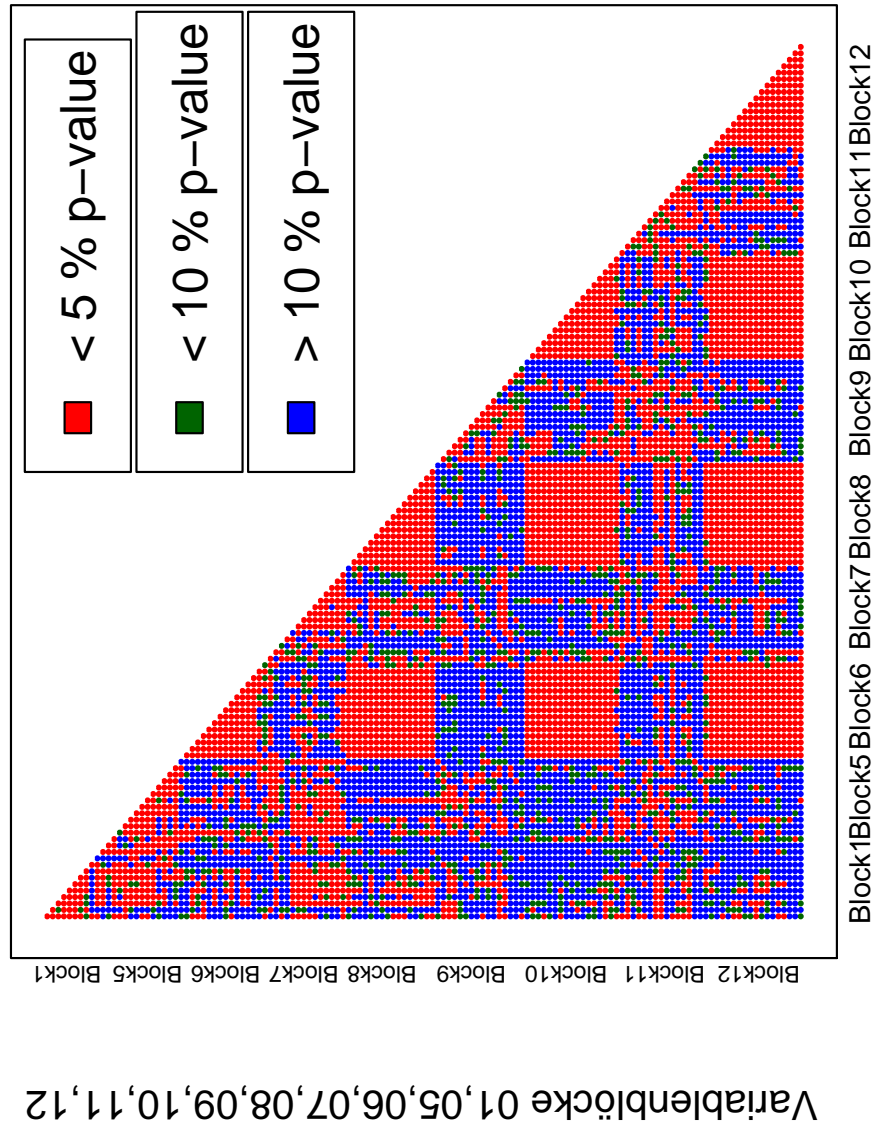


Abbildung 11: P-Values des Tests auf bivariate Normalverteilung (rechte und linke Achse: Variablenblöcke); ohne die Hauptdiagonale; PVALindep.R

Die prozentuale Darstellung der signifikanten Anteile (auf dem 5%-Niveau) von Variablenkombinationen zwischen jeweils zwei verschiedenen Variablenblöcken aus dem χ^2 -Unabhängigkeitstest, sei für eine bessere Übersicht in folgender Tabelle 13 dargestellt.

Variablenblock	1	5	6	7	8	9	10	11
5	70.7							
6	44.2	42.3						
7	61.9	73.2	46.8					
8	52.8	33.5	100.0	61.3				
9	41.5	49.1	14.3	53.9	18.8			
10	29.5	28.6	100.0	41.0	100.0	34.4		
11	49.7	55.0	34.0	59.9	46.7	74.6	49.3	
12	30.5	29.0	100.0	43.8	100.0	30.9	100.0	47.8

Tabelle 13: Prozentualer Anteil der signifikanten Variablenkombinationen (auf dem 5%-Niveau) zwischen verschiedenen Blockkombination

Wie zu erwarten, ist anhand der Tabelle abzulesen, dass überwiegend diejenigen Variablenblöcke die höchsten Anteile an signifikanten Variablen miteinander aufweisen, die auch den gleichen inhaltlichen Fragebereichen und Aussagestrukturen angehören.

So lässt sich ablesen, dass die Blöcke 5 und 7, die dem Fragebereich 2 angehören und der Aussagestruktur „...zum gegenwärtigen Stand“, jeweils beieinander mit 73.2% den höchsten signifikanten Anteil aufweisen.

Ebenso Block 9 und 11: diese weisen einen signifikanten Anteil von 74.6% auf, gehören dem Fragebereich 3 an und besitzen die Aussagestruktur „...zum gegenwärtigen Stand“.

Es ist jedoch auch zu erkennen, dass die Abhängigkeit zwischen den Variablenblöcken, die geradzahlig sind, durchgängig hoch signifikant ist. So weisen die Blöcke 6, 8, 10 und 12 untereinander zu 100% signifikante Anteile auf,

obwohl jeweils nur die Blöcke 6 und 8 oder 10 und 12 gemeinsamen Fragebereichen angehören. Dieses Ergebnis kann aus der künstlichen Veränderung dieser Daten abgeleitet werden.

5.3 Erklärungsverlust der Daten

In den Simulationsstudien konnte dargestellt werden, dass die Anwendung von statistischen Methoden abhängig vom Skalenniveau (und der Verteilung) der Daten zu anderen Ergebnissen führen kann.

Um auf statistischer Basis nachvollziehen zu können, mit welchem Ausmaß an nachteilhaften Ergebnissen man zu rechnen hat, dadurch, dass das metrische Skalenniveau der Daten für eine EFA nicht gegeben ist, seien nachfolgende Formeln und deren Ergebnisse aufgeführt:

Als Basis für die nachfolgenden Erklärungen sei eine Formel verwendet, die besagt, dass wenn die Summe der unwesentlichen Eigenwerte (nach dem Kaiser-Kriterium Eigenwerte < 1) klein ist, der Approximationsfehler des Modells der Explorativen Faktoranalyse ebenfalls klein ist, nach Härdle und Simar(2003, S.288):

$$\sum_{j,h} (S - \hat{A}\hat{A}^\top - \hat{E}\hat{E})_{h,j}^2 \leq \lambda_{g+1}^2 + \dots + \lambda_m^2 \quad (37)$$

mit der Kovarianzmatrix S und $\lambda_1, \dots, \lambda_g$ vergleichsweise großen Eigenwerten.

Zusätzlich gilt: $\lambda_1 \geq \dots \geq \lambda_m$.

Auf das zugrundeliegende Problem abgeändert, ergibt sich nun Formel 38:

$$\sum_{j,h} (R - R^*)_{h,j}^2 \leq \lambda_m^2 + \dots + \lambda_e^2 \quad (38)$$

mit der Bravais-Pearson Korrelationsmatrix R und der polychorischen Korrelationsmatrix R^* , dabei sei $\sum_{j=e}^m \lambda_j^2$ die notwendige quadrierte aufsum-

mierte Zahl aus Eigenwerten der Bravais-Pearson Korrelationsmarix, die notwendig ist, damit die Gleichung 38 erfüllt ist.

In der folgenden Tabelle 14 sei versucht darzustellen, inwieweit ein Erklärungsverlust in der Analyse der ordinalen Daten durch die Verwendung der Bravais-Pearson Korrelation eintritt.

Formel/Variablenblock	1	5+7	6+8	9+11	10+12
$\sum_{h,j} (R - R^*)_{hj}^2$	0.30	0.53	1.39	1.23	0.94
$\sum_{j=e}^m \lambda_j^2$	0.35	0.57	1.51	1.26	0.99
Anzahl relev. Eigenwerte(e)	2.00	5.00	12.00	10.00	13.00
Variablen insgesamt (m)	11.00	30.00	30.00	33.00	33.00
$\sum_{j=e}^m \lambda_j$	0.83	1.67	4.08	3.45	3.51
$\frac{\sum_{j=e}^m \lambda_j}{\sum_{j=1}^m \lambda_j}$ in %	7.58	5.58	13.61	10.46	10.64

Tabelle 14: Verlorene Erklärungsanteile berechnet über Eigenwerte und Zusammenhangsmaße

Anhand der Tabelle 14 ist zum Beispiel abzulesen, dass sich hinsichtlich Variablenblock 1, eine quadrierte und aufsummierte Differenz von 0.30 ergibt zwischen der Bravais-Pearson Korrelationsmatrix und der polychorischen Korrelationsmatrix.

Entsprechend müssen für die Erfüllung der Gleichung 38 die zwei kleinsten Eigenwerte quadriert und aufsummiert werden. Unquadriert und aufsummiert erlangen die zwei Eigenwerte einen prozentualen Anteil an den 11 Variablen (also an der Gesamtvarianz von 11) von Block 1 in Höhe von 7.58%.

Entsprechend muss ein Erklärungsverlust von 7.58% einkalkuliert werden, wenn auf die vorliegenden ordinalen Daten die Bravais-Pearson Korrelation angewandt wird, anstelle der für die Daten geeigneten polychorischen Korre-

lation.

Den größten Erklärungsverlust weisen die gemeinsam untersuchten Variablenblöcke 6 und 8 auf, mit 13.61%.

6 Faktorladungen und Faktorwerte

Die Interpretation der Faktorwerte und Faktorladungen stellt das Ergebnis der Analysen dar.

So war die ursprüngliche Idee der Analyse, dass wenn die Ergebnisse der Faktorwerte und der Faktorladungen der EFA und UVA Modelle ähnlich sind, dies (unter Vorbehalt) bedeuten kann, dass die Explorative Faktoranalyse trotz der Anwendung auf ordinale Daten zu interpretierbaren Ergebnissen führt.

Es werden je Faktor jeweils die Variablen als relevant für den jeweiligen Faktor eingestuft, wenn ihre Ladungen im Vergleich zu allen anderen Ladungen dieses Faktors hoch sind. Üblicherweise spricht man dann von Variablen, die auf einem Faktor hochladen, wenn die Ladung mindestens 0.5 beträgt, siehe Backhaus u.a.(2003). Aber im Rahmen dieser Analyse hat sich gezeigt, dass es häufiger für die Interpretation notwendig war, auch Variablen mit kleineren Ladungen zuzulassen.

Im folgenden seien die Faktorwerte beider Modelle in einem Plot gegenübergestellt. Diese Abbildungen sollen aufzeigen, in welchem Verhältnis die Ergebnisse der Faktoren beider Modelle zueinander stehen. Zusätzlich werden die Bravais-Pearson Korrelationen zwischen den Faktorwerten aus UVA und EFA berechnet, sollten die Plots einen linearen Zusammenhang widerspiegeln.

Daneben seien auch die Variablen mit ihren Faktorladungen beider Modelle aufgeführt, die auf dem jeweiligen Faktor hoch laden.

Faktorwerte werden, so weit dies möglich ist, interpretiert.

6.1 Fragebereich 1

Im Rahmen beider Analysen wurden jeweils 2 Eigenwerte der Korrelationsmatrizen ausgewiesen, die größer als eins sind. Da das Kaiser-Kriterium verwendet wird, bedeutet dies für beide Modelle, dass zwei Faktoren aus den Korrelationsmatrizen geschätzt wurden.

Bereich 1, Variablenblock 1 Es werden 2 Eigenwerte in der UVA und EFA ausgewiesen, die größer als 1 sind. Daher erfolgt eine Bildung von 2 Faktoren.

Modell	EFA-	UVA-
Block 1	Eigenwert	Eigenwert
	3.715	4.218
	1.259	1.315

Tabelle 15: Eigenwerte > 1 , im Block 1

Interpretation Faktor 1: klare offene Strukturen und eine harmonische Gemeinschaft.

Modell		EFA-	UVA-
1.Faktor	Variable	Ladung	Ladung
	f0105	0.758	-0.800
	f0107	0.719	-0.706
	f0108	0.711	-0.708
	f0104	0.675	-0.662
	f0103	0.666	-0.656
	f0111	0.621	-0.574
	f0106	0.524	-0.548

Tabelle 16: Faktorladungen des 1.Faktors des Variablenblocks 1, Bereich 1

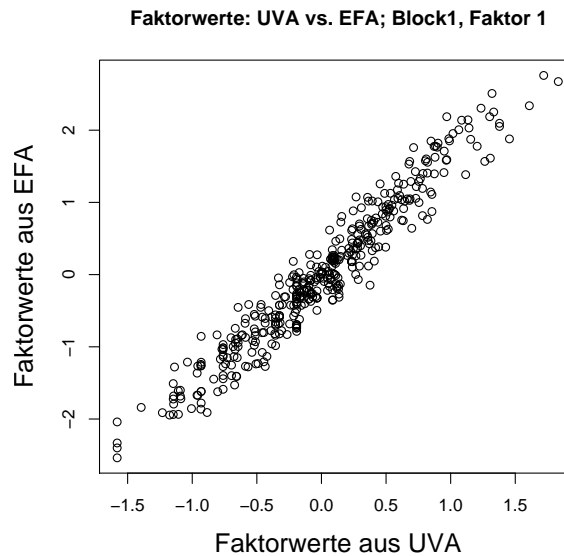


Abbildung 12: Plot der Werte des 1. Faktors des Variablenblocks 1, Bereich 1, aus UVA(rechte Achse) und EFA(linker Achse)

Anhand der Abbildung 12 ist abzulesen, dass ein linearer Zusammenhang zwischen den Faktoren der EFA und der UVA besteht. Die Bravais-Pearson Korrelation der Faktorwerte zeigt eine starke Korrelation von 0.97 auf.

Interpretation Faktor 2: Freie Entscheidungen in den eigenen Handlungen als Lehrer.

Modell		EFA-	UVA-
2.Faktor	Variable	Ladung	Ladung
	f0110	0.771	-0.411
	f0101	-0.700	0.745

Tabelle 17: Faktorladungen des 2. Faktors des Variablenblocks 1, Bereich 1

Anmerkung zu der Bildung von Faktorwerten Die Bildung eines Faktorwertes des 2. Faktors des Variablenblocks 1 war aus technischen Gründen nicht möglich (bzw. eine Confirmatorische Faktoranalyse(CFA) über Mplus war nicht möglich).

In nachfolgenden Darstellungen seien auch weitere Plots der Faktorwerte anderer Variablenblöcke nicht darstellbar, da sich auch bei ihnen die Faktorwerte der UVA über Mplus nicht bilden lassen. Es seien jedoch alle verfügbaren Plots abgebildet.

6.2 Fragebereich 2

Fragebereich 2, Variablenblock 5 und 7 In beiden Analysen weisen die Korrelationsmatrizen der verschiedenen Modelle 7 Eigenwerte auf, die größer als 1 sind. Es werden 7 Faktoren geschätzt.

Modell	EFA-	UVA-
Block 5+7	Eigenwert	Eigenwert
	7.117	7.688
	3.636	3.887
	1.491	1.535
	1.437	1.466
	1.283	1.294
	1.149	1.162
	1.029	1.028

Tabelle 18: Eigenwerte > 1 , im Block 5 und 7

Interpretation Faktor 1: Gefühl von aufgezwungener Unterordnung seitens der Lehrer, mit zu wenig Selbstbestimmung.

Die Bravais-Pearson Korrelation zwischen den Faktoren der beiden Modelle beträgt 0.92 und kann als starker linearer Zusammenhang eingestuft

werden.

Anmerkung: Es sei darauf hinzuweisen, dass nicht stets die gleichen Variablen beider Analysen auch bei den gleichen Faktoren hochladen. Der Grund dafür ist meist die Tatsache, dass einige Variablen bei mehreren Faktoren hochladen.

So können diese Variablen inhaltlich zur Zusammensetzung verschiedener Faktoren gleicher Maßen „gut“ beitragen. Sollte in den nachfolgenden Tabellen der Faktorladungen eine Variable anstelle eines Wertes ein „ * “ aufweisen, so bedeutet dies, dass die Variable bei einem anderen Faktor eine höhere Ladung besitzt.

Modell		EFA-	UVA-
1.Faktor	Variable	Ladung	Ladung
	f0505	0.735	-0.690
	f0716	0.726	-0.768
	f0709	-0.705	0.678
	f0708	0.629	-0.593
	f0508	-0.586	0.581
	f0710	0.584	-0.544
	f0504	0.576	-0.527
	f0711	0.507	-0.476
	f0707	0.476	-0.486
	f0715	*	-0.522

Tabelle 19: Faktorladungen des 1. Faktors der Variablenblöcke 5 und 7, Bereich 2

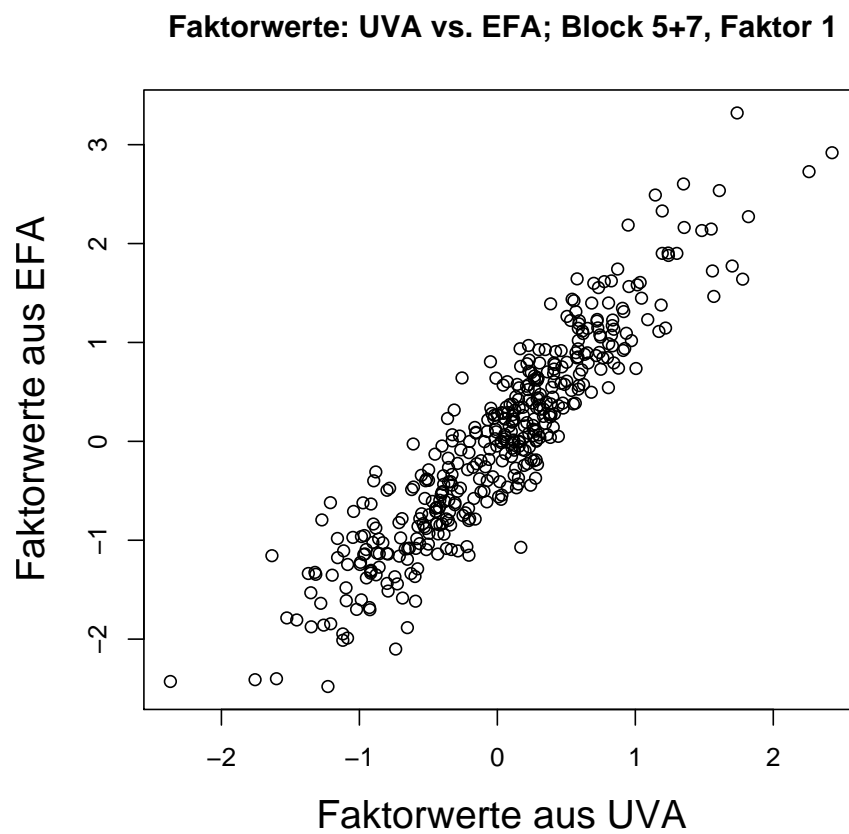


Abbildung 13: Plot der Werte des 1. Faktors der Variablenblöcke 5 und 7, Bereich 2, aus UVA(rechte Achse) und EFA(linke Achse)

Interpretation Faktor 2: Einheitliche von allen akzeptierte und angestrebte Leistungsziele für die Schüler, die eindeutig und transparent sind.

Modell		EFA-	UVA-
2.Faktor	Variable	Ladung	Ladung
	f0509	0.783	0.745
	f0510	0.699	0.658
	f0507	0.694	0.647
	f0501	0.689	0.555
	f0512	0.654	0.593

Tabelle 20: Faktorladungen des 2. Faktors der Variablenblöcke 5 und 7, Bereich 2

Es lässt sich eine Bravais-Pearson Korrelation von 0.9 für die Faktorwerte des Faktors 2 berechnen.

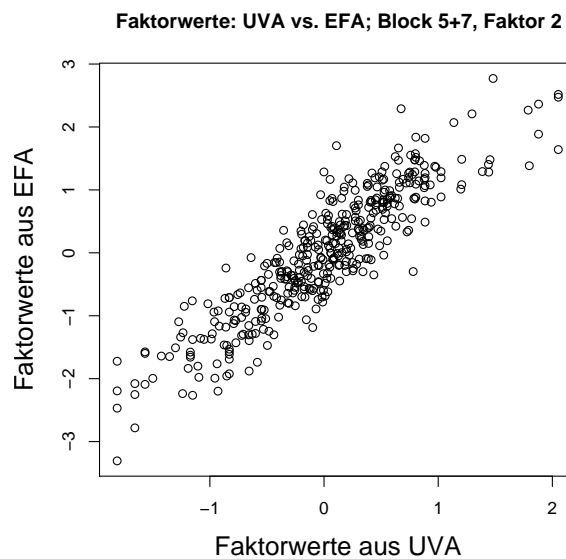


Abbildung 14: Plot des Faktors 2 in Block 5+7 aus UVA(rechte Achse) und EFA(linke Achse)

Interpretation Faktor 3: Es existieren gemeinsam beschlossene und akzeptierte Evaluationsmaßnahmen, die Einfluss auf die eigene pädagogische Arbeit haben.

Modell		EFA-	UVA-
3.Faktor	Variable	Ladung	Ladung
	f0701	0.772	0.710
	f0703	0.736	0.689
	f0513	0.668	0.684
	f0511	0.596	0.562
	f0704	0.523	0.399

Tabelle 21: Faktorladungen des 3.Faktors der Variablenblöcke 5 und 7, Bereich 2

Die Bravais-Pearson Korrelation des 3. Faktors beider Modelle beträgt 0.91.

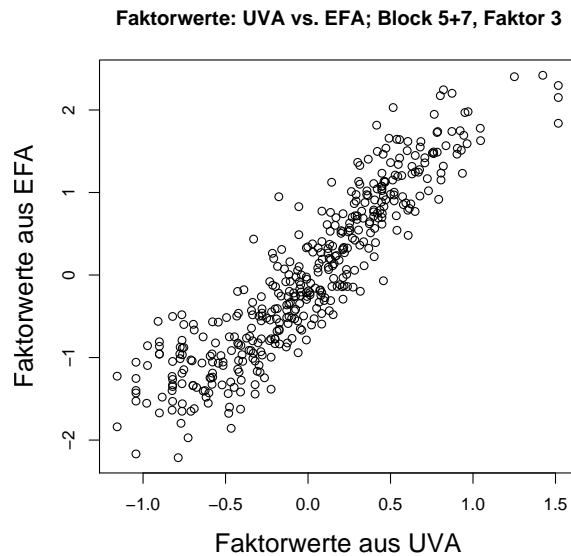


Abbildung 15: Plot der Werte des 3. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linken Achse)

Interpretation Faktor 4: Eine Schulmodernisierung wird gewünscht bzw. sich dafür positiv ausgesprochen.

Modell		EFA-	UVA-
4.Faktor	Variable	Ladung	Ladung
	f0506	0.684	-0.547
	f0702	-0.621	0.409
	f0514	0.601	-0.521

Tabelle 22: Faktorladungen des 4.Faktors der Variablenblöcke 5 und 7, Bereich 2

Der Stärke des linearen Zusammenhangs wird hier mit 0.79 über den Bravais-Pearson Koeffizienten gemessen.

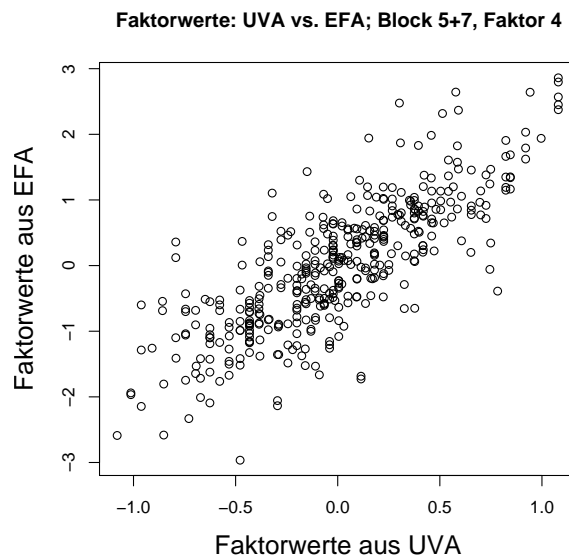


Abbildung 16: Plot der Werte des 4. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linker Achse)

Interpretation Faktor 5: Einzelne Lehrkräfte haben (größeren) Einfluss auf die Schulentwicklung, da sie sich engagieren(; diese Tatsache weckt Unzufriedenheit bei anderen Lehrern).

Zwischen den Faktoren konnte eine Bravais-Pearson Korrelationsbeziehung von 0.78 gemessen werden.

Modell		EFA-	UVA-
5.Faktor	Variable	Ladung	Ladung
	f0715	0.640	*
	f0712	0.586	-0.502
	f0705	0.568	-0.462
	f0714	0.532	-0.376
	f0713	0.527	-0.414

Tabelle 23: Faktorladungen des 5.Faktors der Variablenblöcke 5 und 7, Bereich 2

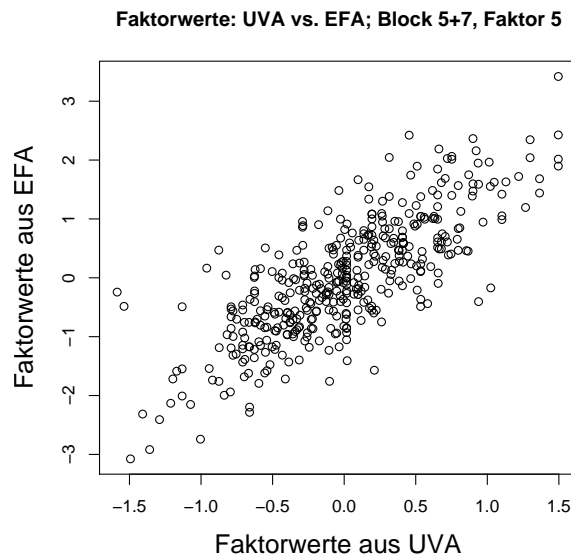


Abbildung 17: Plot des 5. Faktors in Block 5+7, Bereich 2, aus UVA(rechte Achse) und EFA(linker Achse)

Interpretation Faktor 6: Lehrer werden in Veränderungsprozesse der Schule miteinbezogen.

Aus technischen Gründen war eine Bildung von Faktorwerten nicht möglich (siehe Fragebereich 1, Faktor 2).

Modell		EFA-	UVA-
6.Faktor	Variable	Ladung	Ladung
	f0502	0.746	0.571
	f0503	0.513	0.492

Tabelle 24: Faktorladungen des 6.Faktors der Variablenblöcke 5 und 7, Bereich 2

Interpretation Faktor 7: Nicht möglich. Zwar wurden 7 Faktoren ausgewiesen, doch tatsächlich weisen beide Modelle nur eine Variable (f0706) aus, die auf diesen Faktor hochlädt. Da jedoch ein gemeinsamer Faktor nach dem UVA und EFA Modell nur dann gebildet werden kann, wenn der Faktor auf mindestens 2 Variablen wirkt, ist eine Faktorinterpretation nicht möglich. Eine genauere Betrachtung der Eigenwerte in Tabelle 18 lässt erkennen, dass der 7. Eigenwert nur noch geringfügig größer als 1 ist.

Da dieser Sachverhalt auch bei weiteren Faktoren auftritt, wird an dieser Stelle nur noch ein kurzer Hinweis auf die nicht erfüllten Voraussetzungen des Faktors gegeben sein.

Modell		EFA-	UVA-
(7.Faktor)	Variable	Ladung	Ladung
	f0706	0.764	0.457

Tabelle 25: Faktorladungen des (7.Faktors) der Variablenblöcke 5 und 7, Bereich 2

Im folgenden seien nur noch die Interpretationen der Faktoren, soweit dies möglich war, aufgeführt. Im Appendix seien die dazugehörigen Abbildungen und Tabellen der Faktorwerte und Faktorladungen zu finden.

Fragebereich 2, Variablenblöcke 6 und 8 Da sich aus den Korrelationsmatrizen beider Modelle UVA und EFA 5 Eigenwerte mit Werten größer als 1 ableiten lassen, ergeben sich 5 geschätzte Faktoren. (Für nachfolgende zugehörige Tabellen siehe Tabelle 32 bis 36 im Appendix.)

Interpretation Faktor 1: Offenheit gegenüber jeglicher Art von Evaluation, die auch eine Schulmodernisierung fördern sollte

Interpretation Faktor 2: Gefühl des positiven Zusammenhalts zwischen Kollegen, aber besonders als stärkendes Gemeinschaftsgefühl gegenüber an sie gestellte Regeln

Interpretation Faktor 3: Einheitliche von allen akzeptierte Leistungsziele an die Schüler, die eindeutig und transparent sind

Interpretation Faktor 4: Gefühl der Beengtheit bzw. der Überforderung durch Verwaltungsaufgaben

Interpretation Faktor 5: Zu wenig gemeinsame Absprachen unter den Lehrern, so dass sich nur einige Lehrer verstärkt in der Schule engagieren

6.3 Fragebereich 3

Fragebereich 3, Variablenblöcke 9 und 11 Es können jeweils 9 Eigenwerte aus der polychorischen Korrelationsmatrix und aus der Bravais-Pearson Korrelationsmatrix gebildet werden, die größer als 1 sind. Es werden 9 Faktoren

in der Explorativen Faktoranalyse und 9 Faktoren in der Underlying Variable Analyse gebildet.

(Für nachfolgende zugehörige Tabellen siehe Tabelle 37 bis 44 im Appendix.)

Interpretation Faktor 1: Verwaltungsaufgaben behindern die Lehrer in der effizienten Ausübung ihrer Tätigkeiten

Interpretation Faktor 2: Schüler fördern die eigene pädagogische Tätigkeit der Lehrer durch „Feedback“

Interpretation Faktor 3: Befürwortung und Interesse an der Mitwirkung von Schulinnovationen

Interpretation Faktor 4: Eigenverantwortliche Bestimmung des Unterrichtsablaufs

Interpretation Faktor 5: Konferenzbeschlüsse werden vom jeweiligen Lehrer stets durchgeführt, unabhängig von seiner pers. Einstellung demgegenüber

Interpretation Faktor 6: Keine Interpretation möglich, da die Voraussetzungen an die gemeinsamen Faktoren nicht erfüllt sind

Interpretation Faktor 7: Empfundene Überforderung im Beruf durch zu viele Aufgaben

Interpretation Faktor 8: Gefühl der Fremdbestimmtheit in der eigenen pädagogischen Arbeit

Interpretation Faktor 9: Leistungen der Schüler von den eigenen pädagogischen Leistungen abhängig machen

Fragebereich 3, Variablenblöcke 10 und 12 Die Analysen haben beide 6 Eigenwerte der Korrelationsmatrizen ausgewiesen, die größer als 1 sind. Es wurden daher in der EFA als auch in der UVA 6 Faktoren gebildet. (Für nachfolgende zugehörige Tabellen siehe Tabelle 45 bis 49 im Appendix.)

Interpretation Faktor 1: Leistungen der Schüler von dem eigenen selbstbestimmten Unterrichtsablauf abhängig machen.

Interpretation Faktor 2: Überforderung im Beruf durch zu viele Aufgaben

Interpretation Faktor 3: Gefühl der Fremdbestimmtheit in eigener pädagogischen Arbeit

Interpretation Faktor 4: Einhalten von Regeln und Beschlüssen

Interpretation Faktor 5: Befürwortung und Interesse an der Mitwirkung von Schulinnovationen

Interpretation Faktor 6: Keine Interpretation möglich, da die Voraussetzungen an die gemeinsamen Faktoren nicht erfüllt sind

Korrelation der Faktorwerte aus EFA und UVA Im folgenden sei eine Tabelle angeführt, die einen Überblick über die Bravais-Pearson Korrelationsergebnisse gibt. Die Korrelationen wurden, so weit dies möglich war, aus den Faktorwerten der verschiedenen Modelle gebildet. Konnte ein Faktorwert aus der UVA nicht gebildet werden, sei an dieser Stelle in Tabelle 26 ein „ - “ aufgeführt.

Frage- bereich	Variablen- block	Faktor	B.-Pearson Korr. zw. den Faktorwerten aus EFA und UVA
1	1	1	0.97
	1	2	-
2	5+7	1	0.92
	5+7	2	0.90
	5+7	3	0.91
	5+7	4	0.79
	5+7	5	0.78
	5+7	6	-
	5+7	7	-
2	6+8	1	0.80
	6+8	2	0.84
	6+8	3	-
	6+8	4	0.82
	6+8	5	0.80
3	9+11	1	0.93
	9+11	2	0.88
	9+11	3	0.94
	9+11	4	0.90
	9+11	5	-
	9+11	6	-
	9+11	7	-
	9+11	8	-
	9+11	9	0.86
3	10+12	1	0.93
	10+12	2	0.92
	10+12	3	0.81
	10+12	4	0.83
	10+12	5	0.82
	10+12	6	-

Tabelle 26: Korrelationen der jeweiligen Faktoren aus EFA und UVA

Empfehlung für die Verwendung der Faktorwerte Wie bereits zu Beginn der Analyse erklärt, wird die Faktorenanalyse auch durchgeführt, um mit einer verminderten Variablenzahl, also den generierten Faktorwerten weitere statistische Analysen durchzuführen.

Es ist anhand der oben angeführten Plots und Bravais-Pearson Korrelation zu zeigen, dass man einen starken Zusammenhang zwischen den Faktorwerten aus EFA und UVA ableiten kann. Jedoch sollten die Ergebnisse der Faktorwerte der EFA auf ordinale Daten nur unter Vorbehalt weiter verwendet werden. Dass heißt, dass nicht die Werte der Faktoren selbst verwendet werden, sondern vielleicht nur ihre Rangordnung bei den jeweiligen Beobachtungen.

Anmerkung: Die Ergebnisse der Faktoranalyse unter Mplus und SPSS sind auf beiliegender CD verfügbar.

7 Fazit

Die Analyse der ordinalen Daten über die Modelle der Explorativen Faktoranalyse sowie durch die Underlying Variable Analyse hat folgendes Ergebnis:

Es lässt sich zeigen, dass die gebildeten Faktoren beinahe vollständig in beiden Analysen über die gleichen beobachtbaren Variablen gebildet wurden. Der Grund dafür war, dass trotz verschiedener Zusammenhangsmaße ähnliche Korrelationsmatrizen aus den Daten entstanden sind.

Grundsätzlich lässt sich jedoch von diesem Ergebnis nicht darauf schließen, dass die EFA trotz Anwendung auf ordinale Daten stets zu sinnvollen Ergebnissen führt, wenn die Kategorienanzahl nur hinreichend groß genug ist (> 4 Kategorien).

So wurde zu Beginn der Untersuchung aufgezeigt, dass die Modellannahmen der Explorativen Faktoranalyse nur in einem geringen Umfang von dem vorliegenden Datensatz erfüllt werden konnte. De facto konnte nur 1 von 4 Modellannahmen (großer Stichprobenumfang) an die Daten erfüllt werden.

Die Voraussetzung der metrischen Daten wurde für die Analyse als einzige Voraussetzung bewusst verletzt. Aber darüber hinaus konnte weder die approximative Normalverteilungsannahme, noch die Anforderung unabhängiger Beobachtungen gesichert werden.

Es sei anzumerken, dass ebenfalls die einzige Voraussetzung der Underlying Variable Analyse an die Daten, normalverteilte Underlying Variables, nicht vollständig erfüllt werden konnte.

Zwar wurden in der Simulation der Zusammenhangsmaße jeweils verschiedene Aspekte von Abweichungen von Modellannahmen untersucht, doch um eine zuverlässige Aussage machen zu können mit welchen negativen Ergebnissen man bei der Verletzung von mehreren Modellannahmen rechnen muss, müssten Simulationen ganzer Faktoranalysen durchgeführt werden.

Das bedeutet, man müsste in den simulierten Daten eine latente Struktur

vorgeben, anschließend einzelne Voraussetzungen (und verschiedene Kombinationen von Voraussetzungen) bewusst verletzen und dann die Ergebnisse erneut untersuchen. Dann wären Aussagen über die Robustheit der Explorativen Faktoranalyse eventuell möglich.

Daher ist an dieser Stelle keine generelle Aussage über die Qualität der Ergebnisse bei Anwendung von EFA bei Verletzung mehrerer Modellannahmen möglich.

8 Quellen

- Allison, P.D.(2001). Missing Data, in Series: Quantitative Applications in the social sciences, Sage University Paper.
- Backhaus, K.; Erichson, B.; Plinke, W. und Weiber, R.(2003). Multivariate Analysemethoden: Eine anwendungsorientierte Einführung, 10 Ed., Springer, Berlin.
- Bartholomew, D. J.; Steele, F.; Moustaki, I.; und Galbraith, J.I.(2002). The Analysis and Interpretation of Multivariate Data for Social Scientists, Chapman & Hall/CRC.
- Bollen, K.A.(1989). Structural Equations with latent Variables, Wiley Series in: Probability and Mathematical Statistics, Wiley, New York.
- Bollen, K.A. und Barb, K.H.(1981). Pearson's R and coarsely categorized measures. American Sociological Review, 46: 232-239.
- Green, P.E. und Rao, V.R.(1970). Rating Scales and Information Recovery - How many Scales and Response Categories to use?. Journal of Marketing, 34: 33-39.
- Härdle, W. und Simar, L.(2003). Applied Multivariate Statistical Analysis, Springer.
- Jöreskog, K. und Sörbom, D.(1993)(nach Reinecke). LISREL 8: Structural equation modeling with the SIMPLIS command language. Scientific Software International, Erlbaum.
- Kölling, A. und Rässler, S. Effekte der Multiplen Imputation fehlender Werte am Beispiel von Produktivitätsschätzungen mit dem IAB-Betriebspanel.
www.statistik.wiso.uni-erlangen.de/diskussionspapiere/d0040.pdf
(Dieses Dokument befindet sich auf beiliegender CD und im Anhang)
- Kotz, S.; Balakrishnan, N. und Johnson, N.L.(2000). Continuous Multivariate Distributions, Volume 1: Models and Applications, 2nd Ed., Wiley Series in Probability and Statistics, Wiley.
- Olsson, U.; Drasgow, F. und Dorans, N.J.(1982). The Polyserial Correlation Coefficient. Psychometrika, 47(3): 337-347.
- Olsson, U.(1979). Maximum Likelihood estimation of the polychoric correlation coefficient. Psychometrika, 44: 443-460.

- Raghunathan, T.E.; Lepkowski, J.M.; van Hoewyk, J. und Solenberger, P.(2001). A Multivariate Technique for Multiply Imputing Missing Values using a Sequence of Regression Models. *Survey Methodology*, 27(1): 85-88.
- Reinicke, J. Statistische Grundlagen für Strukturgleichungsmodelle.
www.uni-trier.de/uni/fb4/soziologie/faecher/empirik/fund.pdf
(Dieses Dokument befindet sich auf beiliegender CD und im Anhang)
- Rönz, B.(2000). Skript Computergestützte Statistik II.
- Rönz, B. Skript Verallgemeinerte Lineare Modelle.
- Schafer, J.(1997). *Analysis of incomplete Multivariate Data*. Chapman and Hall/-CRC Press, London.

A Appendix

³Die Variable ist im UVA - Modell äußerst niedrig geladen:(-0.18)

⁴Die Variable ist im UVA - Modell äußerst niedrig geladen:0.03

⁵Die Variable ist im UVA - Modell äußerst niedrig geladen:0.12

Anzahl der Kategorien	Abstände in Prozent	Aufteilung
2	50%	$[-\text{Inf}; 0] = 1$ $[0; \text{Inf}] = 2$
3	33.3%	$[-\text{Inf}; -0.43] = 1$ $[-0.43; 0.43] = 2$ $[0.43; \text{Inf}] = 3$
4	25%	$[-\text{Inf}; -0.67] = 1$ $[-0.67; 0] = 2$ $[0; 0.67] = 3$ $[0.67; \text{Inf}] = 4$
5	20%	$[-\text{Inf}; -0.84] = 1$ $[-0.84; -0.25] = 2$ $[-0.25; 0.25] = 3$ $[0.25; 0.84] = 4$
		$[0.84; \text{Inf}] = 5$
6	16.6%	$[-\text{Inf}; -0.97] = 1$ $[-0.97; -0.43] = 2$ $[-0.43; 0] = 3$ $[0; 0.43] = 4$
		$[0.43; 0.97] = 5$ $[0.97; \text{Inf}] = 6$
7	14.3%	$[-\text{Inf}; -1.07] = 1$ $[-1.07; -0.57] = 2$ $[-0.57; -0.18] = 3$ $[-0.18; 0.18] = 4$
		$[0.18; 0.57] = 5$ $[0.57; 1.07] = 6$ $[1.07; \text{Inf}] = 7$
8	12.5%	$[-\text{Inf}; -1.15] = 1$ $[-1.15; -0.67] = 2$ $[-0.67; -0.32] = 3$ $[-0.32; 0] = 4$
		$[0; 0.32] = 5$ $[0.32; 0.67] = 6$ $[0.67; 1.15] = 7$ $[1.15; \text{Inf}] = 8$
9	11.11%	$[-\text{Inf}; -1.22] = 1$ $[-1.22; -0.76] = 2$ $[-0.76; -0.43] = 3$ $[-0.43; -0.14] = 4$
		$[-0.14; 0.14] = 5$ $[0.14; 0.43] = 6$ $[0.43; 0.76] = 7$ $[0.76; 1.22] = 8$
		$[1.22; \text{Inf}] = 9$
10	10%	$[-\text{Inf}; -1.28] = 1$ $[-1.28; -0.84] = 2$ $[-0.84; -0.52] = 3$ $[-0.52; -0.25] = 4$
		$[-0.25; 0] = 5$ $[0; 0.25] = 6$ $[0.25; 0.52] = 7$ $[0.52; 0.84] = 8$
		$[0.84; 1.28] = 9$ $[1.28; \text{Inf}] = 10$

Tabelle 27: Kategorienaufteilung(symmetrisch)

Kategorie	Wahre Korrelation	Bravais-Pearson Korrelation auf metrische Daten	Stdabw.	Bravais-Pearson Korrelation auf kategoriale Daten	Stdabw.
2	0.2	0.199	0.048	0.128	0.049
2	0.4	0.399	0.042	0.262	0.048
2	0.5	0.499	0.038	0.333	0.047
2	0.6	0.599	0.032	0.409	0.046
2	0.8	0.800	0.018	0.590	0.041
2	0.9	0.899	0.010	0.712	0.035
3	0.2	0.199	0.049	0.159	0.049
3	0.4	0.399	0.042	0.323	0.046
3	0.5	0.500	0.037	0.409	0.043
3	0.6	0.599	0.032	0.496	0.040
3	0.8	0.799	0.018	0.690	0.028
3	0.9	0.899	0.009	0.798	0.020
4	0.2	0.198	0.047	0.170	0.048
4	0.4	0.400	0.042	0.348	0.045
4	0.5	0.499	0.038	0.437	0.042
4	0.6	0.599	0.032	0.531	0.037
4	0.8	0.799	0.018	0.728	0.024
4	0.9	0.899	0.009	0.835	0.016
5	0.2	0.199	0.048	0.178	0.048
5	0.4	0.399	0.042	0.359	0.044
5	0.5	0.500	0.037	0.453	0.041
5	0.6	0.599	0.031	0.548	0.036
5	0.8	0.799	0.018	0.748	0.023
5	0.9	0.899	0.009	0.854	0.014

Tabelle 28: Korrelation der simulierten kategorisierten Daten, n=400, 10.000 Durchläufe

Kategorie	Wahre Korrelation	Bravais-Pearson Korrelation auf metrische Daten	Stdabw.	Bravais-Pearson Korrelation auf kategoriale Daten	Stdabw.
6	0.2	0.200	0.048	0.182	0.049
6	0.4	0.399	0.041	0.366	0.044
6	0.5	0.499	0.038	0.461	0.040
6	0.6	0.599	0.032	0.557	0.036
6	0.8	0.799	0.018	0.759	0.023
6	0.9	0.899	0.010	0.865	0.013
7	0.2	0.199	0.048	0.184	0.049
7	0.4	0.399	0.042	0.370	0.044
7	0.5	0.499	0.037	0.466	0.039
7	0.6	0.599	0.032	0.563	0.035
7	0.8	0.799	0.018	0.765	0.021
7	0.9	0.899	0.010	0.872	0.013
8	0.2	0.199	0.048	0.185	0.048
8	0.4	0.399	0.042	0.373	0.043
8	0.5	0.500	0.037	0.470	0.039
8	0.6	0.600	0.032	0.568	0.035
8	0.8	0.799	0.018	0.770	0.021
8	0.9	0.899	0.010	0.876	0.012
9	0.2	0.199	0.048	0.187	0.049
9	0.4	0.399	0.042	0.376	0.044
9	0.5	0.499	0.038	0.471	0.039
9	0.6	0.599	0.032	0.570	0.035
9	0.8	0.799	0.018	0.773	0.021
9	0.9	0.899	0.009	0.879	0.012
10	0.2	0.199	0.048	0.187	0.048
10	0.4	0.399	0.042	0.378	0.043
10	0.5	0.499	0.037	0.474	0.039
10	0.6	0.599	0.032	0.572	0.035
10	0.8	0.799	0.018	0.775	0.021
10	0.9	0.899	0.009	0.881	0.012

Tabelle 29: Korrelation der simulierten kategorisierten Daten, n=400, 10.000 Durchläufe
(Fortsetzung)

Klassen- aufteilung in %	Wahre Korrelation	Bravais-Pearson Korrelation auf kategoriale Daten	Stdabw.
10-90	0.2	0.079	0.062
10-90	0.4	0.184	0.070
10-90	0.5	0.249	0.072
10-90	0.6	0.321	0.074
10-90	0.8	0.513	0.072
10-90	0.9	0.654	0.064
20-80	0.2	0.104	0.055
20-80	0.4	0.226	0.058
20-80	0.5	0.295	0.057
20-80	0.6	0.369	0.057
20-80	0.8	0.555	0.052
20-80	0.9	0.687	0.046
30-70	0.2	0.120	0.052
30-70	0.4	0.248	0.052
30-70	0.5	0.318	0.0510
30-70	0.6	0.395	0.050
30-70	0.8	0.577	0.045
30-70	0.9	0.703	0.039
40-60	0.2	0.126	0.049
40-60	0.4	0.259	0.050
40-60	0.5	0.330	0.047
40-60	0.6	0.406	0.047
40-60	0.8	0.587	0.041
40-60	0.9	0.711	0.036
50-50	0.2	0.128	0.050
50-50	0.4	0.262	0.048
50-50	0.5	0.334	0.046
50-50	0.6	0.410	0.045
50-50	0.8	0.590	0.040
50-50	0.9	0.713	0.035

Tabelle 30: Korrelation der simulierten asymmetrisch klassifizierten Daten in 2 Kategorien, n=400, 10.000 Durchläufe

Modell	EFA- Eigenwert	UVA- Eigenwert
Block 6+8	10.934	12.046
	2.114	2.173
	1.338	1.342
	1.225	1.202
	1.109	1.089
Block 9+11	6.853	7.558
	3.657	4.237
	1.924	1.992
	1.589	1.649
	1.345	1.357
	1.183	1.179
	1.143	1.149
	1.081	1.093
	1.051	1.041
Block 10+12	10.914	11.760
	4.478	4.398
	1.875	1.840
	1.337	1.312
	1.148	1.125
	1.052	1.034

Tabelle 31: Eigenwerte der jeweils gemeinsam untersuchten Blöcke mit Werten > 1

Modell		EFA-	UVA-
1.Faktor	Variable	Ladung	Ladung
	f0611	0.724	0.659
	f0804	0.633	0.564
	f0803	0.615	0.538
	f0613	0.671	0.639
	f0805	0.627	*
	f0614	0.447	0.550

Tabelle 32: Faktorladungen des 1.Faktors der Variablenblöcke 6 und 8, Bereich 2

Modell		EFA-	UVA-
2.Faktor	Variable	Ladung	Ladung
	f0608	0.656	0.589
	f0604	0.652	0.596
	f0605	0.649	0.598
	f0603	0.601	0.555
	f0808	0.573	0.505
	f0809	0.556	0.483

Tabelle 33: Faktorladungen des 2.Faktors der Variablenblöcke 6 und 8, Bereich 2

Modell		EFA-	UVA-
3.Faktor	Variable	Ladung	Ladung
	f0609	0.779	0.739
	f0607	0.694	0.549
	f0610	0.610	*
	f0612	0.534	*
	f0601	0.524	0.425

Tabelle 34: Faktorladungen des 3.Faktors der Variablenblöcke 6 und 8, Bereich 2

Modell		EFA-	UVA-
4.Faktor	Variable	Ladung	Ladung
	f0806	0.764	0.446
	f0811	0.622	0.706
	f0802	0.585	0.498

Tabelle 35: Faktorladungen des 4.Faktors der Variablenblöcke 6 und 8, Bereich 2

Modell		EFA-	UVA-
5.Faktor	Variable	Ladung	Ladung
	f0812	0.731	0.634
	f0807	0.633	0.514
	f0815	0.612	0.576
	f0813	0.559	0.439
	f0814	0.539	0.466
	f0816	0.512	0.493
	f0805	*	0.509

Tabelle 36: Faktorladungen des 5.Faktors der Variablenblöcke 6 und 8, Bereich 2

Modell		EFA-	UVA-
1.Faktor	Variable	Ladung	Ladung
	f1103	0.810	0.841
	f0905	0.788	0.765
	f1108	0.775	0.780
	f1105	0.716	0.733
	f0914	0.697	0.694
	f0913	0.496	0.508
	f0903	*	0.598

Tabelle 37: Faktorladungen des 1.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
2.Faktor	Variable	Ladung	Ladung
	f0912	0.680	0.612
	f0901	0.700	0.554
	f1104	0.547	0.357

Tabelle 38: Faktorladungen des 2.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
3.Faktor	Variable	Ladung	Ladung
	f1101	0.800	0.778
	f1111	0.747	0.655
	f0910	0.711	0.658
	f1117	0.616	0.601

Tabelle 39: Faktorladungen des 3.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
4.Faktor	Variable	Ladung	Ladung
	f0917	0.761	0.665
	f0906	0.735	0.629
	f1106	0.681	0.692

Tabelle 40: Faktorladungen des 4.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
5.Faktor	Variable	Ladung	Ladung
	f0916	0.792	0.718
	f0907	0.757	0.537

Tabelle 41: Faktorladungen des 5.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
7.Faktor	Variable	Ladung	Ladung
	f0902	0.696	0.606
	f0904	0.624	0.651
	f0903	0.595	*

Tabelle 42: Faktorladungen des 7.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
8.Faktor	Variable	Ladung	Ladung
	f1114	0.658	0.678
	f1113	0.609	0.606
	f1112	0.533	** 3
	f0911	-0.503	** 4
	f1109	0.471	** 5

Tabelle 43: Faktorladungen des 8.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
9.Faktor	Variable	Ladung	Ladung
	f0908	0.703	0.536
	f1115	0.613	0.480
	f1107	0.610	0.582

Tabelle 44: Faktorladungen des 9.Faktors der Variablenblöcke 9 und 11, Bereich 3

Modell		EFA-	UVA-
1.Faktor	Variable	Ladung	Ladung
	f1008	0.806	0.795
	f1216	0.785	0.749
	f1206	0.736	0.728
	f1207	0.722	0.715
	f1012	0.708	0.668
	f1215	0.657	0.632
	f1015	0.634	0.640
	f1017	0.647	0.638
	f1202	0.619	0.639
	f1001	0.619	0.630
	f1006	0.586	0.600
	f1011	0.582	0.553

Tabelle 45: Faktorladungen des 1.Faktors der Variablenblöcke 10 und 12, Bereich 3

Modell		EFA-	UVA-
2.Faktor	Variable	Ladung	Ladung
	f1004	0.848	0.841
	f1002	0.801	0.745
	f1003	0.797	0.761
	f1005	0.635	0.638
	f1208	0.580	0.587
	f1203	0.560	0.592
	f1013	0.527	0.498

Tabelle 46: Faktorladungen des 2.Faktors der Variablenblöcke 10 und 12, Bereich 3

Modell		EFA-	UVA-
3.Faktor	Variable	Ladung	Ladung
	f1214	0.759	0.653
	f1213	0.686	0.610
	f1212	0.484	0.399
	f1205	0.480	0.428*** ⁶

Tabelle 47: Faktorladungen des 3.Faktors der Variablenblöcke 10 und 12, Bereich 3

⁶tatsächlich im 4. Faktor höher geladen(-0.471), aber für Bildung von Faktorwerten in der CFA hier eingesetzt

Modell		EFA-	UVA-
4.Faktor	Variable	Ladung	Ladung
	f1014	0.680	-0.657
	f1007	0.677	-0.561
	f1016	0.641	-0.555

Tabelle 48: Faktorladungen des 4.Faktors der Variablenblöcke 10 und 12, Bereich 3

Modell		EFA-	UVA-
5.Faktor	Variable	Ladung	Ladung
	f1211	0.725	0.679
	f1201	0.665	0.579
	f1217	0.602	0.470
	f1210	0.575	0.568
	f1209	0.558	0.515
	f1010	0.554	0.458

Tabelle 49: Faktorladungen des 5.Faktors der Variablenblöcke 10 und 12, Bereich 3

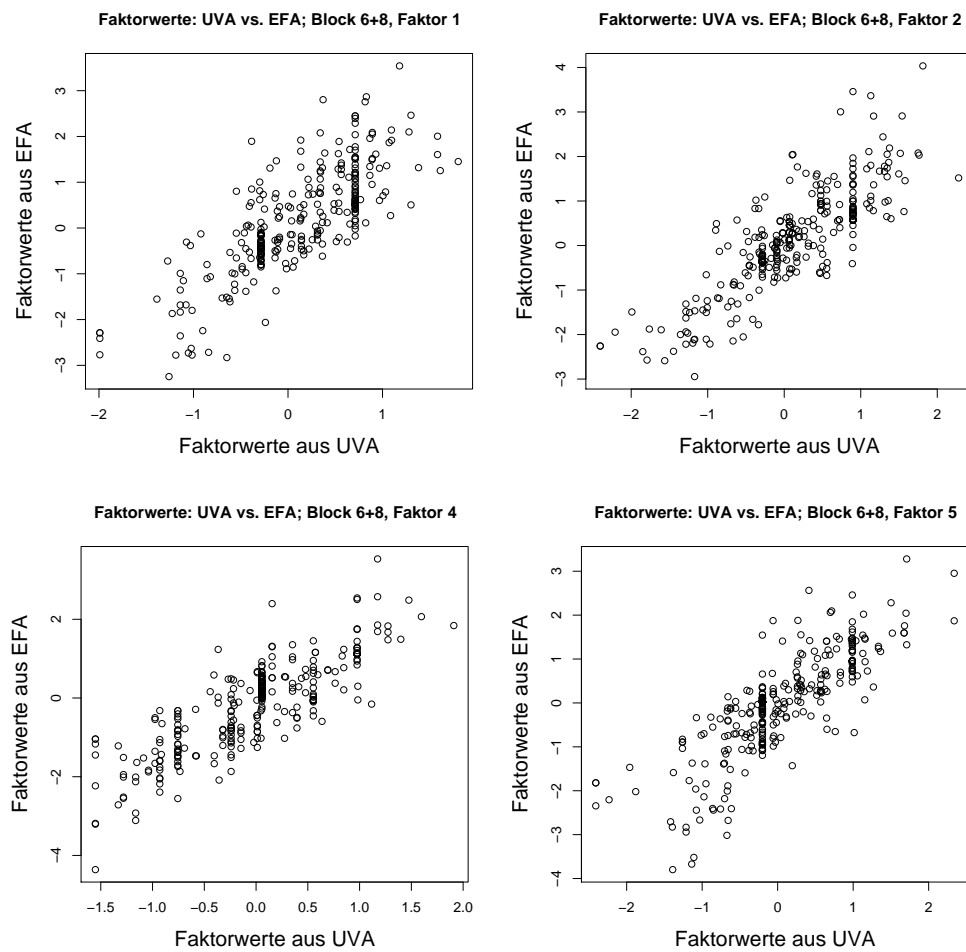


Abbildung 18: Plots des Faktorwerte der Variablenblöcke 6+8

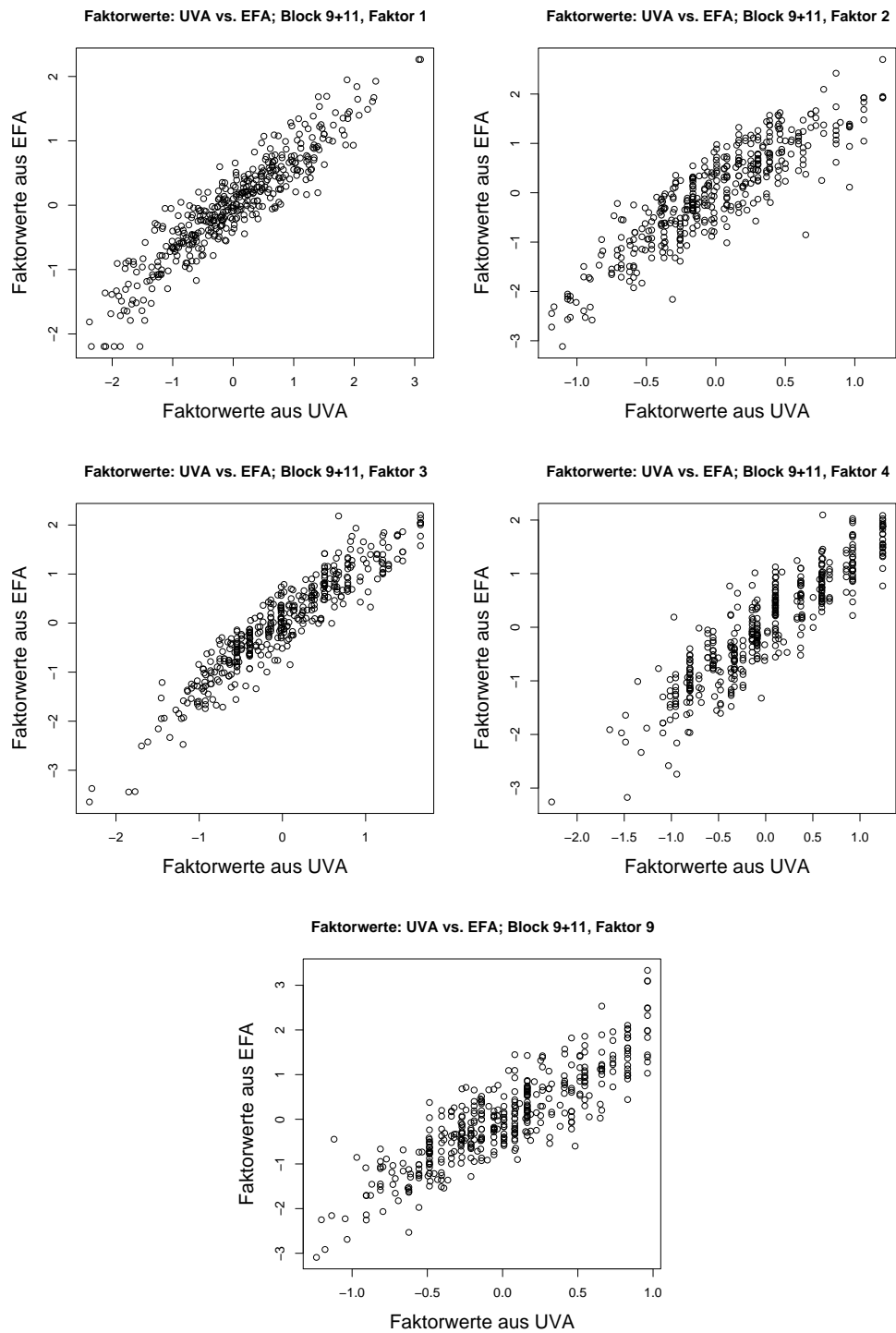


Abbildung 19: Plots des Faktorwerte der Variablenblöcke 9+11

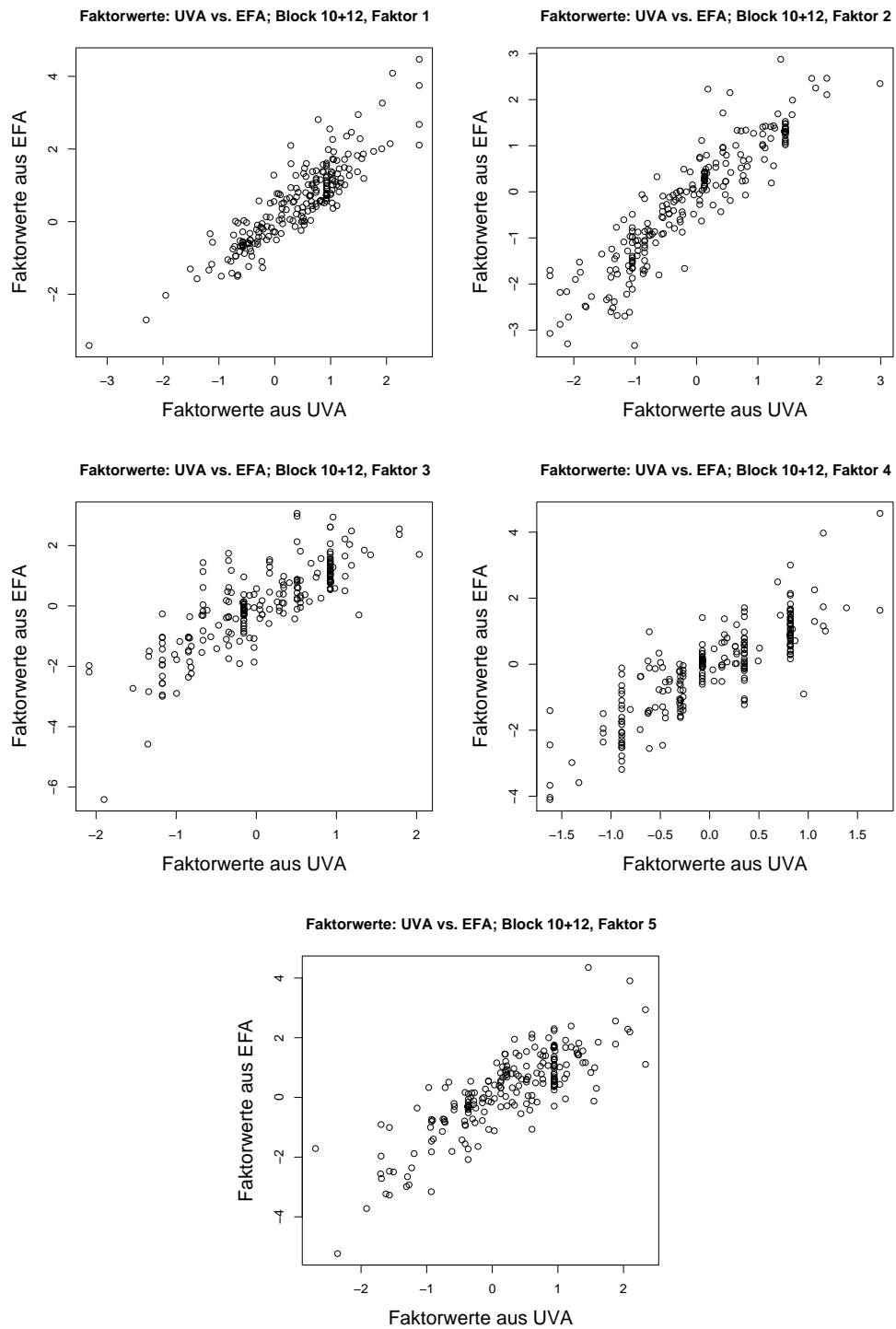


Abbildung 20: Plots des Faktorwerte der Variablenblöcke 10+12

Humboldt-Universität zu Berlin
 Philosophische Fakultät IV
 Institut für Erziehungswissenschaften
 Abteilung Wirtschaftspädagogik
 Geschwister-Scholl-Str. 7
 10117 Berlin



Tel.: 2093-4189
 Fax: 2093-4165
 Ansprechpartner:
 Dr. Olga Zlatkin-Troitschanskaia
 in Zusammenarbeit mit
 Prof. Dr. Jürgen van Buer

ZUR SCHULENTWICKLUNG AN DEN BERLINER BERUFSBILDENDEN SCHULEN
- BEFRAGUNG DER LEHRKRÄFTE -
(DEZEMBER 2004)

Sehr geehrte Lehrerinnen und Lehrer,

der Fokus dieser (Teil-)Studie liegt auf der Erfassung und Analyse Ihrer Einschätzungen und Einstellungen zu Ihrer pädagogischen Arbeit sowie zu Ihrer Schule zum gegenwärtigen Zeitpunkt. Um gezielt in Ihrer Schule Verbesserungsmaßnahmen realisieren zu können, sind genaue Informationen über die gegenwärtig wahrgenommene Situation vom gesamten Kollegium erforderlich. Die Ergebnisse der Studie werden dem *Schulentwicklungsausschuss* gemeldet und sollen zur *Förderung der Qualität* der pädagogischen Prozesse sowie zur Schulentwicklung in Ihrer Schule beitragen.

Da in Ihrer Schule bereits mehrere Erhebungen durchgeführt und ausgewertet wurden, bezieht sich diese (Teil-)Studie auf die ausgewählten Aspekte, die bislang noch nicht in dieser Form erfasst sind. In den zurückliegenden Jahren wurden an Ihrer Schule die Jugendlichen als Adressaten von Unterricht befragt. Die nachfolgenden Fragen werden an Sie als diejenigen gestellt, die diese Lehr-Lern-Prozesse ausgestalten und steuern.

Die Teilnahme an der Befragung ist *freiwillig*. Wir versichern Ihnen, dass Ihre Angaben *anonym* behandelt werden. Wir garantieren, dass zum einen *kein Dritter Einblick in die und Zugang zu den Fragebögen* erhält und zum anderen Ihre Angaben nicht in individualisierter Form an Dritte weitergegeben werden. Datenschutzrechtliche Bestimmungen werden strengstens eingehalten.

Bitte füllen Sie die nachfolgenden Fragen vollständig aus, legen Sie den Fragebogen in den beiliegenden Briefumschlag und geben Sie diesen verschlossen im Schulsekretariat ab.

Die schriftliche Rückmeldung dieser (Teil-)Studie erfolgt ab Ende Januar 2005. Bei der Rückmeldung der Daten werden keinerlei Aussagen zu einzelnen Fragebögen dargestellt, sondern lediglich Gruppendaten, die Rückschlüsse auf Einzelpersonen ausschließen. Bitte erkundigen Sie sich nach einer Kopie der Rückmeldung in Ihrem Schulsekretariat.

Vielen herzlichen Dank für Ihre Mitarbeit!

Codierung

Wie versichert, werden Ihre Angaben anonym behandelt. Um jedoch eine eindeutige Zuordnung bei eventuell zukünftig stattfindenden Untersuchungen zu ermöglichen, ist eine Codierung der Fragebögen erforderlich. Wir haben uns bemüht, einen Code zu finden, der einerseits jederzeit verfügbar ist und andererseits keine Rückschlüsse auf Einzelpersonen zulässt. Dieser Code dient *ausschließlich einer datentechnischen Zuordnung* bei möglichen weiteren Befragungen. Wir garantieren, dass *keinem Dritten* Ihr Code mitgeteilt wird bzw. dass keiner Zugang zu diesem erhält. Nach der Zuordnung der einzelnen Fragebögen in den Gesamtdatensatz wird die Codierung gelöscht.

- 1. Geburtsdatum der Mutter (nur Tag und Monat):** . . f 0001
- 2. Die ersten drei Buchstaben des Vornamens der Mutter:** f 0002

FRAGEBEREICH 1, BLOCK 1		(1)	(2)	(3)	(4)
Die Arbeitsbedingungen in der Abteilung, in der ich hauptsächlich arbeite, kann ich wie folgt beschreiben:		ja	eher ja	eher nein	nein
f 0101	1. Meine pädagogische Freiheit ist zu stark eingeengt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0102	2. Ich bin mit Aufgaben und Zuständigkeiten belastet – aber angemessen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0103	3. Die Aufgaben sind in dieser Abteilung gerecht und kollegial verteilt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0104	4. Für das Erreichen der Arbeitsziele gibt es klare Vorgaben und Termine.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0105	5. Vereinbarte Arbeitsziele werden in dieser Abteilung i. d. R. eingehalten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0106	6. Meine Arbeit in der Abteilung wird von meinem Kollegium anerkannt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0107	7. Interessengegensätze werden offen diskutiert und konsensorientiert aufgearbeitet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0108	8. Wir verfolgen in dieser Abteilung dieselben pädagogischen Absichten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0109	9. Von meinen KollegInnen erhalte ich regelmäßig Rückmeldung über meine Arbeit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0110	10. Eigentlich kann mir bei meiner pädagogischen Arbeit faktisch Keiner etwas vorschreiben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f 0111	11. Unser Kollegium in der Abteilung hat ein starkes Innovationspotenzial, so dass Neuerungen immer durchgesetzt werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Was wünschen Sie sich vor allem für ihre Tätigkeit als Lehrerin/Lehrer?

Wählen Sie bitte aus den nachfolgenden Vorschlägen zwei (2) aus und kreuzen Sie diese bitte an:

f 0201	mehr Freiräume für meine pädagogische Arbeit	<input type="checkbox"/>
f 0202	mehr Informationen über neuere Entwicklungen in Schule und Unterricht	<input type="checkbox"/>
f 0203	mehr Kompetenzen für pädagogische Aufgaben	<input type="checkbox"/>
f 0204	mehr Kompetenzen für Verwaltungsaufgaben	<input type="checkbox"/>
f 0205	mehr Beratung und Unterstützung bei pädagogischen Aufgaben	<input type="checkbox"/>
f 0206	mehr Beratung und Unterstützung bei Verwaltungsaufgaben	<input type="checkbox"/>
f 0207	mehr Transparenz bei der Schulorganisation und -entwicklung	<input type="checkbox"/>
f 0208	mehr Transparenz bei der Schulleitung	<input type="checkbox"/>
f 0209	mehr Koordination bei der Schulorganisation und -entwicklung	<input type="checkbox"/>
f 0210	mehr Zeit für pädagogische Aufgaben	<input type="checkbox"/>
f 0211	mehr Zeit für Verwaltungsaufgaben	<input type="checkbox"/>
f 0212	mehr Zeit für den Informationsaustausch und die Zusammenarbeit mit KollegInnen	<input type="checkbox"/>

Allgemeine Angaben zu Ihrer Person und Ihrer Tätigkeit:

f 0301 **Tätigkeit:** (1) Diplom-Lehrer/-in bzw. Studienrat/-rätin (2) Lehrkraft in Funktionsstelle (A14/A15)
(3) Referendar/-in (4) Lehrkraft für Fachpraxis

Wie lange sind Sie bereits (einschließlich Referendariatszeit) Lehrerin/Lehrer?

f 0302 **im Schuldienst:** weniger als 1 Jahr (1) 1 - 5 Jahre (2) 6 - 10 Jahre (3) 11 - 20 Jahre (4) über 20 Jahre (5)
f 0303 **in diesem OSZ:** weniger als 1 Jahr 1 - 5 Jahre 6 - 10 Jahre 11 - 20 Jahre über 20 Jahre

f 0304 **derzeitige Beschäftigung:** Vollzeit Teilzeit

In welchen Bildungsgängen sind Sie im aktuellen Schuljahr 2004/2005 eingesetzt?		Einsatz in ff. Bildungsgängen	Bildungsgang mit dem höchsten Deputat (bitte <u>nur 1</u> Kreuz!)
Berufsvorbereitende Lehrgänge (z. B. MDQM I, BQL, BB 10, BVL)			
Berufsschule (auch Doppelqualifizierung, Verbundausbildung)			
f 0403	Berufsfachschule (OBF 1, OBF 2, OBF 3)		
f 0404	Fachoberschule		
	Fachschule		
	Gymnasiale Oberstufe		

f 0407

FRAGEBEREICH 2, BLOCK 5 UND 6		... zum gegenwärtigen Stand					... wahrgenommene Veränderung ¹					
	<i>Für meine Schule als gesamte Institution kann ich Folgendes konstatieren. ...:</i>	trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver	
f 0501	In meiner Schule gibt es einheitliche Bewertungsmaßstäbe für die Leistungsergebnisse der Schüler.											f 0601
f 0502	Meine Schule führt derzeit viele Veränderungen in verschiedenen Bereichen durch.											f 0602
f 0503	In meiner Schule werden die meisten KollegInnen in die Entscheidungsprozesse einbezogen.											f 0603
f 0504	In meiner Schule gibt es zu viele starre Regelungen im Bereich der Organisation des Unterrichts.											f 0604
f 0505	In meiner Schule wird oft über meinen Kopf hinweg entschieden.											f 0605
f 0506	Meine Schule bemüht sich zu wenig um Erneuerung und Entwicklung in Schulorganisation und Unterricht.											f 0606
f 0507	In meiner Schule gibt es klare, von allen Lehrkräften akzeptierte Zielvereinbarungen hinsichtlich der Unterrichtsprozesse.											f 0607
f 0508	In meiner Schule herrscht ein kollegialer Führungsstil.											f 0608
f 0509	In meiner Schule gibt es klare, von allen Lehrkräften akzeptierte Mindeststandards hinsichtlich der zu erreichenden Leistungsergebnisse bei den Schülern.											f 0609
f 0510	In meiner Schule gibt es klare Entwicklungsziele, die auch die Planungen meiner eigenen pädagogischen Arbeit bestimmen.											f 0610
f 0511	Wir – die Lehrkräfte in meiner Schule – sind verpflichtet, an Tests, Befragungen, Erhebungen zur Evaluation unserer Schule und der pädagogischen Arbeit teilzunehmen.											f 0611
f 0512	Meine Schule hat ihr eigenes pädagogisches Handlungskonzept, das praktisch von allen weitgehend akzeptiert ist.											f 0612
f 0513	Meine Schule hat von allen Lehrkräften akzeptierte Evaluationskriterien ausgearbeitet, mit denen die Qualität pädagogischer Arbeit beurteilt wird.											f 0613
f 0514	Meine Schule benötigt einen klaren Modernisierungsschub.											f 0614

¹ seitdem die Schule an der Entwicklung des Schulprogramms arbeitet

Abbildung 23: Fragebogen, Seite 3

FRAGEBEREICH 2, BLOCK 7 UND 8		Für meine Schule als gesamte Institution kann ich Folgendes konstatieren ...:					... zum gegenwärtigen Stand					...wahrgenommene Veränderung			
		trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver				
f 0701	In meiner Schule werden regelmäßig interne Evaluationen durchgeführt (z. B. in Form von Schülerbefragungen).														f 0801
f 0702	Ich halte den derzeitigen Veränderungsdruck für meine Schule für deutlich übertrieben.														f 0802
f 0703	Meine Schule hat ein Evaluationsprogramm, das in der Schulkonferenz beschlossen worden ist.														f 0803
f 0704	Meine Schule ist für Fremdbeurteilung und für Leistungsvergleiche mit anderen Schulen offen.														f 0804
f 0705	Die Weiterentwicklung meiner Schule basiert primär auf der Initiative und dem Engagement der einzelnen Lehrkräfte.														f 0805
f 0706	Die Verwaltungsvorschriften und Beschlüsse für die Berliner Schulen behindern die wünschenswerten Veränderungen an meiner Schule.														f 0806
f 0707	In meiner Schule habe ich oft das Gefühl, dass sich Niemand für Etwas verantwortlich fühlt.														f 0807
f 0708	In meiner Schule gibt es zu viele starre Regelungen im Bereich der Schulorganisation.														f 0808
f 0709	In meiner Schule habe ich hinreichend Mitbestimmungsmöglichkeiten.														f 0809
f 0710	In meiner Schule gibt es eine klar spürbare Hierarchie zwischen den Lehrkräften und den Mitgliedern der Schulleitung.														f 0810
f 0711	In meiner Schule gelten so viele Vorschriften und Beschlüsse, dass ich sie nicht alle behalten kann.														f 0811
f 0712	In meiner Schule gibt es nur wenige Lehrkräfte, die aktiv an unserer Schulentwicklung arbeiten.														f 0812
f 0713	Die Konferenzen in meiner Schule haben immer den gleichen Ablauf.														f 0813
f 0714	Die Veränderungen an meiner Schule müssen miteinander besser abgestimmt sein.														f 0814
f 0715	In meiner Schule bestimmen fast immer dieselben KollegInnen, was getan werden soll.														f 0815
f 0716	In meiner Schule wird das Schulleben von oben geregelt.														f 0816

Abbildung 24: Fragebogen, Seite 4

FRAGEBEREICH 3, BLOCK 9 UND 10		... zum gegenwärtigen Stand					... wahrgenommene Veränderung				
	<i>Zu meiner alltäglichen <u>Arbeitsfähigkeit</u> an der Schule kann ich Folgendes sagen:</i>	trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver
f 0901	Die SchülerInnen meiner Klasse bewirken meine persönliche pädagogische (Weiter-)Entwicklung.										f 1001
f 0902	Durch meine alltägliche Unterrichtstätigkeit fühle ich mich überlastet.										f 1002
f 0903	Aufgrund der kaum zu bewältigenden Aufgabenfülle verliere ich die Lust an meinem Beruf.										f 1003
f 0904	Die Anforderungen, die an mich als LehrerIn gestellt werden, sind so umfangreich, dass ich all diesen Ansprüchen nicht mehr gerecht werden kann.										f 1004
f 0905	Ich benötige viel Zeit für administrative Tätigkeiten, die eigentlich nicht zu meinen Berufsaufgaben gehören.										f 1005
f 0906	Für mein pädagogisches Handeln trage ich allein die Verantwortung. Selbst dann, wenn ich die Konferenzbeschlüsse nicht vernünftig finde, folge ich diesen in meinem professionellen Handeln.										f 1006
f 0907	Den Erfolg meiner Unterrichtstätigkeit mache ich an den Lern- und Entwicklungsschritten meiner Schüler fest.										f 1007
f 0908	Vor Elternabend habe ich häufig ein ungutes Gefühl.										f 1008
f 0910	Ich bin aufgeschlossen gegenüber Initiativen zur Schulentwicklung.										f 1010
f 0911	Ich tausche mich regelmäßig mit meinen KollegInnen über meine Probleme im Unterricht bzw. mit meinen Schülern aus.										f 1011
f 0912	Meine Schüler sind meine besten Kritiker.										f 1012
f 0913	Die vielfältigen Ansprüche von Schülern, KollegInnen, Schulleitung sind so <i>widersprüchlich</i> , dass ich diesen nicht gerecht werden kann.										f 1013
f 0914	Bei Verwaltungsaufgaben bin ich durch formale Bestimmungen extrem reglementiert.										f 1014
f 0915	Als LehrerIn muss ich mich ständig auf verschiedene Situationen einstellen und Entscheidungen selbstständig treffen.										f 1015
f 0916	In meinem pädagogischen Handeln bin ich an die Konferenzbeschlüsse gebunden.										f 1016
f 0917	Ich bestimme allein, wie ich meinen Unterricht gestalte.										f 1017

Abbildung 25: Fragebogen, Seite 5

FRAGEBEREICH 3, BLOCK 11 UND 12		Zu meiner alltäglichen <u>Arbeitsfähigkeit</u> an der Schule kann ich Folgendes sagen:										... zum gegenwärtigen Stand				... wahrgenommene Veränderung			
		trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft über- wiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver								
f 1101	Ich bin motiviert, mich an meiner Schule verstärkt für Veränderungen/Innovationen zu engagieren.														f 1201				
f 1102	Wenn ich versuche im Unterricht etwas Neues zu machen, ziehen meine Schüler i. d. R. mit.														f 1202				
f 1103	Die alltäglichen Verwaltungsaufgaben verhindern die systematische Verbesserung der Qualität meiner Unterrichtsarbeit.														f 1203				
f 1104	Als Lehrerin fange ich alle Probleme (wie Familienprobleme etc.) bei den SchülerInnen auf.														f 1204				
f 1105	Durch die vielen formalen Bestimmungen bin ich in meiner pädagogischen Freiheit eingeschränkt.														f 1205				
f 1106	Als Lehrerin handle ich in meinem Unterricht immer eigenverantwortlich.														f 1206				
f 1107	Die Qualität meiner pädagogischen Arbeit überprüfe ich regelmäßig an den Leistungsergebnissen meiner Schüler.														f 1207				
f 1108	Wegen der Vielfalt von Verwaltungs- und Organisationsaufgaben, die ich erledigen muss, reicht die Zeit für eine intensive Unterrichtsvorbereitung nicht aus.														f 1208				
f 1109	Als Lehrerin kann ich in meiner Schule wenig bewirken.														f 1209				
f 1110	Mir fehlen konkrete Hilfen und Anweisungen bei der praktischen Umsetzung von Veränderungen.														f 1210				
f 1111	Ich beteilige mich aktiv am Entwicklungsprozess meiner Schule.														f 1211				
f 1112	Bei meiner Arbeit „versinke“ ich in Routinen.														f 1212				
f 1113	Meine Arbeit als Lehrerin ist zu sehr fremdbestimmt.														f 1213				
f 1114	In meiner Arbeit werde ich zu stark kontrolliert und überwacht.														f 1214				
f 1115	Meine Lernerfolgskontrollen bei den Schülern versehe ich regelmäßig mit förderlichen Hinweisen für die weitere Entwicklung.														f 1215				
f 1116	Wenn die Leistung eines Schülers meinen Anforderungen voll entspricht, erteile ich die Note „sehr gut“.														f 1216				
f 1117	Ich befürworte die meisten Veränderungen in meiner Schule.														f 1217				

Abbildung 26: Fragebogen, Seite 6

FRAGEBEREICH 4, BLOCK 13 UND 14		Hinsichtlich der <u>Schulleitung</u> in meiner Schule kann ich Folgendes konstatieren:						... zum gegenwärtigen Stand					...wahrgenommene Veränderung			
		trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver					
f 1301	Die Qualität meiner pädagogischen Arbeit wird von der Schulleitung regelmäßig überprüft.															f 1401
f 1302	In meiner Schule delegiert die Schulleitung die Funktionsaufgaben an Mitglieder des Kollegiums.															f 1402
f 1303	Die seitens der Schulleitung gesetzten Ziele sind für mich selbst klar und verständlich.															f 1403
f 1304	Die Schulleitung informiert sich regelmäßig über den ordnungsgemäßen Ablauf der Unterrichtsarbeit.															f 1404
f 1305	Die Schulleitung in meiner Schule tut viel für die kontinuierliche Verbesserung der Unterrichtsarbeit.															f 1405
f 1306	Die Schulleitung an meiner Schule sorgt dafür, dass ihre pädagogischen Ziele von Allen ähnlich interpretiert werden.															f 1406
f 1307	Die Schulleitung in meiner Schule fördert meine Mitwirkung und mein Engagement an Veränderungen und Innovationen.															f 1407
f 1308	Die Schulleitung fördert eine breite systematische Information des gesamten Kollegiums über alles Neue in der Schule.															f 1408
f 1309	Meine Schulleitung respektiert und fördert die pädagogische Freiheit der Lehrkräfte und bietet mir hinreichende Freiräume.															f 1409
f 1310	Die Schulleitung sorgt für den reibungslosen Ablauf in meiner Schule.															f 1410
f 1311	Die Schulleitung holt neue Projekte von außen in die Schule ein.															f 1411
f 1312	Die Schulleitung hat einen genauen Überblick dafür, was in der Schule vorgeht.															f 1412
f 1313	Meine Schulleitung ermutigt mich zu eigenen Ideen und Problemlösungen.															f 1413
f 1314	Vor Gesprächen mit einem Mitglied der Schulleitung habe ich ein komisches Gefühl.															f 1414
f 1315	Meine Schulleitung setzt sich systematisch für Neuerungen ein.															f 1415
f 1316	Die Schulleitung hat viel Einfluss auf mein professionelles Handeln in meiner Schule.															f 1416

Abbildung 27: Fragebogen, Seite 7

FRAGEBEREICH 5, BLOCK 15 UND 16		Hinsichtlich der Entwicklung des Schulprogramms in meiner Schule kann ich Folgendes konstatieren:						... zum gegenwärtigen Stand				...wahrgenommene Veränderung			
		trifft gar nicht zu	trifft eher nicht zu	trifft wenig zu	trifft eher zu	trifft überwiegend zu	trifft völlig zu	negativer	eher negativer	eher positiver	positiver				
f 1501	In meiner Schule wird seit einiger Zeit intensiv an einem von allen Lehrkräften akzeptierten Schulprogramm gearbeitet.														f 1601
f 1502	Meine Identifikation mit meiner Schule hat sich durch die Schulprogrammarbeit und die darauf bezogenen Diskussionen erhöht.														f 1602
f 1503	Informationsaustausch und Kommunikation mit meinen KollegInnen haben sich durch die Schulprogrammarbeit und die darauf bezogene Diskussion verbessert.														f 1603
f 1504	Die Schulprogrammarbeit und die darauf bezogene Diskussion tragen zu einem gemeinsamen Verständnis unserer Arbeit in der Schule bei.														f 1604
f 1505	Die Schulprogrammarbeit und die darauf bezogenen Diskussionen fördern das zielorientierte Arbeiten und die Entwicklung klarer inhaltlicher Schwerpunkte der Entwicklung unserer Schule.														f 1605
f 1506	Die Schulprogrammarbeit führt zur Entwicklung wirksamer schulinterner Arbeitsstrukturen.														f 1606
f 1507	Die Schulprogrammarbeit führt zur zusätzlichen Arbeitsbelastung der Lehrkräfte.														f 1607
f 1508	An meiner Schule diskutiere ich mit meinen KollegInnen intensiv über das Schulprogramm und die Schulentwicklung.														f 1608
f 1509	Ich habe wichtigere Probleme, als mich mit dem Schulprogramm zu beschäftigen.														f 1609
f 1510	Die Diskussionen über unser gemeinsames pädagogisches Konzept haben positiven Einfluss auf meine pädagogischen Einstellungen.														f 1610
f 1511	Die Schulprogrammarbeit führt zu mehr Transparenz bei den Entwicklungsprozessen in unserer Schule.														f 1611
f 1512	Die Entwicklung eines anspruchsvollen Schulprogramms ist für meine Schule eine wichtige Herausforderung.														f 1612
f 1513	Mir fehlt die fachliche Kompetenz, um aktiv an der Entwicklung des Schulprogramms mitzuwirken.														f 1613
f 1514	Über Schulprogrammentwicklung in meiner Schule bin ich hinreichend informiert.														f 1614

Abbildung 28: Fragebogen, Seite 8