

Berichte aus der
Abteilung Angewandte Psychologie

**Kurzversionen des MRS-Inventars von Ostendorf (1990)
zur Erfassung der fünf "grossen" Persönlichkeitsfaktoren**

Urs Schallberger und Martin Venetz

Nr. 30, 1999

Zusammenfassung

In jüngster Zeit häufen sich die Befunde, die zeigen, dass Befragungsergebnisse in einem oft unerwarteten Ausmass persönlichkeitsabhängig sind. Auf diesem Hintergrund besteht das Bedürfnis nach einem Instrument, das 1.) den Persönlichkeitsbereich möglichst umfassend abdeckt, 2.) trotzdem so kurz ist, dass es routinemässig in Befragungen eingebaut werden kann und 3.) eine gute Akzeptanz aufweist. Ziel dieses Projektes war es, ein solches Instrument zu entwickeln und in testologischer Hinsicht zu überprüfen.

Ausgangspunkt bildete das 45 bipolare Adjektivpaare umfassende Inventar minimal redundanter Skalen (MRS-Inventar) von Ostendorf (1990). Es erfasst die im Fünf-Faktoren-Modell der Persönlichkeit enthaltenen Dimensionen Extraversion, Verträglichkeit, Gewissenhaftigkeit, Emotionale Stabilität und Kultur. Aufgrund verschiedener empirischer Studien an total über 3500 Untersuchungspersonen werden Kurzversionen von 30, 25 und 20 Items vorgeschlagen. Sie weisen trotz der Kürzung hohe faktorielle Validität und für Forschungszwecke mehr als genügende Reliabilität auf. Ebenso befriedigend sind erste Befunde zur zeitlichen Stabilität (1 Jahr) und zur Beziehung zu Drittvariablen (Validität).

Inhaltsübersicht

1 . Einleitung	7
2 . Untersuchung I: Kontrolluntersuchungen zum MRS-45 (Originalversion)	11
2.1 Zielsetzungen	11
2.2 Instrument, Stichprobe und Durchführung	11
2.3 Methoden und Ergebnisse	11
2.4 Schlussfolgerungen	15
3 . Untersuchung II: Konstruktion des MRS-30 und Kontrolluntersuchungen	17
3.1 Zielsetzung und Vorblick	17
3.2 Die Auswahl der 30 Items	17
3.3 Konsequenzen der Kürzung in den Daten der Untersuchung I	18
3.4 Replikationen der Befunde zum MRS-30 in vier anderen Stichproben	20
3.4.1 Vorbemerkung	20
3.4.2 Stichproben	20
3.4.3 Methoden und Ergebnisse	21
3.5 Schlussfolgerungen	25
4 . Exkurs: Untersuchungen über andere Kurzversionen um 30 Items	27
4.1 Zielsetzung	27
4.2 Stichproben	27
4.3 Ergebnisse	27
4.4 Schlussfolgerungen	30
5. Untersuchung III: Konstruktion von MRS-25 sowie MRS-20 und Kontrolluntersuchungen	31
5.1 Zielsetzung und Vorblick	31
5.2 Die Auswahl der 25 bzw. 20 Items	31
5.3 Konsequenzen der Kürzungen in den Daten der Stichprobe B-E	31
5.4 Replikation der Befunde in anderen Stichproben	34
5.4.1 Vorbemerkung	34
5.4.2 Stichproben	34
5.4.3 Ergebnisse	35
5.5 Schlussfolgerungen	37
6. Bilanz, Ausblick auf weitere Ergebnisse und Schlussbemerkungen	39
6.1 Rückblick	39
6.2 Ausblick auf weitere Befunde (Stabilität, Validität) und Schlussbemerkungen	40
7. Literaturverzeichnis	45
Anhang: Die MRS-Versionen und ihre Anwendung	49

1. Einleitung

In der psychologischen Forschung der letzten Jahre ist zunehmend deutlich geworden, dass es notwendig ist, differentialpsychologische Gesichtspunkte auch in Untersuchungen zu berücksichtigen, die auf den ersten Blick wenig mit persönlichkeitspsychologischen Fragen zu tun haben. Der Grund dafür liegt in der Tatsache, dass Befragungsergebnisse sich als in einem unerwarteten Ausmass persönlichkeitsabhängig erweisen, oft geradezu im Gegensatz zur theoretischen Grundlage, auf der die Befragungen aufbauen. Dies gilt beispielsweise für Themen wie Gesundheit, Belastung und Beanspruchung, Burnout, Coping, Motive und Wertvorstellungen, Arbeits-, Freizeit- und Lebenszufriedenheit sowie Arbeitsplatzbeschreibungen usw. Die Befunde zu solchen Themen – insbesondere auch im Hinblick auf kausale Fragestellungen – erweisen sich auf diesem Hintergrund letztlich nur dann vernünftig interpretierbar, wenn "die Persönlichkeit" der Befragten mitberücksichtigt wird. Andererseits ist es in solchen Untersuchungen aus forschungspraktischen Gründen selten möglich, einen vollständigen Persönlichkeitsfragebogen einzubeziehen. Daher besteht ein Bedürfnis nach einem Instrument, das folgenden drei Kriterien genügt: 1.) Es sollte den Persönlichkeitsbereich möglichst umfassend abdecken, 2.) so kurz sein, dass es sozusagen "routinemässig" in Untersuchungen eingebaut werden kann, und 3.) für die Befragten – unabhängig vom konkreten Untersuchungsziel – eine gute Akzeptanz aufweisen.

Ein solches Instrument wird in diesem Papier vorgeschlagen. Es handelt sich dabei um Kurzversionen der sogenannten MRS-Ratingskalen, die Ostendorf (1990) entwickelt hat (siehe unten). Diese MRS-Skalen basieren auf dem Fünf-Faktoren-Modell (FFM), das nach einer weit verbreiteten Ansicht ein – soweit dies überhaupt mit Befragungsdaten erreichbar ist – "vollständiges Bild" der Persönlichkeit liefert (Kriterium 1). Allerdings geschieht dies auf einem relativ hohen Abstraktionsniveau, das aber oft genügt. Die Länge des Instruments (Kriterium 2) ist variabel; es wird als 30-, 25- oder 20-Item-Version vorgestellt, wobei die Zielsetzung der Untersuchung und der vorhandene Spielraum bei der Fragebogengestaltung darüber entscheiden werden, welche Version eingesetzt wird. Und schliesslich war die ganze Entwicklungsarbeit durch die Auffassung der Autoren motiviert, dass eine Selbsteinschätzung aufgrund von bipolar angeordneten Adjektivlisten wohl eine optimale Akzeptanz aufweist (Kriterium 3): Praktisch bei jedem Befragungsthema ist es mit einer entsprechenden Überleitung möglich, für den Befragten einsichtig zu begründen, dass sich die Untersucher zum besseren Verständnis der gemachten Aussagen auch noch für das Bild des Befragten von sich selbst interessieren. Die Erfahrungen im Rahmen von bis heute über zehn Untersuchungen mit sehr heterogenen Themen (z.B. aus den Bereichen der Arbeits-, Berufs-, Freizeit-, Familien- und Sportpsychologie) bestätigen diese Auffassung.

Wie erwähnt, basiert das Instrument auf dem Fünf-Faktoren-Modell (FFM, auch "Big-Five" genannt). Dieses Modell ist im letzten Jahrzehnt zu einem eigentlichen Standard in der empirischen Persönlichkeitsforschung geworden (vgl. etwa die zusammenfassenden Darstellungen der Geschichte, Theorie und Anwendungen des FFM in Digmann, 1990; McCrae & John, 1992; Ostendorf & Angleitner, 1992). Entsprechend erfreut es sich nicht nur als Basis der Persönlichkeitsdiagnostik grosser Beliebtheit, sondern auch als taxonomischer Bezugsrahmen in Literaturübersichten und Metaanalysen. Trotz aller Kritik und Skepsis (siehe z.B. Block, 1995) muss es wohl als das im Moment empirisch am besten fundierte faktorenanalytische Persönlichkeitsmodell betrachtet werden, das selbst in theoretisch ganz anders begründeten Persönlichkeitsfragebogen repliziert werden kann. Die Interpretationen der fünf Faktoren, insbesondere des fünften (vgl. De Raad & Van Heck, 1994), sind zwar noch Gegenstand von Kontroversen. Im folgenden werden die gut eingebürgerten Bezeichnungen "Extraversion", "Verträglichkeit", "Gewissenhaftigkeit", "Emotionale Stabilität" (bzw. "Neurotizismus") und "Kultur" benutzt.

Eine der Wurzeln des FFM liegt im sogenannten lexikographischen Ansatz der Persönlichkeitsforschung (vgl. John, Angleitner & Ostendorf, 1988). In dieser Tradition wurden eine ganze Reihe bipolarer Adjektivlisten vorgelegt, die zur Selbst- und Fremdbeurteilung benutzt werden können. Ostendorf (1990) hat deutsche Übersetzungen verschiedener "klassischer" Varianten solcher Listen analysiert und in eine umfassende Untersuchung zu deutschsprachigen personbeschreibenden Adjektiven eingebettet. Im Rahmen dieser Untersuchungen wurden auch erste Schritte zur Konstruktion eines ökonomischen Verfahrens zur Messung der fünf Faktoren unternommen. Dabei resultierte unter anderem ein Inventar "minimal redundanter Skalen" (MRS-Inventar) (a.a.O., S. 145ff.). Es umfasst 45 Adjektivpaare und wurde in Anlehnung an die "Minimum Redundancy Item Analysis" von Neil & Jackson (1976) konstruiert. Ziel dieser Analyse war es, eine möglichst klare Faktor- (Einfach-) Struktur zu gewinnen, was ermöglicht, durch einfaches Aufsummieren der Markiertvariablen der Faktoren relativ unabhängige Skalenwerte zu gewinnen. Dies ist vor allem in Untersuchungen, die eine unzureichende Basis für die Bestimmung von Faktorwerten liefern, vorteilhaft, also z.B. für Studien mit kleinen und/oder stark selektierten Gruppen.

Ziel der im folgenden dargestellten Untersuchungen ist es, zu prüfen, ob und wie dieses MRS-Inventar noch weiter gekürzt werden kann. 45 Adjektivpaare ergeben ja eine immer noch relativ lange Liste, die in einem Fragebogen sehr dominant wirken kann. Als Legitimation für diese Zielsetzung kann gelten, dass zum Beispiel die bekannten Ratingskalen von Norman (1963) lediglich 20 Adjektivpaare (4 Items pro Faktor) umfassen. Wie kurz Skalen sein dürfen, ist aber letztlich eine empirische Frage. Die Kriterien, die dabei im gegebenen Zusammenhang von Bedeutung sind, sind v.a. Reliabilität und faktorielle Validität.

Der Aufbau des vorliegenden Berichts gestaltet sich wie folgt: In einem ersten Schritt wird die Originalversion des MRS-Inventars ("MRS-45") daraufhin überprüft, ob sich dessen Eigenschaften, wie sie von Ostendorf (1990) berichtet werden, in einer anderen Stichprobe replizieren lassen (Abschnitt 2). Anschliessend wird eine Version mit 30 Items vorgeschlagen und empirisch überprüft (MRS-30), gefolgt von analogen Untersuchungen zu Versionen mit 25 und 20 Items (MRS-25 bzw. MRS-20) (Abschnitte 3 bis 5). Der Schlussabschnitt (Abschnitt 6) enthält einige allgemeine Überlegungen und Befunde zu den derart erarbeiteten Kurzversionen.

2. Untersuchung I: Kontrolluntersuchungen zum MRS-45 (Originalversion)

2.1 Zielsetzungen

Ziel dieser Untersuchungen ist es, das MRS-Inventar in seiner Originalversion daraufhin zu untersuchen, ob sich dessen Eigenschaften, vor allem die Fünf-Faktoren-Struktur, die faktorielle Einfachstruktur und die relative Unabhängigkeit der Skalensummenwerte, auch in einer neuen Stichprobe replizieren lassen. Weiter wird die von Ostendorf (1990) durchgeführte "Minimum Redundancy Item Analysis" wiederholt, um den Boden für die Erstellung einer Kurzversion mit 30 Items zu bereiten.

2.2 Instrument, Stichprobe und Durchführung

Als **Instrument** dienten die in Ostendorf (1990, Tab. 39, S. 149f.; siehe auch unten, Tab. 1) abgedruckten 45 MRS-Ratingskalen. Ungefähr die Hälfte der Items wurde umgepolt. Die Reihenfolge der Items wurde per Zufall bestimmt.

Die **Stichprobe** (Rohrer, Seidenberg-Jenny & Zeller-Eshel, 1993) umfasste 171 Psychologiestudierende aus einer Grundstudiumsveranstaltung, wobei die Teilnahme an der Testaufnahme und die Abgabe der Testprotokolle freiwillig war. 13 Pbn wurden in der Auswertung nicht berücksichtigt, weil sie angaben, der deutschen Sprache nicht genügend mächtig zu sein. Die resultierende Gruppe von 158 Personen umfasste 42 Männer und 116 Frauen. Das Durchschnittsalter betrug 27.8 Jahre ($s = 6.5$ Jahre). Die jüngste Person ist 21, die älteste Person 55 Jahre alt. Die Testaufnahmen erfolgten kollektiv in einem Hörsaal. Mit dem MRS-45 wurde gleichzeitig auch der Giessen-Test (Beckmann, Brähler & Richter, 1991) vorgelegt, wobei die Reihenfolge der beiden Tests nach der Sitzordnung variiert wurde. Da die Testaufnahmen als Einleitung zum Kapitel "Faktorenanalytische Persönlichkeitsforschung" diente, kann davon ausgegangen werden, dass die Pbn das FFM noch nicht kannten. Weitere Indizien dafür, dass die spezielle Stichprobe die Ergebnisse nicht wesentlich beeinflusst hat, sind die empirischen Befunde zur Faktorkongruenz (siehe unten) sowie die Befunde zu den GT-Skalen (siehe Schallberger & Venetz, 1998).

2.3 Methoden und Ergebnisse

Die Darstellung der Ergebnisse erfolgt in drei Schritten: a) Analyse der Faktorenstruktur, b) Analyse der Skalensummenwerte und c) Minimum Redundancy Item Analysis.

a) Faktorenstruktur: Zur Überprüfung der Faktorenstruktur wurden die Daten einer Hauptkomponentenanalyse mit anschließender VARIMAX-Rotation unterzogen. Dabei ergaben sich 12 Faktoren mit einem Eigenwert grösser 1. Der Eigenwerteverlauf (9.4, 4.5, 3.4, 2.9, 2.2, 1.7, 1.4, 1.3, 1.2, 1.2, 1.1, 1.0, 0.9 ...) spricht zwar nicht ganz eindeutig für eine 5-Faktorenlösung, ist aber damit zumindest verträglich.¹ Das Ergebnis der Varimax-Rotation dieser fünf Faktoren ist in den Spalten (1) bis (5) von Tab. 1 festgehalten. Die Reihenfolge der Items richtet sich dabei nach der Reihenfolge in Ostendorf (1990). Wie eine genauere Inspektion des Ladungsmuster zeigt, lassen sich die Faktoren – trotz einer ganzen Reihe von Anomalien – problemlos im Sinne der Faktoren des FFM interpretieren, deren Bezeichnungen daher im Tabellenkopf festgehalten sind.

Zur Überprüfung des Eindrucks, dass es sich in Tab. 1 tatsächlich um die fünf Faktoren des FFM handelt, wurden Faktorkongruenzkoeffizienten nach Tucker (1951, siehe Pawlik, 1968) zwischen dem Ladungsmuster in Tab. 1 und den entsprechenden Angaben in Ostendorf (1990, Tab. 39, je Spalte 4) berechnet. Das Ergebnis ist in der letzten Zeile von Tab. 1 festgehalten: Für die gesamte Lösung ergab sich ein Kongruenzkoeffizient von .91, für die einzelnen Faktoren Werte zwischen .92 und .95, mit Ausnahme von "Verträglichkeit", wo lediglich .82 resultierte. Bekanntlich ist die Evaluation von Kongruenzkoeffizienten ein nicht ganz einfaches Unterfangen (siehe z.B. Bortz, 1993, S. 513f.). Gemäss ten Berge (1986) ist aber davon auszugehen, dass Faktorenanalyseexperten Faktoren aus Ladungsmatrizen, die zu Koeffizienten von .85 und höher führen, klar als identisch beurteilen. Diese Bedingung ist im vorliegenden Fall für die Gesamtmatrix, aber auch für vier der fünf Faktoren eindeutig erfüllt. Das knappe Verfehlen dieses Kriteriums bei Verträglichkeit (.82) darf möglicherweise nicht überbewertet werden: Broadbooks & Elmore (1987) haben durch Monte Carlo Studien gezeigt, dass bei einem wahren Wert von .90 und höher – eine Annahme, die gemäss den obigen Befunden zumindest nicht abwegig ist – der Stichprobenkongruenzkoeffizient eine Tendenz zur Unterschätzung aufweist. Es kann somit davon ausgegangen werden, dass in den vorliegenden Daten die Faktorenstruktur des MRS-Inventars im Sinne des FFM im wesentlichen repliziert werden konnte.

b) Skalensummenwerte: Ein Anspruch des MRS-Inventars ist, dass durch eine möglichst reine faktorielle Struktur bereits durch blosses Aufsummieren der je neun Ratingskalen brauchbare Schätzungen der "Faktorwerte" gewonnen werden können. Als Ergebnis der entsprechenden Entwicklungsarbeit wird festgehalten, dass die durchschnittliche Skaleninterkorrelation

¹ Der Scree-Test liesse auch eine Sechs-Faktorenlösung zu. Diese Lösung besteht in einer Aufteilung des Faktors Kultur in die zwei Komponenten "Intelligenz" (höchst ladende Items: intelligent, modern, gebildet) und "Kreativität" (künstlerisch, kreativ, phantasievoll). Diese Aufteilung findet sich auch in anderen Anwendungen des MRS-Inventars (z.B. Becker, 1996), ist aber über dieses Instrument hinaus von Bedeutung (vgl. diverse Beiträge in De Raad & Van Heck, 1994).

Tab. 1: Faktorladungen des MRS-45, Faktorkongruenz- und Item-Effizienz-Koeffizienten

	(1)	(2)	(3)	(4)	(5)	(6)	
	Extra- version	Verträ- glichkeit	Gewissen- haftigkeit	Emot. Stabilität	Kultur	Item- Effizienz- Koeff.	
Extraversion							
1	gesprächig - schweigsam	66	-14	08	30	18	-2.1
2	direkt - taktierend	10	-43	06	25	23	-3.8
3	kontaktfreudig - zurückhaltend	75	-18	-01	25	14	-1.5
4	freimütig - gehemmt	36	-06	04	52	39	-3.7
5	gesellig - zurückgezogen	70	-09	10	36	05	-1.7
6	anschlussbedürftig – einzelgängerisch	65	-02	07	12	-15	-0.7
7	aktiv - passiv	37	-23	34	30	36	-4.6
8	offen - zugeknöpft	64	-16	11	41	18	-2.8
9	impulsiv - selbstbeherrscht	43	-39	-10	09	18	-2.6
Verträglichkeit							
10	zustimmend - gegensätzlich	20	43	09	-14	-23	-2.2
11	vertrauensvoll - misstrauisch	62	20	-13	19	15	-4.1
12	gutmütig - reizbar	-18	64	-01	28	-20	-2.1
13	nachsichtig - barsch	-11	52	-01	09	-06	-0.6
14	selbstlos - selbstsüchtig	-23	49	-10	06	05	-1.3
15	friedfertig - streitsüchtig	-03	75	10	-04	01	0.0
16	weichherzig - rücksichtslos	22	49	09	-12	13	-1.7
17	leichtgläubig - zynisch	43	28	05	-29	-11	-3.3
18	höflich - grob	08	45	17	10	37	-2.4
Gewissenhaftigkeit							
19	gründlich - unsorgfältig	-07	09	80	-01	19	-0.6
20	ordentlich - unachtsam	04	07	81	18	03	-0.5
21	geordnet - ungeordnet	05	11	72	14	-06	-0.8
22	strebsam - ziellos	16	-08	54	14	02	-1.0
23	geschäftstüchtig - verspielt	-03	-17	35	27	-14	-2.1
24	übergenu - ungenau	-20	09	74	-04	03	-0.7
25	gewissenhaft - nachlässig	06	07	80	-10	09	-0.5
26	fleißig - faul	11	-18	66	-07	08	-1.1
27	sparsam - verschwenderisch	00	15	49	05	-25	-1.3
Emotionale Stabilität							
28	robust - verletzlich	08	05	-03	71	-03	-0.0
29	selbstzufrieden - selbstmitleidig	31	10	-02	72	05	-1.2
30	entspannt - überempfindlich	19	18	02	76	15	-1.4
31	mutig - furchtsam	23	-13	-06	54	36	-2.6
32	ruhig - ängstlich	-03	01	11	69	25	-0.9
33	überzeugt - selbstachtunglos	20	-15	23	59	35	-3.2
34	selbstvertrauend - hilflos	27	-16	12	68	33	-2.9
35	gefühlsstabil - labil	16	08	16	78	-08	-1.1
36	sicher - unsicher	24	-12	03	76	26	-1.9
Kultur							
37	künstlerisch - unkünstlerisch	07	-09	04	-10	64	-0.5
38	kreativ - unkreativ	00	-05	-01	10	75	0.1
39	komplex - einfach	-24	-16	15	-24	36	-2.8
40	originell - konventionell	04	-05	-14	23	66	-1.2
41	liberal - konservativ	11	-07	-23	15	39	-1.9
42	phantasievoll - phantasielos	17	02	01	09	72	-0.4
43	intelligent - unintelligent	-04	00	13	18	66	-0.7
44	gebildet - ungebildet	-06	11	12	26	62	-1.6
45	modern - traditionell	24	-05	-06	18	42	-1.7
Faktorkongruenzkoeffizienten (total: 91)		92	82	95	94	93	

auf .20 gesenkt werden konnte, bei einem mittleren Cronbach-Alpha von .85 (Ostendorf, 1990, S. 146).

Die entsprechenden Befunde in den vorliegenden Daten sind in Tab. 2 festgehalten. Die Reliabilitäten bewegen sich zwischen .63 und .90, wobei wiederum "Verträglichkeit" durch den klar tiefsten Wert (.63) auffällt. Dieser Einzelwert ist auch dafür verantwortlich, dass der (via die Fishersche z-Transformation) gewonnene Durchschnittswert mit .82 etwas tiefer ausfällt als bei Ostendorf.

Betrachtet man die Interkorrelationen in Tab. 2 fällt v.a. die hohe Korrelation zwischen Extraversion und Emotionaler Stabilität auf (.60). Die mittlere Interkorrelation (bestimmt via die Fishersche z-Transformation) beträgt aber trotzdem nur .21, entspricht also der bei Ostendorf (1990) angegebenen Grösse.

Tab. 2: Reliabilitäten (Cronbach-Alpha) und Interkorrelationen der Skalensummenwerte

Faktor:	r_{tt}	(2)	(3)	(4)	(5)
(1) Extraversion	.85	-.06	.15	.60	.35
(2) Verträglichkeit	.63		.11	.10	-.06
(3) Gewissenhaftigkeit	.84			.18	.01
(4) Emot. Stabilität	.90				.34
(5) Kultur	.78				

c) Minimum Redundancy Item Analysis: Wie bereits weiter oben erwähnt, basiert das MRS-45 auf einer Variante einer "Minimum Redundancy Item Analysis" (MRA). Leider sind die Detailbefunde in Ostendorf (1990) nicht publiziert, sodass eine eigentliche Replikation nicht möglich ist. Im Hinblick auf die hier leitende Absicht, das Instrument weiter zu kürzen, wurde trotzdem eine MRA durchgeführt. Dabei wurde der von Ostendorf (1990) vorgeschlagene Item Effizienz-Index (IEI) benutzt, der als

$$IEI_i = |a_{ig}| - 4 * \sum |a_{ih}|$$

definiert ist, wobei a_{ig} die Ladung des Items i im theoretisch erwarteten Faktor g kennzeichnet und a_{ih} die Ladungen in den anderen Faktoren. Der IEI ist also umso kleiner, je grösser die Nebenladungen eines Items im Verhältnis zu seiner (theoretischen) Hauptladung sind. Die faktoriell relativ "reinen" Items zeichnen sich somit durch relativ hohe IEI aus. Die Ergebnisse sind in der Spalte (6) von Tab. 1 festgehalten. Die IEI streuen zwischen -4.6 (Item 7) und 0.1 (Item

38), bei einem Mittelwert von -1.7 ($s = 1.2$). Es sind somit noch gewaltige Qualitätsunterschiede zwischen den Items festzustellen, wie sie sich ja bereits in den Faktorladungsmustern ausdrücken.

2.4 Schlussfolgerungen

Ziel der Untersuchungen zum MRS-45 war, die Hauptbefunde von Ostendorf (1990) zu diesem Inventar zu replizieren, nämlich die Fünf-Faktoren-Struktur, die Reliabilität und die relative Unabhängigkeit der Skalensummenwerte sowie – via eine Minimum Redundancy Item Analysis – die faktorielle Einfachstruktur. Global gesehen ergaben sich dabei recht befriedigende Ergebnisse. Im einzelnen zeigten sich aber einige Unschönheiten, vor allem Items mit abweichenden Faktorladungen und entsprechend tiefen Item-Effizienz-Koeffizienten, einzelne sehr hohe Skaleninterkorrelationen sowie problematische Kennwerte der Skala "Verträglichkeit". Auf diesem Hintergrund ergeben sich somit eine Reihe von Ansatzpunkten für eine Kürzung des MRS-45.

3. Untersuchung II: Konstruktion des MRS-30 und Kontrolluntersuchungen

3.1 Zielsetzung und Vorblick

Hauptziel dieses Kapitels ist es, eine auf sechs Items pro Faktor gekürzte Version des MRS-Inventars (MRS-30) vorzuschlagen. Die Kürzung wird aufgrund der in Kapitel 2 beschriebenen Untersuchung vorgenommen (Abschnitte 3.2 und 3.3) und in neuen Stichproben überprüft (Abschnitt 3.4).

3.2 Die Auswahl der 30 Items

Bei der Auswahl der Items werden zwei Informationsquellen berücksichtigt: Das erste Kriterium bilden die in der sechsten Spalte von Tab. 1 enthaltenen Item-Effizienz-Indizes (IEI). Um die faktorielle Einfachstruktur zu optimieren, sollen Items mit tiefen IEIs eliminiert werden. Das zweite Kriterium bilden die (erfragten) Rückmeldungen der Pbn über aus ihrer Sicht irritierende Items. Damit sollte die Akzeptanz des Instruments erhöht werden. Das Ergebnis dieser Überlegungen sieht wie folgt aus:

Extraversion: Hier fallen in Tab. 1 vor allem drei Items mit extrem tiefen IEI auf (Item 2, 4 und 7, $IEI \leq -3.7$, alle anderen Items ≥ -2.8). Da die übrigen Items nicht speziell kritisiert worden sind, werden diese drei Items gestrichen.

Verträglichkeit: In dieser Gruppe sind die Items 11 und 17 negativ auffällig ($IEI \leq -3.3$), die zudem klar auf dem "falschen" Faktor (Extraversion) laden. Weitere drei Items haben einen $IEI \leq -2.1$ (Item 18, 10 und 12). Davon wurde vor allem Item 10 ("zustimmend – gegensätzlich") als ungewöhnlicher Gegensatz empfunden und daher gestrichen.

Gewissenhaftigkeit: Die Items dieser Gruppe weisen fast durchgängig relativ hohe IEI auf (≥ -2.1). Ausgeschlossen werden die beiden Items mit den tiefsten Werten (Item 23 und 27). Von den Items mit den nächst höheren Werten (Item 22 und 26) wird der gängigere Gegensatz beibehalten, somit Item 22 ausgeschlossen.

Emotionale Stabilität: Von den IEI her gesehen ist die Situation eigentlich recht klar: Die Items 31, 33 und 34 fallen mit Werten unter -2.6 negativ auf. Andererseits wurde das Item 32 ("ruhig - ängstlich") von mehreren Pbn als "schiefer" Gegensatz kritisiert. So wird dieses Item ausgeschlossen und von den negativ auffälligen Items dasjenige mit der höchsten Ladung im "Zielfaktor" (Item 34) beibehalten.

Kultur: Hier werden wiederum die Items mit den tiefsten IEI (die aber fast alle relativ hoch ausfielen) ausgeschlossen (Items 39, 41 und 45).

3.3 Konsequenzen der Kürzung in den Daten der Untersuchung I

Da die Selektion der 30 Items (MRS-30) aufgrund der Daten der Untersuchung I vorgenommen wurden, sollte sich die Kürzung positiv auf die relevanten Kennwerte in diesen Daten auswirken. In welchem Ausmass dies tatsächlich der Fall ist, wurde mit einer neuen Faktorenanalyse der ausgewählten Items (Hauptkomponentenanalyse mit Varimax-Rotation) überprüft. Der Eigenwerteverlauf (6.5, 3.8, 2.9, 2.7, 1.6, 1.3, 1.0, 0.9 ...) spricht nun interessanterweise deutlich für eine Vier-Faktorenlösung. Die Inspektion der entsprechenden Faktorladungen zeigt, dass dafür ein relativ klares Zusammenfallen der Faktoren Emotionale Stabilität (bzw. Neurotizismus) und Extraversion verantwortlich ist (alle Ladungen über .54). Lediglich das Extraversions-Item 9 ("impulsiv - selbstbeherrscht") lädt nur mit .31 in diesem Superfaktor, hingegen mit -.48 auf dem Faktor Verträglichkeit. Die übrigen Faktoren: Gewissenhaftigkeit, Offenheit für Erfahrung und auch Verträglichkeit (mit der oben genannten Ausnahme) werden hingegen auch in dieser Vier-Faktorenlösung klar repliziert. Dieses Zusammenfallen von Extraversion und Neurotizismus hat seinen Grund wohl in jenem Sachverhalt, der auch in der relativ hohen Korrelation der entsprechenden Skalensummenwerte des MRS-45 (.60, vgl. Tab. 2) zum Ausdruck kommt.

Die Ergebnisse einer forcierten Fünf-Faktorenlösung sind aus Tab. 3 ersichtlich. Das Bild, das die Faktorkongruenzkoeffizienten liefern, hat sich erwartungsgemäss leicht verbessert. (Sie basieren wiederum auf einem Vergleich mit den Spalten 4 aus der Tab. 39 aus Ostendorf, 1990.) Die Gesamt-Kongruenz beträgt jetzt .94 (gegenüber .91 bei der Originalversion), der tiefste Einzelkoeffizient .86 (Verträglichkeit; gegenüber .82 in der Originalversion). Auch die Item-Effizienz-Indizes (Spalte 6) haben sich in 19 (von 30) Fällen erhöht und sind in den übrigen Fällen zumindest von vergleichbarer Grössenordnung.

Eine analoge Verbesserung ergibt sich bei den Reliabilitäten (vgl. Tab. 4; alle Koeffizienten über .67, versus .63) und den Interkorrelationen der Skalensummenwerte (höchste Interkorrelation .51 versus .60; durchschnittliche Interkorrelation .17 vs. .21).

Tab. 3: Faktorladungen des MRS-30, Faktorkongruenz- und Item-Effizienz-Koeffizienten (letztere im Vergleich zu jenen des MRS-45)

	(1) Extra- version	(2) Verträ- glichkeit	(3) Gewissen- haftigkeit	(4) Emot. Stabilität	(5) Kultur	(6) IEI (30/45)	
Extraversion							
1	gesprächig - schweigsam	73	-08	06	25	16	-1.5/ -2.1
3	kontaktfreudig - zurückhaltend	83	-10	-04	19	13	-1.0/ -1.5
5	gesellig - zurückgezogen	80	-01	05	31	05	-0.9/ -1.7
6	anschlussbedürftig - einzelgängerisch	69	03	05	12	-19	-0.9/ -0.7
8	offen - zugeknöpft	71	-09	10	35	14	-2.0/ -2.8
9	impulsiv - selbstbeherrscht	45	-33	-12	04	22	-2.4/ -2.6
Verträglichkeit							
12	gutmütig - reizbar	-22	63	02	32	-19	-2.4/ -2.1
13	nachsichtig - barsch	-06	56	00	07	-12	-0.4/ -0.6
14	selbstlos - selbstsüchtig	-20	57	07	02	03	-0.7/ -1.3
15	friedfertig - streitsüchtig	-11	72	10	-01	-01	-0.2/ 0.0
16	weichherzig - rücksichtslos	25	55	06	-21	13	-2.0/ -1.7
18	höflich - grob	14	53	16	08	32	-2.3/ -2.4
Gewissenhaftigkeit							
19	gründlich - unsorgfältig	-03	08	83	00	15	-0.2/ -0.6
20	ordentlich - unachtsam	06	04	84	19	00	-0.4/ -0.5
21	geordnet - ungeordnet	09	12	72	15	-09	-1.1/ -0.8
24	übergenu - ungenau	-15	07	78	-04	-01	-0.3/ -0.7
25	gewissenhaft - nachlässig	06	06	82	-08	05	-0.2/ -0.5
26	fleissig - faul	09	-18	63	-05	09	-1.0/ -1.1
Emotionale Stabilität							
28	robust - verletzlich	14	08	-04	69	01	-0.4/ 0.0
29	selbstzufrieden - selbstmitleidig	25	05	-02	76	07	-0.8/ -1.2
30	entspannt - überempfindlich	18	16	00	78	19	-1.4/ -1.4
34	selbstvertrauend - hilflos	30	-18	14	68	35	-3.2/ -2.9
35	gefühlsstabil - labil	16	07	12	80	00	-0.6/ -1.1
36	sicher - unsicher	29	-11	02	72	30	-2.1/ -1.9
Offenheit für Erfahrung							
37	künstlerisch - unkünstlerisch	07	-12	04	-15	73	-0.8/ -0.5
38	kreativ - unkreativ	04	-02	01	06	80	0.3/ 0.1
40	originell - konventionell	03	-07	-14	22	69	-1.1/ -1.2
42	phantasievoll - phantasielos	18	05	-01	07	76	-0.5/ -0.4
43	intelligent - unintelligent	-01	01	16	17	63	-0.8/ -0.7
44	gebildet - ungebildet	00	16	13	24	61	-1.5/ -1.6
Faktorkongruenzkoeffizienten:							
	MRS-30 (total: 94)	96	86	96	95	93	
	MRS-45 (total: 91)	92	82	95	94	93	
Skalensummenwerte (MRS-30/MRS45):							
	Cronbach-Alpha	85/85	67/63	87/84	88/90	81/78	
	Skaleninterkorrelationen \bar{x}	17/21					

Erläuterungen: IEI (MRS-30/MRS-45).

Tab. 4: Reliabilitäten (Cronbach-Alpha) und Interkorrelationen der Skalensummenwerte (MRS-30)

Faktor:	r_{tt}	(2)	(3)	(4)	(5)
(1) Extraversion	.85	-.18	.06	.51	.23
(2) Verträglichkeit	.67		.10	.11	-.01
(3) Gewissenhaftigkeit	.87			.10	.06
(4) Emot. Stabilität	.88				.30
(5) Offenheit	.81				

3.4 Replikationen der Befunde zum MRS-30 in vier anderen Stichproben

3.4.1 Vorbemerkung

Die in Abschnitt 3.3 berichteten Verbesserungen der Skaleneigenschaften durch die Kürzung sind – wie bereits erwähnt – nicht sehr aussagekräftig, da die Kürzung auf denselben Daten basierte und auf eben diese Verbesserungen ausgerichtet war. Daher wird im folgenden die in 3.2 beschriebene Analyse des MRS-30 in anderen Stichproben repliziert.

3.4.2 Stichproben

Bei den Stichproben handelt es sich um folgende Gruppen, wobei der Vollständigkeit halber auch die Stichprobe aus Untersuchung I (als Stichprobe A) nochmals aufgeführt wird:

Stichprobe A (Rohrer et al., 1993) umfasst 158 Psychologiestudierende (siehe Abschnitt 2.2).

Stichprobe B (Nussbaum, 1998) bilden 530 Turn- und Sportlehrer in der Deutschschweiz. Davon sind 328 männlichen und 202 weiblichen Geschlechts. Das Altersmittel liegt bei 39.2 Jahren ($s = 8.7$ Jahre). Der jüngste Proband ist 24, der älteste 68 Jahre alt. Die Rücklaufquote der postalischen Befragung betrug 53%.

Stichprobe C (Basler-Melbourne & Jutz, 1996) setzt sich aus total 543 Freizeitsportlern (291) und -sportlerinnen (252) zusammen, die in einem Verein entweder Karate, Badminton oder Volleyball ausüben. Das Durchschnittsalter beträgt 28.4 Jahre ($s = 8.2$ Jahre), die Spannweite reicht von 16 bis 62 Jahren. Es handelt sich um 53% der angeschriebenen Personen.

Stichprobe D (Büchi, Hegemann & Krapf, 1996) umfasst 135 Turn- und Sportstudenten der ETH Zürich im 6. und 8. Semester (81 Männer, 54 Frauen). Das Durchschnittsalter beträgt 24.6 Jahre ($s = 1.8$ Jahre). Der jüngste Proband ist 22, der älteste 31 Jahre alt. Die Rücklaufquote der postalischen Umfrage betrug 72%.

Stichprobe E (Chalkiopoulos, Nierop & Spörri, 1997) umfasst 100 Personen (43 Männer, 57 Frauen) aus dem Bekanntenkreis der Autorinnen. Das durchschnittliche Alter beträgt 40.8 Jahre ($s = 15.4$ Jahre). Die Spannbreite reicht von 15 bis 83 Jahren. 62.5% der angeschriebenen Personen hatten an der Untersuchung teilgenommen.

Stichprobe B-E besteht aus der Zusammenfassung der Stichproben B, C, D und E. Sie umfasst total 1308 Personen. Von diesen 1308 Personen sind 743 männlichen und 565 weiblichen Geschlechts. Das Altersmittel liegt bei 33.3 Jahren ($s = 10.6$ Jahre).

Die Stichproben B, C, D, E und B-E werden im folgenden als Replikationsstichproben bezeichnet.

3.4.3 Methoden und Ergebnisse

Die Daten aus den verschiedenen Stichproben wurden je für sich faktoranalytisch und denselben Analysen unterzogen, wie sie in Abschnitt 3.3 vorgenommen wurden. Im folgenden werden jedoch nur Eigenwertverläufe, Faktorkongruenzkoeffizienten, Item-Effizienz-Indizes und die Kennwerte der Skalensummenwerte ausführlicher dargestellt und diskutiert.

a) Eigenwertverläufe: Die Eigenwertverläufe sind in Tab. 5 dargestellt. In der ersten Zeile (Stichprobe A) ist zu Vergleichszwecken nochmals der Befund aus Abschnitt 3.3 aufgeführt. Der Scree-Test spricht nur in der Stichprobe E relativ deutlich für eine 5-Faktorenlösung. In den Stichproben B, C, D und B-E hingegen ergeben sich eher Indizien für eine 6-Faktorenlösung. Eine Kontrolle dieser Lösungen ergab, dass immer die bereits in Fussnote 1 erwähnte Aufspaltung von Kultur in einen Kreativitätsaspekt (Item 37, 38, 40 und 42) und einen Intelligenz-/Bildungsaspekt (Item 43, 44) dafür verantwortlich ist, während die anderen Faktoren intakt bleiben.

Für die weiterführenden Analysen wurde in allen Fällen eine 5-Faktorenlösung forciert. Die Ladungsmatrizen werden im folgenden nicht mehr aufgeführt. Ein auffälliges Ergebnis war dabei, dass das Item 9 "impulsiv - selbstbeherrscht" in den Stichproben B, C, D, E und B-E seine höchste Ladung nicht - wie theoretisch erwartet - auf dem Faktor Extraversion, sondern auf dem Faktor Verträglichkeit aufweist.

Tab. 5: Die Eigenwertverläufe aus den fünf Stichproben

Stichprobe	Beginn des Eigenwertverlaufs										
	1	2	3	4	5	6	7	8	9	10	11
A:	6.5	3.8	2.9	2.7	1.6	1.3	1.0	0.9
B:	5.2	3.9	3.1	2.5	2.0	1.4	0.9	0.9	0.8
C:	5.1	3.9	2.6	2.5	2.1	1.5	1.0	0.9	0.8
D:	5.0	4.0	3.3	2.7	1.8	1.4	1.1	1.0	0.9
E:	5.0	4.1	2.9	2.2	2.1	1.5	1.4	1.1	1.1	0.9	...
B-E:	5.0	3.9	2.8	2.5	2.0	1.4	0.9	0.9	0.8

b) Faktorkongruenzkoeffizienten: Tab. 6 enthält die Faktorkongruenzkoeffizienten aus den verschiedenen Replikationsstichproben, wobei wiederum als Vergleichswerte auch die Ergebnisse aus Stichprobe A aufgeführt sind.

Die Kongruenzkoeffizienten für die verschiedenen Gesamtlösungen liegen zwischen .91 (Stichprobe E) und .96 (Stichprobe C und B-E). Diese hohen Werte deuten darauf hin, dass das MRS-45 von Ostendorf (1990) und das MRS-30 eine identische faktorielle Gesamt-Struktur aufweisen.

Tab. 6: Die Faktorkongruenzkoeffizienten

Stichprobe	Faktorkongruenzkoeffizienten					
	Total	Extra- version	Verträ- glichkeit	Gewissen- haftigkeit	Emotionale Stabilität	Kultur
A:	.94	.96	.86	.96	.95	.93
B:	.95	.96	.96	.98	.98	.85
C:	.96	.97	.96	.97	.98	.94
D:	.92	.92	.92	.93	.95	.88
E:	.91	.94	.86	.92	.91	.91
B-E:	.96	.97	.96	.98	.98	.92

Betrachtet man die Einzelkoeffizienten, so ergeben sich für die Dimensionen Extraversion, Gewissenhaftigkeit und Emotionale Stabilität über alle Replikationsstichproben hinweg gute Faktorkongruenzkoeffizienten ($\geq .91$). Auch für den Faktor Verträglichkeit, für den in der Untersuchung I (Stichprobe A) ein Koeffizient von (nur) .86 ermittelt wurde, resultieren in den Replikationsstichproben – ausser in Stichprobe E - hohe Faktorkongruenzkoeffizienten ($\geq .92$).

Eine kleine Unschönheit zeigt sich beim Faktor Kultur: Hier liegt der Kongruenzkoeffizient mit .85 und .88 gleich bei zwei Stichproben (B und D) unterhalb von .9. Eine mögliche Erklärung für diesen Befund liefert die Inspektion der (hier nicht wiedergegebenen) Faktorladungsmatrizen: In diesen Replikationsstichproben weisen die beiden Kultur-Items "intelligent – unintelligent" (Item 43) und "gebildet – ungebildet" (Item 44) ihre höchsten Ladungen auf dem Faktor Emotionale Stabilität auf.

Tab. 7: Die Item-Effizienz-Indizes aus den verschiedenen Stichproben im Vergleich

	MRS 45 A	MRS 30 A	MRS 30 B	MRS 30 C	MRS 30 D	MRS 30 E	MRS 30 B-E	
<i>Extraversion</i>								
1	gesprächig - schweigsam	-2.1	-1.5	-1.0	-1.2	-1.9	-0.4	-1.0
3	kontaktfreudig - zurückhaltend	-1.5	-1.0	-0.7	-0.3	-1.6	-1.6	-0.6
5	gesellig - zurückgezogen	-1.7	-0.9	-0.6	-0.8	-1.3	-0.6	-0.7
6	anschlussbedürftig - einzelgängerisch	-0.7	-0.9	-1.2	-0.6	-1.4	-1.9	-1.0
8	offen - zugeknöpft	-2.8	-2.0	-1.7	-1.8	-2.3	-1.1	-1.7
9	impulsiv - selbstbeherrscht	-2.6	-2.4	-3.8	-3.4	-2.7	-4.1	-3.5
<i>Verträglichkeit</i>								
12	gutmütig - reizbar	-2.1	-2.4	-0.6	-0.6	-0.2	-1.5	-0.4
13	nachsichtig - barsch	-0.6	-0.4	0.3	0.0	-0.8	-1.3	0.4
14	selbstlos - selbstsüchtig	-1.3	-0.7	-0.2	-0.4	-0.6	-2.1	-0.1
15	friedfertig - streitsüchtig	0.0	-0.2	-0.5	-0.4	-0.8	-2.4	-0.4
16	weichherzig - rücksichtslos	-1.7	-2.0	-1.4	-1.5	-1.5	-0.3	-1.4
18	höflich - grob	-2.4	-2.3	-1.4	-2.3	-0.9	-1.4	-1.5
<i>Gewissenhaftigkeit</i>								
19	gründlich - unsorgfältig	-0.6	-0.2	0.0	-0.2	-0.1	-2.7	-0.1
20	ordentlich - unachtsam	-0.5	-0.4	-0.3	-0.5	-0.1	-1.6	-0.4
21	geordnet - ungeordnet	-0.8	-1.1	-0.3	0.2	0.0	-2.5	-0.2
24	übergenu - ungenau	-0.7	-0.3	-0.6	-0.6	0.1	-0.9	-0.3
25	gewissenhaft - nachlässig	-0.5	-0.2	0.3	0.0	-0.1	-0.1	0.3
26	fleißig - faul	-1.1	-1.0	-1.1	-0.5	-1.2	-2.3	-0.8
<i>Emotionale Stabilität</i>								
28	robust - verletzlich	0.0	-0.4	-0.9	-0.1	-0.6	-0.1	-0.2
29	selbstzufrieden - selbstmitleidig	-1.2	-0.8	-0.8	-0.3	-1.8	-0.6	-0.5
30	entspannt - überempfindlich	-1.4	-1.4	-1.1	-0.5	-1.8	0.0	-0.7
34	selbstvertrauend - hilflos	-2.9	-3.2	-1.8	-2.3	-1.1	-4.3	-2.0
35	gefühlsstabil - labil	-1.1	-0.6	-0.4	-0.7	0.0	-0.9	-0.3
36	sicher - unsicher	-1.9	-2.1	-2.0	-2.0	-1.5	-2.9	-2.0
<i>Kultur</i>								
37	künstlerisch - unkünstlerisch	-0.5	-0.8	0.3	0.1	-1.4	-0.8	0.2
38	kreativ - unkreativ	0.1	0.3	-0.1	0.3	-0.8	-0.8	0.2
40	originell - konventionell	-1.2	-1.1	-1.7	-0.8	-1.0	-1.0	-1.0
42	phantasievoll - phantasielos	-0.4	-0.5	-0.6	-0.3	-1.0	-0.4	-0.3
43	intelligent - unintelligent	-0.7	-0.8	-3.3	-2.0	-2.1	-2.6	-2.3
44	gebildet - ungebildet	-1.6	-1.5	-3.1	-1.9	-3.2	-2.2	-2.0

c) Item-Effizienz-Indizes: Tab. 7 enthält die Item-Effizienz-Indizes aus den verschiedenen Replikationsstichproben, zusammen mit den Werten aus der Stichprobe A (MRS-45 und MRS 30). Der Tabelle kann zunächst entnommen werden, dass auch in den Replikationsstichproben - wie bereits in der Stichprobe A - beträchtliche Qualitätsunterschiede zwischen den einzelnen Items bestehen: die Item-Effizienz-Indizes streuen zwischen -4.1 (Item 9, Stichprobe E) und 0.4 (Item 13, Stichprobe B-E). Eine eingehendere Betrachtung zeigt, dass über die verschiedenen Stichproben hinweg weitgehend dieselben Items als problematisch erscheinen. Wie in Stichprobe A liegen auch in den Replikationsstichproben die IEI der Items 8 ("offen – zugeknöpft") und 9 ("impulsiv – selbstbeherrscht") der Dimension Extraversion, die Items 34 ("selbstvertrauend – hilflos") und 36 ("sicher – unsicher") des Faktors Emotionale Stabilität sowie Item 43 ("intelligent – unintelligent") und 44 ("gebildet – ungebildet") des Faktors Kultur meist klar tiefer als diejenigen der übrigen Items. Beim Faktor Verträglichkeit sind die Übereinstimmungen nicht ganz so eindeutig: Item 12 ("gutmütig – reizbar") weist in der Stichprobe A (MRS-30) den tiefsten IEI-Wert (-2.4) auf, in den Replikationsstichproben hingegen einen durchschnittlichen IEI von -0.7. Die Items 16 ("weichherzig – rücksichtslos") und 18 ("höflich – grob") sind in Stichprobe E unauffällig, zeigen in allen anderen Stichproben jedoch die tiefsten Werte der Verträglichkeits-Items. Schliesslich fallen die Gewissenhaftigkeits-Items durch generell recht hohe IEI auf.

d) Kennwerte der Skalensummenwerte: Aus Tab. 8 wird ersichtlich, dass sich die Reliabilitäten in den Replikationsstichproben zwischen .71 und .88 bewegen. In der Zusammenfassung der Replikationsstichproben (B-E) zeigt sich im Vergleich zur Stichprobe A in der Dimension Verträglichkeit ein höherer Reliabilitätskoeffizient (.75 vs. .67), in der Dimension

Tab. 8: Reliabilitäten (Cronbach-Alpha) und Skaleninterkorrelation aus verschiedenen Stichproben

	MRS-40 A	MRS-30 A	MRS-30 B	MRS-30 C	MRS-30 D	MRS-30 E	MRS-30 B-E
Extraversion	.85	.85	.81	.80	.78	.80	.80
Verträglichkeit	.63	.67	.75	.76	.79	.68	.75
Gewissenhaftigkeit	.84	.87	.87	.87	.88	.81	.87
Emotionale Stabilität	.90	.88	.79	.79	.81	.78	.79
Kultur	.78	.81	.71	.76	.76	.78	.75
Skaleninterkorrelationen							
Maximum	.60	.51	.35	.28	.28	.36	.31
\bar{x}	.21	.17	.14	.12	.17	.17	.13

Gewissenhaftigkeit ein vergleichbarer (.87 vs. .87), in den übrigen hingegen tiefere Reliabilitäten (Extraversion: .80 vs. .85; Emotionale Stabilität: .79 vs. .88; Kultur: .75 vs. .81).

Die höchsten Skaleninterkorrelationen liegen in den Replikationsstichproben mit einem (via Fishersche z-Transformation gewonnenen) Durchschnittswert von .32 deutlich tiefer als in Stichprobe A (.51), und die gemittelten Skaleninterkorrelationen erzielen niedrigere oder höchstens gleich hohe Koeffizienten (.17 vs. .12 bis .17).

3.5 Schlussfolgerungen

Ziel dieses Kapitels war es, die aus Untersuchung I gewonnene Version des MRS-30 in weiteren Stichproben zu überprüfen. Dazu wurden die faktorielle (Einfach-)Struktur sowie die Kennwerte der Skalensummenwerte analysiert. Obwohl die Eigenwertverläufe in den Kontrolluntersuchungen eher für eine 6-Faktorenlösung sprechen, liefert eine forcierte 5-Faktorenlösung in allen Stichproben Faktoren, die mit jenen des MRS-45 praktisch übereinstimmen. Weiter korrelieren die Skalensummenwerte nur geringfügig miteinander und auch die Reliabilitätskoeffizienten erzielen befriedigende Werte. Gesamthaft gesehen kann also davon ausgegangen werden, dass sich das MRS-30 gut zur Messung der 'Big Five' eignet. Gleichzeitig erbrachte aber die Minimum Redundancy Item Analysis auch deutliche Hinweise darauf, dass zwischen den einzelnen Items zum Teil immer noch grosse Qualitätsunterschiede bestehen und dass diese im allgemeinen nicht stichprobenabhängig zu sein scheinen.

4. Exkurs: Untersuchungen über andere Kurzversionen um 30 Items

4.1 Zielsetzung

Parallel zu der oben beschriebenen Version des MRS-30 wurden in zwei weiteren Untersuchungen zwei etwas andere Kurzversionen (vgl. dazu Tab. 12) des MRS-Inventars eingesetzt, die im folgenden MRS-30a und MRS-31 genannt werden. Diese werden denselben Kontrolluntersuchungen unterzogen wie in Abschnitt 3.4.3 beschrieben, also hinsichtlich Eigenwerten, Faktorkongruenzkoeffizienten, Item-Effizienz-Indizes, Reliabilitäten und Interkorrelationen der Skalensummenwerte untersucht.

4.2 Stichproben

Die zwei Stichproben lassen sich wie folgt charakterisieren:

Stichprobe F (MRS-30a) umfasst 413 junge Hochschulabsolventen und -absolventinnen der Universität Zürich. Davon sind 253 weiblichen und 160 männlichen Geschlechts. Das Durchschnittsalter beträgt 32.0 Jahre ($s = 5.9$ Jahre). Die Spannbreite reicht von 25 bis 70 Jahre. Die Rücklaufquote dieser Studie (Amacker & Niggli, 1996) betrug 58%.

Stichprobe G (MRS-31) bilden 130 erwerbstätige Personen - 69 Männer und 61 Frauen - aus dem Bekanntenkreis der AutorInnen (Buzzi, Kunz & Schmid, 1994). Das Altersmittel liegt bei 36.1 Jahren ($s = 10.7$). Die jüngste Person ist 18, die älteste 61 Jahre alt.

4.3 Ergebnisse

Der Aufbau des Ergebnisabschnittes richtet sich nach Abschnitt 3.4.3.

a) Eigenwertverläufe: Die Eigenwertverläufe sind in Tab. 9 festgehalten. Es ergaben sich 7 (Stichprobe F) bzw. 8 (Stichprobe G) Faktoren mit einem Eigenwert grösser 1. Der Eigenwertverlaufs spricht aber in beiden Stichproben klar für eine 5-Faktorenlösung.

Tab. 9: Die Eigenwertverläufe aus den Stichproben F (MRS-30a) und G (MRS-31)

Stichprobe	Beginn des Eigenwertverlaufs										
	1	2	3	4	5	6	7	8	9	10	11
F (MRS-30a)	4.7	3.7	2.6	2.4	2.2	1.3	1.1	0.9	0.9
G (MRS-31)	6.1	3.9	2.8	2.2	1.9	1.3	1.1	1.0	0.9	0.8	...

b) Faktorkongruenzkoeffizienten: In Tab.10 sind die Faktorkongruenzkoeffizienten aus den Stichproben F und G aufgeführt. Zum Vergleich sind auch die Koeffizienten der Replikationsstichproben B-E in die Tabelle aufgenommen worden.

Die Kongruenzkoeffizienten für die Gesamtlösungen der Versionen MRS-30a und MRS-31 sind mit .91 und .96 recht hoch. Dies gilt auch für die Koeffizienten der einzelnen Faktoren (.91 bis .97), ausser für denjenigen der Dimension Kultur in Stichprobe G (.86). Ein Vergleich der Koeffizienten mit dem MRS-30 zeigt, dass das MRS-30a vergleichbare Kongruenzkoeffizienten, das MRS-31 aber durchwegs leicht tiefere Werte erzielt.

Tab. 10: Die Faktorkongruenzkoeffizienten aus den Stichproben F (MRS-30a) und G (MRS-31)

Stichprobe	Faktorkongruenzkoeffizienten					
	Total	Extra-version	Verträglichkeit	Gewissenhaftigkeit	Emotionale Stabilität	Kultur
B-E (MRS-30):	.96	.97	.96	.98	.98	.92
F (MRS-30a):	.96	.97	.95	.96	.96	.96
G (MRS-31):	.91	.94	.92	.92	.93	.86

c) Item-Effizienz-Indizes: Tab. 11 enthält die Item-Effizienz-Indizes der Stichproben F (MRS-30a) und G (MRS-31), zusammen mit den Werten aus der Stichprobe B-E (MRS-30). Die IEI der von der MRS-30-Version 'abweichenden' Items sind in der Tabelle zur Hervorhebung kursiv gedruckt.

Die IEI der Stichprobe F (MRS-30a) liegen zwischen -2.6 und 0.1, diejenigen der Stichprobe G (MRS-31) zwischen -3.5 und 0.3 bei einem Mittelwert von -1.0 ($s = 0.6$) bzw. -1.7 ($s = 1.3$). Sie fallen also – vor allem in Stichprobe G – leicht tiefer aus als in der Replikationsstichprobe B-E (MRS-30; Durchschnittswert = -0.8; $s = 0.7$). Betrachtet man nun die IEI der von dem MRS-30 'abweichenden' Items, so kann der Tabelle zunächst entnommen werden, dass

Tab. 11: Item-Effizienz-Indizes aus den Stichproben F und G im Vergleich zu B-E

	MRS-30 B-E	MRS-30a F	MRS-31 G	
Extraversion				
1	gesprächig - schweigsam	-1.0	-1.2	-1.7
3	kontaktfreudig - zurückhaltend	-0.6	-0.3	-1.9
5	gesellig - zurückgezogen	-0.7	-0.6	-0.8
6	anschlussbedürftig - einzelgängerisch	-1.0	-0.8	-1.0
8	offen - zugeknöpft	-1.7	-2.3	-2.4
9	impulsiv - selbstbeherrscht	-3.5	-2.6	-4.2
Verträglichkeit				
10	zustimmend - gegensätzlich		-2.0	
11	vertrauensvoll - misstrauisch		-2.0	-1.8
12	gutmütig - reizbar	-0.4	-1.0	-1.8
13	nachsichtig - barsch	0.4		-0.2
14	selbstlos - selbstsüchtig	-0.1	-0.6	-1.5
15	friedfertig - streitsüchtig	-0.4	-0.3	-1.8
16	weichherzig - rücksichtslos	-1.4	-2.9	-3.0
17	leichtgläubig - zynisch			-3.1
18	höflich - grob	-1.5		
Gewissenhaftigkeit				
19	gründlich - unsorgfältig	-0.1		-0.2
20	ordentlich - unachtsam	-0.4	-1.2	0.0
21	geordnet - ungeordnet	-0.2	-1.4	0.0
22	strebsam - ziellos		-0.9	
24	übergenu - ungenau	-0.3	-0.1	0.3
25	gewissenhaft - nachlässig	0.3	0.1	-1.6
26	fleißig - faul	-0.8	-0.3	
Emotionale Stabilität				
28	robust - verletzlich	-0.2	-0.4	-3.4
29	selbstzufrieden - selbstmitleidig	-0.5	-1.0	-1.7
30	entspannt - überempfindlich	-0.7		-2.3
31	mutig - furchtsam		-1.9	
32	ruhig - ängstlich			-2.2
34	selbstvertrauend - hilflos	-2.0	-0.8	-2.9
35	gefühlsstabil - labil	-0.3	-0.3	-2.0
36	sicher - unsicher	-2.0	-0.9	-0.9
Kultur				
37	künstlerisch - unkünstlerisch	0.2	-0.4	-0.9
38	kreativ - unkreativ	0.2	0.0	-1.6
40	originell - konventionell	-1.0	-0.7	-1.2
41	liberal - konservativ		-1.6	
42	phantasievoll - phantasielos	-0.3	-0.2	-1.5
43	intelligent - unintelligent	-2.3	-1.8	-3.4
44	gebildet - ungebildet	-2.0		-3.5

die Extremwerte bei -0.9 (Item 22) und -3.1 (Item 17) liegen. Ferner nehmen die IEI der 'neuen' Items innerhalb der entsprechenden Faktoren keinen 'vorderen Rangplatz' ein, d.h. keines dieser Items zeichnet sich als ausgesprochene Markiervariable aus. Wenn die durchschnittlichen IEI eines Faktors in den verschiedenen Kurzversionen verglichen werden, ergeben sich z.T. analoge Befunde: Der durchschnittliche IEI der Verträglichkeits-Items des MRS-30a beträgt

-1,5 ($s = 1.0$), der des MRS-31 -1.9 ($s = 1.0$), derjenige des MRS-30 aber -0.6 ($s = 0.7$). Ähnlich verhält es sich mit der Dimension Gewissenhaftigkeit: Hier liegt der mittlere IEI bei -0.3 ($s = 0.4$; MRS-30), -0.6 ($s = 0.6$; MRS-30a) und -0.4 ($s = 0.7$; MRS-31). In den übrigen Faktoren ergeben sich für MRS-30 und MRS-30a vergleichbare IEI (die maximale Differenz beträgt 0.1). Die IEI des MRS-31 liegen hingegen deutlich tiefer (minimal -0,6 und maximal -1.2).

d) Kennwerte der Skalensummenwerte: Die Reliabilitäten des MRS-30a und MRS-31 (vgl. dazu Tab. 12) bewegen sich zwischen .63 und .85. Abgesehen von "Verträglichkeit", welche im MRS-30a ein deutlich tieferes Cronbach-Alpha von .63 hat, erzielen die übrigen Skalen ähnliche Koeffizienten wie die des MRS-30. Die berechneten Durchschnittswerte liegen mit .77 (MRS-30a) und .78 (MRS-31) aber dennoch etwas tiefer als derjenige von MRS-30 (.80).

Tab. 12: Reliabilitäten (Cronbach-Alpha) und Skaleninterkorrelation aus den Stichproben B-E, F und G

	MRS-30 B-E	MRS-30a F	MRS-31 G
Extraversion	.80	.82	.73
Verträglichkeit	.75	.63	.74
Gewissenhaftigkeit	.87	.82	.85
Emotionale Stabilität	.79	.80	.78
Kultur	.75	.73	.77
Skaleninterkorrelationen			
Maximum	.31	.21	.40
\bar{x}	.13	.12	.24

4.4 Schlussfolgerungen

Neben dem MRS-30 wurden in anderen Untersuchungen noch zwei weitere Kurzversionen (MRS-30a und MRS-31) eingesetzt. In diesem Kapitel ging es darum, die verschiedenen Versionen einander bezüglich Faktorenstruktur und Skaleneigenschaften gegenüberzustellen. Hauptbefund der Analysen ist, dass einerseits die MRS-30-Varianten im Vergleich zum MRS-30 keinen Gewinn darstellen. Andererseits sind die Qualitätsunterschiede der drei verglichenen Versionen aber auch nicht sehr gross, was natürlich mit der grossen Zahl der gemeinsamen Items zusammenhängt.

5. Untersuchung III: Konstruktion des MRS-25 sowie MRS-20 und Kontrolluntersuchungen

5.1 Zielsetzung und Vorblick

Ziel dieses Kapitels ist es, das MRS-30-Inventar weiter zu kürzen, und zwar auf eine Version mit fünf (MRS-25) bzw. mit vier (MRS-20) Items pro Faktor. Ausgangspunkt für die weiteren Kurzversionen bilden die Daten der Stichprobe B-E, die mit 1'308 Probanden eine sehr verlässliche Basis bilden. Als Kriterium zur Auswahl der Items werden die Item-Effizienz-Indizes herangezogen (Abschnitt 5.2). In Abschnitt 5.3 werden die Konsequenzen der Kürzungen auf die Faktorenstruktur und die Reliabilität in der Stichprobe B-E analysiert. Anschliessend soll geprüft werden, ob sich die Ergebnisse in insgesamt drei weiteren Stichproben replizieren lassen (Abschnitt 5.4).

5.2 Die Auswahl der 25 bzw. 20 Items

Werden die Items mit den höchsten Item-Effizienz-Indizes (vgl. Tab.7, siebte Spalte) ausgewählt, setzen sich MRS-25 und MRS-20 aus folgenden Items zusammen (in Klammern steht das für MRS-20 wegzulassende Item):

Extraversion:	1, 3, 5, 6 und (8)
Verträglichkeit:	12, 13, 14, 15 und (16)
Gewissenhaftigkeit:	19, 21, 24, 25 und (20)
Emotionale Stabilität:	28, 29, 30, 35 und (36) ²
Kultur:	37, 38, 40, 42 und (44)

5.3 Konsequenzen der Kürzungen in den Daten der Stichprobe B-E

Die Eigenwertverläufe des MRS-25 und MRS-20 in der Stichprobe B-E sind in Tab.13 festgehalten. In beiden Kurzversionen spricht der Eigenwerteverlauf deutlich für eine 5-Faktorenlösung.

² Item 36 wird Item 34 vorgezogen, weil Item 34 in Einzelstichproben A & E (Tab. 7) sehr tiefe IEI aufweist, was sich auch in Stichprobe G (Tab. 11) nochmals bestätigt.

Tab. 13: Die Eigenwertverläufe aus der Stichprobe B-E

Stichprobe	Beginn des Eigenwertverlaufs							
	1	2	3	4	5	6	7	8
B-E (MRS-25):	4.2	3.5	2.5	2.3	1.9	1.0	0.9	...
B-E (MRS-20):	3.4	2.9	2.4	2.0	1.6	0.9	0.7	...

Tab. 14: Faktorladungen, Faktorkongruenz-, Item-Effizienz-Koeffizienten und Reliabilitäten des MRS-25 und MRS-20

	Extra- version (25/20)	Verträg- lichkeit (25/20)	Gewissen- haftigkeit (25/20)	Emot. Stabilität (25/20)	Kultur (25/20)	IEI (25/20)
Extraversion						
1	79/ 77	-16/ -18	05/ 06	06/ 06	14/ 16	-0.8/ -1.1
3	85/ 83	-05/ -07	02/ 03	10/ 10	18/ 20	-0.5/ -0.8
5	79/ 82	04/ 03	-06/ -05	15/ 15	11/ 13	-0.6/ -0.6
6	71/ 75	14/ 14	-10/ -09	-11/ -11	-09/ -06	-1.1/ -0.9
8	67	04	11	20	22	-1.6
Verträglichkeit						
12	-03/ -01	72/ 74	01/ 00	19/ 17	03/ 04	-0.3/ -0.1
13	00/ 00	72/ 73	01/ 00	-01/ -04	03/ 04	0.5/ 0.4
14	-05/ -04	54/ 55	07/ 07	03/ 00	-02/ -02	-0.1/ 0.1
15	-01/ 01	71/ 75	08/ 07	20/ 18	-01/ 00	-0.5/ -0.3
16	14	68	04	-22	09	-1.3
Gewissenhaftigkeit						
19	-01/ -01	07/ 08	84/ 85	07/ 07	04/ 04	0.1/ 0.1
20	04	12	83	10	-03	-0.3
21	01/ 01	06/ 08	80/ 79	12/ 12	-05/ -05	-0.1/ -0.3
24	-09/ -08	-05/ -03	77/ 79	-12/ -13	00/ -01	-0.3/ -0.2
25	02/ 01	04/ 04	79/ 81	03/ 03	01/ 01	0.4/ 0.4
Emotionale Stabilität						
28	02/ 02	-07/ -06	-08/ -07	74/ 76	-04/ -02	-0.1/ 0.1
29	10/ 09	09/ 06	08/ 09	63/ 66	05/ 06	-0.6/ -0.5
30	02/ 03	22/ 25	-09/ -09	74/ 75	01/ 03	-0.7/ -0.9
35	03/ 03	08/ 08	14/ 15	73/ 74	01/ 01	-0.3/ -0.4
36	22	-11	21	63	14	-2.1
Offenheit für Erfahrung						
37	-03/ -04	02/ 03	-02/ -02	-07/ -08	81/ 82	0.3/ 0.2
38	11/ 09	02/ 02	04/ 05	03/ 03	86/ 87	0.1/ 0.1
40	21/ 20	00/ -01	-09/ -09	10/ 09	68/ 68	-0.9/ -0.9
42	17/ 15	04/ 02	04/ 05	05/ 05	81/ 82	-0.4/ -0.3
44	07	15	18	12	16	-1.9
Faktorkongruenzkoeffizienten:						
MRS-25	(total: 97)	97	97	98	97	93
MRS-20	(total: 97)	96	97	97	95	96
Skalensummenwerte (MRS-25/MRS20):						
Maximum				29/25		
Skaleninterkorrelationen	\bar{x}			11/09		

Die Faktorladungsmuster, die Faktorkongruenzkoeffizienten sowie die Item-Effizienz-Indizes des MRS-25 und MRS-20 sind in Tab.14 aufgeführt. Aus dieser Tabelle wird ersichtlich, dass sich sowohl für die Gesamtlösung als auch für die einzelnen Faktoren sehr hohe Faktorkongruenzkoeffizienten ergeben (.97 für die Gesamtlösungen und .93 bis .98 für die einzelnen Faktoren). Die Koeffizienten des MRS-25 und MRS-20 - ausser vielleicht jene für die Dimension Kultur - unterscheiden sich kaum voneinander. Für den etwas niedrigeren Kongruenzkoeffizienten des MRS-25 in Kultur dürfte Item 44 ("gebildet – ungebildet") verantwortlich sein, welches tiefe und zudem Mehrfach-Ladungen aufweist. Dennoch: Mit einer Gesamtkongruenz von .96 und .97 zeigen beide Versionen noch höhere Werte als das MRS-30 (vgl. Tab. 6).

Die Item-Effizienz-Indizes des MRS-25 streuen zwischen -2.1 (Item 36) und 0.5 (Item 13) bei einem Mittelwert von -0.5 ($s = 0.8$), die des MRS-20 zwischen -1.1 (Item 1) und 0.4 (Item 13). Hier liegt der Durchschnitt gar bei -0.3 ($s = 0.6$). Diese Befunde machen deutlich, dass zum einen das MRS-25 und das MRS-20 aus deutlich 'effizienteren' - d.h. faktoriell 'reineren' - Items besteht als das MRS-45 oder das MRS-30, bei welchen das Mittel bei -1.7 bzw. -0.8 liegt. Zum anderen bestehen zwischen den einzelnen Items des MRS-25 und MRS-20 - ausgedrückt in einer relativ geringen Streuung - keine grossen Qualitätsunterschiede mehr.

Tab. 15: Interkorrelationen der Skalensummenwerte (MRS-25/ MRS-20)

Faktor:	r_{tt}	(2)	(3)	(4)	(5)
(1) Extraversion	84/ 82	.03/ -.02	.00/ -.04	.19/ .11	.29/ .25
(2) Verträglichkeit	72/ 68		.11/ .09	.14/ .22	.07/ .03
(3) Gewissenhaftigkeit	87/ 83			.11/ .06	.00/ -.02
(4) Emot. Stabilität	75/ 72				.10/ .05
(5) Kultur	76/ 81				

Auch die Skalensummenwerte liefern recht gute Werte (vgl. dazu Tab. 15): Die Reliabilitäten der beiden Versionen bewegen sich zwischen .68 und .87, wobei wie beim MRS-30 "Verträglichkeit" die tiefsten Werte (.72 bzw. .68) aufweist. Mit einem mittleren Cronbach-Alpha von .80 (MRS-25) bzw. .78 (MRS-20) liegen die gewonnenen Durchschnittswerte etwa in der Höhe des MRS-30 (.80) sowie des MRS-45 (.82). Hingegen fallen mit .11 (MRS-25) und .09 (MRS-20) die mittleren Interkorrelationen tiefer aus als bei den längeren Versionen (.13 für das MRS-30 und .21 für das MRS-45). Die einzelnen Interkorrelationen können Tab. 16 ent-

nommen werden. Gesamthaft zeigt sich, dass die Zusammenhänge zwischen den einzelnen Skalen relativ niedrig ausfallen. Die höchste Korrelation weisen nach wie vor Extraversion mit Emotionaler Stabilität auf (.29 bzw. .25).

5.4 Replikationen der Befunde in anderen Stichproben

5.4.1 Vorbemerkung

Im folgenden soll der Frage nachgegangen werden, ob sich die eben berichteten "Güteeigenschaften" des MRS-25 und MRS-20 auch in anderen Stichproben replizieren lassen. Daher werden im folgenden vier weitere Stichproben denselben Analysen unterzogen, wobei zum Vergleich auch die Stichprobe A hinzugenommen wird.

5.4.2 Stichproben

Die Replikationen der in Abschnitt 5.3 berichteten Befunde basieren auf folgenden Gruppen:

Stichprobe A (Rohrer et al., 1993) umfasst 158 Psychologiestudierende (siehe Abschnitt 2.2).

Stichprobe H (Honisch, 1998) setzt sich aus 192 ehemaligen Eiskunstläuferinnen aus dem Spitzensport zusammen. Das Durchschnittsalter beträgt 30.4 Jahre ($s = 10.7$ Jahre). Unter den Eiskunstläuferinnen befinden sich 16- bis 60-Jährige. Die Rücklaufquote der Befragung lag bei 59%.

Stichprobe I (Wagenbreth, 1997) besteht aus 521 Studentinnen und 217 Studenten ($n=738$). Das mittlere Alter beträgt 30.1 Jahre ($s = 7.1$ Jahre). Die jüngste Person ist 20, die älteste Person 61 Jahre alt. 62% der angeschriebenen Personen hatten an der Befragung teilgenommen.

Stichprobe J (Bulaty, 1998) umfasst Studenten und Studentinnen der Philosophischen Fakultät I der Universität Zürich. Von den 642 Studierenden sind 221 männlichen und 421 weiblichen Geschlechts. Das Altersmittel liegt bei 30.1 Jahren ($s = 5.3$ Jahre). Der jüngste Proband ist 24, der älteste 59 Jahre alt. Die Rücklaufquote der Befragung betrug 50%.

Die zusammengefasste **Stichprobe H-J** umfasst total 1572 Personen. Von diesen sind 1134 weiblichen und 438 männlichen Geschlechts. Das Altersmittel liegt bei 30.1 Jahren ($s = 7.0$ Jahre).

Diese Stichproben werden im folgenden als Replikationsstichproben bezeichnet. Das MRS-25 wurde nur in den Stichproben A, H, und I eingesetzt. Das MRS-25 wird daher in den Stichproben A, H und I analysiert. Für die Untersuchungen zum MRS-20 wird dagegen nur die Zusammenfassung der Stichproben H, I und J ("H-J") benutzt.

5.4.3 Ergebnisse

a) Eigenwertverläufe: Die Eigenwertverläufe der einzelnen Replikationsstichproben sind in Tab. 16 festgehalten. Der Scree test spricht in den Stichproben H, I und H-J klar für eine 5-Faktorenlösung.

Tab. 16: Die Eigenwertverläufe von MRS-20 und MRS-25 in den verschiedenen Stichproben

Stichprobe:	Beginn des Eigenwertverlaufs								
	1	2	3	4	5	6	7	8	9
MRS-25:									
A:	5.5	3.4	2.7	2.4	1.5	1.0	0.9	0.8	...
H:	4.5	3.4	2.4	2.1	2.0	1.2	0.9	0.9	...
I:	4.7	3.4	2.4	2.2	2.1	0.9	0.9
MRS-20:									
H-J:	3.6	2.6	2.1	1.9	1.6	0.9	0.7

Tab. 17: Die Faktorkongruenzkoeffizienten

Stichprobe	Faktorkongruenzkoeffizienten					
	Total	Extra- version	Verträglichkeit	Gewissenhaftigkeit	Emotionale Stabilität	Kultur
MRS-25:						
A:	.95	.96	.91	.96	.96	.93
H:	.94	.95	.94	.98	.95	.89
I:	.96	.97	.96	.98	.98	.95
MRS-20:						
H-J:	.97	.97	.98	.98	.97	.97

b) Faktorkongruenzkoeffizienten: In Tab. 17 sind die Faktorkongruenzkoeffizienten aus den verschiedenen Replikationsstichproben festgehalten. (Vergleichsbasis bilden wieder die Ladungen in Ostendorf, 1990.) Für die Gesamtlösungen liegen die Koeffizienten zwischen .94

und .97 (MRS-25) und bei .96 (MRS-20). Die Kongruenzkoeffizienten der einzelnen Faktoren erreichen - bis auf eine Ausnahme (Kultur in Stichprobe H) - Werte, die höher als .90 sind.

Tab. 18: Die Item-Effizienz-Indizes aus den verschiedenen Replikationsstichproben

	MRS-25	MRS-25	MRS-25	MRS-20	
	A	H	I	H-J	
Extraversion					
1	gesprächig - schweigsam	-1.5	-1.4	-1.0	-1.3
3	kontaktfreudig - zurückhaltend	-1.2	-0.5	-0.8	-0.6
5	gesellig - zurückgezogen	-1.2	-0.2	-0.7	-0.2
6	anschlussbedürftig - einzelgängerisch	-0.9	-1.7	-0.7	-0.5
8	offen - zugeknöpft	-2.8	-1.1	-1.6	
Verträglichkeit					
12	gutmütig - reizbar	-2.2	-0.2	-1.2	-0.8
13	nachsichtig - barsch	-0.6	-0.4	0.0	0.2
14	selbstlos - selbstsüchtig	-0.6	0.2	-0.2	-0.1
15	friedfertig - streitsüchtig	-0.2	-1.8	-0.4	0.1
16	weichherzig - rücksichtslos	-2.1	-0.5	-1.4	
Gewissenhaftigkeit					
19	gründlich - unsorgfältig	-0.3	-0.2	0.2	0.0
20	ordentlich - unachtsam	-0.1	-0.4	0.1	
21	geordnet - ungeordnet	-0.9	-0.1	-0.5	-0.6
24	übergenu - ungenau	-0.1	-0.3	0.0	0.3
25	gewissenhaft - nachlässig	-0.3	0.4	0.3	0.2
Emotionale Stabilität					
28	robust - verletzlich	-0.2	-0.3	-0.2	0.0
29	selbstzufrieden - selbstmitleidig	-0.6	-0.7	-1.0	-1.3
30	entspannt - überempfindlich	-1.1	-0.3	-0.5	-0.6
35	gefühlsstabil - labil	-0.5	-0.8	-0.3	-0.3
36	sicher - unsicher	-2.1	-3.5	-1.7	
Kultur					
37	künstlerisch - unkünstlerisch	-0.3	-0.2	0.5	0.4
38	kreativ - unkreativ	0.3	-0.6	0.1	0.0
40	originell - konventionell	-1.3	-2.0	-1.2	-0.9
42	phantasievoll - phantasielos	-0.5	-0.5	-0.3	-0.5
44	gebildet - ungebildet	-1.2	-1.5	-2.0	

c) **Item-Effizienz-Indizes:** Tab. 18 enthält die Item-Effizienz-Indizes der Replikationsstichproben. Die IEI des MRS-25 streuen zwischen -3.5 (Item 36 in Stichprobe H) und 0.5 (Item 37 in Stichprobe I) bei Durchschnittswerten von -0.9 ($s = 0.8$; Stichprobe A), -0.7 ($s = 0.8$; Stichprobe H) und -0.6 ($s = 0.7$; Stichprobe I). Beim MRS-20 bewegen sich die IEI zwischen -1.3 (Item 1 und 29) bei einem Mittelwert von -0.3 ($s = 0.6$). Das MRS-25 und das MRS-20 erzielen somit in den Replikationsstichproben - ausser vielleicht in Stichprobe A - ähnlich hohe durchschnittliche IEI wie in der Stichprobe B-E, auf welcher die Kürzungen beruhen (vgl. dazu Abschnitt 5.3).

d) Kennwerte der Skalensummenwerte: Tab. 19 enthält die Reliabilitäten und Skalenerkorrelationen aus den Replikationsstichproben. Die Reliabilitätskoeffizienten des MRS-25 bewegen sich zwischen .65 und .87, die des MRS-20 zwischen .68 und .84. Die mittleren Cronbach-Alpha des MRS-25 liegen bei .83 (Stichprobe A), .78 (Stichprobe H) und .81 (Stichprobe I), das des MRS-20 (Stichprobe H-J) beträgt .79. Dieses Ergebnis erhärtet die schon in Abschnitt 5.3 gemachte Feststellung, dass sich die Kürzungen kaum negativ auf die Reliabilitäten auswirken. Analog verhält es sich mit den Interkorrelationen: Gesamthaft gesehen korrelieren die einzelnen Skalen umso weniger miteinander, je kürzer die MRS-Version ist. So betragen z.B. die mittleren Interkorrelationen der Stichprobe A beim MRS-45 .21, beim MRS-30 .17 und beim MRS-25 noch .16.

Tab. 19: Reliabilitäten (Cronbach-Alpha) und Skalenerkorrelationen aus den Replikationsstichproben

	MRS-25 A	MRS-25 H	MRS-25 I	MRS-20 H-J
Extraversion	.87	.85	.86	.84
Verträglichkeit	.65	.71	.72	.68
Gewissenhaftigkeit	.88	.87	.84	.80
Emotionale Stabilität	.86	.73	.82	.76
Kultur	.80	.68	.79	.82
Skalenerkorrelationen				
Maximum	.49	.25	.26	.26
\bar{x}	.16	.14	.12	.13

5.5 Schlussfolgerungen

Im Zentrum dieses Kapitels stand die Überprüfung einer auf fünf (MRS-25) und einer auf vier (MRS-20) Items pro Faktor gekürzten Version in weiteren Stichproben. Gesamthaft gesehen ergaben sich sehr gute Ergebnisse. Beide Kurzversionen weisen eine hohe faktorielle Validität sowie recht gute Reliabilitätskoeffizienten auf.

6. Bilanz, Ausblick auf weitere Ergebnisse und Schlussbemerkungen

6.1 Rückblick

Ausgangspunkt dieser Untersuchungen bildete die Feststellung, dass bei Befragungen oft mit einer substantiellen Persönlichkeitsabhängigkeit der Antworten zu rechnen ist. Um diese Abhängigkeit überhaupt in den Blick zu bekommen, ist es notwendig, "die" Persönlichkeit der Befragten mitzuerfassen. Auf diesem Hintergrund besteht das Bedürfnis nach einem Instrument, das 1.) so kurz ist, dass es gleichsam 'routinemässig' in Untersuchungen eingesetzt werden kann, 2.) den Persönlichkeitsbereich möglichst umfassend abdeckt und 3.) in Umfragen zu ganz verschiedenartigen Themen eine gute Akzeptanz aufweist. Ziel war, ein solches Instrument zu entwickeln und in testologischer Hinsicht zu überprüfen. Als Ausgangsbasis diente das 45 Items umfassende MRS-Inventar von Ostendorf (1990) zur Erfassung der im Fünf-Faktoren-Modell der Persönlichkeit enthaltenen Dimensionen Extraversion, Verträglichkeit, Gewissenhaftigkeit, Neurotizismus und Kultur.

Methodisch wurde so vorgegangen, dass die ursprüngliche Menge von 45 bipolaren Rating-skalen sequentiell aufgrund der Kriterien faktorielle Einfachstruktur und (aus der Sicht der Pbn) nicht irritierende Items auf Mengen von 30 ("MRS-30" genannt), 25 ("MRS-25") und 20 Items ("MRS-20") reduziert und diese anschliessend in jeweils neuen Stichproben testologischen Kontrolluntersuchungen unterzogen wurden. Insgesamt wurden 10 Stichproben von rund 100 bis 650, total über 3500 Personen einbezogen.

Gesamthaft gesehen haben die Kontrolluntersuchungen gezeigt, dass sich die vorgeschlagenen MRS-Kurzversionen durch eine hohe faktorielle Validität und durch eine für Forschungszwecke mehr als genügende Reliabilität auszeichnen. Die Faktorkongruenzkoeffizienten (Kurzversionen im Vergleich zu den Angaben zu den Original-MRS-Ratingskalen in Ostendorf, 1990) sind bis auf wenige Abweichungen hoch, vor allem beim MRS-20 (.95 - .98). Einzig für die Faktoren Verträglichkeit und Kultur des MRS-25 und MRS-30 wurden in einzelnen Stichproben Kongruenzen ermittelt, die unter .90 (aber noch über .85) liegen. Der Befund, dass mit den kürzeren Inventaren eine prägnantere Faktorstruktur erzielt wurde, dürfte auf das Weglassen von faktoriell "unreinen" Items zurückzuführen sein.³ Generell ist aber anzumerken, dass die Fünffaktorenlösung nicht immer die Lösung darstellt, die sich aufgrund des Eigenwertkriteriums und/oder des Scree-Test ganz klar aufdrängt. Sie musste in einem Teil unserer Stichproben bewusst forciert werden.

³ Es gibt in der Literatur allerdings Hinweise, dass eine sinkende Variablenzahl selbst bei Zufallsdaten zu höheren Kongruenzkoeffizienten führen kann (z.B. Nesselroade & Baltes, 1970; Schneewind & Cattell, 1970). Da sich aber in unserem Fall die kürzeren Versionen auch hinsichtlich anderer Qualitätskriterien auszeichneten, ist nicht anzunehmen, dass es sich beim oben berichteten Befund um ein Zufallsprodukt handelt.

Ein anderer Anspruch an das MRS-Inventar ist, dass bereits die Skalensummenwerte brauchbare und weitgehend orthogonale Schätzungen der fünf Persönlichkeitsfaktoren des FFM liefern. Die Güte dieser Summenwerte der verschiedenen MRS-Inventare wurde anhand der internen Konsistenz, der mittleren Skaleninterkorrelationen sowie deren Maxima beurteilt. Gesamthaft betrachtet fielen die entsprechenden Analysen sehr zufriedenstellend aus. Die mittleren Cronbach-Alpha liegen mit .80 (MRS-30 und MRS-25) sowie .78 (MRS-20) nur geringfügig unter dem von Ostendorf (1990, S. 146) ermittelten Durchschnittswert der Originalversion von .85 – dies trotz der zum Teil deutlich geringeren Anzahl Items der Kurzversionen, die ja nach der bekannten Formel von Spearman-Brown eigentlich mit einem bedeutenden Reliabilitätsverlust verbunden sein müsste. Auch hier dürfte sich ausgewirkt haben, dass es sich bei den ausgewählten Items um eine (in anderen Stichproben replizierbare) Positiv-Auswahl aus den Original-Items handelt. Die Auswahl stützte sich dabei (primär) auf eine "Minimum Redundancy Item Analysis" (MRA), die gleichzeitig zur (wiederum in unabhängigen Stichproben reproduzierbaren) Folge hatte, dass die mittleren Skaleninterkorrelationen von .21 (bzw. .20 bei Ostendorf, 1990) im MRS-45 auf .12 - .17 (MRS-30) und .11 - .16 (MRS-25) bzw. .09 - .13 (MRS-20) sanken. Das Analoge gilt für die Maxima der Skaleninterkorrelationen: Die höchste Interkorrelation, jene zwischen Extraversion und Emotionaler Stabilität, konnte kontinuierlich von .60 (MRS-45) auf etwa .26 (MRS-20) gesenkt werden.

In testologischer Hinsicht weisen die MRS-Kurzversionen somit gute bis sehr gute Güteeigenschaften auf, und es gibt keinen Grund, der gegen ihre Verwendung im eingangs skizzierten Sinne spricht. Praktische Hinweise für den Einsatz der verschiedenen Versionen finden sich im Anhang.

6.2 Ausblick auf weitere Befunde (Stabilität, Validität) und Schlussbemerkungen

Wie bereits in der Auflistung der verschiedenen Stichproben in den Kapiteln 2 bis 5 dieses Berichts zum Ausdruck kam, haben wir in unserer Forschungsgruppe bereits eine Fülle von Erfahrungen mit den hier dargestellten MRS-Skalen machen können. Die forschungspraktisch wichtigste Erfahrung ist wohl, dass – unabhängig vom Befragungsziel – die Items von den Befragten weitgehend problemlos akzeptiert werden.⁴ Dasselbe wird uns von andern Forschern berichtet, die eine der beschriebenen Kurzversionen in ihren Untersuchungen eingesetzt haben.

⁴ Allerdings liessen sich einzelne Items durchaus verbessern. Auf einen entsprechenden Versuch haben wir bisher verzichtet, weil wir möglichst unmittelbar an der umfangreichen lexikographischen Vorarbeit von Ostendorf (1990) anknüpfen wollten.

In all diesen Untersuchungen wurde natürlich eine Fülle inhaltlicher Einsichten in die Zusammenhänge zwischen Persönlichkeit und anderen Befragungsdaten gewonnen. Sie können hier nicht resümiert werden. Jedoch seien vier Einzelpunkte herausgegriffen, die einige wichtige zusätzliche Rahmeninformationen für den Einsatz der MRS-Skalen liefern können:

1.) Ein erster Punkt betrifft die **Höhe der Korrelationen** der MRS-Faktor- oder Skalen-sammenwerte zu Befragungsdaten. Anknüpfend an Mischel (1968) wird oft die Meinung vertreten, "Traitmasse" würden wenig zur Varianzaufklärung anderer Variablen beitragen. Der sog. "Persönlichkeitskoeffizient" von rund .3 bilde zumeist eine obere Grenze für Zusammenhänge mit Drittvariablen. Wie schon andernorts zeigte sich auch in unseren Untersuchungen, dass diese Behauptung ins Reich der Fabeln gehört – auch wenn natürlich die Stärke der Zusammenhänge mit dem Befragungsthema variiert. Höhere Koeffizienten finden sich insbesondere bei Variablen, die – von den Forschern oft unbemerkt – selbst bereits den Charakter von Persönlichkeitsmerkmalen haben, wie z.B. (um an konkreten Befunden anzuknüpfen) pauschale Angaben über Belastungen durch das Studium und Copingstrategien bei Studierenden, Angaben über Mobbingereignisse in den letzten Monaten bei Erwerbstätigen oder über das Ausmass an Burnout bei VertreterInnen sozialer Berufe, usw. Ebenfalls hoch sind die Zusammenhänge naturgemäss bei Selbstcharakterisierungen, und zwar unabhängig von der theoretischen Einbettung der entsprechenden Konstrukte.

Dass Variablen, die im selben Instrument erhoben wurden, oft höher korrelieren als wenn sie zu verschiedenen Zeitpunkten erhoben werden, ist eine altbekannte sozialwissenschaftliche Tatsache, die hie und da den Artefakt-Verdacht weckt. Daher sei hier noch auf ein ganz andersartiges Ergebnis hingewiesen: Im Rahmen von Untersuchungen mit der sog. Experience Sampling Method (ESM) wurden die Pbn während einer Woche zu 30 bis 50 Zufallspunkten mit einem Signal gebeten, einen kurzen Fragebogen zu ihrem aktuellen Befinden auszufüllen. Die Mittelwerte der verwendeten Befindensvariablen über die Untersuchungsperiode korrelieren bis annähernd .5 mit Dimensionen des MRS-30 (Schallberger, 1998). Die höchste Korrelation besagt dabei, dass gewissenhaftere Personen sich in ihrem Alltag – gemessen an der klassischen Valenzdimension des Befindens – wohler fühlen als weniger gewissenhafte.

Diese Hinweise vermögen vielleicht deutlich zu machen, dass der Einbezug von Standard-Persönlichkeitsmerkmalen in Befragungen eine Reihe interessanter Einsichten zu erzeugen vermag – aber natürlich auch neue Probleme. Denn die Frage, wie die Persönlichkeitsabhängigkeit von Befragungsdaten zu erklären ist, ist alles andere als trivial (vgl. dazu z.B. Schallberger, 1995, am Beispiel von arbeitsanalytischen Befragungsdaten).

2.) Aber handelt es sich bei den MRS-Dimensionen tatsächlich um Persönlichkeitsmerkmale? Der wichtigste empirische Beleg dafür, dass ein Merkmal als Persönlichkeitsmerkmal angese-

hen werden kann, ist der Nachweis einer gewissen **Langzeitstabilität**. Noch unveröffentlichte und vorläufige Befunde zu dieser Frage liegen aus dem Projekt "Arbeitserfahrungen und Lebensqualität in der Schweiz" (Schwerpunktprogramm Zukunft Schweiz) vor, in dem das MRS-30 zur Anwendung kam. Die Untersuchung erstreckte sich über ein Jahr und betraf junge Menschen beim Übergang von der Berufsausbildung in den Beruf ($n = 644$). Die entsprechenden (nicht minderungskorrigierten) Stabilitätskoeffizienten liegen zwischen .61 und .79, also in einem für Persönlichkeitsfragebogen, die untersuchte Altersstufe und eine Phase, die durch einen einschneidenden ökologischen Übergang gekennzeichnet ist, durchaus befriedigenden Bereich (Extraversion: .79; Gewissenhaftigkeit: .72; Emotionale Stabilität: .7; Kultur: .67; Verträglichkeit: .61) (Kälin, W., persönliche Mitteilung 14.11.98).

3.) Ein drittes Problem, das angesprochen werden soll, betrifft die **Äquivalenz der verschiedenen Versionen**. Die Befunde über die Faktorkongruenzkoeffizienten, über die früher berichtet wurde, lehren zwar, dass die Dimensionen der verschiedenen MRS-Kurzversionen mit den Originaldimensionen des MRS-45 sehr verwandt sind. Weiter wurde gezeigt, dass die Reliabilität der verschiedenen Kurzversionen von annähernd vergleichbarer Grössenordnung ist. Andererseits unterscheiden sich die verschiedenen Versionen nicht nur in der Zahl der Items; durch das Weglassen bestimmter Items in den kürzeren Versionen ergeben sich möglicherweise doch auch bestimmte Bedeutungsverschiebungen. Auf diesem Hintergrund stellt sich die Frage, ob empirische Befunde, die mit verschiedenen MRS-Versionen gefunden werden, wirklich vergleichbar sind, oder ob die Wahl einer bestimmten Version nicht auch bestimmte inhaltliche Konsequenzen nach sich zieht.

Diese Frage ist empirisch nicht leicht zu beantworten. Möglichkeiten dazu bietet die Stichprobe A (Rohrer et. al., 1993; vgl. Abschnitt 2.2), in der das MRS-45 vorgelegt wurde. Ein naheliegender Gedanke ist, in dieser Stichprobe die Skalensummenwerte für die verschiedenen MRS-Versionen zu bilden und diese zu korrelieren. Dies ergibt extrem hohe Koeffizienten (Streuung von .79 bis .99, wobei 5/6 der Koeffizienten über .9 und knapp 2/3 über .95 liegen). Sie sind in aller Regel umso grösser, je grösser der Anteil der Items ist, den die korrelierten Skalen gemeinsam haben, was bereits aufgrund der Natur solcher Part-Whole-Korrelationen zu erwarten ist. Eine zweite Möglichkeit besteht darin, in dieser Stichprobe für die Versionen MRS-45, MRS-30, MRS-25 und MRS-20 je eine Faktorenanalyse mit je fünf Faktoren zu rechnen und die entsprechenden Faktorwerte untereinander zu korrelieren. Bei diesem Vorgehen tragen somit auch die unabhängig voneinander durchgeführten VARIMAX-Rotationen zu einer zusätzlichen Senkung der Interkorrelationen bei. Trotzdem ergaben sich Korrelationen in derselben Grössenordnung (.83 bis .99), wobei wiederum eine Abhängigkeit von der anteilmässigen Grösse des gemeinsamen Itempools festzustellen ist.

Tab. 20: Der Zusammenhang der Faktorwerte aus den verschiedenen MRS-Versionen mit den Giessen-Test-Skalen

	Giessen-Test-Dimensionen					
	Resonanz	Dominanz ^{a)}	Kontrolle	Grundstimmung ^{a)}	Durchlässigkeit ^{a)}	Potenz ^{a)}
Extraversion						
MRS-45	.31*	-.06	-.26*	-.29*	-.68*	-.37*
MRS-30	.35*	-.13	-.25*	-.32*	-.65*	-.43*
MRS-25	.37*	-.10	-.23*	-.32*	-.65*	-.42*
MRS-20	.34*	-.05	-.22*	-.29*	-.59*	-.39*
Verträglichkeit						
MRS-45	-.06	.46*	.04	.05	.12	.05
MRS-30	-.03	.39*	.02	.05	.12	.00
MRS-25	-.10	.41*	.03	.07	.12	.02
MRS-20	-.14	.44*	.02	.11	.18*	.05
Gewissenhaftigkeit						
MRS-45	.12	.19*	.71*	.10	.00	-.12
MRS-30	.12	.20*	.69*	.12	.03	-.07
MRS-25	.12	.20*	.68*	.12	.06	-.06
MRS-20	.11	.20*	.69*	.13	.06	-.05
Emotionale Stabilität						
MRS-45	.36*	-.02	-.18*	-.69*	-.29*	-.41*
MRS-30	.34*	.06	-.16*	-.66*	-.28*	-.36*
MRS-25	.32*	.06	-.18*	-.65*	-.26*	-.37*
MRS-20	.33*	.04	-.16*	-.66*	-.29*	-.37*
Kultur						
MRS-45	.43*	-.32*	-.10	-.15	-.17*	-.29*
MRS-30	.41*	-.33*	-.10	-.21*	-.18*	-.35*
MRS-25	.39*	-.32*	-.07	-.20*	-.19*	-.35*
MRS-20	.35*	-.30*	-.08	-.17*	-.19*	-.34*

Erläuterungen:

Die kleinen numerischen Unterschiede beim MRS-45 im Vergleich zu Schallberger & Venetz (1998) resultieren aus einer anderen Form der Behandlung fehlender Werte

* = $p < .05$

^{a)} = Diese Giessen-Test-Dimensionen sind anders gepolt, als man von der Bezeichnung her erwarten würde

Etwas aussagekräftiger ist auf diesem Hintergrund ein dritter Untersuchungstyp, bei dem die Korrelationen der verschiedenen Versionen mit Drittvariablen verglichen werden. Die Ergebnisse einer solchen Analyse finden sich in Tab. 20. Sie basieren auf den eben beschriebenen Faktorenanalysen in Stichprobe A. In dieser Stichprobe wurde – wie in Abschnitt 2.2 erwähnt – auch der Giessen-Test (Beckmann et al., 1991) durchgeführt. Dies ermöglicht, die Faktorwerte aus den verschiedenen MRS-Versionen mit denselben Aussenkriterien in Beziehung zu setzen. Es zeigt sich dabei ein weitestgehend identisches Befundmuster: Welche MRS-Version man auch zugrundelegt, man käme aufgrund der empirischen Befunde zu denselben Schlüssen.

Das darf wohl als starker Hinweis auf die inhaltliche Äquivalenz der verschiedenen MRS-Versionen angesehen werden.

4.) Ein letztes Problem, das hier aufgegriffen werden soll, betrifft die Bedeutung der aus den MRS-Skalen gewonnenen "Messwerte" für die **Big-Five**. Es ist in den letzten Jahren üblich geworden, von "dem" Fünf-Faktoren-Modell (FFM) bzw. "den" Big Five zu sprechen. Dies ist auf einer gewissen Abstraktionsstufe sicher zulässig. Andererseits wurde bereits in der Einleitung angedeutet, dass die genaue Bedeutung der fünf Faktoren nicht unumstritten ist.⁵ Dies zeigt sich bereits in der Benennungsfrage: der fünfte Faktor wird z.B. als "Culture", "Intellect" oder "Openness to experience" bezeichnet; der erste Faktor "Extraversion" oder "Surgency". Theoretisch manifestieren sich die Unklarheiten vor allem in der Diskussion um das "richtige" hierarchische Modell, d.h. die Differenzierung der fünf Faktoren in Facetten. Auch empirisch zeigen sich analoge Probleme: Zwar ergeben sich beim Vergleich verschiedener Operationalisierungen des FFM in aller Regel klare Hinweise dafür, dass sie analoge Konstrukte erfassen. Die Korrelationen sind aber von derart unterschiedlicher Höhe, dass man nicht einfach generell von identischen Konstrukten ausgehen kann. Einzelne Autoren sehen hier ein grundlegendes Problem des FFM. Nach Meinung der Verfasser dieses Papiers sollte man dieses Problem aber nicht überbewerten: Geht man vom "Faktorenchaos" aus, das die Persönlichkeitsforschung bis weit in die achtziger Jahre hinein kennzeichnete, ist die zwar nicht perfekte, aber immerhin recht hohe konvergente Validität bereits ein gewaltiger Fortschritt!

Trotzdem bleibt natürlich die Frage, nach den Spezifitäten verschiedener Operationalisierungen des FFM, hier vor allem der MRS-Skalen, wichtig. Eine nicht nur spekulative Antwort auf diese Frage könnte nur durch einen direkten empirischen Vergleich der MRS-Skalen mit anderen gängigen FFM-Operationalisierungen gewonnen werden. Eine derartige Studie ist in Vorbereitung.

⁵ Ein weiteres Problem, das hier aber völlig ausgeklammert wird, ist die Frage nach der Notwendigkeit eines sechsten oder eines sechsten und siebten etc. Faktors. Die entsprechenden Vorschläge in der Literatur geben (noch?) kein konsistentes Bild.

7. Literaturverzeichnis

- Amacker, B. & Niggli, E. (1996). *Belastungen und Bewältigungsversuche von Studierenden während der Lizentiatsarbeit. Eine Befragung von ehemaligen Lizentiandinnen und Lizentianden der Philosophischen Fakultät I der Universität Zürich*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Basler-Melbourne, M. & Jutz, H. (1996). *Persönlichkeit und Motivation in der sportlichen Aktivität*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Becker, P. (1996). Wie big sind die Big Five? *Zeitschrift für Differentielle und Diagnostische Psychologie*, 17, 209-221.
- Beckmann, D., Brähler, E. & Richter, H.E. (1991). *Der Giessen-Test (GT). Ein Test für Individual- und Gruppendiagnostik* (vierte, überarbeitete Auflage mit Neustandardisierung). Bern: Huber.
- Block, J. (1995). Five Factor Model: A contrarian view. *Psychological Bulletin*, 117, 187-215.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler* (4. Aufl.). Berlin: Springer.
- Broadbents, W.J. & Elmore, P.B. (1987). A Monte Carlo study of the sampling distribution of the congruence coefficient. *Educational and Psychological Measurement*, 47, 1-11.
- Büchi, Ch., Hegemann, U. & Krapf, Ch. (1996). *Sportliche Identität von Studierenden in Turnen & Sport der ETH Zürich*. Unveröffentlichte Forschungsseminararbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Bulaty, H. (1998). *Befürchtungen hinsichtlich der beruflichen Zukunft und ihr Einfluss auf die Studiengestaltung*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Buzzi, C., Kunz, S. & Schmid, H. (1994). *Die Rolle von Persönlichkeitsmerkmalen im Erleben von Arbeit und Freizeit*. Unveröffentlichte Forschungsseminararbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Cellar, D. F., Miller, M. L., Doverspike, D. D. & Klawnsky, J. D. (1996). Comparison of factor structures and criteria-related validity coefficients for two measures of personality based on the Five Factor Model. *Journal of Applied Psychology*, 81, 694-704.
- Chalkiopoulos, I., Nierop, A. & Spörri, K. (1997). *Flow-Erfahrung und Persönlichkeit*. Unveröffentlichte Forschungsseminararbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.

- De Raad, B. & Van Heck, G. L. (Eds.). (1994). The fifth of the Big Five. *European Journal of Personality [Themenheft]*, 8, 225 - 356.
- Digmann, J.M. (1990). Personality structure: Emergence of the five-factor model. In M.R. Rosenzweig & L.W. Porter (Hrsg.), *Annual Review of Psychology*, 41 (S.417-440). Palo Alto: Annual Reviews.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4, 26-42.
- Honisch, L. (1998). *Die psychische Entwicklung im Leistungssport aus der Sicht von Eiskunstläuferinnen: Eine retrospektive Untersuchung*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- John, O. P., Angleitner, A. & Ostendorf, F. (1988). The lexical approach to personality: a historical review of trait taxonomic research. *European Journal of Personality*, 2, 171-203.
- McCrae, R.R. & John, O.P. (1992). An introduction to the five-factor model and its application. *Journal of Personality*, 60, 175-215.
- Mischel, W. (1968). *Personality and assessment*. New York: Wiley.
- Neil, J.A. & Jackson, D.N. (1976). Minimum redundancy item analysis. *Educational and Psychological Measurement*, 36, 123-134.
- Nesselroade, J.R. & Baltes, P.B. (1970). On a dilemma of comparative factor analysis: a study of factor matching based on random data. *Educational and Psychological Measurement*, 30, 935-948.
- Norman, W.T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66, 574-583.
- Nussbaum, P. (1998). *Der Einfluss der Persönlichkeit auf die Beschreibung der eigenen Arbeitssituation. Eine Untersuchung bei Schweizer Turn- und Sportlehrerinnen und -lehrern*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Ostendorf, F. & Angleitner, A. (1992). On the generality and comprehensiveness of the Five-Factor model of personality. Evidence for five robust factors in questionnaire data. In G.V. Caprara and G.L. van Heck (Hrsg.), *Modern personality psychology. Critical reviews and new directions* (S.73-109). Harvester: Wheatsheaf.
- Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur. Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit*. Regensburg: Roderer.
- Pawlik, K. (1968). *Dimensionen des Verhaltens*. Bern: Huber.

- Rohrer, A., Seidenberg-Jenny, M. & Zeller-Eshel, R. (1993). *Der Giessen-Test und das Fünf-Faktorenmodell*. Unveröffentlichte Forschungsseminararbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.
- Schallberger, U. & Venetz, M (1998). Der Giessen-Test und das Fünf-Faktoren Modell. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19 (3), 204-210.
- Schallberger, U. (1995). Die Persönlichkeitsabhängigkeit der Beschreibung der eigenen Arbeitssituation. *Zeitschrift für experimentelle Psychologie*, 42, 111-131.
- Schallberger, U. (1998). *Persönlichkeit, Tätigkeit und Befinden im Alltag*. Referat, gehalten am 41. Kongress der Deutschen Gesellschaft für Psychologie, Dresden, 27.9. – 1.10.98.
- Schneewind, K.A. & Catell, R.B. (1970). Zum Problem der Faktoridentifikation: Verteilungen und Vertrauensintervalle von Kongruenzkoeffizienten für Persönlichkeitsfaktoren im Bereich objektiv-analytischer Tests. *Psychologische Beiträge*, 12, 214-226.
- ten Berge, J.M.F. (1986). Rotation to perfect congruence and the cross-validation of component weights across populations. *Multivariate Behavioral Research*, 21, 41-64.
- Tucker, L.R. (1951). A method for synthesis of factor analysis studies. Personnel research section report no. 984. Washington, D.C.: Department of the army.
- Wagenbreth, G. (1997). *Studienzufriedenheit: Eine Analyse der personalen Einflüsse in verschiedenen Studienphasen*. Unveröffentlichte Lizentiatsarbeit, Psychologisches Institut der Universität Zürich, Abteilung Angewandte Psychologie.

((komplex – einfach))
 ((liberal – konservativ))
 ((modern – traditionell))

Bei der Herstellung konkreter Skalen sind wir in unseren Untersuchungen bisher meist in zwei Schritten vorgegangen: Erstens wurde jedes zweite Item pro Faktor umgepolt, und zwar beim ersten, dritten und fünften Faktor beginnend mit dem zweiten Item, bei den anderen Faktoren beginnend mit dem ersten Item. Zweitens wurden die Items der verschiedenen Faktoren gemischt, indem zuerst die jeweils ersten Items pro Faktor, dann die zweiten etc. aufgelistet wurden.

Am Rande sei bemerkt, dass Goldberg (1992) bei einem ähnlichen Set bipolarer Adjektivpaare auch ein sog. transparentes Format verwendet hat: Die Items wurden pro Faktor gruppiert, wobei durchgängig der jeweils sozial positiv bewertete Pol rechts stand. Ferner wurden die Namen der Faktoren angegeben (Beispiel a.a.O., S. 42). Er fand bei dieser Präsentationsform bei Studierenden als Pbn eine klarere Faktorenstruktur (vgl. aber die Untersuchung von Cellar, Miller, Doverspike und Klawnsky, 1996, die dafür spricht, dass mindestens im Anwendungskontext damit auch vermehrte Antwortverfälschungen induziert werden können).

Als Instruktion und Antwortformat benutzten wir in der Regel (sinngemäss) folgende Vorgaben:

Wie siehst Du Dich selber? Im folgenden findest Du eine Anzahl von Eigenschaftswörtern, die in Form von Gegensatzpaaren angeordnet sind. Bitte versuche mit Hilfe dieser Liste zu beschreiben, wie Du selber *im allgemeinen* bist.

Beispiel:

sehr ziemlich eher eher ziemlich sehr
 aktiv ₁ ₂ ₃ ₄ ₅ ₆ passiv

Falls Du Dich als *im allgemeinen* ziemlich passiv empfindest, kreuze wie im Beispiel das Kästchen mit der kleinen 5 an. Falls Du Dich eher im Mittelbereich siehst, hast Du die Wahl zwischen «eher aktiv» (Kästchen 3) oder «eher passiv» (Kästchen 4), je nachdem, welche Tendenz bei Dir *im allgemeinen* eher stärker ist.

Bitte versuche so spontan und offen als möglich zu antworten – auch wenn manche Eigenschaftswörter vielleicht etwas «direkt» formuliert sind und gewisse Gegensätze für Dich nicht ganz stimmen.

Bei der Auswertung sind wenn immer möglich Faktor-Werte zu benutzen. (Die Fünf-Faktoren-Lösung muss dabei unter Umständen forciert werden.) Das Interesse richtet sich ja (meist) nicht

auf die konkreten Skalenwerte, sondern auf die Konstrukte des Fünf-Faktorenmodells. Bei zu kleinen oder hoch speziellen Untersuchungsgruppen kann aber auch mit Skalensummenwerten gearbeitet werden.

Für Hinweise auf Anwendungen der und Erfahrungen mit den MRS-Skalen sind die Autoren dankbar.