

Arbeitszufriedenheitsmessung im interkulturellen Vergleich

Inauguraldissertation
zur
Erlangung des Doktorgrades
der
Wirtschafts- und Sozialwissenschaftlichen Fakultät
der
Universität zu Köln

2009

vorgelegt

von

Diplom-Psychologe, Master of IR&HRM Christian Bosau

aus

Mönchengladbach

Referent: Prof. Dr. Lorenz Fischer

Korreferent: Prof. Dr. Egon Stephan

Tag der Promotion: 3. Juli 2009

Dank

Zur erfolgreichen Durchführung eines Dissertationsvorhabens ist neben dem eigenen Engagement auch die Unterstützung vieler weiterer Personen notwendig. So wäre auch diese Arbeit hier nicht ohne die Hilfe anderer Menschen möglich gewesen, denen ich an dieser Stelle für ihre Unterstützung danken möchte.

Zunächst möchte ich mich bei dem Betreuer meiner Dissertation, Herrn Prof. Dr. Lorenz Fischer, bedanken. Dank seiner Bereitschaft, dieses Thema zu betreuen, ist diese Arbeit überhaupt erst möglich geworden. Er stand jederzeit für mich für Fragen zur Verfügung und hat durch zahlreiche fruchtbare Anregungen und Kommentare zum Erfolg dieser Arbeit beitragen. In vielen konstruktiven Diskussionen unterstützte er mich sowohl mit inhaltlichen als auch mit methodischen Ideen.

Ebenfalls gilt mein Dank meinem Zweitgutachter Herrn Prof. Dr. Egon Stephan für die Zeit und Mühe, die er in meine Arbeit investiert hat.

Darüber hinaus bedanke ich mich für die vielen inhaltlichen Anregungen bei allen meinen Kollegen und Kolleginnen, die mich während der Dissertationsphase unterstützt haben. Dieser Dank gilt einerseits den direkten Kollegen des wirtschafts- und sozialpsychologischen Instituts und andererseits allen Kollegen aus meinem Forschungsumfeld, die mich dank zahlreicher inhaltlicher und methodischer Ratschläge in meiner Arbeit weitergebracht haben.

Zuletzt und ganz besonders danke ich meiner Familie – im Speziellen meinen Eltern – und meiner Freundin für die motivationale und emotionale Unterstützung. Dank dieser Menschen, die mir auch in schwierigeren Zeiten zur Seite standen, war es mir möglich, diese Arbeit erfolgreich abzuschließen.

Christian Bosau

Juni, 2009

Inhaltsverzeichnis

Abbildungsverzeichnis III

Tabellenverzeichnis V

1 Einleitung 1

1.1 Hintergrund der Arbeit und ihr Praxisbezug 1

1.2 Ziel der Arbeit 6

1.3 Aufbau der Arbeit 10

THEORETISCHER TEIL

2 Arbeitszufriedenheit 14

2.1 Bedeutung von Arbeitszufriedenheit im Arbeitskontext 14

2.2 Messung von Arbeitszufriedenheit 19

2.3 Arbeitszufriedenheitsmessung als Managementtool 28

3 Die Bedeutung von Kultur 32

3.1 Kulturelle Rahmenmodelle 32

3.2 Kultur als Sozialisationsinstanz für kommunikatives Verhalten 45

4 Arbeitszufriedenheit im interkulturellen Kontext 49

4.1 Bisherige Ergebnisse zur Höhe der Arbeitszufriedenheit im interkulturellen Vergleich 49

4.2 Einflußgrößen auf das Arbeitszufriedenheitsniveau 54

5 Methodische Aspekte interkultureller Forschung 66

5.1 Kulturabhängigkeit vs. Kulturunabhängigkeit 66

5.2 Nationale Einflußfaktoren 72

5.3 Forschungsdesign und Stichprobenproblematik 77

5.4 Experimentelles Vorgehen in der interkulturellen Forschung 85

5.5 Äquivalenz im interkulturellen Vergleich 92

5.6 Antworttendenzen 105

5.6.1 Arten von Antworttendenzen 105

5.6.2 Umgang mit Antworttendenzen und deren inhaltliche Interpretation .. 115

6 Antworttendenzen und Kultur 122

6.1 Deskriptive Ergebnisse der bisherigen Forschung 122

6.2 Kulturelle Werte und Antworttendenzen 125

6.3 Kulturelle Werte, individuelle Unterschiede, Antworttendenzen und deren Wechselwirkung 132

7 Arbeitszufriedenheit, Kultur und Antworttendenzen 140

7.1 Die Beziehung von Kultur und Arbeitszufriedenheit 141

7.2	Die Beziehung von Kultur und Antworttendenz	143
7.3	Die Beziehung von Antworttendenz und Arbeitszufriedenheit.....	145
7.4	Die Beziehung von Kultur, Antworttendenz und Arbeitszufriedenheit auf nationalem Level	154

EMPIRISCHER TEIL

8	Methodisches Vorgehen.....	158
9	Länderbezogene Sekundäranalysen.....	160
9.1	Die Stichproben	160
9.2	Die Meßgrößen	163
9.3	Ergebnisse der bivariaten Analysen.....	168
9.3.1	Deskriptive Ergebnisse zum Arbeitszufriedenheitsniveau	168
9.3.2	Der Zusammenhang von Kultur und Arbeitszufriedenheitsniveau	172
9.3.3	Der Zusammenhang von Kultur und Antworttendenz-Norm	178
9.3.4	Der Zusammenhang von Antworttendenz-Normen und Arbeitszufriedenheitsniveau	183
9.3.5	Zusammenfassung der Ergebnisse der bivariaten Analysen.....	190
9.4	Die Mediationsanalysen.....	191
9.4.1	Vorgehen der Mediationsanalysen.....	192
9.4.2	Ergebnisse.....	193
9.4.3	Zusammenfassung der Ergebnisse der Mediationsanalysen.....	202
9.5	Die Mehrebenenanalyse.....	204
9.5.1	Vorgehen der Mehrebenenanalyse.....	206
9.5.2	Stichprobe und Meßgrößen.....	209
9.5.3	Ergebnisse.....	210
9.5.4	Zusammenfassung der Ergebnisse der Mehrebenenanalysen.....	220
10	Experimentelle Überprüfung.....	222
10.1	Vorgehen des Experiments	222
10.2	Stichprobe und Meßgrößen.....	224
10.3	Ergebnisse.....	229
10.4	Die Diskussion der experimentellen Ergebnisse	236
11	Diskussion	241
11.1	Erkenntnisse aus den empirischen Analysen	242
11.2	Schlußfolgerungen für die Unternehmenspraxis	251
11.3	Schlußfolgerungen für die Forschung.....	257
11.4	Einschränkungen der Arbeit	264
11.5	Fazit	267
	Literatur	270
	Anhang	291

Abbildungsverzeichnis

Abbildung 1: Das Modell des semantischen Prozesses des Arbeitszufriedenheitsurteils (gemäß Fischer & Belschak, 2006).....	23
Abbildung 2: Das Werte-Modell nach Schwartz (1999).....	36
Abbildung 3: Eindimensionaler Vergleich interkultureller Forschungsansätze.....	68
Abbildung 4: Forschungsansätze im interkulturellen Kontext, gemäß Huang (2002).....	68
Abbildung 5: Hypothetischer Vergleich einer angefallenen Stichprobe mit einer Gesamtpopulation.....	82
Abbildung 6: Zusammenhang von interdependentem Selbstkonzept und Zustimmungstendenz in Kulturen mit hohen und niedrigen Kollektivismuspraktiken (aus Smith & Fischer, 2008).....	135
Abbildung 7: Zusammenhang von interdependentem Selbstkonzept und Zustimmungstendenz in Ländern mit hoher und niedriger Intellektueller sowie Affektiver Autonomie (aus Smith & Fischer, 2008).....	136
Abbildung 8: Zusammenhang von Interdependentem Selbstkonzept und Extremtendenzen in Ländern mit hohem und niedrigem Kollektivismus (aus Smith & Fischer, 2008).....	138
Abbildung 9: Angenommener bivariater Zusammenhang von Kultur, Antworttendenz und Arbeitszufriedenheit.....	154
Abbildung 10: Angenommener Gesamtzusammenhang von Kultur, Antworttendenz und Arbeitszufriedenheit.....	155
Abbildung 11: Antworttendenz als möglicher mediiender Faktor zwischen Kultur und Arbeitszufriedenheit.....	157
Abbildung 12: Die Arbeitszufriedenheitsniveaus vom WVS und EB im Vergleich....	169
Abbildung 13: Das Arbeitszufriedenheitsniveau gemäß den Daten des World-Values-Surveys.....	170
Abbildung 14: Das Arbeitszufriedenheitsniveau gemäß den Daten des Eurobarometers.....	171
Abbildung 15: Zusammenhang des Arbeitszufriedenheitsniveaus im WVS und der Individualismus-Dimension (nach Hofstede).....	175
Abbildung 16: Zusammenhang der Zustimmungsnorm und des Arbeitszufriedenheitsniveaus in armen und reichen Ländern.....	188
Abbildung 17: Bestätigter Gesamtzusammenhang von Individualismus, Zustimmungsnorm und Arbeitszufriedenheitsniveau.....	191
Abbildung 18: Mediationsanalyse mit Individualismus (Hofstede) in reichen Ländern; N=21.....	194
Abbildung 19: Mediationsanalyse mit InGroup-Kollektivismus (GLOBE-Studie) in reichen Ländern; N=18.....	195

Abbildung 20: Mediationsanalyse mit Individualismus (Triandis) in reichen Ländern; N=21.....	195
Abbildung 21: Mediationsanalysen mit Individualismus-Dimensionen und Zustimmungsnorm anhand der Eurobarometer-Daten, $N_1 = 16 / N_2 = 14 / N_3 = 16$	197
Abbildung 22: Mediationsanalyse mit Individualismus (Hofstede) und Zustimmungsnorm (Harzing) in reichen Ländern, N=12.....	197
Abbildung 23: Mediationsanalyse mit Individualismus (Hofstede) und Zustimmungsnorm (Harzing) anhand der Eurobarometer-Daten, N=10.....	198
Abbildung 24: Mediationsanalyse mit Machtdistanz (Hofstede) in reichen Ländern, N=21.....	199
Abbildung 25: Mehrebenenanalytisches Modell des Einflusses von Zustimmungsnorm und Zustimmungstendenz.....	209
Abbildung 26: Zusammenhang der Zustimmungstendenz mit dem Arbeitszufriedenheitsurteil in den einzelnen reichen Ländern.....	213

Tabellenverzeichnis

Tabelle 1:	Korrelationen der Wertedimensionen und der Praktiken im GLOBE-Projekt auf Länderebene (gemäß House et al., 2004, S. 736)	40
Tabelle 2:	Korrelationen der Wertedimensionen und der Praktiken im GLOBE-Projekt – auf Länderebene – mit Dimensionen von Hofstede (gemäß House et al., 2004, S. 140).....	43
Tabelle 3:	Korrelationen der Wertedimensionen im GLOBE-Projekt – auf Länderebene – mit den inhaltlich ähnlichen Dimensionen von Schwartz (gemäß House et al., 2004, S. 143).....	44
Tabelle 4:	Die drei Arten von Bias und ihre möglichen Gründe im Überblick (nach van de Vijver & Poortinga, 1997, S. 34)	98
Tabelle 5:	Zusammenhang des Arbeitszufriedenheitsniveaus mit kulturellen Dimensionen auf Länderniveau.....	173
Tabelle 6:	Zusammenhang des Arbeitszufriedenheitsniveaus mit sozioökonomischen Indikatoren.....	177
Tabelle 7:	Zusammenhang der Antworttendenz-Normen untereinander	178
Tabelle 8:	Zusammenhang von Antworttendenz-Normen und Kultur	180
Tabelle 9:	Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau	183
Tabelle 10:	Zusammenhang der weiteren Antworttendenz-Normen mit dem Arbeitszufriedenheitsniveau	184
Tabelle 11:	Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau bei armen und reichen Ländern des WVS	187
Tabelle 12:	Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau unter Kontrolle der sozioökonomischen Bedingungen.....	190
Tabelle 13:	Ergebnisse der Mehrebenenanalyse – reiche Länder	212
Tabelle 14:	Ergebnisse des mehrebenenanalytischen destruktiven Testens – reiche Länder	216
Tabelle 15:	Ergebnisse der Mehrebenenanalyse – arme Länder	218
Tabelle 16:	Ergebnisse des mehrebenenanalytischen destruktiven Testens – arme Länder.....	219
Tabelle 17:	Ergebnisse des SJT Manipulationschecks	230
Tabelle 18:	Ergebnisse des TenST Manipulationschecks	231
Tabelle 19:	Ergebnisse der Antworttendenz-Maße	233
Tabelle 20:	Ergebnisse der Zufriedenheitsurteile.....	234
Tabelle 21:	Kovarianzanalyse der Zufriedenheitsurteile.....	236

1 Einleitung

1.1 Hintergrund der Arbeit und ihr Praxisbezug

„Das Engagement am Arbeitsplatz in Deutschland sinkt weiter (12%). [...] Ein geringeres Engagement am Arbeitsplatz als hierzulande weisen lediglich die Länder Japan (9 %), Frankreich (6 %) und Singapur (4 %) auf. Auf Amerika (30 %) an der Spitze des Rankings folgen Kanada (24 %), Israel (20 %), Australien (18 %) und Großbritannien (17 %).“ (The Gallup Organization, 2003)

„Lediglich 13 Prozent der Beschäftigten verspüren [in Deutschland] eine echte Verpflichtung gegenüber ihrem Unternehmen und arbeiten hoch engagiert. Im internationalen Vergleich belegt Deutschland damit nur einen Platz im unteren Mittelfeld. Zum Vergleich: In Großbritannien fühlen sich 20 Prozent der Beschäftigten loyal verbunden und hoch motiviert, in den USA sogar 29 Prozent. Nur in Frankreich und Japan ist der Anteil emotional stark gebundener Mitarbeiter mit zwölf beziehungsweise sieben Prozent geringer als in Deutschland.“ (The Gallup Organization, 2008)

Mit diesen Pressemeldungen versucht die Gallup Organisation aufzuzeigen, wie es um das Arbeitsengagement in Deutschland bestellt sei und welchen Platz Deutschland im internationalen Vergleich diesbezüglich einnehme. Es klingt demnach so, als sei es eher schlecht um die deutsche Arbeitswelt bestellt. Zusätzlich sieht es so aus, daß Deutschland beim Vergleich mit anderen Ländern nicht wirklich positiv dasteht. Solche oder ähnliche Studien versuchen durch den Vergleich zwischen vielen Ländern, international tragfähige Aussagen über die Arbeitsplatzqualität in den einzelnen Ländern zu machen. Aufgrund internationaler Wirtschaftsmärkte und global operierender Unternehmen erhalten solche Informationen auch zunehmend größere Bedeutung. Die Internationalisierung, die zu Beginn des 21. Jahrhunderts inzwischen aus der Wirtschaftswelt nicht mehr wegzudenken ist (Deresky, 1994; Deutscher Bundestag, 2002; Koopmann & Franzmeyer, 2003), führt dazu, daß Unternehmen ihre Aktivitäten auf den ganzen Globus ausdehnen und in immer zahlreicheren, weltweit verstreuten Standorten tätig sind.

Um diese internationalen Standorte effektiv aus dem Heimatland leiten zu können, werden Informationen über die Qualität der Arbeit und der Arbeitsplätze vor Ort

benötigt. Ein dafür oft eingesetztes Mittel sind Befragungen der Mitarbeiter zur Arbeitszufriedenheit und zum Engagement am Arbeitsplatz. Solche internationalen Mitarbeiterbefragungen sind heutzutage zu einem Standardtool der Human Resources Abteilungen in weltweit tätigen Unternehmen geworden (van de Vijver & Leung, 1997). In Fragebögen, die zumeist im Heimatland des Unternehmens entwickelt und dann in die jeweiligen Landessprachen übersetzt werden, werden die Mitarbeiter in den einzelnen weltweiten Standorten dabei nach ihrer Einschätzung der Arbeitsqualität und der Zufriedenheit mit ihrem Arbeitsplatz gefragt. Die Ergebnisse werden dann zusammengetragen und oft in einem Gesamtbericht oder Ranking im Ländervergleich gegenüber gestellt; ähnlich der zu Beginn zitierten Studie der Gallup Organisation (2003). Solche Gegenüberstellungen sollen dazu dienen, bessere und schlechtere Niederlassungen im Gesamtkonzern zu identifizieren. Nicht selten folgen aus den positiven wie negativen Ergebnissen solcher Befragungen direkte Konsequenzen, sowohl struktureller als auch personeller Art. So werden Veränderungsmaßnahmen geplant und Change-Prozesse in Gang gesetzt, oder aber Manager der einzelnen Länderstandorte müssen sich sogar persönlich für ihre schlechten Werte verantworten. Zumeist werden die Ergebnisse aus einzelnen Ländern direkt und eins-zu-eins in Beziehung gesetzt; d.h. der Mittelwert aus einem Land wird einfach mit den Mittelwerten der anderen Länder oder des Heimatlandes verglichen. Schlechtere Werte auf der gemessenen Skala sollen demnach gleichbedeutend mit schlechteren Ausprägungen im Sinne des zugrunde liegenden Konstrukts sein. Auch die Gallup Organisation mißt in dieser Form in der zu Beginn genannten Pressemitteilung den prozentualen Anteil der engagierten Mitarbeiter und vergleicht den Anteil aus Deutschland direkt mit den Anteilen anderer Nationen. Anhand der Äußerungen zu Fragen und Items des eingesetzten Gallup Instruments werden diese Anteile berechnet und als international vergleichbar dargestellt.

Eine Frage wird bei solchen Vergleichen jedoch nie thematisiert: Inwieweit sind eigentlich Werte, die in verschiedenen Nationen, d.h. in verschiedenen Sprachen und Kulturen, gemessen wurden, überhaupt miteinander vergleichbar? Wenn auch diese Frage auf Anhieb etwas überflüssig erscheinen mag, ist es jedoch keineswegs selbstverständlich, daß Werte und Ergebnisse aus verschiedenen Ländern direkt miteinander in Beziehung gesetzt werden können (van de Vijver & Leung, 1997). In vielen Studien wird oftmals stillschweigend davon ausgegangen, daß die gleiche Frage,

z.B. nach dem Engagement oder der Arbeitszufriedenheit, auch die gleiche Bedeutung in verschiedenen Ländern hat und auch gleich verstanden wird. Etwaige Unterschiede in den gemessenen Werten sollen somit einen unterschiedlichen Grad von Engagement oder Zufriedenheit ausdrücken und nicht auf unterschiedliches kulturelles oder nationales Verständnis des Frageninhalts zurückzuführen sein. Basierend auf dieser Annahme wird in internationalen Mitarbeiterbefragungen oder Studien ähnlich der Gallup Studie die Frage nach der internationalen Anwendbarkeit von Fragen zumeist ausgeblendet und als nicht diskussionswürdig betrachtet.

Selbst in der Forschung wird diese mögliche Problematik oft in vielen Bereichen nicht beachtet. Daten aus den Befragungen verschiedener Länder werden oft unmittelbar verglichen, ohne daß die oben genannte Frage thematisiert wird. Jedoch folgen auch nicht alle Forschungsbereiche widerspruchlos dieser Annahme. Speziell im Bereich der interkulturellen Forschung hat sich explizit ein Bewußtsein herauskristallisiert, daß es notwendig und sinnvoll ist, sich mit der Frage der internationalen Anwendbarkeit von Fragen und der Vergleichbarkeit der dazugehörigen Antworten auseinanderzusetzen (van de Vijver & Leung, 1997). In diesem Teilbereich der Forschungslandschaft beschäftigt man sich schon seit längerem mit Faktoren, die Einfluß auf die Höhe des Ergebnisses haben können und die jedoch nicht mit Unterschieden in der Ausprägung des zugrundeliegenden Konstrukts zusammenhängen. Dazu zählen insbesondere Fragen des international unterschiedlichen Verständnisses von ein und derselben Frage sowie interkulturell unterschiedliches Antwortverhalten. Unterschiede hinsichtlich dieser Aspekte können dazu führen, daß Länder sich im gemessenen Mittelwert einer Thematik, z.B. von Arbeitszufriedenheit und Engagement, zwar unterscheiden könnten, dieser Unterschied jedoch nicht auf den eigentlichen Inhalt der Frage zurückzuführen ist, sondern stattdessen vielmehr durch das unterschiedliche semantische Verständnis in den verschiedenen Ländern bedingt ist oder aber beispielsweise durch eine Art von Antwortverhalten verzerrt wird, wenn beispielsweise in Befragungen grundsätzlich den gestellten Fragen eher zugestimmt wird.

Ergebnisse aus internationaler und interkultureller Forschung deuten darauf hin, daß es keineswegs selbstverständlich ist, daß ein und dieselbe Frage in gleicher Form in verschiedenen Nationen gestellt werden kann (Baumgartner & Steenkamp, 2001; Steenkamp & Baumgartner, 1998; van Herk, Poortinga & Verhallen, 2004). Ein Grund dafür ist, daß unterschiedliche Kulturen sich anscheinend in ihrem Antwortverhalten

unterscheiden (Harzing, 2006; Smith, 2004). Allein dieses interkulturell unterschiedliche Antwortverhalten kann somit die Höhe von gemessenen Werten verändern und dazu führen, daß der gemessene Wert nicht mehr eindeutig zu interpretieren ist. Diese bisherigen Ergebnisse lassen erkennen, daß eine direkte Vergleichbarkeit von gemessenen Werten aus verschiedenen Ländern fragwürdig und diskussionswürdig ist. Vor diesem Hintergrund ist auch die eingangs benannte Aussage der Gallup Organisation in dieser Form nicht zwingend haltbar und muß hinterfragt werden. Sind die Unterschiede im prozentualen Anteil tatsächlich ein Indikator für unterschiedliches Engagement am Arbeitsplatz und unterschiedliche Arbeitszufriedenheit oder spielen auch hier methodische und inhaltliche Einwände eine Rolle?

Die Frage nach der Vergleichbarkeit von Daten der Arbeitszufriedenheit aus verschiedenen Ländern stellt in diesem Gesamtkontext einen kleinen aber wichtigen Teilbereich dar und ist neben dem direkten Praxisbezug speziell auch aus wissenschaftlicher Sicht von Bedeutung, da Arbeitszufriedenheit weiterhin das am meisten untersuchte Konzept im Bereich der Arbeits- und Organisationspsychologie ist (Judge, Parker, Colbert, Heller & Ilies, 2001). Dennoch lassen sich so gut wie keine Aussagen speziell darüber finden, ob Maße der Arbeitszufriedenheit überhaupt über Ländergrenzen hinweg verglichen werden können. Selbst Überblicksartikel zur Arbeitszufriedenheit (Judge et al., 2001; Six & Felfe, 2006) sprechen diese Thematik – wenn überhaupt – nur am Rande und in sehr allgemeiner Form an. Sie diskutieren dabei vor allem nicht, welche spezifischen Zusammenhänge es zwischen dem zu messenden Konstrukt Arbeitszufriedenheit und dem allgemeinen Antwortverhalten gibt. Neben der Frage, ob Maße der Zufriedenheit generell beeinflußt sind durch Antworttendenzen, ist ebenso unklar, wie hoch dieser mögliche Zusammenhang überhaupt wäre und ob es sich dabei um einen positiven oder negativen Zusammenhang handeln könnte.

Darüber hinaus ist von Bedeutung, daß traditionell die Arbeits- und Organisationspsychologie im Allgemeinen und die Arbeitszufriedenheitsforschung im Speziellen vornehmlich durch Untersuchungen und Überlegungen aus dem amerikanischen Kulturraum dominiert sind. Die Zusammenstellung einer Meta-Analyse von Judge et al. (2001) zur Arbeitszufriedenheit spiegelt diesen Sachverhalt ebenfalls wider, da die dort verglichenen Studien sich sehr stark auf den nordamerikanischen Raum beziehen. Für eine überblicksartige Darstellung von Six & Felfe (2006) gilt im Prinzip dasselbe, da sie lediglich die bereits bei Judge et al. (2001)

zusammengetragenen Studien replizieren. Es ist somit nachvollziehbar, daß die Betrachtung von kulturellen Einflußfaktoren von Bedeutung ist, wenn man denn ein genaueres Verständnis über die globale Verallgemeinerbarkeit von Erkenntnissen der Arbeitszufriedenheitsforschung erreichen will, denn schließlich ist der amerikanische Kulturraum bei weitem nicht in der Lage, die kulturelle Vielfaltigkeit aller weltweiten Nationen widerzuspiegeln.

Noch bedeutsamer wird diese Fragestellung vor dem Hintergrund der weiten Verbreitung von Befragungsverfahren vor allem im Bereich der Mitarbeiterzufriedenheit, was im Rahmen der Erläuterungen zum Management in internationalen Unternehmen bereits kurz angesprochen wurde. Während eine Mitarbeiterbefragung (MAB) oft einer Vielzahl von Zielen dient, bleibt dennoch die Messung der Ausprägungshöhe der Arbeitszufriedenheit die Kernaufgabe (Schneider, Ashworth, Higgs & Carr, 1996). Damit bleibt das vorrangige Ziel einer MAB, Informationen zu generieren, die für das internationale Human Resource Management (IHRM) zur Steuerung der Mitarbeiterproduktivität wichtig sind. In international tätigen Unternehmen bietet dieses Instrument somit eine besonders wichtige Möglichkeit zur Kommunikation zwischen Management und Mitarbeiterschaft und dient dazu, in Kontakt mit der globalen Mitarbeiterschaft zu bleiben (Schneider et al., 1996). Ein zentraler Aspekt ist jedoch auch dabei, durch die Befragung der Mitarbeiterschaft wichtige Botschaften des Managements zu transportieren und somit eine sogenannte top-down-Kommunikation zu ermöglichen; z.B. durch die Betonung und Aufnahme wichtiger Themen in die Befragung. Noch bedeutsamer wird im internationalen Kontext jedoch, daß das Management auf Informationen aus den weit entfernten Standorten angewiesen ist und somit eine bottom-up-Kommunikation ermöglicht wird. Das Management ist so in der Lage, aus dem Mutterhaus heraus zu erfahren, wie die Situation in den ausländischen Niederlassungen vor Ort ist. Um die erhaltenen Informationen jedoch qualitativ sinnvoll einschätzen zu können, ist der Vergleich mit anderen relevanten Werten notwendig. Dazu werden in internationalen MABs oft die Werte anderer Länder herangezogen, in denen das Unternehmen tätig ist. Dies jedoch wäre nur möglich, wenn geklärt ist – wie bereits oben angemerkt – ob die Befragungswerte auf dem gleichen Verständnis der Frage beruhen und nicht durch das grundsätzliche Befragungsverhalten der Mitarbeiter aus verschiedenen Kulturen verzerrt sind.

Eine eindeutige Antwort auf diese Frage läßt sich jedoch bisher weder in Bezug zu den Erkenntnissen der forschungsorientierten Literatur noch auf Basis der Erfahrung aus praktischen Anwendungskontexten geben. Die bereits kurz angesprochenen Ergebnisse der interkulturellen Forschung legen jedoch nahe, daß auch die Befragung von Mitarbeitern in unterschiedlichen Ländern hinsichtlich ihrer Zufriedenheit mit ihrem Job unterschiedliche Ergebnisse aufgrund des kulturell bedingten Antwortverhaltens erzeugen müßte und nicht grundsätzlich aufgrund der Tatsache, daß Mitarbeiter aus dem einen Land tatsächlich zufriedener mit ihrer Arbeit sind als Mitarbeiter aus einem anderen Land. Wie diese Zusammenhänge sich jedoch genau gestalten, kann aus bisherigen Untersuchungen nicht ersehen werden. Daher soll diese Arbeit einen Beitrag dazu liefern, diese Fragestellung näher zu beleuchten und die Problematik genauer zu analysieren.

1.2 Ziel der Arbeit

Im Rahmen dieser Arbeit sollen Erkenntnisse geliefert werden, die die gegenwärtige Diskussion in der interkulturellen Forschung weiter vertiefen können, inwieweit Ergebnisse aus Befragungen über Ländergrenzen hinweg verglichen werden können. Dabei liegt der Fokus explizit auf der Messung von Arbeitszufriedenheit in verschiedenen Nationen und Kulturen. Die Ergebnisse lassen sich darüber hinaus aber auch in Teilen auch auf andere Felder der interkulturellen Anwendung von Befragungsverfahren ausdehnen.

Um jedoch die Vergleichbarkeit von Arbeitszufriedenheitswerten überhaupt thematisieren zu können, muß zuerst einmal gezeigt werden, inwieweit sich Arbeitszufriedenheitswerte in verschiedenen Ländern voneinander unterscheiden. Dazu sollen in einem ersten Schritt die deskriptiven Ergebnisse international gemessener Arbeitszufriedenheit dargestellt werden. Hierbei werden einerseits bereits veröffentlichte Daten berichtet, die jedoch ebenfalls mit eigenen Berechnungen anhand von Daten aus aktuellen und international umfangreichen Studien verglichen werden. Während diese Darstellung der deskriptiven Ergebnisse lediglich die Basis der zugrunde liegenden Fragestellungen ausmacht, liegt der Hauptfokus auf der Überprüfung der Frage, inwieweit ein Vergleich dieser Ergebnisse interkulturell überhaupt zulässig ist oder aber aufgrund verschiedener methodischer und inhaltlicher Einflußgrößen besser nicht durchgeführt werden sollte.

Die möglichen Einflußfaktoren werden dabei zuerst theoretisch ausführlich diskutiert und dann im weiteren Vorgehen anhand empirischer Daten auf ihre Bedeutung hin untersucht. Dabei soll ein Schwerpunkt auf der Untersuchung von Antworttendenzen liegen, da diese sich als sehr bedeutsam und einflußreich in anderen international vergleichenden Untersuchungen herausgestellt haben (siehe Baumgartner & Steenkamp, 2001; Fischer, 2004; Fischer, Fontaine, van de Vijver & van Hemert, 2006; Harzing, 2006; Smith, 2004; Smith & Fischer, 2008; Steenkamp & Baumgartner, 1998; van Herk et al., 2004). Die verschiedenen Formen dieses speziellen kommunikativen Antwortverhaltens werden dazu getrennt hinsichtlich ihrer Bedeutsamkeit untersucht. Des weiteren sollen die Ergebnisse in die bisherigen Forschungsergebnisse zu kulturellen Werten eingebunden werden. Da sich in der Forschung bereits zeigen ließ, daß einerseits kulturelle Werte in Beziehung zur Arbeitszufriedenheit stehen (Hofstede, 2001) und andererseits kulturelle Werte auch bedeutsam für das Auftreten von Antworttendenzen sind (Harzing, 2006; Smith, 2004), soll gezeigt werden, wie sich die Ergebnisse zum Zusammenhang von Antworttendenzen und der Messung von Arbeitszufriedenheit in diese Dreiecksbeziehung einordnen lassen.

Zur Betrachtung dieser Problemstellungen sind umfangreiche Daten notwendig. Die Forschung im interkulturellen Bereich lebt davon, möglichst viele Daten aus möglichst vielen Nationen analysieren zu können. Nur anhand einer großen Anzahl von Ländern, die in den Studien erfaßt werden, lassen sich tragfähige Ergebnisse berichten. Oftmals werden in Forschungsvorhaben jedoch nur wenige (zumeist westliche) Länder miteinander verglichen. Die Übertragbarkeit solcher Ergebnisse ist dabei verständlicherweise nicht immer eindeutig. In der vorliegenden Studie sollen daher im Gegensatz zum oft eingesetzten Zwei- oder Drei-Länder-Vergleich möglichst viele Länder erfaßt werden und somit einerseits Analysen an umfangreichen, repräsentativen Datensätzen durchgeführt werden, um die globale Gültigkeit der Ergebnisse besser belegen zu können. Diesem enormen Vorteil solcher sekundäranalytischer Untersuchungen steht jedoch andererseits auch durchaus ein bedeutsamer Nachteil gegenüber. So ist es in solchen Untersuchungen zumeist nicht möglich, alle relevanten Störeinflüsse weder methodisch noch statistisch zu kontrollieren. Aus diesem Grund sollen die in der Sekundäranalyse gewonnenen Erkenntnisse durch weitere experimentelle Studien auf ihre Bedeutsamkeit hin überprüft werden. Erst die Überprüfung anhand von Experimenten läßt es zu, weitere Alternativerklärungen so weit wie möglich auszuschließen. Die vorliegende Arbeit bedient sich somit bewußt

eines Methodenmixes, um den Erkenntnisgewinn zu optimieren. Da es sich bei der hier untersuchten Fragestellung zusätzlich um eine Situation handelt, in der Daten auf mehreren Ebenen (d.h. Individualdaten und Länderkennwerte) zur Verfügung stehen bzw. analysiert werden sollen, greifen die vorliegenden Untersuchungen auch neuere methodische Entwicklungen auf und setzen ebenfalls die noch recht neuen Mehrebenenanalysen ein (Bryk & Raudenbush, 1992; Engel, 1998; Hox, 1995; Nezlek, Schröder-Abé & Schütz, 2006). Insgesamt werden somit Ergebnisse aus verschiedenen methodischen Herangehensweisen, d.h. Sekundäranalyse, Mehrebenenanalyse und experimentelle Designs, herangezogen, um die Problematik zu beleuchten.

Neben der Frage nach der optimalen methodischen Vorgehensweise muß bei interkultureller Forschung auch jeweils die Entscheidung getroffen werden, ob es sich bei den durchgeführten Studien um ausprägungsorientierte oder aber auch strukturorientierte Forschung handelt (van de Vijver & Leung, 2000). Ausprägungsorientierte Fragestellungen – auch oft Level-orientierte Fragestellungen genannt – befassen sich dabei mit der Frage, welche Einflußfaktoren es auf die absolute Höhe von Werten gibt, die in verschiedenen Ländern gemessen wurden. Im Blickpunkt stehen somit das absolute Ergebnis hinsichtlich des untersuchten Konstrukts und die Überlegung, inwieweit diese Ausprägungshöhe über Ländergrenzen hinweg verglichen werden kann. Bei strukturorientierten Fragestellungen hingegen ist der relative Bezug von Faktoren untereinander von Bedeutung. Es handelt sich dabei dann zumeist um Untersuchungen, die prüfen, ob die Binnenstruktur eines Konstruktes sich in verschiedenen Ländern und Kulturen replizieren läßt. Untersuchungen zur internationalen Generalisierbarkeit spezifischer Arbeitszufriedenheitsmodelle fallen ebenfalls unter diese Thematik. Ein konkretes Beispiel solch einer strukturorientierten Betrachtungsweise wäre beispielsweise die Frage, ob sich die klassische Unterteilung von Arbeitswerten in extrinsische, intrinsische und soziale Werte bzw. kognitive, affektive und instrumentelle Werte bei allen Kulturen finden läßt oder aber ein Ergebnis ist, welches sich rein in westlichen Kulturen ergibt (siehe Borg, 1986). Studien, die zeigen können, daß sich in der Tat die Struktur von Arbeitswerten in verschiedenen kulturellen Umgebungen gleich gestalten (Borg, 1986; Schwartz, 1999), können jedoch keine Aussagen darüber machen, ob die unterschiedlichen Arbeitswerte auch hinsichtlich des absoluten Levels in verschiedenen Kulturen hinsichtlich ihrer Bedeutsamkeit gleich hoch bewertet werden. Sie sagen lediglich aus, daß es drei verschiedene Arten von Arbeitswerten gibt, die interkulturell generalisierbar sind. Auf

die in dieser Arbeit vorliegende Betrachtung von Arbeitszufriedenheit übertragen bedeutet dies, daß eine strukturorientierte Betrachtungsweise von Arbeitszufriedenheit im interkulturellen Kontext darin bestünde, zu prüfen, ob beispielweise eine Gesamtzufriedenheit bezüglich der eigenen Arbeit sich in verschiedenen Kulturen jeweils aus den gleichen Arbeitszufriedenheitsfacetten erklären läßt (vgl. McFarlin & Rice, 1991; Müller, 2006; Staples & Higgins, 1998; Wanous & Lawler, 1972). Diese strukturorientierte Betrachtung soll hier jedoch nicht verfolgt werden. Hingegen handelt es sich bei den in dieser Arbeit durchgeführten Studien um eine rein ausprägungsorientierte Untersuchung. Geprüft wird somit, ob die gleiche absolute Höhe von Arbeitszufriedenheit, die beispielsweise in zwei Ländern gemessen wurde, auch tatsächlich als gleiches Arbeitszufriedenheitsniveau interpretiert werden kann, oder ob nicht vielmehr bestimmte Faktoren Einfluß darauf ausüben, daß per se die Höhe des Niveaus in verschiedenen Ländern unterschiedlich ist.

Neben der Unterscheidung in ausprägungs- und strukturorientierte Betrachtungsweisen (van de Vijver & Leung, 2000) finden sich in der Literatur auch weitere Bezeichnungen für die gleiche Thematik. So unterscheiden andere Autoren zwischen Typ-I- und Typ-II-Fragestellung (Brett, Tinsley, Janssens, Barsness & Lytle, 1997) oder aber zwischen Gruppenmittelwerts- und Korrelationsanalysen (House, Wright & Aditya, 1997). Diese Unterscheidungen behandeln jedoch die gleiche Thematik und sind somit nur als weitere Bezeichnungen für das gleiche Problemfeld anzusehen.

Ein Vergleich speziell der Ausprägungshöhe von Arbeitszufriedenheit in verschiedenen Ländern und der Einflußfaktoren auf das Arbeitszufriedenheitsniveau liefert neben dem Erkenntnisgewinn innerhalb der internationalen Arbeits- und Organisationsforschung ebenfalls einen Beitrag zur generellen Diskussion der Anwendbarkeit von Meßverfahren und -instrumente in verschiedenen Kulturen. Speziell bei der Durchführung von Mitarbeiterbefragungen im internationalen Kontext ist jedoch bedeutsam, inwieweit die gleichen Instrumente und Fragen zu äquivalenten Ergebnissen kommen. Die Äquivalenz von Fragebögen und anderen Meßverfahren kann dabei im interkulturellen Kontext nicht grundsätzlich angenommen werden (Ryan, Chan, Ployhart & Slade, 1999; Ryan, Horvath, Ployhart, Schmitt & Slade, 2000) und stellt eines der grundlegenden methodischen Probleme in der interkulturellen Forschung dar (van de Vijver & Leung, 1997). Die verschiedenen Formen von Äquivalenz werden dabei sowohl in der Forschung als auch in der Praxis nur selten explizit betrachtet und vor dem Einsatz von Meßverfahren geprüft. Nicht selten werden dabei gerade in

multinationalen Unternehmen Maßnahmen aus den Ergebnissen abgeleitet, die als unangemessen und falsch zu bezeichnen sind, wenn die Datenbasis keine Vergleichbarkeit zuläßt. Sollte nämlich in der Tat keine Äquivalenz bestehen, da die Ausprägungshöhe der Zufriedenheit von Faktoren beeinflusst wird, die nicht auf das zugrundeliegende Konstrukt zurückzuführen sind, würden Unterschiede in der Höhe der Arbeitszufriedenheit lediglich die unterschiedliche Eignung der Zufriedenheitsfrage in verschiedenen Kulturen darlegen (van de Vijver & Leung, 1997 & 2000). Ein Rückschluß auf die tatsächlich empfundene Zufriedenheit der Befragten wäre damit dann nicht mehr möglich. Das Ergebnis der Messung von Arbeitszufriedenheit wäre dann stattdessen ein Indikator für unterschiedliche methodische Einflußgrößen in verschiedenen Kulturen oder aber für das interkulturell unterschiedliche Kommunikationsverhalten der Mitglieder der jeweiligen Kultur. Ob ein Land – bzw. im praxisbezogenen Anwendungsfall ein internationaler Standort eines Unternehmens – im Mittel eine höhere Zufriedenheit aufweist als sein Vergleichspartner, kann dann nicht mehr ohne weiteres aus den Meßergebnissen geschlossen werden.

Die vorliegende Arbeit hat somit insgesamt das Ziel zu prüfen, inwieweit „nicht-inhaltliche“ Einflüsse bei der Messung von Arbeitszufriedenheit im interkulturellen Kontext eine Rolle spielen. Als „nicht-inhaltliche“ Faktoren sollen dabei speziell methodische Aspekte und das kommunikative Verhalten der Befragten im Befragungsprozeß beleuchtet werden. Diese Faktoren werden hinsichtlich ihres Einflusses auf den absoluten Wert der Höhe der Arbeitszufriedenheit untersucht. Die Erkenntnisse dieser Untersuchung helfen sowohl weiteren Forschungsvorhaben im interkulturellen Kontext als auch praxisbezogenen Umsetzungen im Unternehmensalltag, bei denen Informationen aus verschiedenen Nationen und Kulturen zusammengetragen werden und miteinander in Beziehung gesetzt werden sollen.

1.3 Aufbau der Arbeit

Der Aufbau der Arbeit gliedert sich in zwei übergeordnete Teile. Während im ersten Teil die relevanten theoretischen Überlegungen und Modelle diskutiert werden, enthält der zweite Teil die eigentlichen empirischen Untersuchungen.

Zum Verständnis der hier vorliegenden Arbeit sind vor allem drei Konstrukte bedeutsam: a) Arbeitszufriedenheit, b) Kultur und c) Antworttendenzen. Daher wird zu Beginn der Ausführungen im theoretischen Teil zuerst einmal das Konstrukt der

Arbeitszufriedenheit dargestellt (Kapitel 2). Dabei wird einerseits die Bedeutung dieses Konstrukts im Arbeitsleben erläutert, um die hohe Relevanz der in dieser Arbeit bearbeiteten Fragestellung für die unternehmerische Praxis zu betonen (Kapitel 2.1). Des weiteren wird die Thematik der Messung von Arbeitszufriedenheitsurteilen problematisiert (Kapitel 2.2) sowie die Bedeutsamkeit von gemessenen Zufriedenheitsurteilen für das Managen von Organisation erläutert (Kapitel 2.3). Im daran anschließenden Kapitel 3 wird der Kulturbegriff eingeführt. Nach einer knappen Darstellung der bislang in der Forschung prominenten Kulturmodellen (Kapitel 3.1) wird ebenfalls erläutert, inwieweit die Kultur eines Landes als bedeutsamer Sozialisationsfaktor, vor allem für das Kommunikationsverhalten der Menschen, angesehen werden kann (Kapitel 3.2). Während zu Beginn der theoretischen Ausführungen das Konstrukt der Arbeitszufriedenheit in allgemeiner Weise dargestellt wurde, werden im dritten Kapitel die Ausprägungen dieses Konstrukts im interkulturellen Vergleich thematisiert. Vor allem werden die bisherigen empirischen Ergebnisse interkultureller Studien zur Messung von Arbeitszufriedenheit zusammengetragen (Kapitel 4.1). Dabei wird einerseits thematisiert, welche verschiedenen Einflußfaktoren es auf die Höhe der Arbeitszufriedenheit geben könnte, und andererseits problematisiert, daß das kulturelle bedingte Kommunikationsverhalten und die damit verbundenen Antworttendenzen bisher als bedeutsamer Einflußfaktor in der Forschung vernachlässigt wurden (Kapitel 4.2). Da die hier vorliegende Arbeit insbesondere einen methodischen Fokus beinhaltet, werden die dafür notwendigen theoretischen Grundlagen im Kapitel 5 gelegt. Diese Darstellung beinhaltet vor allem die Diskussion der allgemeinen methodischen Probleme der interkulturellen Forschung, die insgesamt die folgenden bedeutsamen Aspekte enthält: a) die Gegenüberstellung von kulturabhängiger und –unabhängiger Forschung und der damit verbundenen Annahmen (Kapitel 5.1), b) die Probleme bei der Erfassung nationaler Einflußfaktoren (Kapitel 5.2), c) die im Vergleich zur rein nationalen Forschung größere Bedeutsamkeit des Forschungsdesigns und der Auswahl geeigneter Stichproben (Kapitel 5.3), d) die Möglichkeit von experimentellen Studien in der interkulturellen Forschung (Kapitel 5.4) sowie e) die allgemeine Frage nach der Äquivalenz interkulturell erhobener Daten (Kapitel 5.5). Alle diese Aspekte stellen Fragen dar, die grundsätzlich bei interkultureller Forschung in jedem inhaltlichen Forschungsgebiet beantwortet werden müssen und daher auch für die Vergleichbarkeit von Arbeitszufriedenheitsurteilen über Ländergrenzen hinweg von besonderer Bedeutung sind. Die in diesen Kapiteln

diskutierten Aspekte dienen als Grundlage für die hier behandelte Fragestellung und die in dieser Arbeit problematisierte Sichtweise, daß Arbeitszufriedenheitsurteile aus verschiedenen Ländern problemlos verglichen werden könnten. Das spezielle methodische Problem interkultureller Forschung, welches in dieser Arbeit genauer beleuchtet werden soll, nämlich der Einfluß von Antworttendenzen wird schließlich in Kapitel 5.6 in detaillierter Weise dargestellt. Dabei werden sowohl die in der Forschung unterschiedenen Formen von Antworttendenzen erläutert (Kapitel 5.6.1) sowie die Interpretation dieses speziellen Antwortverhaltens vorgestellt, so wie sie derzeit in der interkulturellen Forschung in weiten Teilen verstanden wird (Kapitel 5.6.2).

Die erste Zusammenführung von zwei der zentralen Konstrukte dieser Arbeit findet in Kapitel 6 statt, wenn erläutert wird, wie welche Art von Kommunikationsverhalten mit welcher Ausprägung von Kultur zusammenhängt. Dazu werden in einem ersten Schritt allgemein die deskriptiven Ergebnisse dahingehend dargestellt, in welchen Ländern welche Art von Antworttendenz gemessen werden konnten (Kapitel 6.1). Erst wenn im nächsten Kapitel 6.2 die konkrete Ausprägung kultureller Werte mit in Betracht gezogen wird, werden die deskriptiven Ergebnisse um einen erklärenden Faktor für die gemessenen Unterschiede erweitert. Abschließend wird in diesem Kapitel 6 dann noch auf den Unterschied zwischen länderbezogenen und individuellen Einflüssen hinsichtlich des Kommunikationsverhaltens eingegangen (Kapitel 6.3), da auch ein Fokus dieser Arbeit auf der Prüfung von mehrebenen-bedingten Einflüssen liegt.

Die letztendliche Verbindung der drei Konstrukte Arbeitszufriedenheit, Kultur und Antworttendenzen aus theoretischer Sicht wird dann im Kapitel 7 vollzogen. Dabei werden nacheinander die dieser Arbeit zugrunde liegenden Hypothesen zu den Zusammenhängen der drei Konstrukte dargestellt. Während die Erläuterungen zum Zusammenhang von Kultur und Arbeitszufriedenheit (Kapitel 7.1) sowie zum Zusammenhang von Kultur und Antworttendenzen (Kapitel 7.2) in Teilen den Versuch einer Replikation bisheriger Erkenntnisse der Forschung und damit die Voraussetzung für weitere Analysen dieser Arbeit darstellen, werden in den darauf folgenden Kapiteln explizit die Hypothesen zu den in dieser Arbeit neu betrachteten Zusammenhängen von Antworttendenz und Arbeitszufriedenheit (Kapitel 7.3) sowie zum Gesamtzusammenhang der drei Konstrukte (Kapitel 7.4) hergeleitet.

Die eigentlichen empirischen Untersuchungen werden ab Kapitel 8 beschrieben. Dieser Teil der Arbeit läßt sich aufgrund der unterschiedlichen Untersuchungen, die durchgeführt wurden, in insgesamt drei größere Bereiche untergliedern. Ein

bedeutsamer Teil der Analysen bezieht sich dabei auf die Untersuchung bereits veröffentlichter Daten und stellt somit ein sekundäranalytisches Vorgehen dar (Kapitel 9). Dieses Vorgehen bezieht sich zu Anfang auf die Prüfung der Zusammenhänge der drei theoretischen Konstrukte auf dem Länderniveau. Diese werden wiederum sowohl auf ihre jeweils rein bivariaten Zusammenhänge untersucht (Kapitel 9.3), jedoch ebenso anhand von Mediationsanalysen auf ihre Erklärungskraft in einer Dreiecksbeziehung geprüft (Kapitel 9.4). Neben diesen Berechnungen, die ausschließlich auf dem Länderniveau durchgeführt werden, werden jedoch im daran anschließenden Kapitel das Vorgehen und die Ergebnisse der Mehrebenenanalysen dargestellt, die als der nächste empirische Schritt sowohl das Länderniveau als auch das Individualniveau der vorhandenen Datensätze mit einbeziehen (Kapitel 9.5). Sowohl die Analysen zu den bivariaten Zusammenhängen als auch die mediations- und mehrebenenanalytischen Berechnungen stellen dabei reine Sekundäranalysen dar. Die Ergebnisse, die sich dabei ergaben, wurden schließlich in einem dritten Schritt durch ein eigens durchgeführtes Experiment auf ihre Tragfähigkeit hin geprüft, dessen Vorgehen und Ergebnisse im Kapitel 10 der Arbeit erläutert werden.

Im abschließenden Kapitel 11 der Arbeit werden die empirischen Ergebnisse in einer Gesamtübersicht reflektiert sowie etwaige Einschränkungen und weiterer Forschungsbedarf diskutiert.

Theoretischer Teil

2 Arbeitszufriedenheit

2.1 Bedeutung von Arbeitszufriedenheit im Arbeitskontext

Das psychologische Erleben von Mitarbeitern im Arbeitskontext wird in der Forschung im Rahmen der Arbeits- und Organisationspsychologie betrachtet. Dabei beschäftigt sich dieser Bereich der psychologischen Forschung mit einer Vielzahl von unterschiedlichen Themenfeldern; wie z.B. Führung, Personalauswahl, Organisationskultur, Mobbing, etc. – um nur einige zu nennen. Ein zentrales Feld jedoch stellt dabei in der organisationspsychologischen Forschung die Beschäftigung mit der Arbeitszufriedenheit dar. Schon bis in die 70er Jahre ließen sich zu diesem Themengebiet über 3000 Artikel finden (Locke, 1976). Dieses starke Interesse der Forschung an dieser Thematik ließ keineswegs in den Folgejahren nach. Judge et al. (2001) fanden allein in ihrer erneuten ausführlichen Recherche weitere ca. 8000 Artikel für die folgenden Jahrzehnte bis zum Jahr 2000. Diese umfassende Anzahl von Studien zu diesem Thema lassen Judge et al. (2001) dann auch zu dem Schluß gelangen, daß das Thema Arbeitszufriedenheit wohl als das bedeutsamste und wichtigste Themenfeld in der Arbeits- und Organisationspsychologie anzusehen ist. Zu keinem anderen Themenfeld lassen sich solch eine Vielzahl von Untersuchungen finden. Jedoch wird dieses Themenfeld – wie auch viele andere Felder der Organisationspsychologie – in weiten Teilen vom amerikanischen Verständnis und von europäischen Sichtweisen dominiert. Die meisten der durchgeführten Studien stützen sich dabei auf Stichproben, die in diesem kulturellen Umfeld erhoben wurden. Im Kontext einer interkulturellen Betrachtungsweise stellt sich somit direkt die Frage der Generalisierbarkeit. Lassen sich gefundene Erkenntnisse zur Arbeitszufriedenheit tatsächlich auch auf nicht-westliche Kulturen übertragen?

Bevor in weiteren Kapiteln dieser Arbeit speziell auf die bisherigen interkulturellen Studien zur Arbeitszufriedenheit eingegangen wird, soll in einem ersten Schritt hier erst einmal deutlich gemacht werden, was unter diesem Konstrukt in der Forschungslandschaft verstanden wird. Dabei scheint ein sehr vielfältiges Verständnis zu existieren. Neuberger (1974) listet bereits eine Vielzahl von Definitionen auf, die den

Schluß nahe legen, jeder Forscher habe sein eigenes Verständnis von dem zu untersuchenden Gegenstand. Insgesamt haben sich verschiedene Ansätze entwickelt, die Arbeitszufriedenheit einerseits mehr als kognitives oder andererseits mehr als emotionales und affektives Phänomen verstehen. Darüber hinaus finden sich jedoch auch ebenso beide Formen des Verständnisses in vielen Definitionen. So versteht Locke (1976) Arbeitszufriedenheit als „a pleasurable or positive emotional state resulting from the appraisal of one’s job or job experience“ (S. 1304) und bezieht sich somit sowohl auf die kognitive Bewertung als auch auf das affektive Empfinden. Ein ähnliches Verständnis findet sich bei Cranny, Smith und Stone (1992), die ebenfalls darstellen, daß Arbeitszufriedenheit eine affektive Reaktion auf einen kognitiven Vergleich von Erwartungen und tatsächlichen Resultaten am Arbeitsplatz darstellt. Sie schlußfolgern in ihrer Darstellung, in der sie ebenfalls verschiedene Definitionsansätze vergleichen, daß dies ein Konsens in der Definitionsfrage zur Arbeitszufriedenheit sei. Jedoch existiert an anderer Stelle das Verständnis, daß es sich bei Arbeitszufriedenheit um eine Einstellung handelt (siehe z.B. Weiss, 2002). Demnach sei Arbeitszufriedenheit „an attitude towards one’s job“ (Brief, 1998, S. 10) bzw. „generally equivalent to job attitudes“ (Miner, 1992, S. 116). Weiss (2002) kommt in seiner Darstellung, in der er kognitive und affektive Perspektiven zur Definition von Arbeitszufriedenheit gegenüber stellt, schließlich zu dem Schluß, daß „thus, job satisfaction is an attitude [... and] is a positive (or negative) evaluative judgement one makes about one’s job or job situation“ (S. 175). Diesem Verständnis folgend, jedoch ebenfalls die affektive Komponente nicht vernachlässigend, gehen Wegge und van Dick (2006) davon aus, daß Arbeitszufriedenheit „gewöhnlich als Einstellung aufgefaßt [wird], die sowohl kognitive als auch emotionale Komponenten umfaßt“ (S. 13). Demnach könne somit davon ausgegangen werden, daß in weiten Teilen der Forschung Arbeitszufriedenheit klassischer Weise als eine Einstellung gemäß dieser Art betrachtet wird (Six & Kleinbeck, 1989). Ein weiteres prominentes Modell, welches dem kognitiven Ansatz folgt, findet sich im deutschsprachigen Raum im prominenten Zürcher Modell der Arbeitszufriedenheit (Baumgartner & Udriș, 2006; Bruggemann, Groskurth & Ulich, 1975). Insgesamt kann somit festgehalten werden, daß es durchaus verschiedene Ansichten gibt, in welcher Form genau Arbeitszufriedenheit verstanden werden kann. Es herrscht in weiten Teilen der Forschung jedoch Einigkeit darüber, daß Arbeitszufriedenheit als Einstellungskonstrukt verstanden werden sollte, welches das

Endprodukt eines Bewertungsprozesses ist, der sowohl kognitive als auch emotionale Aspekte umfaßt.

Betrachtet man neben den Definitionsversuchen die verschiedenen theoretischen Ansätze, die sich zum Verständnis von Arbeitszufriedenheit entwickelt haben, dann zeigt sich auch dort je nach Theorie ein etwas unterschiedliches Verständnis. Insgesamt stellen viele der Theorien jedoch auch die kognitiven Aspekte in der Entstehung von Arbeitszufriedenheit in den Vordergrund. Innerhalb dieser eher kognitiven Theoriengruppe läßt sich nach Judge et al. (2001) insgesamt eine Dreiteilung vornehmen: a) situationale Theorien, b) dispositionale Theorien und c) interaktive Theorien. Situationale Theorien versuchen dabei die Arbeitszufriedenheit aus externen Aspekten in der Umwelt und somit außerhalb des Individuums zu erklären. Nach ihrer Einteilung fallen darunter vor allem die Zwei-Faktoren-Theorie von Herzberg (1967) und das Job Characteristics Model von Hackman und Oldham (1976). Dispositionale Theorien sehen die Ursache von Arbeitszufriedenheit dagegen eher in der jeweiligen Persönlichkeit des Mitarbeiters bedingt. Unabhängig von äußeren Umständen spiegelt die Äußerung zur Zufriedenheit mit der Arbeit somit eher die jeweilige Persönlichkeitsstruktur des Befragten wider. Interaktive Theorien bilden nach Judge et al. (2001) letztendlich dann das Zusammenspiel aus beidem ab, der sozialen und situativen Umwelt und der Persönlichkeit.

In ähnlicher Form unterscheiden Weiss und Cropanzano (1996) ebenfalls drei Arten theoretischer Ansätze voneinander: a) dispositionale Ansätze, b) kognitive Ansätze und c) Theorien des sozialen Einflusses. Während sich die erste Kategorie der dispositionalen Ansätze mit der Einteilung nach Judge et al. (2001) deckt, unterscheiden sich jedoch die anderen beiden Gruppen geringfügig. Kognitive Ansätze legen dabei den Fokus auf einen Vergleichsprozeß, in dem die Aspekte der Umwelt, so wie sie von dem Individuum wahrgenommen werden, mit einem persönlichen Soll-Standard verglichen werden. Die Arbeitszufriedenheit ergibt sich dann zumeist aus dem jeweiligen Verhältnis von wahrgenommenem Ist-Zustand und dem zu erreichenden Soll-Niveau. Entscheidend ist somit die Diskrepanz zwischen den eigenen Erwartungen und der Erfüllung dieser Erwartungen (Lawler, 1973). Dabei wird davon ausgegangen, daß das kognitive Urteil relativ stabil ist und somit auch das Arbeitszufriedenheitsurteil nicht oder kaum von externen momentanen Aspekten beeinflusst wird (Weiss & Cropanzano, 1996). Es mache demnach keinen Unterschied, zu welchem Zeitpunkt genau ein Urteil über die Zufriedenheit mit der Arbeit abgefragt wird, denn diese

Aussage greife dann auf das bereits gebildete Urteil über die Soll- und Ist-Aspekte der Arbeitssituation zurück. Dieses liegt somit sozusagen im kognitiven Speicher auf Abruf bereit. Im Gegensatz dazu geht die dritte Theoriengruppe nach Weiss und Cropanzano (1996) – die Theorien des Sozialen Einflusses – von einer etwas anderen Sichtweise aus. Als prominentestes Beispiel ist dabei die Social Information Processing Theory (SIP) (Salancik & Pfeffer, 1977, 1978) zu nennen. Dieser Ansatz geht davon aus, daß Informationen aus der Umwelt wichtig sind für das Urteil zur Arbeitszufriedenheit und die soziale Umwelt in zweierlei Weise Einfluß darauf nimmt. Einerseits gibt es zwar auch einen direkten Einfluß – in gleicher Weise wie bei den kognitiven Theorien – in Form der Wahrnehmung der Arbeitsumgebung (Adler, Skov & Salvemini, 1985), andererseits jedoch existieren ebenfalls indirekte Einflüsse auf das Zufriedenheitsurteil (Weiss & Shaw, 1979). Diese indirekten Einflüsse bestehen aus Informationen aus der sozialen Umwelt, auf die sich Befragte bei ihrer Urteilsfindung beziehen. So beziehen sie Hinweise von Kollegen und die Interpretation ihres eigenen Verhaltens mit ein und sogar die Art, wie die Zufriedenheitsfrage gestellt wird, kann ebenfalls Einfluß auf das Zufriedenheitsurteil haben (Judge et al., 2001). Nach Ansicht dieses Ansatzes wird ein Zufriedenheitsurteil nicht nach umfangreichen Abwägungen über das Arbeitsumfeld gebildet, sondern stattdessen vielmehr in jeder neuen Befragungssituation ad hoc und auf Anforderung generiert (Judge et al., 2001; Weiss & Cropanzano, 1996). Die Einstellung zur Arbeit wird sozusagen erst gebildet, wenn danach konkret gefragt wird. Dabei ist dieses Urteil dann den vielfältigen sozialen Einflüssen ausgesetzt, die in der jeweiligen Befragungssituation vorhanden sind. Im Gegensatz zu den kognitiven Erklärungsansätzen kann somit bei diesem Ansatz eher nicht davon ausgegangen werden, daß es sich bei dem Urteil zur Arbeitszufriedenheit immer um eine stabile Einstellung im klassischen Sinne handelt. Die jeweilige Situation des sozialen Befragungsumfeldes kann hierbei als eine Art „Störeinfluß“ auf das Zufriedenheitsurteil mehr oder weniger stark einwirken. Neben den bereits genannten sozialen Informationen, die einwirken können, fallen darunter zum Beispiel auch die Art der Befragungssituation, der Aufforderungscharakter einer Befragung und somit im allgemeinen die kommunikative Struktur der jeweiligen Situation. Insgesamt wird in diesem Ansatz somit ein Arbeitszufriedenheitsurteil vielmehr als sozial konstruierte Realität anstatt als kognitiver Ist-Soll-Vergleich angesehen (Salancik & Pfeffer, 1977). Während einzelne Autoren von einem Rückgang des Interesses an dieser Theorie berichten (vgl. Judge et al., 2001), beenden Weiss und Cropanzano (1996) ihren

Theorievergleich hingegen mit der Schlußfolgerung, daß vielmehr inzwischen davon ausgegangen wird, daß Einstellungen nicht grundsätzlich abgespeichert sind und dann abgerufen werden, sondern im Gegensatz dazu oft erst spontan auf Nachfrage hin gebildet werden (siehe auch Wilson & Hodges, 1992). Die in der akuten Situation gebildeten Arbeitszufriedenheitsurteile werden dann sowohl von im Gedächtnis abgelegten Informationen über die Arbeitssituation als auch von kontextuellen und sozialen Informationen aus der jeweiligen Befragungssituation beeinflusst. Ein Arbeitszufriedenheitsurteil würde somit auch erst in der Befragungssituation abschließend gebildet werden und lediglich von im Gedächtnis abgespeicherten Inhalten mit beeinflusst werden. Folglich handelt es sich bei diesem theoretischen Ansatz eher um eine situative Sichtweise, so daß die SIP-Theorie folglich auch bei Judge et al. (2001) zu den situativen Theorien gezählt wird.

Gerade für die hier in dieser Arbeit behandelte Fragestellung nimmt die Argumentation der SIP-Theorie eine besondere Rolle ein. Die „nicht-inhaltlichen“ Einflüsse auf das Arbeitszufriedenheitsurteil, die in der Arbeit untersucht werden, lassen sich genau in diese Theorie einordnen. Da Mitarbeiter, die nach der Zufriedenheit mit ihrem Arbeitsplatz gefragt werden, bei diesem Urteil gemäß dieser Theorie sowohl durch objektive Kriterien als auch durch kontextuelle Bedingungen beeinflusst sind, ließe sich somit theoretisch begründen, warum allein aus der kommunikativen Sicht der Befragungssituation sich die Urteile der Personen aus verschiedenen Kulturen unterscheiden könnten. Kulturell bedingte Antworttendenzen würden folglich genau solch einen kontextuellen Faktor darstellen.

Während die Theorien zur Arbeitszufriedenheit versuchen zu erklären, wie ein Urteil überhaupt zustande kommt, lassen sie hingegen keine nachhaltigen Schlußfolgerungen zur spezifischen Messung der Zufriedenheit in der konkreten Erhebungssituation zu. Sie versuchen lediglich die kausalen Ursachen des Zufriedenheitsurteils konzeptuell genauer zu erfassen. Lediglich die SIP-Theorie (Salancik & Pfeffer, 1977, 1978) rückt die Erhebungssituation durch die Betonung der sozialen Einflüsse auf die spontan gebildeten Urteile stärker in den Fokus. Daher muß diskutiert werden, wie in einer Erhebungssituation überhaupt ein valides Maß zur Arbeitszufriedenheit erhoben werden kann. Wie Arbeitszufriedenheit gewöhnlich gemessen wird und inwieweit es generelle Schwierigkeiten bei der Erhebung des Zufriedenheitsurteils geben mag, soll im nächsten Abschnitt erläutert werden.

2.2 Messung von Arbeitszufriedenheit

Bei der Betrachtung der Frage nach der bestmöglichen Messung von Arbeitszufriedenheit gibt es verschiedene Ansätze, die für gewöhnlich angewendet werden. Während Judge et al. (2001) nach ihrer ausführlichen Literatur-Recherche zu dem Urteil kommen, daß von den meisten Forschern Arbeitszufriedenheit als ein globales Konstrukt mit verschiedenen Facetten angesehen wird und auch so gemessen werden sollte, konstatieren Fischer und Belschak (2006), daß in den meisten tatsächlichen Untersuchungen lediglich eine simple Frage gemäß „Wie zufrieden sind Sie mit Ihrer Arbeit?“ gestellt wird. Diese beiden Ansätze stellen auf den ersten Blick einen Widerspruch zwischen der theoretischen Konzeption von Arbeitszufriedenheit und der Umsetzung in der jeweiligen Befragungssituation dar. Jedoch widersprechen sich diese vermeintlich gegensätzlichen Aussagen nicht zwingend, da es letztendlich auf das spezifische Meßmodell ankommt, welches zum Verständnis und zur Erfassung der Arbeitszufriedenheit herangezogen wird. Der Auffassung von Judge et al. (2001) folgend wenden in der Tat viele Verfahren eine facetten-orientierte Messung des Konstruktes an. Zu den bekanntesten dieser Verfahren zählen der ‚Job Descriptive Index‘ (JDI; Smith, Kendall & Hulin, 1969) und der ‚Minnesota Satisfaction Questionnaire‘ (MSQ; Weiss, Davis, England & Lofquist, 1967). Dabei werden gewöhnlich folgende Facetten unterschieden: Bezahlung, Aufstieg/Weiterentwicklung, Kollegen, Führung/Leitung und die Arbeit selbst (Judge et al. 2001); so auch im JDI (vgl. Smith et al. 1969). Zum Teil werden von anderen Autoren weitere Facetten ergänzt, z.B. Arbeitsbedingungen, Anerkennung, das Unternehmen bzw. das Management (Locke, 1976). Jedoch ist keinesfalls eindeutig, wie diese einzelnen Facetten zu einem Gesamturteil zusammengefaßt werden können. In der Literatur werden verschiedene Ansätze diskutiert, inwieweit ein Gesamtmaß der Arbeitszufriedenheit als zusammengesetzt aus Facetten oder als Globalurteil zu messen ist. Judge et al. (2001) schlagen sich dabei eindeutig auf die Seite der Zufriedenheitsmessung anhand von Facetten, die dann für gewöhnlich zu einem Gesamtmaß aufsummiert werden. Zwar gestehen sie auch zu, daß es kritischere Position zur Facettenmessung gäbe (vgl. Scarpello & Campbell, 1983), ziehen jedoch ihre eigenen Untersuchungen als Gegenbeispiel heran (vgl. Judge & Hulin, 1993), um gemäß ihrer Überzeugung zu zeigen, daß ein aufsummiertes Facettenurteil sehr wohl

ein valides Maß für eine Gesamtzufriedenheit darstellen kann. Jedoch wird an anderer Stelle im Gegensatz dazu kritisch diskutiert, ob solch ein Aufsummieren einzelner Aspekte überhaupt zulässig ist (Law & Wong, 1999; Law, Wong & Mobley, 1998; Weiss, 2002). Dieses Aufsummieren stellt nämlich letztendlich eine reine lineare Kombination des Gesamturteils aus den Einzelaussagen dar. Somit wird das Gesamtkonstrukt erst durch die Aggregation dieser einzelnen Facetten bestimmt. Unterschiedliche Zusammenstellungen von Facetten würden somit auch unterschiedliche Gesamturteile der Arbeitszufriedenheit zur Folge haben. Fehlt eine bedeutsame Facette in der Messung, würde sich folglich auch das Gesamtkonstrukt anders gestalten. Statt dieser linearen Kombination von Einzelfacetten befürworten diese Autoren eher das Verständnis eines Arbeitszufriedenheitsurteils als latentes Konstrukt, welches lediglich durch die einzelnen gemessenen Aspekte möglichst gut abgebildet werden soll (Law & Wong, 1999; Law, Wong & Mobley, 1998; Weiss, 2002). Im ersteren Fall der linearen Aufsummierung ist es sehr entscheidend, welche Facetten letztendlich gemessen werden. Beim letzteren Fall hingegen, wo Arbeitszufriedenheit als latentes Konstrukt hinter den eigentlichen Meßwerten der Facetten liegt, ist dies weniger bedeutsam, da die gemessenen Aspekte bloß eine mehr oder weniger gute statistische Manifestation des dahinter liegenden Konstrukts darstellen müssen. Bei der Aggregation und Addition einzelner gemessener Zufriedenheitsurteile von Arbeitsfacetten hingegen kann sich der Gesamtwert der Arbeitszufriedenheit sehr stark ändern, je nachdem welche weitere Facette mit aufgenommen wird oder nicht. Dieser Punkt läßt die facetten-orientierte Meßtechnik somit eher in einem problematischen Licht erscheinen.

Darüber hinaus besteht in der Forschung auch bisher Unklarheit darüber, inwieweit es einen eindeutigen und klaren Zusammenhang zwischen der Arbeitszufriedenheit gemessen an Facetten und einer gesondert gemessenen Gesamtarbeitszufriedenheit gibt. Weiss (2002) geht in seinen Darstellungen davon aus, daß es keinen klaren Bezug zwischen Facettenurteilen und einem Gesamturteil gäbe. Hinsichtlich dieser unklaren Ausgangslage geht er sogar so weit zu sagen: „there are no „facets of job satisfaction“; there are only discriminable objects in the work environment that we evaluate” (S. 186). Diese unterschiedlichen Objekte der Arbeitsumgebung werden nach Weiss (2002) herangezogen, um sich ein Urteil über die Arbeitssituation zu bilden. Demnach können auch viele weitere denkbare Aspekte als Urteilsobjekt dienen. Neben dem Arbeitsplatz und der Bezahlung könnte sogar „the toupee of your boss“ (Weiss, 2002, S. 186) die

Urteilsbasis bilden. Eine klare Hierarchie, welche Aspekte als besonders wichtig anzusehen sind und somit in ein Meßinstrument gemäß dem Facettenansatz integriert werden müßten, ließe sich demnach nur schwer finden. Ein Gesamtzufriedenheitsurteil scheint folglich insgesamt etwas anderes zu sein als eine reine Aufsummierung einzelner Urteile seitens des Individuums. Wie solche Einzelurteile in ein Gesamturteil überführt werden bzw. welchen Einfluß sie auf das Gesamturteil haben, dazu könne gemäß Weiss (2002) die bisherige Forschung keine genauere Aussage machen.

Eine etwas andere Sichtweise vertritt jedoch Borg (2003), wenn er in seiner Sandwich-Theorie versucht zu erläutern, wie ein Gesamturteil die Urteile zu einzelnen Facetten beeinflussen kann. Er geht somit davon aus, daß es sehr wohl einen Zusammenhang zwischen den Einzelurteilen und einem Gesamturteil gäbe. Dieser Zusammenhang stellt sich jedoch eher als eine Art top-down Beeinflussung dar, während zentral für die Überlegungen von Weiss (2002) eher die Problematik einer bottom-up-Verrechnung ist. Letztendlich stellen jedoch auch seine Überlegungen eine eher kritische Position zur Anwendung der facetten-orientierten Messung von Arbeitszufriedenheit dar. Seine Theorie befaßt sich im Spezifischen mit der Brauchbarkeit der gemessenen Zufriedenheitswerte von Facetten. Sie beschreibt die Problematik, daß Urteile zu mehreren einzelnen Objekten und Facetten der Arbeitsumgebung nicht zwingend völlig einzeln und unabhängig voneinander stattfinden. Die jeweiligen Zufriedenheitswerte hängen gemäß Borg (2003) vielmehr stark miteinander zusammen. Statt der Unabhängigkeit einzelner Teilurteile geht die Theorie von einem sogenannten g-Faktor aus, der auf jedes Urteil zu einzelnen spezielleren Aspekten der Arbeit Einfluß ausübt. Dieser g-Faktor, was gleichbedeutend mit dem Gesamtzufriedenheitsurteil ist, schlägt somit auf jedes einzelne Teilurteil durch. Dies führt unter anderem dann auch dazu, daß die jeweiligen Facetten hoch miteinander korreliert sind und das jeweils abgegebene Urteil eben keinesfalls unabhängig von den anderen Urteilen zu bewerten ist; ein Ergebnis welches sich in der Tat empirisch feststellen läßt (z.B. zur Problematik beim JDI siehe Kinicki, McKee-Ryan, Schriesheim & Carson, 2002). Personen, die nach ihrem Urteil zu spezifischen Aspekten des Arbeitslebens gefragt werden, bilden somit nach Borg (2003) nicht komplett unabhängige Urteile über einzelne Aspekte, sondern haben zumeist eine allgemeine Einschätzung über ihre Arbeitsumgebung im Kopf und beurteilen vor dem Hintergrund dieses Urteils dann erst die jeweiligen einzelnen Facetten. Der Zufriedenheitswert der Facette besteht somit zwar einerseits aus einem Urteil über diese Facette, jedoch ist er andererseits ebenso von dem

Gesamtzufriedenheitsurteil im Sinne eines Halo-Effektes überlagert. Bei der empirischen Messung sollte man nach Borg (2003) dann auch aus den Einzelurteilen jeweils den Einfluß dieses Gesamturteils bzw. g-Faktors herauspartialisieren. Um dies zu ermöglichen sollte man daher die Urteile zu den Einzelaspekten und die Gesamtzufriedenheit getrennt voneinander erfassen, um deren gegenseitige Beziehung im Anschluß berücksichtigen zu können. Im Gegensatz zur facettenorientierten Messung müßte demnach hier das Gesamturteil folglich auf jeden Fall gesondert erhoben werden. Meßtechnisch gesehen ist das Gesamturteil nach Borg (2003) mehr bzw. etwas anderes als lediglich die Summe der Einzelurteile. Es überstrahlt die Einzelurteile und verwischt sozusagen deren genaue Bedeutung. Borg (2003) weist jedoch ebenso darauf hin, daß die Einzelurteile, die für das Gesamturteil korrigiert wurden, nicht grundsätzlich „richtiger“ als die unkorrigierten Werte sind. Sie sind lediglich vom Hintergrundrauschen befreit und somit unter Umständen besser für die Organisationsdiagnose und die Entwicklung von Maßnahmen als Folge einer Zufriedenheitsbefragung brauchbar. Dennoch zeigen diese Überlegungen, daß eine Facetten-Messung von Zufriedenheit nicht unbedingt frei von Problemen ist und der aufsummierte Gesamtwert der Facetten nicht immer gleichzusetzen ist mit einer gesonderten allgemeinen Gesamtfrage zur Arbeitszufriedenheit.

Eine Messung anhand eines Globalurteils gemäß Fischer und Belschak (2006) muß somit nicht grundsätzlich als unterlegen oder schlechter angesehen werden als eine dezidierte Messung anhand von Einzelaspekten bzw. Facetten, die dann wiederum aufsummiert werden müssen. Bei der Messung und Frage nach einem Gesamturteil fällt diese lineare mathematische Bildung des Gesamtwerts weg. Vielmehr wird die Bildung des Gesamturteils wieder auf den Befragten zurück delegiert, der dann gemäß seiner eigenen Situation alle verschiedenen Urteilsobjekte in den Gesamtwert irgendwie einfließen läßt. Zwar ist diese Vorgehensweise durchaus subjektiv geprägt, jedoch kann ja gerade die bisherige Forschung nicht zeigen, daß es einen klaren und eindeutigen Zusammenhang zwischen den jeweiligen Urteilsobjekten oder auch Facetten und einer globalen Arbeitszufriedenheit gibt. Eine Messung anhand standardisierter Verfahren wie dem JDI oder MSQ ist somit nicht zwingend besser. Für manche Studien mag somit auch die einfache Abfrage nach einem Globalurteil zur Arbeitszufriedenheit für die jeweiligen Analysen völlig ausreichen.

Die bisher diskutierten Überlegungen betreffen die Probleme, die sich ganz allgemein beim Messen von Arbeitszufriedenheit stellen. Ein viel bedeutsamerer Aspekt, der vor

allem im Rahmen dieser Arbeit von großer Bedeutung ist, jedoch durch die bisherigen Überlegungen nicht beantwortet werden kann, ist die Frage, inwieweit man überhaupt in der Lage ist, brauchbare Zufriedenheitswerte zu messen. Messen wir wirklich „Objektive Arbeitszufriedenheit? Oder: Was messen wir, wenn wir nach der Zufriedenheit mit der Arbeit fragen, (Fischer & Belschak, 2006, S. 80). Diese Frage ist vor allem auch zentral für die vorliegende Studie. Sie wird jedoch in weiten Feldern der Arbeitszufriedenheitsmessung überhaupt nicht diskutiert. Auch beim Einsatz der bekannten Meßverfahren (wie z.B. JDI oder MSQ, sowie im Deutschen die SAZ - ‚Skala zur Messung der Arbeitszufriedenheit‘ - von Fischer & Lück, 1972) gehen Anwender sicherlich oftmals davon aus, objektive Zufriedenheitswerte durch die Anwendung der jeweiligen Skala zu messen. Diese oft nicht hinterfragte Annahme ist jedoch keinesfalls selbstverständlich. Dies wird umso offensichtlicher, wenn man sich das Entstehen eines Arbeitszufriedenheitsurteils als semantischen Prozeß vorstellt (Fischer, 1989; 1991), bei dem aus der objektiven Arbeitssituation nach mehreren Stufen am Ende erst ein Wert auf einer Arbeitszufriedenheitsskala erscheint (vgl. Abbildung 1). Innerhalb dieses Prozesses gibt es zahlreiche Einflußgrößen, die es fraglich erscheinen lassen, inwieweit der Wert auf der Skala eine direkte Repräsentation der objektiven Situation im Arbeitsumfeld darstellt (Fischer & Belschak, 2006).

Semantische Stufen der Abbildung von Arbeitszufriedenheit

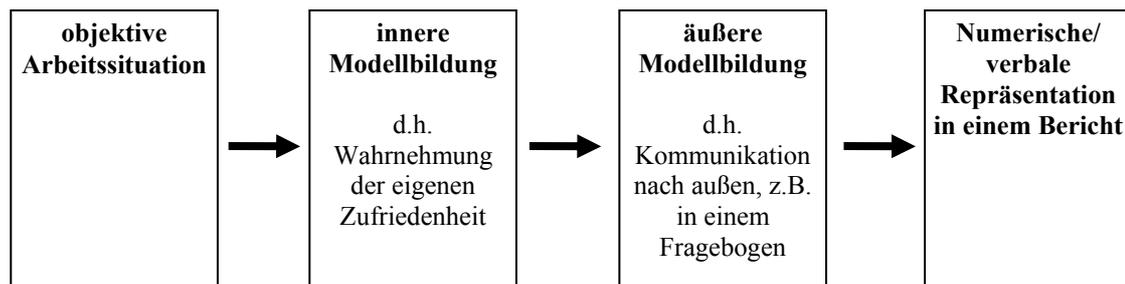


Abbildung 1: Das Modell des semantischen Prozesses des Arbeitszufriedenheitsurteils (gemäß Fischer & Belschak, 2006)

Das Modell des semantischen Prozesses des Arbeitszufriedenheitsurteils (siehe Abbildung 1) unterscheidet dabei zwischen drei verschiedenen Stufen der Urteilsbildung: a) die innere Modellbildung, b) die äußere Modellbildung und c) die numerische/verbale Repräsentation. Während sich der Teil der inneren Modellbildung

mit der Frage beschäftigt, inwieweit ein Individuum überhaupt in der Lage ist, Zugang zu seinen Empfindungen und Zufriedenheitsurteilen zu bekommen, beschreibt die äußere Modellbildung die Phase, in der das innere Urteil in ein von außen wahrnehmbares Urteil überführt wird, z.B. durch die Mitteilung in einem Fragebogen oder einem Interview. Die dritte Phase befaßt sich dann schließlich mit dem Bericht an weitere Dritte, z.B. in einem Bericht einer Mitarbeiterbefragung, und der Darstellung der gemessenen Werte seitens des Forschers. In jeder dieser Phasen gibt es unterschiedliche Einflußgrößen, die auf den Prozeß von der objektiven Situation zur numerischen Repräsentation einwirken.

Da für die vorliegende Untersuchung vor allem die Phase 2 der Urteilsbildung von Bedeutung ist, sollen die anderen Phasen nur kurz hinsichtlich ihrer bedeutsamsten Einflußfaktoren erläutert werden. In Phase 1 – der inneren Modellbildung – stellt sich vor allem erst einmal die Frage, wie das Individuum seine zurückliegenden Erfahrungen zu einem Gesamtzufriedenheitswert hinsichtlich der Arbeitsaspekte für gewöhnlich integriert. Fischer und Belschak (2006) merken dabei Zweifel an, inwieweit tatsächlich von einer Gesamtbetrachtung und Aufsummierung aller relevanten vergangenen Erfahrungen seitens des Individuums ausgegangen werden kann und inwieweit somit ein „Integral von Intensität und Zeit einzelner Empfindungen“ (S. 88) die passende Repräsentation zur Verrechnung zurückliegender Zufriedenheitserfahrungen darstellt. Stattdessen werden vielfach nicht alle Erfahrungen von Individuen gleichberechtigt zusammenaddiert, sondern es spielen Prozesse der Kontrastbildung und der Gewöhnung sowie möglicherweise Endowment-Effekte bei der inneren Modellbildung des Zufriedenheitsurteils eine Rolle. Bekannte Effekte in dieser Hinsicht sind beispielsweise, daß Lottogewinner nach einer gewissen Zeitspanne nicht mehr grundsätzlich zufriedener sind als Personen, die mit sehr negativen Ereignissen (z.B. Querschnittslähmung) konfrontiert waren (siehe z.B. Brickman, Coates & Janoff-Bulman, 1978). Individuen passen sich stattdessen nach einer gewissen Zeit an objektive Begleitumstände an und ändern somit ihr Anspruchsniveau, welches dann wiederum Einfluß auf das entsprechende Zufriedenheitsurteil nimmt. In gleicher Weise besagt auch das sogenannte Easterlin-Paradoxon (Easterlin, 1974), daß beispielsweise zunehmender Wohlstand nicht automatisch zu größerer Zufriedenheit führt; auch hier ist der Grund in Anpassungseffekten in der inneren Modellbildung der Individuen zu sehen. Zusammengefaßt werden solche Effekte oft als „hedonistische Tretmühle“ (Brickman & Campbell, 1971) bezeichnet. Neben diesen klassischen

Gewöhnungseffekten spielen ebenso Kontrasteffekte eine Rolle, wenn Individuen sich bei der Urteilsbildung anhand repräsentativer Momente (Kahneman, 1999) vor allem auf die jeweilige Veränderung der Situation stützen. Auch dabei werden eben nicht alle vorhandenen Informationen gleichermaßen in das Gesamturteil mit einbezogen, sondern vielmehr nimmt allein die Veränderung eines über längere Zeit stabilen Zustands eine umso bedeutsamere Rolle ein. Die Vernachlässigung zahlreicher vergangener Informationen wird noch deutlicher bei der Betrachtung der sogenannten Peak-End-Regel, die oft als valide Verrechnungsformel für die Bildung von Gesamturteilen über eine konkrete Zeitspanne von Erfahrungen fungieren kann (siehe Fischer & Belschak, 2006; Kahneman, 1999). Demnach sind scheinbar nur zwei bestimmte Informationen aus den zahlreichen vergangenen Erfahrungen notwendig und hilfreich, um ein Gesamturteil relativ gut vorherzusagen. Einerseits wird die extremste affektive Empfindung (Peak) und andererseits die letzte affektive Empfindung (End) kurz vor Ende des zu beurteilenden Zeitraums vom Individuum zu einem Gesamtwert ‚verrechnet‘. Alle anderen Empfindungen sowie vor allem die Dauer und Ausdehnung der gesamten Zeitspanne scheinen überhaupt keine bedeutsame Rolle zu spielen.

Neben diesen anpassungs- und erinnerungsbezogenen Problemaspekten bei der Messung von Arbeitszufriedenheit lassen sich auch Fragen hinsichtlich der Gültigkeit eines allgemeinen bipolaren Gut-Schlecht-Kontinuums sowie überhaupt der allgemeinen Einschätzbarkeit zukünftiger Affekte stellen. Jedoch sollen diese Punkte hier nicht näher dargestellt werden (für eine ausführliche Darstellung siehe Fischer & Belschak, 2006). Auch diese Aspekte problematisieren jedoch die Annahme, daß Individuen eine Einschätzung über ihre Zufriedenheit auf einem Kontinuum von ‚sehr zufrieden‘ bis ‚überhaupt nicht zufrieden‘ ohne Schwierigkeiten abgeben könnten. Trotz dieser Einwände und Probleme werden dennoch in der Arbeitszufriedenheitsforschung zumeist Meßverfahren angewendet, welche retrospektiv die Zufriedenheit im Sinne einer ‚erinnerten Nützlichkeit‘ (Kahnemann, 1999) bzw. ggf. sogar einer Augenblicksempfindung (Fischer & Belschak, 2006) messen; d.h. die Individuen geben lediglich in einer konkreten Befragungssituation ihre Zufriedenheit an, die jedoch dann eher eine Art Versuch des zeitlichen Rückblicks darstellt. Die bisherige Forschung läßt jedoch kein abschließendes Urteil zu, inwieweit solch eine Form der Messung einen validen Wert des tatsächlichen Zufriedenheitsgefühls darstellt. Andererseits existieren auch vielfach keine besseren Meßverfahren, so daß oft diese Form der Messung die gängige Praxis bei Arbeitszufriedenheitsuntersuchungen darstellt.

Über die Prozesse der inneren Modellbildung hinaus, die sich rein auf die kognitive und affektive Welt der jeweiligen Individuen bezieht, treten gerade in Phase 2 – der äußeren Modellbildung – weitere Faktoren der zwischenmenschlichen Kommunikation in den Vordergrund, die besonders für die hier vorliegende Untersuchung von Bedeutung sind. Die Phase 2 befaßt sich damit, wie das vom Individuum in Phase 1 gebildete Urteil an ein jeweiliges Gegenüber – oftmals der Forscher – kommuniziert wird. Somit sind hier vor allem Aspekte bedeutsam, die in der Befragungssituation wirken. Fischer und Belschak (2006) erwähnen hierbei unter anderem die Bedeutung von Konversationsnormen gemäß Grice (1975). Innerhalb einer Kommunikationssituation wird gemäß diesen Normvorstellungen automatisch von dem Sender der Informationen abgeschätzt, inwieweit eine Information Neuigkeitswert für den Empfänger hat. Jeder der beteiligten Kommunikationspartner setzt nach Grice (1975) voraus, daß man maximal informativ sein sollte. Die Folgen dieser Annahme zeigen sich dann unter anderem in der Abhängigkeit von Befragungsergebnissen von der Reihenfolge der Fragen. Es können somit unterschiedliche Antwortwerte gemessen werden, je nachdem ob eine Frage vor oder nach einer weiteren bedeutsamen Frage gestellt wird. Beispielsweise würde die eine Frage zur Zufriedenheit mit einem konkreten Arbeitsplatzmerkmal zu unterschiedlichen Ergebnissen führen, je nachdem, ob sie vor oder nach einer allgemeinen Arbeitszufriedenheitsfrage gestellt wurde. Ein Grund dafür kann darin gesehen werden, daß die Antwort auf eine übergreifende Frage nach einer konkreteren Teilfrage gemäß solchen Konversationsnormen für die Befragten keinen Neuigkeitswert mehr beinhaltet.

Ein weiterer Einflußfaktor, der in einer Befragungssituation wirksam sein kann, ist die Stimmung bzw. die affektive Lage der befragten Person in der konkreten Situation (Fischer & Belschak, 2006). Der augenblickliche affektive Zustand des befragten Individuums mag dann dabei einen größeren Einfluß auf die berichtete Zufriedenheit haben als das eigentlich wahre Zufriedenheitsurteil, welches das Resultat der Phase 1 sein sollte. In der Tat können empirische Untersuchungen (Schwarz & Clore, 1983; Schwarz & Strack, 1991) diesen Effekt nachweisen.

Ein ebenso bedeutsamer Aspekt, der von Fischer und Belschak (2006) jedoch nicht explizit erwähnt wird und der aber für die hier vorliegende Arbeit eine große Bedeutung hat, ist der Einfluß des Antwortverhaltens von Befragten im Sinne von Antworttendenzen. In der Forschung werden dabei klassischer Weise drei verschiedene Arten von Antworttendenzen unterschieden: a) eine Zustimmungstendenz, b) eine

Extremtendenz und c) Soziale Erwünschtheit (Fischer et al., 2006). Während unter einer Zustimmungstendenz verstanden wird, daß die Befragten den Fragen und Äußerungen des Interviewers ungeachtet des Inhalts der Frage immer zustimmen, d.h. auf einer ‚stimme-zu vs. stimme-nicht-zu‘-Skala fast ausnahmslos die ‚stimme-zu‘-Seite wählen, tendieren bei einer Extremtendenz die Befragten dazu, die äußeren Skalenpunkte – sowohl positiv wie negativ – häufiger auszuwählen als die Punkte in der Mitte der Skala. Bei Sozialer Erwünschtheit schließlich berichtet der Befragte dann nur das, was nach seiner Meinung nach im jeweiligen sozialen Kontext als angemessen und normgerecht empfunden wird. In der Forschung ist die Existenz dieser Antwortmuster in weiten Teilen bestätigt worden (vgl. Harzing, 2006; Smith, 2004). Eine nähere Diskussion dieser Thematik findet dann speziell in einem der späteren Kapitel statt. Da solches Kommunikationsverhalten von seinem Verständnis her per definitionem als inhaltsunabhängig angesehen wird und somit bei jeglicher Befragungsthematik auftreten sollte, müßte folglich auch diese Art von kommunikativem Verhalten auf die Abgabe eines Arbeitszufriedenheitsurteil einen Einfluß ausüben. Jedoch wurde dieser Zusammenhang in der gesamten Forschung zur Arbeitszufriedenheitsmessung und zu Mitarbeiterbefragungen bisher weitestgehend ausgeblendet; ein Tatbestand, der mit dieser Untersuchung insbesondere genauer beleuchtet werden soll. Antworttendenzen gehören somit generell zu einflußreichen Faktoren, die der Phase 2 zuzurechnen sind und dort an Bedeutung gewinnen können. Selbst wenn die eigentliche Zufriedenheit, die ein Individuum gemäß Phase 1 verspürt, hoch wäre, könnte somit aufgrund von Zustimmung- oder Extremtendenzen ein völlig anderer Wert von dieser Person auf einer Zufriedenheitsskala berichtet werden.

In Phase 3 gemäß dem Modell von Fischer (1989) sind dann schließlich noch Einflußfaktoren bedeutsam, die sich beispielsweise auf unterschiedliche Anwendungsmöglichkeiten von statistischen Verfahren beziehen. Auch hier können unterschiedliche Analysen (Mittelwertsvergleiche, Korrelationsanalysen, etc.) je nach ihrer Vorgehensweise die Darstellung der eigentlich zu messenden Einschätzung der objektiven Arbeitssituation aus Phase 1 beeinflussen, da auch sie immer eine Form der Aggregation und Zusammenfassung bedeuten und somit Teile von Informationen verloren gehen können. Neben statistischen Verfahren ist ebenso die Interpretation seitens des Forschers bezüglich der gemessenen Zufriedenheitswerte mit all ihren möglichen kognitiven Verzerrungsmöglichkeiten in dieser 3. Phase von Bedeutung.

Bei der Messung von Arbeitszufriedenheit und dem Versuch der Abbildung von objektiven Arbeitssituationen läßt sich somit anhand dieses Modells und seiner 3 Phasen erkennen, welche verschiedenen Faktoren alle Einfluß auf den Abbildungsprozeß nehmen können. Ungeachtet dessen, daß allein die angesprochenen Probleme der Phase 1 des Modells von Fischer (1989) keineswegs als gelöst angesehen werden können, ist diese hier vorliegende Untersuchung, die sich explizit mit der Frage nach dem Zusammenhang von Antworttendenzen und Arbeitszufriedenheitsurteilen beschäftigt – und dies zusätzlich im interkulturellen Kontext –, eindeutig der Phase 2 des Modells zuzuordnen. In diese Phase fallen dann darüber hinaus auch die Rahmenbedingungen von Arbeitszufriedenheitsbefragungen, die in Mitarbeiterbefragungen in Unternehmen vorliegen (Fischer, 1991; Borg, 2000). Deren Bedeutsamkeit für die Interpretation der gemessenen Befunde wird auch von Fischer und Belschak (2006) besonders herausgestellt, soll jedoch hier nicht näher beleuchtet werden.

2.3 Arbeitszufriedenheitsmessung als Managementtool

Bei der Messung von Arbeitszufriedenheit in Unternehmen werden die Rahmenbedingungen der eigentlichen Befragung zumeist eher selten explizit bei der Auswertung der Ergebnisse in den Unternehmen thematisiert. Dennoch ist die Zufriedenheitsmessung in Mitarbeiterbefragungen eines der wichtigsten Tools im Managementkontext. Dabei kann sie eine Vielzahl von Aufgaben übernehmen. Während klassischer Weise eine Mitarbeiterbefragung im Sinne eines bottom-up-Prozesses die Meinung und Zufriedenheitseinschätzungen der Mitarbeiter an das Management weiterleitet, kann sie ebenfalls im Sinne eines top-down-Prozesses dazu genutzt werden, die Aspekte, die dem Management wichtig sind, durch explizite Aufnahme als Befragungsaspekte gesondert herauszustellen und an die Mitarbeiterschaft zu transportieren. Neben diesen kommunikativen Zielen werden zum Teil auch weitergehende Ziele genannt: die Aufdeckung von Problemen, die Entwicklung von Verbesserungsmaßnahmen, die allgemeine Steigerung der Kommunikation zwischen Mitarbeiter und Management sowie die Kontrolle von Veränderungsprozessen (vgl. Bungard, 1997).

Bei der Durchführung der Befragung wird aus pragmatischen Gesichtspunkten oftmals ein möglichst einfach gehaltener Fragebogen erstellt, der die Mitarbeiter nach einer

Einschätzung zu verschiedenen Aspekten ihrer Arbeitsumgebung befragt. Dabei wird von den Befragten ein Urteil abgegeben, was als Maßzahl für die Zufriedenheit über die Arbeitssituation in einer zurückliegenden Periode (zumeist das letzte Jahr) herhalten soll, jedoch oftmals lediglich als Augenblicksurteil anzusehen ist. Probleme der inneren Modellbildung (vgl. Kapitel 2.2) werden dabei in der Praxis der Unternehmensbefragungen völlig ausgeblendet. Bei der anschließenden Analyse der gemessenen Werte wird zumeist ebenfalls einer einfachen Auswertungsroutine gefolgt und in weiten Teilen werden die Zufriedenheitswerte in rein deskriptiven Darstellungen zusammengefaßt. Dabei wird oftmals eine Gegenüberstellung unterschiedlicher Unternehmenseinheiten (z.B. Teams, Abteilungen, Geschäftstellen, Standorte, etc.) vorgenommen. Tabellarische Auflistungen oder Rankings mit den einzelnen Unternehmenseinheiten sollen dazu dienen, eine leicht verständliche Übersicht über gute und schlechte Einheiten zu erhalten. Vor dem Hintergrund der globalen Entwicklung der Wirtschaft und der Internationalisierung von Unternehmen und somit dem Ausbau von internationalen Standorten erweitert sich dieser Vergleich unterschiedlicher Unternehmenseinheiten, der ansonsten zumeist innerhalb eines Landes stattfindet, um den Vergleich über Nationengrenzen hinweg. Zusätzlich werden nun auch Standorte aus verschiedenen Ländern verglichen und die Zufriedenheitswerte aus den in den jeweiligen Ländern unterschiedlichen Kulturen miteinander in Beziehung gesetzt. Dabei kann erneut an dieser Stelle darauf hingewiesen werden, daß dieser internationale Vergleich nicht ohne Probleme ist oder sogar ganz in Frage zu stellen ist (vgl. Kapitel 5 für eine genauere Darstellung). Bei diesem Vergleich ist dann in der Regel die absolute Höhe der Arbeitszufriedenheit in den einzelnen Standorten von Interesse. Das Interesse der Unternehmen liegt somit in einem Vergleich der Zufriedenheitsniveaus und ist somit eine Level-orientierte Fragestellung. Die Struktur des Arbeitszufriedenheitsurteils selbst ist nicht von Interesse. Ob somit in jedem Land extrinsische, intrinsische sowie soziale Arbeitszufriedenheitsaspekte unterschieden werden können, steht in aller Regel nicht im Fokus. Ebenso wenig ist in der Regel die strukturelle Abhängigkeit des Zufriedenheitsurteils von Interesse, d.h. ob die Zufriedenheit in jedem Land von den gleichen Faktoren erzeugt wird. Stattdessen soll der Vergleich der Zufriedenheitsniveaus in den einzelnen internationalen Standorten Hinweise dafür liefern, wo es aus Sicht der Belegschaft bisher gut läuft und wo Verbesserungsbedarf besteht. Eine Mitarbeiterbefragung kann im internationalen Kontext ebenfalls dazu dienen, das kulturelle Verständnis für die verschiedenen

multinationalen Standorte zu erhöhen. Sie ist eines der wenigen Managementtools, welches in der Lage ist, die Stimmung der Belegschaft über nationale Grenzen hinweg zu erfassen und somit der Zentrale des Unternehmens trotz großer geographischer Entfernungen Informationen über die Arbeitsqualität vor Ort zu liefern. Die Wirksamkeit von Unternehmenspraktiken kann so ebenfalls vom Management in den weltweiten Standorten überprüft werden (Dunham & Smith, 1979). Im Gegenzug ist eine weltweite Mitarbeiterbefragung jedoch auch in der Lage, den Mitarbeitern vor Ort eine Möglichkeit zu bieten, ihre Vorstellungen über die Angemessenheit ihrer Arbeitsplätze dem weit entfernten Management zurückzumelden. Vor allem große multinationale Unternehmen wenden daher dieses Managementtool bei ihrer Steuerung der weltweiten Standorte an, um die in den verschiedenen Kulturen verankerten Arbeitnehmerschaften effektiv führen zu können (Johnson, 1996).

Selbst innerhalb eines Landes kann die Berücksichtigung von kulturellen Einflüssen bei der Zufriedenheitsmessung von zunehmender Bedeutung sein, da durch die Internationalisierung des Arbeitsmarktes oftmals auch verschiedenste Nationen innerhalb eines nationalen Standortes zusammenarbeiten.

Zusätzlich steigt die Bedeutung von Zufriedenheitseinschätzungen vor allem auch im Bereich von Zusammenschlüssen von Unternehmen. Wenn auch bei vielen solcher Merger der Einbezug von Mitarbeiterereinschätzungen oftmals eine sehr untergeordnete Rolle spielt, so ist dennoch nicht zu leugnen, daß die Messung solcher ‚weicher‘ Faktoren zunehmend von Wichtigkeit ist und gerade auch viele Fusionen aufgrund der Inkompatibilität dieser scheinbar unwichtigen Faktoren scheitern (Cartwright & Cooper, 1992). Im internationalen Kontext kommt der Berücksichtigung der Mitarbeiterereinschätzung dann unter Umständen eine umso größere Rolle zu. Zwar werden bei internationalen Akquisitionen und Zusammenschlüssen vielfach nur die ‚harten‘ Faktoren, wie beispielsweise die Marktposition, die finanzielle Ausstattung, die Technologie des anderen Unternehmens, beleuchtet, jedoch sind es gerade die Gemeinsamkeiten und die Passung der Belegschaften, die den langfristigen Erfolg von neu zusammengeschlossenen Unternehmen bedingen (Jöns, 2002; Perlitz & Seger, 1999).

Die Passung der Belegschaften wird verständlicherweise umso mehr erschwert, je größer die kulturelle Distanz zwischen den jeweiligen Mitarbeitern ist. Die kulturellen Unterschiede könnten dann auch zu unterschiedlicher Einschätzung der ansonsten gleichen Unternehmensprozesse führen. Ebenso mag das Verständnis über den Sinn und

Zweck einer Mitarbeiterbefragung sowie das Antwortverhalten der Mitarbeiter in solch einer Befragung sich je nach Kultur unterscheiden. In manchen Kulturen mag eine extreme Zufriedenheitsäußerung als ein Akt der Höflichkeit angesehen werden, da es als nicht angemessen erscheinen könnte, Kritik zu äußern. In entgegengesetzter Form mag es jedoch ebenso in anderen Kulturen als eher negativ eingeschätzt werden, sich extrem positiv über seine eigene Arbeitssituation zu äußern, da man sich dadurch zu sehr in den Vordergrund spielen würde und sich gegenüber seinen Kollegen distanzieren und als etwas Besonderes herausstellen würde. Die Erwartungen, was in einer Mitarbeiterbefragung eine angemessene Äußerung ist, mögen somit je nach Kultur variieren. Ein direkter Vergleich der Äußerungen der Mitarbeiterschaften aus verschiedenen Kulturen könnte dann nur vor dem Hintergrund dieser unterschiedlichen kulturellen Ansichten durchgeführt werden.

Um somit Mitarbeiterbefragungen als ein effektives Managementtool im internationalen Kontext einsetzen zu können, ist es von großer Notwendigkeit, den Einfluß von kulturellen Variablen auf die Ergebnisse einer Befragung zu bestimmen. Trotz dieser offensichtlichen Bedeutsamkeit finden sich in der Forschung so gut wie keine Studien, die sich explizit mit der Messung von Arbeitszufriedenheit in internationalen Organisationen beschäftigen (Ryan et al., 1999; Erez, 1994). Neben einer Vielzahl von Einflußgrößen, die auf nationaler Ebene anzusiedeln sind (z.B. Wohlstandsniveau, politisches System, etc.), ist in der konkreten Befragungssituation vor allem dann das kommunikative Verhalten der Befragten bezüglich der Äußerung zur Arbeitszufriedenheit bedeutsam, da gerade die Einflußfaktoren in Phase 2 des semantischen Modells zur Entstehung des Zufriedenheitsurteils (Fischer, 1991) bei Befragungen innerhalb von Unternehmen von Bedeutung ist (Fischer & Belschak, 2006).

3 Die Bedeutung von Kultur

Mitarbeiterbefragungen und Studien, die sich im internationalen Kontext abspielen und über Landesgrenzen hinweg durchgeführt werden, kommen nicht umhin, sich mit einem zentralen Einflußfaktor zu beschäftigen: der Kultur eines Landes, denn die Ergebnisse solcher Studien lassen sich in der Regel nur vor dem Hintergrund der kulturellen Besonderheiten jedes Erhebungslandes sinnvoll und vernünftig interpretieren. In weiten Teilen der Forschung ist diese Herangehensweise inzwischen als Notwendigkeit akzeptiert oder sogar explizit als Forschungsinteresse in den Fokus gesetzt worden; vor allem im Bereich der interkulturellen Psychologie (vgl. zahlreiche Beiträge im JCCP – Journal of Cross-Cultural Psychology). In der Praxis jedoch wird dieser Aspekt noch oft vernachlässigt, was die Bedeutung der hier vorliegenden Untersuchung umso deutlicher macht.

Um verschiedene Kulturen miteinander vergleichbar zu machen und die Unterschiede und Gemeinsamkeiten untersuchen zu können, haben eine Vielzahl von Ansätzen sich mit der Quantifizierung sogenannter Wertedimensionen befaßt. Diese Unterteilung in Dimensionen, die die verschiedenen Wertausprägungen innerhalb einer Kultur erfassen sollen, läßt es zu, die unterschiedlichen Nationen beziehungsweise ihre Kulturen voneinander abzugrenzen. Dabei existiert jedoch kein allgemein gültiges Set von Dimensionen. Stattdessen gibt es vielmehr verschiedene Rahmenmodelle, die versuchen, die relevanten Bestandteile von Kultur zu erfassen. Während sich die Modelle teilweise in ihren Dimensionen überlappen, gibt es jedoch auch Unterschiede in der Anzahl und der Art der erfaßten Dimensionen. Die derzeit bedeutsamsten Modelle sind das Modell von Hofstede (1980, 2001), das Modell von Schwartz (1994) sowie die Zusammenstellung der Dimensionen aus dem GLOBE-Projekt (House, Hanges, Javidan, Dorfman & Gupta, 2004). Diese sollen daher im Folgenden kurz dargestellt werden.

3.1 Kulturelle Rahmenmodelle

Das Modell von Hofstede

Die bisher größte Resonanz im Forschungsumfeld haben sicherlich die Arbeiten von Hofstede (1980; 2001) und sein Modell kultureller Dimensionen hervorgerufen. Anhand einer großen weltweiten Befragung bei Mitarbeitern des IBM Konzern entwickelte er in den 70er Jahren des letzten Jahrhunderts ein Modell, welches anfänglich vier

Wertedimensionen erfaßte und später um eine fünfte Dimension ergänzt wurde. Mit Hilfe der Daten von ungefähr 88.000 Beschäftigten, deren Antworten sowohl 1968 als auch 1972 erfaßt wurden und somit ungefähr 117.000 Datensätze ergaben, ermittelte er durch statistische Berechnungen und inhaltliche Überlegungen seine anfänglichen Dimensionen, anhand derer sich Kulturen voneinander unterscheiden lassen: a) Individualismus vs. Kollektivismus, b) Maskulinität vs. Femininität, c) Unsicherheitsvermeidung und d) Machtdistanz.

Dabei resultierte die Dimension Machtdistanz aus den Überlegungen und den Ergebnissen zu der Frage, wie verschiedene Kulturen mit hierarchischen Beziehungen umgehen. In ähnlicher Form ergab sich die Dimension Unsicherheitsvermeidung aus Hofstedes Ergebnissen zu der anfänglichen Frage, wie verschiedene Berufsgruppen mit Streß umgehen (2001). Jedoch variierten die Ergebnisse viel stärker zwischen Ländern anstatt zwischen Berufsgruppen, was ihn dann dazu veranlaßte, diese Dimension in sein Modell mit aufzunehmen. Die Dimensionen Individualismus/Kollektivismus und Maskulinität/Feminität waren schließlich das Ergebnis faktorenanalytischer Berechnungen von Items zu der Wichtigkeit von Arbeitswerten, die in dem Fragebogen erfaßt wurden.

Hinsichtlich der methodischen Vorgehensweise bei der Ermittlung der Dimension ist im Kontext der hier vorliegenden Arbeit bedeutsam zu erwähnen, daß Hofstede die Entstehung von kulturell bedingten Antworttendenzen berücksichtigte und versuchte, diese in seinen Analysen durch methodische Verfahren zu kontrollieren (siehe Hofstede, 2001). Berücksichtigt wurde dieser Einflußfaktor bei Hofstede speziell bei der Messung von Arbeitswerten, welche letztendlich die Grundlage für die Ermittlung zweier seiner Dimensionen, nämlich Individualismus/Kollektivismus und Maskulinität/Femininität, dienten.

Insgesamt konnte Hofstede anhand umfassender statistischer Analysen die Stabilität seiner vier verschiedenen Dimensionen untermauern (siehe Hofstede, 2001). Ebenso wurden diese in zahlreichen Untersuchungen in den Folgejahren bestätigt (Sondergaard, 1994), so daß insgesamt von folgendem Verständnis ausgegangen wird:

- Individualismus/Kollektivismus beschreibt, inwieweit Personen ihre eigene Entwicklung in den Vordergrund stellen und sich um sich selbst kümmern, oder wie stark sie im Gegensatz dazu in Gruppen eingebettet sind. Dabei nimmt vor allem die Rolle der Familie eine hohe Bedeutung ein.

- Machtdistanz erfaßt die Akzeptanz der Ungleichverteilung von Macht in Institutionen und inwieweit diese von den Mitgliedern der Kultur toleriert akzeptiert wird. Personen mit großem Einfluß, die höher in der Hierarchie stehen, werden dabei beispielsweise in Kulturen mit großer Machtdistanz ohne Vorbehalte als Autoritätspersonen legitimiert.
- Unsicherheitsvermeidung beschreibt die Tendenz, wie stark die Mitglieder einer Kultur sich in unstrukturierten und unsicheren Situationen unwohl fühlen und diesen Zustand vermeiden wollen. Kulturen mit hoher Unsicherheitsvermeidung versuchen dabei sozusagen das Unkontrollierbare kontrollierbar zu machen.
- Maskulinität/Femininität erfaßt die Wertigkeit von Leistung, eigenem Fortkommen und Wettbewerb im Gegensatz zu Werten wie Hilfsbereitschaft, Solidarität und der Sorge um andere Personen. Diese Unterscheidung korrespondiert gemäß Hofstede mit der Art und Weise der Rollenteilungen der Geschlechter, wobei unter Femininität verstanden wird, daß beide Geschlechter die gleichen Rollen einnehmen können, während bei Maskulinität eher eine klare Rollenteilung vorherrscht.

Zusätzlich zu diesen anfänglichen vier Dimensionen wurde das Modell auf Basis von späteren Arbeiten von Bond (Hofstede & Bond, 1988) um eine fünfte Dimension erweitert, da sich vor allem anhand von Studien im asiatischen Kontext zeigte, daß es die folgende weitere Dimension gibt, anhand derer Kulturen deutlich voneinander unterschieden werden können: kurzfristige vs. langfristige Orientierung. Nationen mit langfristiger Orientierung orientieren sich dabei vor allem an übergeordneten und weit in die Zukunft reichenden Effekten ihres Verhaltens. Diese Wertvorstellungen fanden sich traditionell im asiatischen Kulturraum. Im Gegensatz dazu sind Kulturen mit kurzfristiger Orientierung stärker am Wohlbefinden im Hier und Jetzt und der Befriedigung akut vorhandener Bedürfnisse interessiert.

Alle diese Dimensionen fanden in zahlreichen Studien Eingang in die empirische kulturvergleichenden Forschung (vgl. Hofstede, 2001, S. 503ff), jedoch in unterschiedlich starkem Maße. Da Hofstede für jede seiner Dimensionen Kennwerte für die einzelnen Länder ermittelte, werden seine Kulturwerte oft in interkulturellen Studien herangezogen. Speziell für die zuletzt hinzugefügte Dimension der zeitlichen Orientierung gibt es jedoch derzeit noch nicht für alle von Hofstede untersuchten

Länder entsprechende Kennwerte, was ihre Anwendbarkeit in manchen Kontexten einschränkt. Darüber hinaus haben aber auch die anderen vier Dimensionen recht unterschiedliches Interesse hervorgerufen, obwohl für alle diese Dimensionen Maßzahlen für die inzwischen insgesamt untersuchten 53 Länder vorliegen. Sehr häufig wird in der Forschung die Unterscheidung in individualistische und kollektivistische Nationen angewendet. Ihre Bedeutsamkeit konnte vergleichsweise am häufigsten in empirischen Untersuchungen bestätigt werden (vgl. Hofstede, 2001; Triandis, 1995).

Trotz der überwältigenden Zahl an in der einschlägigen Literatur zu findenden Studien, die Hofstedes Dimensionen nutzten, darf nicht unerwähnt bleiben, daß seine Arbeiten auch kritisiert wurden (Spector & Cooper, 2002; Spector, Cooper & Sparks, 2001). Neben der allgemeinen methodischen Kritik einzelner Forscher hinsichtlich der Frage, inwieweit kulturelle Werte überhaupt als der Mittelwert individueller Daten gebildet werden können (vgl. Sego, Hui & Law, 1997), beschäftigten sich mehrere Arbeiten vor allem mit der Frage der Dimensionalität der einzelnen Aspekte, wobei speziell die Dimension Individualismus/Kollektivismus im Fokus der Aufmerksamkeit stand und zum Teil versucht wurde, sie weiterzuentwickeln (Schwartz, 1990; Smith, Dugan & Trompenaars, 1996; Triandis & Gelfand, 1994). Darüber hinaus finden sich vor allem noch folgende Kritikpunkte: die Erhebung der Daten in nur einem bestimmten internationalen Unternehmen und die zum Teil fehlende Validität und Reliabilität seiner Dimensionen und des daraus entwickelten Meßinstruments (VSM 94 – Hofstedes Values Survey Module 1994). Trotz alledem konnte Hofstede viele dieser Einwände selbst entkräften (u.a. Hofstede, 2002) und beispielsweise darlegen, daß gerade die Erhebung in nur einem Unternehmen im Sinne der Parallelisierung der Stichproben viele potentielle Einflußgrößen konstant halten konnte. Insgesamt konnte er und anhand zahlreicher Studien zeigen, daß sein Modell Gültigkeit beanspruchen kann und zur Erklärung und Vorhersage kultureller Unterschiede fähig ist (Hofstede, 2001).

Das Modell von Schwartz

Als eine Weiterentwicklung der Gedanken von Hofstede, aber auch als eine Art konkurrierendes Modell, kann der Ansatz von Schwartz (1994, 2004; siehe auch Schwartz & Ros, 1995) verstanden werden. In ersten Überlegungen kritisierte Schwartz (1990) die Darstellung von Hofstede und entwickelte anhand eigener Daten ein Modell, welches zwar Ähnlichkeiten zu Hofstedes Kulturmodell besitzt, jedoch ebenfalls weitere Dimensionen zur Beschreibung von Kulturen mit aufnimmt. Anhand von Daten

aus bis zu 46 Ländern entwickelte er insgesamt sieben verschiedene Dimensionen, anhand derer sich Kulturen differenzieren lassen. Dabei enthielt seine Ursprungsstichprobe einerseits Lehrer, da diese als besonders bedeutsam für die Sozialisation und als prototypische Träger von Kultur angesehen wurden (Schwartz, 1994). Andererseits bestand die Stichprobe aus Studenten aus ebenso vielen Ländern, um die gefundenen Ergebnisse verifizieren und ihnen eine größere Allgemeingültigkeit geben zu können. Mit Hilfe von Verfahren der multidimensionalen Skalierung konnte Schwartz schließlich drei Wertedimensionen mit jeweils zwei Ausprägungen ermitteln, welche sich in einer kreisförmigen Anordnung positionieren ließen (siehe Abbildung 2): Konservatismus/Eingebettetsein vs. Autonomie (wobei sich diese in Intellektuelle und Affektive Autonomie unterteilen läßt), Hierarchie vs. Gleichheitsdenken/Egalitarismus und Beherrschung vs. Harmonie.

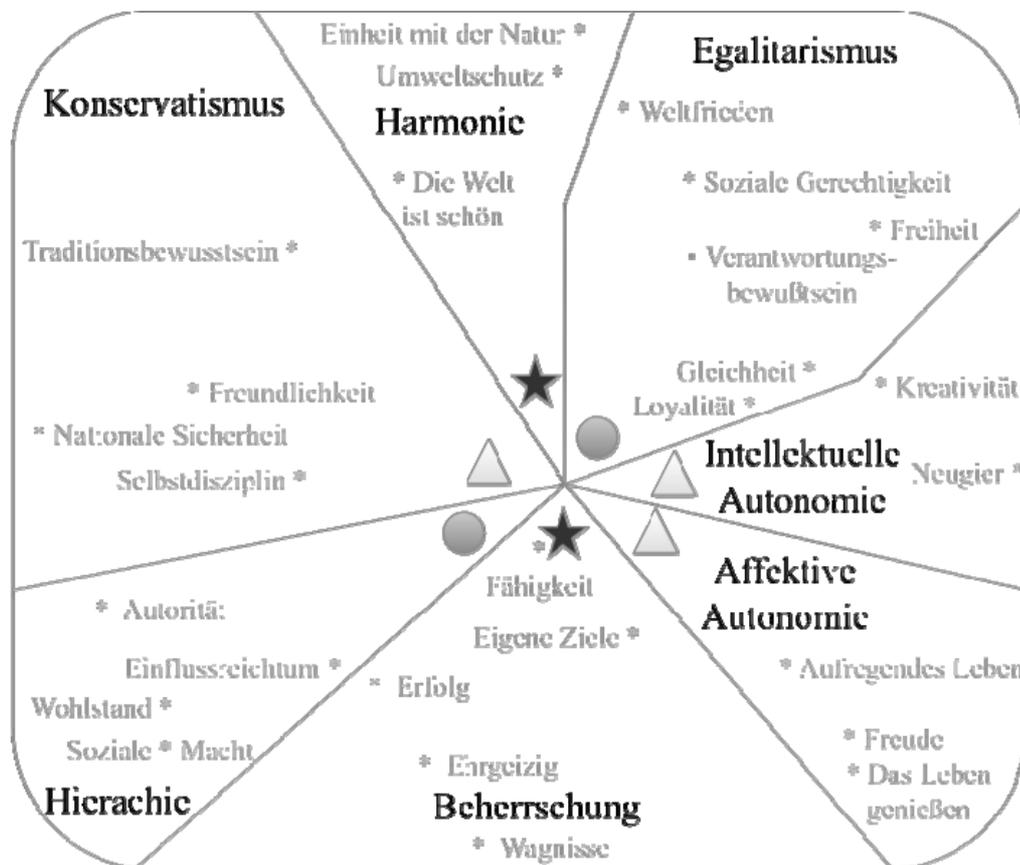


Abbildung 2: Das Werte-Modell nach Schwartz (1999)

Bei dieser Anordnung bilden Aspekte, die sich innerhalb des Kreises gegenüber stehen, entgegengesetzte Pole einer Dimension, während die Aspekte, die nebeneinander stehen, inhaltliche Ähnlichkeiten und Überlappungen aufweisen. So stehen beispielsweise

Konservatismus und Autonomie im Widerstreit miteinander – Kulturen lassen sich entweder eher durch den einen oder aber den anderen Wert charakterisieren –, während Autonomie und Beherrschung in Teilen Ähnlichkeiten aufweisen, nämlich beispielsweise hinsichtlich der Herausstellung und Betonung des Individuums gegenüber seiner Umwelt.

Unter Konservatismus wird dabei insbesondere das Bestreben von Kulturen verstanden, den Status Quo und die bisherigen Traditionen beizubehalten. Jegliche Veränderungen, die die soziale Ordnung in Unordnung bringen könnten, sind hierbei nicht erwünscht. Im Gegensatz dazu bewerten Kulturen mit hoher Autonomie Personen als einzigartig, und die Weiterentwicklung des Individuums wird als zentral und wichtig angesehen. Jede Person darf und soll ihren eigenen Interessen und Motiven folgen, soziale Bindungen und Abhängigkeiten stehen daher dort nicht im Vordergrund. Der Autonomie-Aspekt läßt sich weiter aufgliedern in einerseits eine Form von Intellektueller Autonomie, d.h. die Eigenständigkeit hinsichtlich der eigenen Gedanken und die Verfolgung eigener intellektueller Interessen, und andererseits das Bestreben nach Affektiver Autonomie, d.h. die Unabhängigkeit hinsichtlich der Maximierung eigener positiver und hedonistischer Gefühle (Spaß, aufregendes Leben, etc.).

Die zweite Dimension Hierarchie vs. Gleichheitsdenken beschäftigt sich mit der Frage nach der sozialen Struktur von Kulturen. Bei starker Hierarchie wird innerhalb einer Kultur die ungleiche Verteilung von sozialer Macht und Autorität akzeptiert. Hierarchische Beziehungen werden in solchen Kulturen als Notwendigkeit zur Aufrechterhaltung der sozialen Ordnung angesehen. In Kulturen, in denen Gleichheitsdenken favorisiert wird, streben Menschen hingegen eher nach ausgeglichenen sozialen Strukturen. Die eigenen Interessen werden hierbei stärker zurückgestellt und die freiwillige Verpflichtung, sich um das Wohl der Anderen zu kümmern, wird von den Mitgliedern der Kultur erwartet.

Die dritte Dimension Beherrschung vs. Harmonie schließlich beschreibt die – gemäß Schwartz (1994) – ebenfalls für das Zusammenleben von Menschen zentrale Frage, wie Menschen ihre Beziehung zu ihrer Umgebung und der Welt, in der sie leben, sehen. Wenn Werte von Beherrschung dominieren, strebt eine Kultur danach, die Welt zu kontrollieren, sie gemäß den eigenen Wünschen zu verändern und sie mehrheitlich als reines Objekt zu sehen, das der Erfüllung der eigenen Interessen dienen soll. Bei der Verfolgung der Wertvorstellungen von Harmonie streben Kulturen im Gegensatz dazu

danach, im Einklang mit der Umwelt zu leben und die Welt zu schützen. Die Menschen in solchen Kulturen sehen sich vielmehr in Einheit mit der Natur.

Die von Schwartz entwickelten Dimensionen ließen sich nicht nur in mehreren empirischen Untersuchungen bestätigen, sondern basieren darüber hinaus auch auf umfassenden theoretischen Überlegungen (siehe Schwartz, 1994, 1999, 2004). Im Sinne der konvergenten Validität der Dimensionen im Vergleich zu ähnlichen Konstrukten kann Schwartz (1994) zeigen, daß diese sieben neuen Dimensionen die erwarteten Korrelationen – beispielsweise mit den Dimensionen von Hofstede – zeigen. So korreliert Konservatismus erwartungsgemäß negativ mit Individualismus und positiv mit Machtdistanz, während Autonomie im Gegensatz dazu genau umgekehrt positiv mit Individualismus und negativ mit Machtdistanz zusammenhängt. Ebenso zeigte sich, daß der kulturelle Wert von Beherrschung positiv mit Maskulinität zusammenhängt, während Harmonie positiv mit Unsicherheitsvermeidung korrelierte. Trotz dieser bestätigten Parallelen ergeben sich jedoch nicht bei allen Vergleichen von Schwartz (1994) die Zusammenhänge, die man erwartet hätte. So zeigt sich beispielsweise keine Korrelation von Hierarchie (bei Schwartz) und Machtdistanz (bei Hofstede). Darüber hinaus ist in der Tat überraschend – wie Schwartz (1994) bereits selbst anmerkte – daß Gleichheitsdenken interessanterweise mit Individualismus positiv korreliert ist. Dies ist für ihn jedoch lediglich ein Indikator dafür, daß die Idee, sich um andere zu kümmern, durchaus vereinbar ist mit der Definition des Individuums als eigenständige Person.

Anhand dieser Zusammenhänge wird somit deutlich, daß die Dimensionen von Schwartz als ein eigenständiger Ansatz angesehen werden können, der weitere bedeutsame Aspekte erfaßt, um Kulturen voneinander abgrenzen zu können, und somit nicht nur eine reine Replikation der von Hofstede postulierten Wertedimensionen darstellt.

Auch Schwartz (1994) beachtete bei der Messung seiner Dimensionen die Problematik von kulturell bedingten Antworttendenzen. So versucht auch er seine Werte zu bereinigen, indem er den jeweiligen Gesamtmittelwert der einzelnen Kulturen mit in die Berechnung der Kulturwerte integrierte (nähere Informationen dazu finden sich in den Auswertungsanweisungen seiner beiden Meßverfahren: SVS – Schwartz Value Survey, PVQ – Portrait Value Questionnaire). Durch Berechnung der Kulturwerte jeder einzelnen Dimension in Relation zum Gesamtmittelwert in der Gesamtkultur (d.h. eine Form von Gruppenzentrierung von Werten, vgl. Fischer, 2004) versucht er so für die allgemeine Tendenz zu korrigieren, ob die jeweilige Kultur im Mittel allen Items eher

zustimmt oder ihnen widerspricht. Die gemessenen Kulturwerte sollen demnach frei von verzerrenden Antworttendenzen sein.

Das GLOBE-Modell

Die neueste Studie, die sich umfassend mit der Messung von Kulturunterschieden beschäftigt, ist das Projekt GLOBE – Global Leadership and Organizational Behavior Effectiveness – (House et al., 2004). Zwar war das Hauptziel dieses Projektes der Vergleich des Führungsverhaltens in verschiedenen Kulturen, jedoch entwickelten die beteiligten Forscher zur Messung der Kulturunterschiede ein neues Kulturmodell, welches versucht auf den Arbeiten von Hofstede und Schwartz aufzubauen, und derzeit als das umfassendste und aktuellste Modell angesehen werden kann. Insgesamt wurden 62 Länder in das Projekt integriert, aus denen insgesamt 17.370 Manager der mittleren Hierarchieebenen aus 951 Organisationen teilnahmen. Mit Hilfe von mehreren Pilotstudien und durch umfassende statistische Analysen (explorative Faktorenanalysen, Reliabilitätsanalysen, Aggregationsanalysen, etc.) gelang es dem Forscherteam, insgesamt 9 verschiedene Dimensionen zur Messung von Kulturunterschieden zu entwickeln:

- a) Machtdistanz (Power Distance)
- b) Unsicherheitsvermeidung (Uncertainty Avoidance)
- c) Institutioneller Kollektivismus (Institutional Collectivism)
- d) InGroup-Kollektivismus (InGroup Collectivism)
- e) geschlechterbezogenes Gleichheitsdenken (Gender Egalitarianism)
- f) Bestimmtheit oder auch Selbstbehauptung (Assertiveness)
- g) Humanitäre Orientierung (Humane Orientation)
- h) Zukunftsorientierung (Future Orientation)
- i) Leistungsorientierung (Performance Orientation)

Ein entscheidender Unterschied zu den vorangegangenen Modellen zeigt sich in der methodischen Vorgehensweise der GLOBE-Studie, denn die 9 Dimensionen wurden in dieser Studie in zweierlei Weise erfaßt. Einerseits zielte die GLOBE-Studie darauf ab, das aktuell vorhandene Verhalten in einer Kultur zu erfassen (sog. ‚Practices‘, erfragt als ‚as is‘). Andererseits sollten klassische kulturelle Werte (sog. ‚Values‘) gesondert erfaßt werden (erfragt als ‚should be‘). Somit existieren für jede der Dimensionen jeweils zwei Werte, einmal für die deskriptive Beschreibung der Ist-Situation in einem Land und darüber hinaus für die von den Menschen gewünschte Form der Kultur. Diese

Vorgehensweise sollte die bisherige Vermischung dieser Aspekte in der methodischen Erhebung (was speziell bzgl. Hofstede kritisiert wurde) verhindern. Diese getrennte Erfassung ist jedoch auch im Gegenzug durchaus wieder kritisiert worden (Heine, Lehman, Peng & Greenholz, 2002), da bei 7 der 9 Dimensionen die gemessenen ‚Practices‘ negativ mit den erfaßten ‚Values‘ korrelierten (vgl. House et al., 2004); wie in Tabelle 1 dargestellt. Möglicherweise könnte diese spezielle methodische Vorgehensweise zu einer Art Kontrasteffekt geführt haben, in dem den Befragten durch eine ‚should be‘-Frage nach einer ‚as is‘-Frage suggeriert wurde, daß es Unterschiede zwischen ‚Practices‘ und ‚Values‘ geben müßte und sie daher auf beide Fragen unterschiedlich antworten müßten. Dennoch kann aus einer anderen Perspektive diese unterschiedliche Erfassung durchaus als sinnvoll angesehen werden, denn eine negative Korrelation kann genauso inhaltlich interpretiert werden und bedeutet dann, daß Menschen zwar sehr wohl die gegenwärtige Lage korrekt einschätzen können, sich jedoch genau das Gegenteil davon eigentlich wünschen.

Tabelle 1: Korrelationen der Wertedimensionen und der Praktiken im GLOBE-Projekt auf Länderebene (gemäß House et al., 2004, S. 736)

Werte (‚should be‘)	Praktiken (‚as is‘)
a) Machtdistanz	-.43*
b) Unsicherheitsvermeidung	-.62*
c) Institutioneller Kollektivismus	-.61*
d) InGroup-Kollektivismus	.21
e) geschlechterbezogenes Gleichheitsdenken	.32*
f) Bestimmtheit	-.26*
g) Humanitäre Orientierung	-.32*
h) Zukunftsorientierung	-.41*
i) Leistungsorientierung	-.28*

Anmerkung: Korrelation auf Länderebene; * = $p < .05$

Betrachtet man die Inhalte der einzelnen Dimensionen läßt sich anhand der jeweiligen Namen der Dimensionen die Ähnlichkeit zu den Ansätzen von Hofstede und Schwartz erkennen. Dimensionen, die von diesen Autoren schon früher erfaßt wurden, fanden auch Eingang in die Forschung des GLOBE-Projekts. So entwickelte die GLOBE-Gruppe gemäß den eigenen Angaben ihre Maße zu Unsicherheitsvermeidung, Machtdistanz und Kollektivismus explizit in Anlehnung an die Maße von Hofstede

(House et al, 2004, S. 138f). Jedoch wurde speziell die Kollektivismus-Dimension um einen wichtigen zentralen Punkt erweitert, nämlich die Teilung des Individualismus-Kollektivismus-Konstruktes in eine vertikale und eine horizontale Ausprägung (in Anlehnung an die Ergebnisse von Triandis & Gelfand, 1998). So unterscheidet das GLOBE-Modell Institutionellen Kollektivismus von InGroup-Kollektivismus. Erstere Dimensionen bildet dabei die vertikale Perspektive ab und erfaßt die Bedeutsamkeit von kollektivem Verhalten und der Belohnung, sozialen Normen innerhalb der Gesellschaft zu folgen. Gemeinsames Handeln und die Akzeptanz sozialer Institutionen in einer Gesellschaft werden dabei von den Mitgliedern der Kultur erwartet. InGroup-Kollektivismus hingegen dient der Messung der horizontalen Perspektive und somit der Sichtweise von kollektivistischem Denken, wie sie sicherlich am meisten in der Kulturforschung verbreitet ist und die eher dem Verständnis von Hofstede entspricht. Diesbezüglich wird mit dieser Dimension daher erfaßt, inwieweit sich ein Individuum als autonomes Wesen oder aber eher als eingebunden in und abhängig von seiner zentralen Bezugsgruppe verstehen läßt. Hierbei stehen die Gruppenidentität, der Zusammenhalt von Gruppen, die Bedeutung von Familie und die affektive Bindung an die jeweils unmittelbar umgebenden Personen im Vordergrund.

Die Dimension Zukunftsorientierung ist zwar nicht explizit in Bezug zu Hofstedes Dimension ‚langfristige vs. kurzfristige‘ Orientierung konzeptioniert worden, hat jedoch dennoch Ähnlichkeiten dazu. Sie erfaßt ebenfalls, inwieweit die Mitglieder einer Kultur Aspekte wertschätzen, die weit in die Zukunft reichen. Darüber hinaus erfaßt sie aber auch gemäß House et al. (2004) Unterschiede hinsichtlich einer eher materialistischen oder im Gegensatz dazu spiritualistischen Ausrichtung.

Andere Maße, die zwar auch Ähnlichkeiten hinsichtlich der Namensgebung bei Hofstede und GLOBE besitzen, sind jedoch inhaltlich gesehen unterschiedlicher Natur. Die Dimensionen des geschlechterbezogenen Gleichheitsdenkens und der Bestimmtheit scheinen zwar Ähnlichkeiten mit der Maskulinität von Hofstede zu besitzen, sind jedoch zur Messung anderer Aspekte von den Forschern des GLOBE-Projekts gedacht. Sie erfassen im Gegensatz zu Hofstedes Dimension vielmehr den Ausdruck eigener Gedanken, Gefühle und Ansichten (Bestimmtheit) und das Ausmaß der Gleichbehandlung von Männern und Frauen in einer Gesellschaft (Gleichheitsdenken). Die Aspekte Humanitäre Orientierung und Leistungsdenken gehen schließlich jedoch gänzlich über die Maße von Hofstede hinaus. Die Humanitäre Orientierung einer Kultur mißt das Ausmaß von Menschlichkeit und ist dabei vor allem verbunden mit der

Bedeutung öffentlicher Moral, welche durch entsprechende Gesetze und Normen symbolisiert wird, die moralisches Verhalten fordern. Leistungsdenken schließlich erfaßt die Bedeutung von Herausforderungen und dem Streben nach Verbesserung innerhalb einer Kultur.

Auch zu den Maßen von Schwartz ergeben sich konzeptuelle Überlappungen der Dimensionen der GLOBE-Studie. So gehen die Autoren beispielsweise davon aus, daß Unsicherheitsvermeidung (GLOBE) positiv mit Konservatismus (Schwartz) und negativ mit Intellektueller Autonomie (Schwartz) zusammenhängt (House et al., 2004, S. 143). Ebenso sind Machtdistanz (GLOBE) und Hierarchie (Schwartz) inhaltlich eng miteinander verbunden und auch die Dimension Gleichheitsdenken (GLOBE) soll ähnliche Aspekte wie die Dimension Gleichheitsdenken (Schwartz) erfassen. Darüber hinaus erwarteten die Autoren, daß die Kollektivismus-Dimensionen des GLOBE-Projekts mit den jeweiligen Dimensionen bei Schwartz zusammenhängen, die auch Teile der Individualismus/Kollektivismus-Unterscheidung thematisieren, d.h. Konservatismus und Autonomie.

Vergleich der neuen GLOBE-Dimensionen mit den Modellen von Schwartz und Hofstede

Die empirische Überprüfung der postulierten inhaltlichen Zusammenhänge der im GLOBE-Projekt neu entwickelten Dimensionen mit den bereits bekannten Dimensionen von Schwartz und Hofstede zeigte im Weitesten eine inhaltliche Überlappung und konvergente Validität (House et al., 2004).

Bei der Betrachtung der Gemeinsamkeiten und Unterschiede zwischen den GLOBE-Maßen und den Hofstede und Schwartz Dimensionen muß jedoch erwähnt werden, daß es Unterschiede in den Zusammenhängen gibt hinsichtlich der Unterscheidung des GLOBE-Projektes in ‚Practices‘ und ‚Values‘ (siehe Tabelle 2).

So korrelieren beispielsweise Hofstedes Werte für Machtdistanz zwar stark positiv mit den GLOBE Indizes für Machtdistanz, jedoch nur mit den erfaßten ‚Practices‘, d.h. der Ist-Situation, und nicht mit Machtdistanz als reine Wertvorstellung, wie man es vermutlich eher erwartet hätte. Dies kann daher als Indikator dafür verstanden werden, daß Hofstedes Skala statt Wertvorstellungen in seiner Erhebung eher Praktiken (z.B. des organisationalen Alltags) erfaßt hat. Diesbezüglich merken auch aus theoretischer Sicht House et al. (2004) an, daß dessen Items eher zur Messung von Machtdistanz-Praktiken innerhalb einer Organisationskultur geeignet seien.

Tabelle 2: Korrelationen der Wertedimensionen und der Praktiken im GLOBE-Projekt – auf Länderebene – mit Dimensionen von Hofstede (gemäß House et al., 2004, S. 140)

Dimensionen von Hofstede	Dimensionen von GLOBE	
	Praktiken	Werte
	Machtdistanz	
Machtdistanz	.61**	-.03
	Unsicherheitsvermeidung	
Unsicherheitsvermeidung	-.61**	.32**
	Institutioneller Kollektivismus	
Individualismus	.15	-.55**
	InGroup Kollektivismus	
Individualismus	-.82**	-.20
	Gleichheitsdenken	
Maskulinität	-.16	.11
	Bestimmtheit	
Maskulinität	.42**	-.12

** = $p < .01$

Die Korrelationen von Unsicherheitsvermeidung untereinander (GLOBE und Hofstede) und von Individualismus/Kollektivismus (Hofstede) mit Institutionellem und InGroup Kollektivismus (GLOBE) zeigen hingegen nach House et al. (2004) die von ihnen erwarteten Zusammenhänge, wenn man die Wertedimensionen des GLOBE-Projekts explizit betrachtet. So korrelieren die Dimensionen der Unsicherheitsvermeidung erwartungskonform positiv miteinander (in Tabelle 2 in der rechten Spalte), während die Kollektivismus-Dimensionen des GLOBE-Projekts negativ mit Hofstedes Individualismus-Maß zusammenhängen. House et al. (2004) merken jedoch an, daß dennoch die überraschenderweise negative Korrelation der Praktiken bei Unsicherheitsvermeidung mit der Dimension von Hofstede beachtenswert sind. Eine Erklärung dafür bleiben sie jedoch weitestgehend schuldig und verweisen nur auf das generelle Ergebnis ihrer GLOBE-Studie, daß die erhobenen Praktiken-Indizes zumeist negativ mit den kulturellen Wertedimensionen zusammenhängen, was somit wiederum in der Tat die umgekehrte Richtung dieser speziellen Korrelationen erklären könnte. Maskulinität schließlich korreliert lediglich mit dem GLOBE-Wert für die derzeitigen Praktiken der Bestimmtheit in einer Nation.

Eine mögliche Erklärung für die unterschiedlichen Zusammenhänge zwischen den Hofstede Dimensionen und den jeweiligen GLOBE-Ergebnissen der Praktiken und Wertedimensionen mag auch darin liegen, daß – wie kritischerweise von den Autoren

angemerkt wurde – bei Hofstede diese Dimensionen im Erhebungsinstrument vermischt wurden, während sie in der GLOBE-Studie gezielt getrennt erhoben worden sind (House et al., 2004). Gemäß den ermittelten statistischen Zusammenhängen erfassen die Konstrukte von Hofstede je nach Dimension offensichtlich teilweise mehr die aktuelle Ist-Situation und teilweise mehr die kulturelle Wunschvorstellung der Befragten.

Auch hinsichtlich der Dimensionen von Schwartz zeigen sich weitestgehend die von House et al. (2004) erwarteten Überlappungen (siehe Tabelle 3). So korrelieren beispielsweise die konzeptuell sehr ähnlichen Dimensionen in der erwarteten Richtung: Hierarchie (Schwartz) mit Machtdistanz (GLOBE) positiv, Gleichheitsdenken (Schwartz) mit geschlechterbezogenem Gleichheitsdenken (GLOBE) positiv, Konservatismus (Schwartz) mit Unsicherheitsvermeidung (GLOBE) positiv und Intellektuelle Autonomie (Schwartz) mit Unsicherheitsvermeidung negativ.

Tabelle 3: Korrelationen der Wertedimensionen im GLOBE-Projekt – auf Länderebene – mit den inhaltlich ähnlichen Dimensionen von Schwartz (gemäß House et al., 2004, S. 143)

Dimensionen von Schwartz	Dimensionen von GLOBE - Werte		
	Unsicherheitsvermeidung	Institutioneller Kollektivismus	InGroup Kollektivismus
Konservatismus	.74**	.14	.15
Intellektuelle Autonomie	-.61**	-.13	-.14
Hierarchie	Machtdistanz .33*		
Gleichheitsdenken	Gleichheitsdenken .65**	Bestimmtheit -.44**	
Beherrschung	Leistungsorientierung .12		

* = $p < .05$, ** = $p < .01$

Interessanterweise zeigen sich jedoch keine Zusammenhänge zwischen Kollektivismus-Dimensionen des GLOBE-Projekts und den inhaltlich scheinbar ähnlichen Dimensionen bei Schwartz (Konservatismus und Intellektuelle Autonomie). House et al. (2004) selbst bleiben auch hier eine weitere Erklärung schuldig und beschränken sich rein auf die deskriptive Darstellung der Ergebnisse. Diese bei ihnen nicht gefundenen

Zusammenhänge könnten jedoch somit als Indikator dafür gewertet werden, daß die von Schwartz favorisierten Dimensionen des Konservatismus und der Autonomie, die vordergründig mit dem klassischen Individualismus/Kollektivismus-Konzept zusammenzuhängen scheinen, in der Tat etwas andere inhaltliche Aspekte erfassen.

Insgesamt zeigt sich aber, daß diesen drei kulturellen Rahmenmodellen trotz zum Teil unterschiedlicher Operationalisierungen ein weitestgehend gemeinsames Verständnis zugrunde liegt. Demnach scheint es eine bestimmte Anzahl an Dimensionen zu geben, anhand derer sich Kulturen relativ gut voneinander unterscheiden lassen. Für die hier vorliegende Untersuchung sind die dargestellten Kulturmodelle insofern von großer Bedeutung, als daß sie die theoretische Basis darstellen, aufgrund derer sich auch die Unterschiede in den anderen Faktoren der Fragestellung erklären lassen. Kulturelle Unterschiede sind dabei vor allem dafür verantwortlich, daß Menschen eine unterschiedliche Sozialisierung in ihrer jeweiligen Kultur durchlaufen. Diese unterschiedliche Sozialisierung ist unter anderem dann der Grund dafür, daß Menschen auch ein unterschiedliches Kommunikationsverhalten lernen.

3.2 Kultur als Sozialisationsinstanz für kommunikatives Verhalten

Die kulturelle Umgebung, in die Menschen hineingeboren werden, hat generell einen weitreichenden Einfluß auf die jeweilige Entwicklung der Personen. Von Kindheit an werden die Menschen dabei mit den Werten konfrontiert, die in der jeweiligen Kultur hohe Bedeutung haben. Sie durchlaufen sozusagen eine Art von kultureller Sozialisation. Dabei werden sie mit einer ganzen Anzahl von Erwartungen konfrontiert, die ihnen vorgeben, wie man sich in einer Kultur zu verhalten hat. Für die hier vorliegende Arbeit ist dabei vor allem von Bedeutung, daß Kultur eine Sozialisationsfunktion hinsichtlich des Gebrauchs von Sprache besitzt und daß den Menschen eine bestimmte Art von Kommunikationsverhalten in der jeweiligen Kultur beigebracht wird (vgl. Gudykunst & Matsumoto, 1996). Aufgrund unterschiedlicher kultureller Wertvorstellungen lernen Menschen durch formales und informelles Training sowohl innerhalb der Kindheit aber selbst noch im Erwachsenenalter, wie sie mit Anderen in effektiver Form kommunizieren können (Smith & Fischer, 2008). Eine sehr bedeutende Rolle bei der Sozialisation des Kommunikationsverhaltens nimmt dabei das Erlernen der Sprache als

Kleinkind ein (Schieffelin & Ochs, 1986). Von Klein auf lernen Kinder, was in einer Kultur hinsichtlich kommunikativer Aspekte als gutes und was als schlechtes Verhalten angesehen wird. Kulturelle Sozialisationsroutinen beeinflussen dabei so basale Aspekte wie die Bildung von Worten und das Erlernen der Grammatik (Peters & Boggs, 1987). Welche kulturellen Routinen besonders bedeutsam sind, diskutiert unter anderem auch Hofstede (2001) selbst in seinen Ausführungen. Vor allem in Kulturen mit einer hohen Unsicherheitsvermeidung werden Kinder beispielsweise mit einer Vielzahl von Regeln konfrontiert, die von ihnen einzuhalten sind. Dies gilt nicht nur hinsichtlich des allgemeinen Verhaltens, sondern auch bezüglich der kommunikativen Ausdrucksweise. So lernen Kinder unter anderem sehr deutlich, wann ein „Sie“ und wann ein „Du“ die angemessene Anrede ist, sowie wann sie in der Ich-Form sprechen dürfen oder besser eine Aussage im Passiv formulieren (Hofstede, 2001). Dieses Training hinsichtlich einer angemessenen Ausdrucksweise führt letztendlich zu deutlichen interkulturellen Unterschieden in der Benutzung von Sprache. So zitiert Hofstede (2001) eine interessante Studie von Kashima und Kashima (1998), die einen klaren Bezug zwischen kulturellen Werten und dem Kommunikationsverhalten darlegt. Zwischen Individualismus und dem sogenannten ‚pronoun drop‘ (d.h. die erste Person Singular ‚Ich‘ muß in einem Satz nicht genannt werden) besteht gemäß dieser Studie ein deutlicher Zusammenhang von $r = -.75$. In individualistischen Kulturen benötigen somit gemäß Hofstedes Interpretation Personen die erste Person Singular ‚Ich‘, wenn sie über sich selbst berichten, während in kollektivistischen Kulturen dieses Pronomen auch wegfallen kann (Hofstede, 2001). Darüber hinaus ist bemerkenswert, „that the language spoken in the most individualist countries, English, is the only one [...] that writes *I* with a capital letter“ (Hofstede, 2001, S. 233). In individualistischen Kulturen scheint folglich nicht nur der Inhalt einer Kommunikation oftmals ich-bezogene Aspekte zu enthalten, sondern auch die Struktur der Sprache an sich ist geprägt durch die hohe Bedeutung des Individuums. Kulturelle Erwartungen führen somit insgesamt dazu, daß beispielsweise in manchen Kulturen grundsätzlich stärker in der Ich-Perspektive gesprochen wird, während in anderen Kulturen hingegen eine Unterscheidung in der Sprache in eine direkte Anrede im Vergleich zu einer eher höflicheren und distanzierteren Form der Anrede existiert (z.B. Deutsch: „du“ vs. „Sie“, Französisch: „tu“ vs. „vous“, Niederländisch: „jij“ vs. „u“, Italienisch: „tu“ vs. „Lei“). Kulturelle Gegebenheiten haben somit einen deutlichen Einfluß darauf, wie Menschen miteinander kommunizieren. Die Erwartung, sich möglichst auf eine höfliche Art und

Weise auszudrücken, kann sogar soweit führen, daß „in some cultures and subcultures, being polite to the other person is more important than supplying objectively correct information; in some, respondents will never use *no*.“ (Hofstede, 2001, S. 21). Interessanterweise konnten nämlich beispielsweise Collett und O’Shea (1976) in einem vergleichenden Studie in England und dem Iran zeigen, daß im Iran bis zu 20% der Passanten einem Fremden sogar den Weg zu einem Ort erklärten, der überhaupt nicht existierte, während dies in England niemand getan hat. Dieses anekdotische Beispiel macht deutlich, welch großen Einfluß kulturelle Erwartungen auf die Art und Weise haben, wie Menschen sich in einem Gespräch äußern.

Für die hier vorliegende Studie ist genau dieser kulturelle Einfluß auf die Sozialisation der sprachlichen Ausdrucksweise von großer Bedeutung, denn Antworttendenzen stellen letztendlich nichts anderes dar als eine Art von kulturell antrainiertem Kommunikationsverhalten:

“[...] culture-members’ trained styles of habitual communication are reflected in the way that they respond to the survey questionnaires that we ask them to complete, particularly those that use a Likert-scale response format. Consistently acquiescent responding to a series of survey items corresponds to a habitual emphasis upon inclusion with others, generalized to include the researcher, who has entered the life space of the respondent. On the other hand, consistent use of the extreme response categories to a series of survey items corresponds to a habitual emphasis upon representing oneself as decisive, separate, and distinctive, or at least as exemplifying qualities that do not rest upon a need for endorsement by others” (Smith & Fischer, 2008, S. 289)

Zustimmungstendenzen sind somit beispielsweise ein Ausdruck der eben beschriebenen Art, sich höflich und zurückhaltend in der Kommunikation zu äußern, während Extremtendenzen eher der in individualistischen Ländern favorisierten Kommunikationsweise, die Ansichten der eigenen Person in den Vordergrund zu stellen, entsprechen würden. Vor allem wichtig für die hier untersuchte Fragestellung ist, daß diese Antworttendenzen nicht nur ein Ausdruck der alltäglichen Kommunikationsweise ist, die zwischen den Mitgliedern innerhalb einer Kultur stattfindet, sondern ebenso in Richtung des Forschers gezeigt wird, wie Smith und Fischer (2008) betonen. Man

könnte sogar davon ausgehen, daß dieses Verhalten sich gegenüber einem Forscher verstärkt zeigen müßte, da dieser in der Regel ein Fremder ist (Smith, 2004).

Die Kommunikation zwischen Mitgliedern einer Kultur findet somit insgesamt nicht in einem wertfreien Kontext statt, sondern ist tief verwurzelt in den kulturellen Gegebenheiten des jeweiligen Landes. Die kulturellen Werte eines Landes sozialisieren die Menschen in einer bestimmten Art und Weise, in Befragungen ihre Meinung zu äußern. Kommunikative Aussagen drücken neben einer inhaltlichen Botschaft daher auch immer die Wertvorstellungen der Menschen aus, wie in interpersonalen Beziehungen miteinander umgegangen werden sollte. In unterschiedlichen Kulturen entwickeln sich diesbezüglich unterschiedliche Kommunikationsstile. Inwieweit die in dieser Arbeit betrachteten Antworttendenzen (Zustimmungs- und Extremtendenz sowie Soziale Erwünschtheit) jeweils typisch für ganz spezielle Konstellationen der kulturellen Ausprägungen gemäß Hofstede, Schwartz oder der GLOBE-Studie sind, wird sowohl im theoretischen Teil in einem der späteren Kapitel weiter vertieft als auch in der empirischen Studie einzeln geprüft.

4 Arbeitszufriedenheit im interkulturellen Kontext

4.1 Bisherige Ergebnisse zur Höhe der Arbeitszufriedenheit im interkulturellen Vergleich

Um allgemeine Aussagen über die Arbeitszufriedenheit in verschiedenen Ländern und Kulturen zu erhalten, ist es möglich, sich in einem ersten Schritt anzuschauen, welche Ergebnisse die bisherige Forschung zur Höhe der Arbeitszufriedenheit in verschiedenen Kulturen ergeben hat. So haben einerseits mehrere Studien explizit versucht zu untersuchen, wie Befragte sich in verschiedenen Ländern zu ihrer Arbeitszufriedenheit geäußert haben. Andererseits wird in manchen Studien auch die Höhe der Arbeitszufriedenheit als Nebenprodukt berichtet, selbst wenn der eigentliche Fokus der Untersuchung ein anderer war. Die folgende Darstellung gibt einen Überblick über diese Ergebnisse in den jeweiligen Kulturen.

Viele Studien, daß ähnlich wie in rein nationalen Studien zur Arbeitszufriedenheit generell auch im internationalen Vergleich das Niveau der Arbeitszufriedenheit unabhängig von der Kultur im Allgemeinen hoch zu sein scheint. Das heißt, nicht nur in der rein nationalen Forschung innerhalb einer Kultur sondern auch beim Vergleich verschiedener Länder zeigt sich oft bei Fragen zur Arbeitszufriedenheit eine schiefe Verteilung der Antworten hin zum positiven Ende der Skala. So berichtet beispielsweise Spector (1997) über eine große Studie, die von der Gallup Organisation 1995 durchgeführt wurde und deren Ergebnisse in weltweit 18 verschiedenen Nationen das bekannte Muster der Verteilungsschiefe hin zur Zufriedenheit zeigen. Dabei wird die gesamte Spannbreite der vorgegebenen Zufriedenheitsskala nicht ausgenutzt. Stattdessen antworten alle Befragten in allen Ländern im Mittel eher positiv auf die Frage nach der Zufriedenheit mit ihrer Arbeit. In einer weiteren Gegenüberstellung von Ergebnissen zur Messung der Zufriedenheit mit der Arbeit anhand des Job Satisfaction Surveys (JSS) zeigt Spector (1997), daß auf einer Skala von 4 bis 24 die tatsächlich gemessenen Werte in vier Nationen (Dominikanische Republik, Hongkong, Singapur und Vereinigte Staaten) alle im oberen Bereich von ca. 17 bis 22 liegen. Diese Verschiebung ins Positive betraf sowohl das Urteil der Gesamtzufriedenheit als auch die die Urteil zu allen Subskalen des JSS.

Beispielhaft sei hier zudem eine weitere Studie angeführt, die die repräsentativen Daten des weltweit durchgeführten International Social Survey Programme (ISSP) aus dem Jahre 1997 nutzte (Souza-Poza & Souza-Poza, 2000). Dort ergibt sich ein ähnliches Bild. In den dort untersuchten 21 Ländern zeigt sich, daß auf einer Skala von 1 bis 7 das geringste Zufriedenheitsniveau bei ca. 4,81 (Ungarn) und das höchste Niveau bei ca. 5,7 (Dänemark) lagen. In keinem der Länder lag somit die mittlere Zufriedenheit mit der Arbeit im negativen Bereich der Skala, d.h. hier bei einem Wert von weniger als 3,5. Weitere Studien zeigen ebenso in der Regel die gleichen Ergebnisse hinsichtlich der schiefen Verteilung der Antworten (Six & Felfe, 2006; van de Vliert & Janssen, 2002). Somit bleibt festzuhalten, daß auch bei der interkulturellen Messung von Arbeitszufriedenheit im Mittel alle Befragten sich in der Regel positiv über ihre Zufriedenheit äußern. Das Niveau der Arbeitszufriedenheit auf Länderebene liegt also auch zumeist im positiven Bereich der Meßskala, ähnlich wie es bei der Messung auf individuellem Niveau der Fall ist.

Trotz dieser generellen Verschiebung zum positiven Ende der Skala hin ergeben sich dennoch zum Teil deutliche Unterschiede zwischen den einzelnen Nationen. Oft werden dabei jedoch in der Forschung nur wenige, d.h. 2-3 Länder, miteinander verglichen. So fanden Lincoln und Kalleberg (1985), daß in einer Studie mit über 8000 Befragten aus 52 Werksstandorten in den Vereinigten Staaten und 46 Betriebsstätten in Japan die Japaner insgesamt geringere Zufriedenheitswerte berichteten als die Amerikaner. Dies galt neben der Arbeitszufriedenheit auch für andere Fragen hinsichtlich der Bewertung der Jobs, z.B. die Empfehlungsbereitschaft für den aktuellen Job und die Frage nach der möglichen Wiederaufnahme der jetzigen Stelle. In einer Studie, die ebenfalls amerikanische und japanische Arbeitnehmer verglichen hat – jedoch hier lediglich in japanischen Unternehmen, die einen Sitz in Amerika hatten –, fanden Lincoln, Hanada und Olson (1981), daß Japaner auch hier sich unzufriedener über die Arbeit äußerten als ihre amerikanischen Kollegen. Insgesamt deuten viele Studien daraufhin, daß japanische Mitarbeiter sich in der Tendenz immer weniger zufrieden über ihren Job äußern als Amerikaner (Smith & Misumi, 1989). In ähnlicher Form äußerten sich weitere asiatische Befragte auch in einer anderen Studie, die Krankenschwestern in Hongkong und Singapur mit deren Kolleginnen in Australien und den USA verglich (Chiu & Kosinski, 1999), weniger zufrieden über die Arbeit als die Mitarbeiterinnen in den englischsprachigen Nationen.

Auch eine weitere Studie im mittleren Management, bei der 142 Manager aus Singapur mit 137 australischen Managern verglichen wurden, zeigte, daß sich die Manager im westlichen Kulturkreis zufriedener über ihre Arbeit äußerten (Harrison, 1995). Die Ergebnisse des in dieser Studie eingesetzten und international erprobten Minnesota Satisfaction Questionnaire (MSQ) zeigten dabei ein deutlich höheres Zufriedenheitsniveau bei den Australiern ($M = 70,66$ auf einer 100er-Skala) als bei den Managern aus Singapur ($M = 62,03$).

Neben diesen Vergleichen von amerikanischen mit asiatischen Stichproben zeigt sich auch in einer Studie aus einem anderen kulturellen Umfeld und unter Einbezug einer etwas anderen Art von Stichprobe ein höheres Zufriedenheitsniveau der Amerikaner gegenüber der Vergleichsgruppe. In der Studie von Judge, Locke, Durham und Kluger (1998) berichteten nämlich israelische Studenten geringere Zufriedenheitswerte als amerikanische Studenten und ebenfalls geringere als amerikanische Ärzte, die auch in dieser Studie mit erfaßt wurden.

Während die meisten Studien in der Regel eine höhere Zufriedenheit von Amerikanern zeigen, existieren auch einige wenige Studien, die diesbezüglich bilden eine Art Ausnahme bilden. So zeigt sich in einer Studie (Kirkman & Shapiro, 2001), bei der in zwei multinationalen amerikanischen Unternehmen insgesamt 461 Befragte aus vier Ländern (USA, Belgien, Finnland und den Philippinen) verglichen wurden, ein leicht abweichendes Bild. Dort hatten erstaunlicherweise nicht die Amerikaner die höchsten Zufriedenheitswerte berichtet, sondern stattdessen die Mitarbeiter aus den Philippinen, gefolgt von Belgiern, dann den Amerikanern und an vierter Stelle den Finnen. Jedoch erhielten in dieser Studie die philippinischen Mitarbeiter als einzige nicht den Fragebogen in ihrer Muttersprache, sondern in Englisch. Dies könnte dazu geführt haben, daß diese hier in dieser Untersuchung interessanterweise die höchsten Zufriedenheitswerte berichteten, da die Forschung schon zeigen konnte, daß Ergebnisse von Befragungsstudien sich unterscheiden können, je nachdem ob in der Muttersprache oder ansonsten klassischer Weise in Englisch befragt wird (Harzing, 2005). In der bereits erwähnten Gegenüberstellung von Spector (1997), der auf Daten von vier unabhängigen Länderstudien zurückgriff, zeigte sich, daß die höchsten Zufriedenheitswerte in der Stichprobe der Dominikanischen Republik vorlagen, gefolgt von den USA, Singapur und schließlich Hongkong, die alle drei ein relativ gleiches Zufriedenheitsniveau aufwiesen, jedoch dennoch mit einem vergleichsweise leicht höheren Zufriedenheitsniveau bei den Amerikanern. Während die Daten aus den ersten

drei Nationen einzelnen Befragungen von angefallenen und somit nicht repräsentativen und nicht zufällig ausgewählten Stichproben entstammten, zieht der Autor für die USA hingegen Normwerte des JSS heran. Diese unterschiedliche Stichprobengenerierung mag ein Grund dafür sein, daß die Werte für die USA hier im relativen Vergleich mit den asiatischen Ländern ebenfalls überraschend niedrig erscheinen. In einer frühen Studie von Slocum und Topichak (1972) zeigte sich schließlich noch, daß Mexikaner dort höhere Zufriedenheitswerte berichteten als Amerikaner.

Viele internationale Studien beschränken sich leider – zumeist aus pragmatischen Gründen – nur auf wenige Länder, um Arbeitszufriedenheit in verschiedenen Ländern zu vergleichen. Nur einige wenige Studien hingegen versuchen, Aussagen anhand einer Vielzahl von Ländern zu treffen. In einer der ersten Studien verglichen Haire, Ghiselli und Porter (1966) die Zufriedenheit von Managern in 14 Ländern. Hier wiesen die Manager aus Schweden die höchsten Zufriedenheitswerte auf, gefolgt von Norwegen und Dänemark. Die geringsten Werte waren hingegen in Indien, Argentinien und Spanien zu finden. Souza-Poza und Souza-Poza (2000) konnten in der bereits zitierten Studie anhand der Daten des ISSP immerhin 21 Länder vergleichen. Das höchste Zufriedenheitsniveau berichten hierbei die Dänen und die Zyprioten. Mit etwas Abstand folgen als zweite Gruppe mit relativ ähnlichen Zufriedenheitsniveaus die Schweiz, Israel, die Niederlande, Spanien und die USA. Im breiten Mittelfeld finden sich in absteigender Reihenfolge die folgenden Länder wieder: Neuseeland, Schweden, Norwegen, Italien, Portugal, Deutschland, Tschechische Republik, Großbritannien und Frankreich. Die Schlußlichter bilden Bulgarien, Slowenien, Japan, Rußland und Ungarn. Es scheinen hier westliche Industrienationen eher an der Spitze zu liegen, während ost-europäische und asiatische Länder (als einziger Vertreter hier Japan) eher geringere Zufriedenheitswerte berichtet haben. Das Ergebnis läßt aber insgesamt keine eindeutige Kategorisierung der erfaßten Länder zu. In einer weiteren umfangreichen Studie haben Spector et al. (2004) insgesamt über 2400 Manager aus insgesamt 13 Ländern untersucht. Ihre Ergebnisse, für die sie die Stichprobe in drei Regionen unterteilten, zeigten einen Unterschied in der berichteten Arbeitszufriedenheit zwischen den jeweiligen Regionen. Manager aus Anglo-amerikanischen Kulturen (Australien, Kanada, England, Neuseeland und die USA) berichteten eine höhere Zufriedenheit als Manager aus China. Noch höhere Zufriedenheitswerte kreuzten jedoch in dieser Studie Manager aus Lateinamerika an (Argentinien, Brasilien, Kolumbien, Ecuador, Mexico, Peru und Uruguay). Ebenfalls eine Dreiteilung wurde in einer Studie von Griffeth und Hom

(1987) vorgenommen. Die über 1700 Befragten Manager aus 15 Nationen, die innerhalb eines einzigen amerikanischen weltweit tätigen Unternehmens befragt wurden, konnten in – wie die Autoren es nannten – latein-europäische Länder (Spanien, Italien, Portugal, Griechenland und auch Belgien), in nordisch-europäische Länder (Norwegen, Dänemark, Schweden, Deutschland, Schweiz, Österreich) und in angelsächsische Länder (Kanada, England und auch Holland) geclustert werden. Dabei zeigte das latein-europäische Cluster die geringste Arbeitszufriedenheit und das nordisch-europäische Cluster die höchsten Zufriedenheitswerte. Das angelsächsische Cluster lag hier in der Mitte. In einer weiteren, recht aktuellen Studie, in der über 9000 Mitarbeiter eines global tätigen Unternehmens der Informationstechnologie in 18 Ländern erfaßt wurden, lagen hingegen wiederum mehrheitlich die mittel- und nord-europäischen Länder sowie die USA vorne, während asiatische und süd-europäische Länder wieder die Schlußlichter innerhalb der Zufriedenheitsrankings bildeten (Liu, Borg & Spector, 2004). Das genaue Ranking der Länder sah dabei wie folgt aus: An der Spitze lagen der Reihenfolge nach Schweiz, Italien, Deutschland, USA, Österreich und Kanada. Das Ende des Rankings bestand aus China, Taiwan, Korea, Frankreich, Spanien und dem Schlußlicht Japan. In der Mitte fanden sich Mexiko, Süd-Afrika, Brasilien, Niederlande, England und Australien wieder. Nord- und mitteleuropäische Länder lagen auch in einer Studie von Llorente und Macias (2005) vorn; verständlicherweise, denn sie nutzten für ihre Analysen ebenfalls die Daten des ISSP. Auch in einer Zusammenstellung von Six und Felfe (2006), die die Daten des Eurobarometers nutzten und somit lediglich europäische Länder in ihrem Datensatz betrachten konnten, lagen die nord-europäischen Länder (Niederlande, Dänemark, Schweden, Österreich, Finnland, Belgien) weiter vorne, während die süd-europäischen Länder (Spanien, Griechenland, Portugal, Frankreich, Italien) im Mittel die geringsten Zufriedenheitswerte angaben. Eine bereits früher durchgeführte Studie zeigte ebenfalls ähnliche Ergebnisse beim Vergleich europäischer Länder. Near und Rechner (1993), die in ihrer Studie Befragungsdaten von 1976 aus zehn verschiedenen europäischen Ländern berichten, fanden bereits fast 30 Jahre früher, daß in Dänemark die höchsten Zufriedenheitswerte berichtet wurden. Mit etwas Abstand folgten in ihrer Studie die Niederlande, Nord-Irland und Belgien. Etwas geringere Zufriedenheitsniveaus fanden sich in Irland, England, Deutschland und Luxemburg. Die niedrigsten Werte hatten Frankreich und Italien. Abschließend sei hier noch die Studie von van de Vliert und Janssen (2002) angeführt, in der einerseits wiederum nord- und mittel-europäische Länder die Reihenfolge anführten und

andererseits erneut asiatische Länder eindeutig geringere Zufriedenheitswerte angaben. Die höchste Zufriedenheit lag in dieser Untersuchung in der Schweiz vor, gefolgt von Norwegen, Island, Deutschland und Schweden. Am Ende der Rangfolge finden sich neben der Türkei ausnahmslos asiatische Länder: Singapur, Japan, China, Taiwan, Korea und Hongkong. Somit zeigen sich auch hier hinsichtlich der Rangfolge von Ländern ähnliche Ergebnisse wie in den anderen Vergleichsstudien, auch wenn in dieser Studie Studenten an Stelle von Mitarbeitern eines Unternehmens als Stichprobe dienten. Auf den ersten Blick scheinen nicht alle Ergebnisse ein stringentes Bild abzugeben. Besonders bei den Studien, die nur 2-3 Länder vergleichen, zeigen sich zum Teil unterschiedliche Ergebnisse. Während in der Mehrzahl der Studien Amerikaner höhere Arbeitszufriedenheitswerte berichteten als beispielsweise Asiaten, gibt es jedoch auch einzelne Studien, in denen sich nicht in den USA das höchste Zufriedenheitsniveau finden ließ. Dennoch scheint sich oft zu zeigen, daß im direkten Ländervergleich beispielsweise Japaner sich zumeist geringer über ihre Zufriedenheit äußern als Amerikaner. Der direkte Ländervergleich läßt jedoch grundsätzlich auch wenig Rückschlüsse über eine möglicherweise generelle Tendenz zu. Betrachtet man hingegen die Studien, die viele verschiedene Nationen umfassen und idealerweise sich dabei auch nicht nur auf einen Kulturkreis beschränken, scheint sich durchaus ein mehr oder minder einheitliches Bild zu ergeben. Insgesamt scheinen diese Studien als deskriptives Ergebnis zu zeigen, daß in der Mehrzahl der Studien entweder in nord- und mitteleuropäischen Ländern oder aber in den USA sich höhere Zufriedenheitswerte in Befragungen zur Arbeitszufriedenheit ergeben und im Gegensatz dazu asiatische Länder oftmals das Ende des Zufriedenheitsrankings bilden.

Eine Erklärung, warum die Länder sich unterscheiden, läßt sich jedoch oft in solchen Studien nicht finden. Vielfach belassen es die Autoren bei einem deskriptiven Vergleich. Teilweise werden auch nur Mutmaßungen über die Gründe angestellt, ohne jedoch diese mit gemessenen Kennwerten zu untermauern. Letztendlich ließe sich somit eine Vielzahl von Größen finden, die auf das Arbeitszufriedenheitsniveau einen Einfluß ausüben und der Grund für die Unterschiedlichkeit zwischen den Ländern sein könnten.

4.2 Einflußgrößen auf das Arbeitszufriedenheitsniveau

Die Gegenüberstellung der deskriptiven Unterschiede zwischen den Niveaus der Arbeitszufriedenheit in den einzelnen Ländern deutet darauf hin, daß es vermehrt

westliche Nationen zu sein scheinen, in denen das Arbeitszufriedenheitsniveau höher ausgeprägt ist, während asiatische Länder eher wenig Zufriedenheit mit der Arbeit berichten. Die Kategorisierung in westliche und asiatische Nationen stellt jedoch noch keinen wirklich erklärenden Faktor für die gemessenen Unterschiede dar, denn beide Ländergruppen unterscheiden sich hinsichtlich einer Vielzahl von Faktoren. Daher besteht weiterhin die berechtigte Frage, welche Einflußgrößen auf das Arbeitszufriedenheitsniveau sich finden lassen und die jeweiligen Niveauunterschiede erklären können.

Oftmals wird spontan vermutet, daß die Arbeitsbedingungen in manchen Ländern besser als in anderen seien. So sieht auch Spector (1997) dies als einen möglichen Grund für die Unterschiede in seiner Gegenüberstellung von Ländern an. Genauso gut könne es jedoch nach Meinung des Autors auch an unterschiedlichen Erwartungen in den einzelnen Kulturen gelegen haben. Dies beruht jedoch nur auf Spekulationen, denn explizit gemessen wurden bei ihm keiner der Kennwerte – wie es auch in vielen anderen Studien oft der Fall ist –, weder Arbeitsbedingungen noch Erwartungen. Letztendlich bleibt somit jede solcher Überlegungen lediglich auf einem Plausibilitätsniveau, solange sie nicht mit empirischen Daten unterstützt wird.

Einen Beitrag in dieser Richtung stellt die Studie von Souza-Poza und Souza-Poza (2000) dar, indem die Autoren versuchen, eine Art Kosten-Nutzen-Relation in den einzelnen Ländern aufzustellen. Sie vergleichen in ihrer Analyse die Beziehung von Input- und Outputgrößen innerhalb des Arbeitslebens und setzen die jeweilige Ausprägung der beiden Größen und deren Verhältnis mit dem Arbeitszufriedenheitsniveau in Beziehung. Als Input-Größen verstehen sie a) die Zeit, die jemand in Bildung investiert, b) den Zeitumfang, den man arbeiten muß, c) das Ausmaß, wie anstrengend der Job ist, und d) das Ausmaß, wie gefährlich der Job ist. Diesem Input, den ein Mensch ins Arbeitsleben investiert, steht in ihrem Verständnis als Output gegenüber, wie hoch das Einkommen ist, wie sicher der Job ist, wie interessant der Job ist, welche Entwicklungsmöglichkeiten er bietet, etc., d.h. letztendlich die klassische Auflistung von intrinsischen und extrinsischen Arbeitszufriedenheitsfacetten. Folglich stellen die Input-Größen im Prinzip die Investitionen des Mitarbeiters in die Arbeit dar, während die Output-Größen der Nutzen sind, den Mitarbeiter aus der Arbeit ziehen. In ihrer Analyse unterscheiden sie jeweils, ob in einem Land die entsprechenden Größen hoch oder niedrig ausgeprägt sind und betrachten dann das dort gemessene Zufriedenheitsniveau. So ist beispielsweise nachvollziehbar, daß bei geringem Input

und hohem Output eine hohe Zufriedenheit resultieren sollte. Dies trifft in der Tat für die Länder Dänemark, Schweiz, Niederlande und Norwegen zu. Umgekehrt ist ebenso verständlich, daß ein hohes Maß an Inputs und eine geringe Kompensation durch Outputs zu geringer Zufriedenheit führen sollte. Dies finden Souza-Poza und Souza-Poza (2000) in den Ländern Tschechische Republik, Bulgarien, Rußland und Ungarn wieder. Weniger offensichtlich ist jedoch, zu welcher Höhe von Zufriedenheit die Kombinationen aus geringem Input und geringem Output sowie hohem Input und hohem Output führen sollen. Es gäbe gute Gründe sowohl für die Annahme von hoher Zufriedenheit als auch für geringe Zufriedenheit. Eine klare Begründung liefern auch die Autoren nicht. In den empirischen Analysen der Autoren finden sich fast zwangsläufig auch jeweils mehrere Länder sowohl im Bereich hoher wie auch im Bereich geringer Zufriedenheit wieder. Für Zypern, Israel und die USA scheint zu gelten, daß hoher Input verbunden mit hohem Output zu hoher Zufriedenheit wird. Souza-Poza und Souza-Poza (2000) erklären dies mit einer in ihren Augen möglicherweise vorhandenen „give a lot, get a lot (and be satisfied)-Mentalität“ (S. 527), gehen jedoch nicht näher auf die möglichen Ursachen dieser Mentalität ein. Für Polen und Slowenien führt jedoch die gleiche Kombination aus Input und Output interessanterweise zu geringer Zufriedenheit. In umgekehrter Weise führen niedrige Inputs und niedrige Outputs in Spanien und Schweden zu hoher Zufriedenheit, in England, Frankreich und Japan aber zu geringer Zufriedenheit. Eine nähere Erklärung bleiben die Autoren auch hier schuldig. So gut wie gar nicht mit ihrem Input-Output-Modell können Souza-Poza und Souza-Poza (2000) erklären, warum in Italien bei hohem Input und niedrigem Output eine hohe Zufriedenheit resultiert und warum in Deutschland bei geringem Input und hohem Output eher eine Unzufriedenheit vorherrscht. So plausibel diese Art eines Kosten-Nutzen-Modells auf den ersten Blick auch erscheinen mag, es führt jedoch nur in 8 der untersuchten Länder zu einem schlüssigen Ergebnis. In weiteren 10 Ländern ist die Wirkung der Input-Output-Relation auf das Zufriedenheitsniveau zumindest nicht ganz eindeutig und kann nicht komplett geklärt werden. In zwei Ländern, u.a. Deutschland, kommt das Modell sogar zu kontraintuitiven Ergebnissen. Es scheinen zwar generell die Investitionen, die Mitarbeiter in ihrem Arbeitsumfeld tätigen, in jedem Land eine Rolle zu spielen. Ebenso haben sicherlich auch die Belohnungen, die sie aus dem Arbeitsleben ziehen, einen wichtigen Einfluß auf die resultierende Arbeitszufriedenheit, jedoch kann die

Relation zwischen beiden Größen somit scheinbar nicht für jedes Land das entsprechende Arbeitszufriedenheitsniveau erklären.

Auch Six und Felfe (2006) versuchen in ihrer Analyse der Zufriedenheitsdaten des Eurobarometers Ursachen für die Unterschiede zwischen den einzelnen Ländern zu finden. Ähnlich wie Souza-Poza und Souza-Poza (2000) fokussierten sie sich dabei auf die Arbeitsbedingungen, unterschieden dabei jedoch nicht zwischen Input- und Output-Faktoren. Stattdessen gingen beide Arten von Faktoren (d.h. Bezahlung, Arbeitsinhalt, Entwicklungschancen, Arbeitszeit, etc.) gleichberechtigt als unabhängige Variablen in ihre regressionsanalytischen Berechnungen ein, ohne sie gegenseitig in Bezug zueinander gesetzt zu haben. In ihren Berechnungen, die sie für jedes Land getrennt durchführten, zeigte sich, daß generell vor allem der Aufgabeninhalt, die Beziehungen zum Management sowie die auftretenden Arbeitsbelastungen signifikant mit dem Arbeitszufriedenheitsurteil zusammenhingen. Der Aufgabeninhalt war dabei im direkten Vergleich mit den anderen Aspekten der Arbeitsbedingungen vor allem in Dänemark, den Niederlanden, Finnland und Griechenland von großer Bedeutung, während in Spanien, Italien, Irland, Deutschland, Schweden und Großbritannien das Verhältnis zum Management eine vergleichsweise stärkere Bedeutung von den Befragten zugesprochen bekam. Die Belastungen des Arbeitslebens hatten hingegen vor allem in den Niederlanden und in Deutschland eine besonders negative Auswirkung. Als zusätzlicher Prädiktor zeigte sich lediglich Arbeitssicherheit in Spanien als vergleichsweise bedeutsam. Das Verhältnis zu Kollegen, der Entscheidungsspielraum und die Unterstützung von Seiten des Managements zeigten in den Regressionsanalysen dagegen so gut wie keine erklärende Wirkung. Als Gesamtfazit ihrer Analysen kommen Six und Felfe (2006) zu dem Urteil, daß sich insgesamt ein eher homogenes Bild zwischen den einzelnen europäischen Ländern zeigt und die Arbeitsbedingungen weitestgehend länderunabhängig auf die Arbeitszufriedenheit wirken. Als Prädiktor für die unterschiedlich hohen Arbeitszufriedenheitsniveaus scheinen sie sich demnach auch hier nicht wirklich zu eignen. Zusätzlich müssen zwei weitere Faktoren kritisch hinsichtlich der durchgeführten Analysen angemerkt werden. Auf der einen Seite stützen sich Six und Felfe (2006) nur auf die Selbstberichte der befragten Mitarbeiter. Somit ist eine gegenseitige Beeinflussung der Einschätzungen der Arbeitszufriedenheit und der jeweiligen Bedingungen nicht gänzlich auszuschließen, denn Urteile zur Ausprägung von Arbeitsbedingungen, zur Zufriedenheit mit diesen und zur Gesamtzufriedenheit sind oft nicht unabhängig voneinander (Borg, 2003). Eine

objektive und von den Urteilen der Befragten unabhängige Erfassung der Arbeitsbedingungen wäre hier vorteilhafter gewesen, wenngleich anzumerken ist, daß dies in der Tat durchaus schwer in internationalen Zufriedenheitsstudien zu realisieren ist. Eine umso bedeutsamere Kritik jedoch richtet sich gegen die methodische Vorgehensweise. Durch die für jedes Land getrennt durchgeführten Regressionsanalysen haben die Autoren im Prinzip nicht geprüft, ob unterschiedliche Arbeitsbedingungen in den einzelnen Ländern ein möglicher kausaler Faktor für die unterschiedliche ausgeprägten Arbeitszufriedenheitsniveaus sein können. Dafür wäre eine Analyse über alle Länder hinweg notwendig gewesen. Eine Analyse innerhalb jedes einzelnen Landes überprüft lediglich, ob die in einem Land relevanten Faktoren auch in einem anderen Land von Bedeutung sind. Obwohl ihre Ausgangsfrage eigentlich ausprägungs- bzw. Level-orientiert ausgerichtet war, liefert ihre statistische Analyse hingegen lediglich ein strukturorientiertes Ergebnis. Die Frage danach, warum die absoluten Zufriedenheitswerte über die Länder hinweg so stark variieren, kann somit nicht anhand dieser Studie beantwortet werden. Die Arbeitsbedingungen in den einzelnen Ländern scheinen somit insgesamt zwar durchaus Einflüsse auf das Arbeitszufriedenheitsurteil auszuüben, ihre unterschiedliche Ausprägung zwischen den Ländern kann aber offensichtlich nicht bzw. nicht eindeutig die Unterschiedlichkeit in der Arbeitszufriedenheit zwischen den Ländern erklären.

Judge et al. (2001) kommen in ihrem Überblicksartikel ebenfalls zu dem Schluß, daß ein Großteil der internationalen Forschung zwar durchaus zeigen kann, daß intrinsische Merkmale im Sinne des Job Characteristics Model (Hackman & Oldham, 1976) in verschiedenen Kulturen eine große Rolle für die Arbeitszufriedenheit zu spielen scheinen. Sie geben jedoch gleichzeitig zu bedenken, daß dieses Fazit mit Vorsicht zu genießen ist, da immer noch ein Großteil der Forschung rein amerikanisch oder aber westlich zentriert ist und daher mehr Stichproben, die gerade auch nicht-westliche Länder mit einbeziehen, von Nöten sind. Ein klares Fazit hinsichtlich der Arbeitsbedingungen läßt sich somit nicht ziehen.

Six und Felfe (2006) haben sich daher auch mit dem Einfluß von individuellen Werthaltungen auf das Arbeitszufriedenheitsurteil befaßt. Dabei unterscheiden sie zwischen a) Privat- und Freizeitorientierung, b) Orientierung am Arbeitsinhalt und c) Orientierung an Karriere und Materiellem. In ihrer 15 Länder umfassenden Stichprobe zeigte sich dabei, daß alle drei Werthaltungen auf Länderebene in einem starken Zusammenhang mit dem Arbeitszufriedenheitsniveau standen. Während die

Orientierung am Arbeitsinhalt in einem positiven Zusammenhang mit dem Zufriedenheitsurteil stand (Spearman-Rho = .49), korrelierten sowohl die Freizeitorientierung (Spearman-Rho = -.65) als auch die Orientierung an Karriere und Materiellem (Spearman-Rho = -.65) negativ mit dem allgemeinen Arbeitszufriedenheitsniveau. Individuelle Einstellungen und Wertüberzeugungen scheinen somit durchaus die Unterschiedlichkeit zwischen den Ländern hinsichtlich der Zufriedenheit erklären zu können. Genau genommen handelt es sich hier aufgrund der Mittelung individueller Werturteile auf Länderniveau aber nicht mehr um individuelle Wertüberzeugungen, sondern um eine Art Gesamtüberzeugungen oder auch Norm, wie man in dem jeweiligen Land über die Beziehung zur Arbeit denkt.

Während die Einstellungen zur Arbeit schon einen nicht unwichtigen Bestandteil der Lebenseinstellungen in den jeweiligen Ländern darstellen, ist ein umfassenderer und allgemeinerer Repräsentant für die insgesamt vorherrschende Lebenseinstellung in einer Nation die jeweilige Ausprägung der Kultur. Da bereits Judge et al. (2001) darauf verweisen, daß auf individuellem Niveau positive und negative Affektivität als dispositionelle Faktoren in Zusammenhang mit der persönlichen Arbeitszufriedenheit stehen, ist es nachvollziehbar, daß Müller (2006) versucht, Positivität als kulturellen Repräsentanten dieser Dispositionen mit der Höhe von Arbeitszufriedenheitsniveaus in Verbindung zu bringen. Gemäß dem Fazit seiner Studie ist Positivität im Vergleich mit einer Vielzahl kultureller Variablen der bedeutsamste Prädiktor für den Unterschied im Arbeitszufriedenheitsniveau. In seiner Studie, die drei voneinander unabhängige Länderstichproben vergleicht, scheint sich auf den ersten Blick Positivität als alleiniger bedeutsamer Prädiktor herauszukristallisieren, da dieses kulturelle Maß scheinbar als einzige unabhängige Variable über alle drei Stichproben hinweg bedeutsam mit dem Arbeitszufriedenheitsniveau als abhängige Variable korrelierte, während alle anderen Prädiktoren immer nur in einer oder maximal in zwei der Stichproben eine Korrelation aufwiesen. Leider stellt sich jedoch bei genauerer Prüfung seiner Analysen heraus, daß offensichtlich kleinere Zahlendreher die korrelativen Ergebnisse verzerrt haben. Aus den in seiner Studie aufgeführten Werten und der Darstellung seiner Ergebnisse läßt sich schließen, daß bei 2 der untersuchten 17 Länder in einer seiner Stichproben falsche Werte der Positivität (die aus Diener, Scollon, Oishi, Dzokoto & Suh, 2000 übernommen wurden) in die Korrelation mit dem Arbeitszufriedenheitsniveau eingegangen sein müßten (Italien: falsch = -.10, korrekt = .10; Österreich: falsch = .10, richtig = .01). Eine Korrektur der Werte und eine eigenständig durchgeführte

Neuberechnung anhand der von ihm veröffentlichten Daten ergibt dann hinsichtlich dieser einen Stichprobe nur noch eine geringere und nicht mehr signifikante Korrelation im Vergleich zum von ihm berichteten Zusammenhang von $r = .47$. Über die Korrektheit der Zusammenhänge in den anderen Stichproben kann aufgrund fehlender Veröffentlichungen der Werte leider keine Aussage getroffen werden. Aufgrund dieser falschen Berechnung allein in einer Stichprobe verschwindet jedoch die Eindeutigkeit hinsichtlich des Faktors Positivität, da er nun nicht mehr gleichermaßen über alle drei Stichproben hinweg als signifikanter Prädiktor in Erscheinung tritt. Er reiht sich stattdessen in eine Vielzahl von verschiedenen Zusammenhängen ein, die Müller (2006) zwar mal in einer der Stichproben, aber nicht in den anderen beiden Stichproben finden konnte und daher ebenfalls nicht als stichhaltige Ergebnisse interpretierte. Alles in allem beruht somit das zentrale Fazit dieser Studie zumindest in Teilen auf falschen Berechnungen und sollte daher nicht überinterpretiert werden. Die eindeutige Interpretation der Ergebnisse seitens Müller (2006) läßt somit Zweifel aufkommen. Des weiteren kann neben dieser methodischen Kritik auch in inhaltlicher Hinsicht kritisiert werden, daß Positivität von seinem theoretischen Verständnis nicht allzuweit von Arbeitszufriedenheit entfernt ist. Beide Konstrukte messen im Prinzip positiv konnotierte Empfindungen eines Befragten: bei Arbeitszufriedenheit wird die positive Sicht in Bezug auf ein spezifisch abgegrenztes Feld, d.h. die Arbeitswelt, erfaßt, während Gegenstand des Positivitäts-Konstrukts die Art der Empfindung und der Sichtweise gegenüber dem Leben generell ist, d.h. sozusagen die Stärke einer „rosaroten Brille“. Ein Zusammenhang ist daher nicht wirklich überraschend. Von größerem Erkenntnisgewinn wären daher auch Zusammenhänge mit den grundlegenden Kulturdimensionen, anhand derer sich Länder voneinander unterscheiden lassen.

Diese werden in der Tat in der ein oder anderen Veröffentlichung zur Erklärung der nationalen Unterschiede im Arbeitszufriedenheitsniveau herangezogen. Der beim deskriptiven Vergleich der Arbeitszufriedenheitsniveaus gefundene Unterschied zwischen westlichen und asiatischen Nationen legt dies durchaus auch nahe, denn ein bedeutsamer Unterschied zwischen diesen Gruppen von Ländern ist beispielsweise die unterschiedliche Werte-Ausprägung hinsichtlich Individualismus und Kollektivismus. Demnach scheint es so zu sein, daß gerade in kollektivistischen Nationen die Arbeitszufriedenheit geringer ausgeprägt ist. Das Interesse an der Messung des Zusammenhangs zwischen diesem Wertekonstrukt und Arbeitszufriedenheit hat gemäß der umfassenden Literaturübersicht von Judge et al. (2001) auch im Forschungsbereich

durchaus zugenommen. Mehrere Studien beschäftigten sich bereits Ende der 90er Jahre explizit mit dieser Thematik, sowohl auf dem Level von Nationen (bspw. Chiu & Kosinski, 1999) als auch auf dem Level von Individuen (Hui & Yee, 1999; Hui, Yee & Eastman, 1995). Chiu und Kosinski (1999) fanden diesbezüglich – wie bereits zuvor erwähnt –, daß die Stichprobe des individualistischeren Landes eine höhere Zufriedenheit berichtete als die Vergleichsgruppe aus dem kollektivistischen Land. Hui und Yee (1999) sowie Hui, Yee und Eastman (1995), die in ihren Studien jeweils Arbeitnehmer aus Hongkong befragten, fanden hingegen einen positiven Zusammenhang zwischen Kollektivismus und Arbeitszufriedenheit. Sie betrachteten jedoch auch den Zusammenhang auf individuellem Niveau.

Nach der Durchsicht der in ihrer Literaturübersicht zusammengetragenen Studien kommen Judge et al. (2001) letztendlich zu dem Schluß, daß sich hinsichtlich der Individualismus-Kollektivismus-Dimension interessanterweise ein entgegengesetzter Zusammenhang zu zeigen scheint, je nachdem ob man das individuelle oder das nationale Level betrachtet. Während bei Vergleichen der Mittelwerte von Ländern individualistischerer Nationen oftmals höhere Arbeitszufriedenheitswerte aufweisen, scheint sich das Ergebnis auf individuellem Niveau umzudrehen. Judge et al. (2001) übersehen bei ihrer Interpretation jedoch, daß die genannten Studien, die das Individuallevel untersuchten, auch in einem kulturell kollektivistischen Umfeld durchgeführt wurden. Inwieweit dies die erstaunlicherweise entgegengesetzten Zusammenhänge auf dem individuellen Level beeinflußt hat, kann somit nicht endgültig geklärt werden. Möglicherweise würde dies einfach nur bedeuten, daß sich einfach das kulturell vorherrschende Denken positiv auf die Arbeitszufriedenheit auswirkt und in individualistischen Nationen sich dennoch wieder ein positiver Zusammenhang zwischen individualistischem Denken und der Zufriedenheit mit der Arbeit zeigen könnte. Der Zusammenhang auf individuellem Niveau bzw. auch die Interaktion von kulturellem Umfeld und der individuellen kulturellen Sichtweise kann somit nicht abschließend geklärt werden. Bei der Betrachtung der Zusammenhänge auf dem Nationenlevel scheint sich jedoch anzudeuten, daß dort ein positiver Zusammenhang zwischen Individualismus und Arbeitszufriedenheit besteht. Jedoch gestehen auch Judge et al. (2001) zu, daß es noch weiterer Überprüfungen diesbezüglich bedarf.

Ein oft anzutreffendes Problem bei der Untersuchung von Kultur als ursächlichem Faktor von Unterschieden in den Arbeitszufriedenheitsniveaus ist dennoch, daß sich solche Schlußfolgerungen vermehrt dann antreffen lassen, wenn die untersuchten

Kulturen geographisch sehr weit voneinander entfernt liegen (Six & Felfe, 2006). Dann scheint es intuitiv naheliegend zu sein, daß die scheinbar sehr verschiedene Kultur der Grund für Unterschiede in der abhängigen Variable sein muß. Liegen die Länder jedoch eng beieinander und sind vielleicht sogar Nachbarstaaten, wird oft vorschnell davon ausgegangen, daß sich die Kulturen nicht stark voneinander unterscheiden könnten und Kultur daher kein passender Prädiktor sein könne. Jedoch finden sich allein im innereuropäischen Vergleich trotz direkter Nachbarschaft durchaus starke kulturelle Unterschiede bspw. zwischen Deutschland und den Niederlanden oder auch Deutschland, der Schweiz und Österreich, obwohl letztere sogar die gleiche Sprache miteinander teilen. Um diese Kritik zu entkräften, sollten daher möglichst immer kulturelle Werte explizit erhoben sowie analysiert werden und sich nicht auf die intuitive Wahrnehmung von Kulturen verlassen werden.

Wenn sich der bereits andeutende Zusammenhang von Individualismus und einem hohen Arbeitszufriedenheitsniveau in weiteren Studien bestätigen würde, dann wäre dies zwar für sich genommen ein interessantes Ergebnis. Eine genaue Erklärung für die unterschiedlich hohen Niveaus wäre es hingegen per se nicht. Es würde sich vielmehr die Anschlußfrage stellen, warum genau Kultur dazu führt, daß die abgegebenen Urteile zur Frage der Arbeitszufriedenheit gerade in individualistischeren Nationen höher sind, während sie in kollektivistischen Nationen gemäßiger ausfallen. Es sei an dieser Stelle darauf hingewiesen, daß der positive Zusammenhang von Individualismus und Zufriedenheit keineswegs im Umkehrschluß bedeutet, daß kollektivistische Nationen im Mittel unzufrieden mit der Arbeit wären. Da sich zeigte, daß insgesamt auch auf Länderniveau die Zufriedenheitswerte in den positiven Bereich der Skala verschoben sind, bedeutet es vielmehr, daß sich kollektivistische Nationen zwar auch positiv zur Arbeit äußern, jedoch eben eher gemäßigt und nicht so extrem positiv wie individualistischere Nationen. Warum jedoch sind diese Zufriedenheitswerte in manchen Nationen gemäßiger, während sie in anderen Nationen extremer sind? Solange diese Frage nicht beantwortet ist, steht zwischen dem Kulturwert Individualismus/Kollektivismus auf der einen Seite und dem Arbeitszufriedenheitsniveau auf der anderen Seite im Prinzip immer noch eine Art Blackbox.

Man könnte verschiedene Gründe finden, die gegebenenfalls eine Brücke zwischen den beiden Konstrukten schlagen könnten und vielleicht Licht in das Dunkel der Blackbox bringen könnten. In einem ersten Schritt könnte man vermuten, daß in

individualistischen Nationen die Einstellung zur Arbeit und damit die Motivation eine andere ist. Ein höherer intrinsischer Bezug zur Arbeit (siehe die Ergebnisse von Six & Felfe, 2006) könnte auf der einen Seite dazu führen, daß sich Menschen eher mit ihrer Arbeit identifizieren und daher mit ihr zufriedener sind. Auf der anderen Seite steigt aber auch mit einem höheren Commitment zur Arbeit die Anspruchshaltung an die Arbeit. Daher würde wahrscheinlich auch der Anspruch hinsichtlich Selbstverwirklichungs- und Entfaltungsmöglichkeiten steigen. Ein hohes Zufriedenheitsniveau wäre jedoch dann wiederum nur zu erwarten, wenn im Sinne einer Input-Output-Relation (vgl. Souza-Poza & Souza-Poza, 2000) ein günstiges Verhältnis vorläge und die hohen Ansprüche auch auf eine Breite an Möglichkeiten treffen würde. Eine Argumentationskette in dieser Form würde somit allein zwei Prämissen aufstellen, die erst einmal empirisch überprüft werden müßten. Von einer simplen linearen Beziehung in der angenommenen Richtung ist grundsätzlich dabei nicht auszugehen. Es wäre beispielsweise unangemessen, einfach anzunehmen, daß in individualistischeren Nationen sich grundsätzlich Wege zur Selbstverwirklichung in der Arbeitswelt bieten, während in kollektivistischeren Nationen diese Wege zumeist nicht existent wären. Ebenso bedarf allein die Annahme, daß nur die Menschen in individualistischen Nationen zur Selbstverwirklichung neigen, während Menschen kollektivistischer Herkunft dies nicht täten, erst einer empirischen Verifizierung. Solch eine Argumentation erschiene zwar durchaus plausibel, würde jedoch oftmals einer empirischen Grundlage entbehren. Sie versucht im Prinzip die Blackbox mit mehreren Erklärungsschritten und –variablen zu füllen, die für sich genommen bisher nicht zwingend bestätigt sind und mit sich selbst auch wieder in Wechselwirkung stehen können, so daß damit letztendlich eine Erklärung der Blackbox immer noch nicht stattgefunden hat. Hinzu kommt, daß je komplexer die erdachten Zusammenhänge innerhalb der Blackbox sind, desto schwieriger – wenn nicht sogar unmöglich – wird es, den empirischen Beweis anzutreten. Darüber hinaus würde solch eine Argumentation das Individuallevel mit dem Nationenlevel vermischen. Während die Erklärungsversuche individualpsychologische Zusammenhänge beschreiben, liegt das zu lösende Problem – nämlich Unterschiede im Arbeitszufriedenheitsniveau – aber auf dem Nationenlevel. Es muß somit bei der Suche nach erklärenden Faktoren darauf geachtet werden, daß die Faktoren nationenspezifische Maße darstellen, welche ein inter-kulturelles und kein inter-individuelles Phänomen beschreiben.

Ein interkulturelles Phänomen, welches hier als weiterer naheliegender und ebenso plausibler Faktor herangezogen werden könnte, der zwischen dem Zusammenhang von Individualismus und Arbeitszufriedenheit vermitteln könnte, ist schließlich das Kommunikationsverhalten von Menschen. Da die Kultur Menschen in ihrer Art und Weise sozialisiert, wie sie miteinander umgehen, hat Kultur somit auch einen Einfluß auf die Kommunikation zwischen Menschen. Die Messung von Arbeitszufriedenheit ist aber im Wesentlichen auch nichts anderes als Kommunikation; Kommunikation zwischen dem Forscher und dem Probanden, zwischen dem Interviewer und dem Interviewten bzw. ganz allgemein zwischen dem Untersuchungsteilnehmer und dem (Ab-)Sender der Untersuchungsfrage. In der Forschung zur interkulturellen Arbeitszufriedenheit ist dieser Faktor bisher jedoch überhaupt nicht berücksichtigt worden, obwohl die Forschung im Bereich der interkulturellen Psychologie diese Erklärung nahelegen würde. Aus diesem Grund soll diese hier vorliegende Studie überprüfen, inwieweit tatsächlich Kommunikationsnormen als ein vermittelnder Faktor gelten können, um erklären zu können, warum manche Kulturen sich tendenziell höher und andere Kulturen tendenziell gemäßiger zur Arbeitszufriedenheit äußern.

Wenn sich erneut bestätigen ließe, daß verschiedene Kulturen auch ein unterschiedliches Antwortverhalten zeigen (Smith, 2004), und diese Kulturen auch in gleicher Richtung sich auch bei der Betrachtung von Arbeitszufriedenheit unterschieden, wäre es möglich, daß sich Unterschiede im Arbeitszufriedenheitsniveau in weiten Teilen auf das unterschiedliche, kulturell sozialisierte Antwortverhalten zurückführen ließen. Wenn ein Mensch in einer bestimmten Kultur aufwächst, würde er somit in einer gewissen Art und Weise trainiert und sozialisiert, auf Fragen des Gegenüber zu antworten und würde dann folglich auch eine Zufriedenheitsfrage gemäß seiner Vorstellung über eine angemessene Kommunikation beantworten. Unterschiedliche Arbeitszufriedenheitsniveaus in unterschiedlichen Ländern könnten demnach in Teilen ein Ausdruck von unterschiedlichen Kommunikationsnormen sein, die in dem jeweiligen Land untrennbar mit der entsprechenden kulturellen Denkweise verbunden sind. Solche Ansätze der Interpretation finden sich in Teilen durchaus schon in der einschlägigen Literatur. Beispielsweise argumentiert Marion-Landais (1993), daß die überraschend hohe Ausprägung der Zufriedenheit in der Dominikanischen Republik möglicherweise daran liegen könnte, daß die Mitarbeiter sich nicht trauten, ihre Unzufriedenheit gegenüber dem Chef zu äußern. Diese Interpretationen wurden jedoch bisher nicht konsequent empirisch überprüft, so daß dies ein Ziel der hier vorliegenden

Studie ist. Dabei sollte noch einmal betont werden, daß die Studie vorrangig den Fokus auf die Einflußfaktoren für die Niveauunterschiede zwischen verschiedenen Ländern legt und damit eine Level-orientierte Fragestellung verfolgt. Die zahlreichen Einflußfaktoren, die auf individueller Ebene die persönliche Arbeitszufriedenheit beeinflussen, sollen bewußt nicht vertiefend thematisiert werden. Lediglich die individuelle Kommunikationstendenz soll als ein Individualfaktor aufgrund der spezifischen Thematik dieser Arbeit auf seine Erklärungskraft hin geprüft werden.

5 Methodische Aspekte interkultureller Forschung

Wenn Forschung im Kontext verschiedener Nationen und Kulturen durchgeführt wird, stellen sich in der Regel besondere methodische Probleme. Dies gilt sowohl für die Messung und den Vergleich von Arbeitszufriedenheit über Nationengrenzen hinweg als auch im Prinzip für die Messung jedes beliebigen Konstrukts, für das interkulturelle Vergleiche angestellt werden sollen. Die einschlägige Forschungsliteratur würde diesbezüglich der Aussage von Berry (1997) folgen, wonach „most areas of psychology enquiry are defined by their content; however, cross-cultural psychology is defined primarily by its method“ (S. 1). Zwar finden sich methodische Probleme in jeder Art von Forschung, jedoch sind gerade Messungen im interkulturellen Bereich von speziellen methodischen Problemen betroffen (Triandis, 1994). Dies betrifft verschiedenste Aspekte, angefangen bei der Frage nach der Universalität eines Konstruktes, über die Operationalisierung des Konstruktes in verschiedenen Kulturen, bis hin zur Interpretation der gemessenen Unterschiede.

5.1 Kulturabhängigkeit vs. Kulturunabhängigkeit

Bezüglich der Diskussion über die Universalität von psychologischen Konstrukten gibt es in der Forschung im Allgemeinen zwei verschiedene Sichtweisen. Unterschieden wird dabei zwischen sogenannten Etic- und Emic-Ansätzen (Berry, 1997; Berry, Poortinga, Segall & Dasen, 1992; Triandis, 1972, 1994). Anhänger des Etic-Ansatzes gehen davon aus, daß es Konstrukte gibt, deren Bedeutung international, interkulturell und somit weltweit gültig ist. Alle Menschen ‚funktionieren‘ gemäß diesem Ansatz nach den gleichen Gesetzmäßigkeiten. Diese Idee vertritt somit die Auffassung von der Universalität psychologischer Phänomene. Ziel der Forschung ist beim Etic-Ansatz, anhand des gleichen Konstruktes unterschiedliche Kulturen miteinander zu vergleichen. Es wäre somit möglich, ein Fragebogentool in einer Kultur zu entwickeln, in eine andere Kultur zu übertragen und dort in gleicher Form einzusetzen.

Die Ansichten des Emic-Ansatzes sind dieser Auffassung komplett entgegengesetzt. Gemäß dieser Sichtweise ist es schwer wenn nicht gar unmöglich, ein Konstrukt, welches eine hohe Bedeutsamkeit in einer Kultur haben mag, einfach in eine andere Kultur zu übertragen. Zumindest ist nach diesem Ansatz davon auszugehen, daß es im interkulturellen Bereich nicht angemessen ist, einfach von der grundsätzlichen

Übertragbarkeit auszugehen, ohne diese zu hinterfragen. Der emische Ansatz interkultureller Forschung geht vielmehr davon aus, daß in jeder Kultur spezielle Charakteristika gelten, die zu jeweils spezifischen Konstrukten führen. Diese können sich zwar im Kern überlappen, sind jedoch in der Regel in ihren Teilen unterschiedlich und bestehen aus Sub-Aspekten, die einzigartig für die jeweilige Kultur sind. Die Erfassung des Konstrukts anhand ein und desselben Fragebogens, der nur in einer Kultur entwickelt wurde, wäre demnach nicht möglich. Stattdessen sollte das Verständnis eines Konstruktes aus der jeweiligen Kultur heraus entwickelt werden. Eine extreme Variante dieses Verständnisses würde sogar inhaltlich soweit gehen, daß manche Konstrukte nur in einer Kultur von Bedeutung sein könnten, während sie in anderen Kulturen möglicherweise überhaupt keine Relevanz besitzen und ein Vergleich somit überhaupt gar nicht möglich und sinnvoll wäre.

Neben dieser klassischen dualen Unterscheidung teilen beispielsweise Berry et al. (1992) die verschiedenen Ansätze in insgesamt drei Bereiche auf: in einerseits Absolutismus (=etic) und andererseits Relativismus (=emic) sowie darüber hinaus zusätzlich in Universalismus. Obwohl dieser neue dritte Ansatz aufgrund seiner begrifflichen Ähnlichkeit leicht zu verwechseln wäre mit der etischen Sichtweise, da diese ja von universalen Gesetzmäßigkeiten ausgeht, steckt jedoch ein anderes Verständnis dahinter. Anhänger des Universalismus gestehen demnach zwar durchaus zu, daß es allgemeine Konstrukte geben mag, die für alle Kulturen von Bedeutung sind. Ihre Manifestation jedoch kann sich aber sehr wohl von Kultur zu Kultur unterscheiden. Somit müßten diese verschiedenen Manifestationen auch mit unterschiedlichen Instrumenten erhoben werden, selbst wenn inhaltlich dasselbe gemessen werden soll. Dieses Verständnis findet sich auch bei Triandis (1994) wieder, wenn er darauf hinweist, daß zwar in vielen Fällen von der Existenz von Etic-Konstrukten ausgegangen werden kann, jedoch die Argumente des Emic-Ansatzes dazu genutzt werden können, dieses universelle Konstrukt kulturspezifisch zu messen und zu erfassen. Ein Vergleich der gemessenen Werte wäre dann durchaus über Kulturgrenzen hinweg möglich, selbst wenn das Instrument zur Erfassung sich kulturspezifisch unterscheidet. Für diese Art der Vorgehensweise hat sich in der einschlägigen Literatur der Begriff des ‚derived etic approach‘ durchgesetzt (Berry, 1969).

Diese Unterscheidung der Ansätze zum Forschungsverständnis im interkulturellen Bereich lassen sich somit erst einmal auf einem eindimensionalen Kontinuum anordnen: vom komplett kulturunabhängigen Verständnis auf der einen Seite, über die

Verbindung von kulturunabhängigen Konstrukten mit kulturabhängiger Manifestation, bis hin zum komplett kulturabhängigen Verständnis auf der anderen Seite (vgl. Abbildung 3). Von links nach rechts nimmt somit die Kulturabhängigkeit psychologischer Prozesse zu.

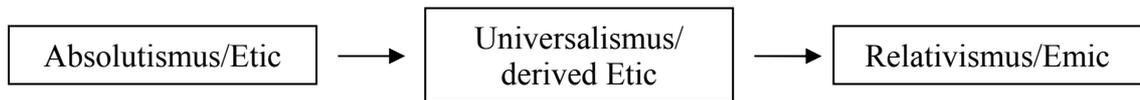


Abbildung 3: Eindimensionaler Vergleich interkultureller Forschungsansätze

Neben diesen drei Ansätzen findet sich bei Huang (2002) noch eine weitere Unterscheidung. Er bezeichnet die Vorgehensweise seiner Studien, die neuerdings auch in anderen Teilen der Forschung an Bedeutung zunimmt, als ‚komparativen Relativismus‘ und erweitert die eindimensionale Kategorisierung zusätzlich um eine zweite Dimension, anhand derer sich die verschiedenen Forschungsansätze einordnen lassen (vgl. Abbildung 4).

Durchführung interkultureller Vergleiche

		Nein	Ja
Gültigkeit psychologischer Prozesse	Universell	<i>Absolutismus</i>	<i>Universalismus</i>
	kultur-spezifisch	<i>Relativismus</i>	<i>komparativer Relativismus</i>

Abbildung 4: Forschungsansätze im interkulturellen Kontext, gemäß Huang (2002)

Während sich vertikal unterscheiden läßt, inwieweit angenommen wird, ob psychologische Prozesse allgemein gültig und somit kulturunabhängig oder aber kultur-spezifisch sind, unterscheidet Huang (2002) horizontal dahingehend, ob davon ausgegangen wird, daß interkulturelle Vergleiche sinnvoll sind und durchgeführt werden oder ob sie in dem jeweiligen Verständnis nicht für lohnenswert erachtet werden. Gemäß dieser Einteilung findet sich folglich der Absolutismus oben links

wieder, da in diesem Verständnis von einer universellen Gültigkeit der Konstrukte ausgegangen wird und daher ein interkultureller Vergleich nicht notwendig und möglicherweise nicht sinnvoll ist, da die Prozesse bei allen Menschen weltweit sowieso gleich funktionieren. Relativismus dagegen findet sich unten links wieder, da auch dort keine interkulturellen Vergleiche angestrebt werden; hier jedoch nun aufgrund der Annahme, daß jedes Konstrukt nur in der jeweiligen Kultur selbst sinnvoll gemessen und aus ihr heraus verstanden werden kann. Universalismus hingegen läßt sich oben rechts einordnen, da trotz des Zugeständnisses universeller Gültigkeit psychologische Konstrukte diese dennoch über Kulturgrenzen hinweg verglichen werden sollen. Die neue Unterscheidung von Huang (2002) in komparativen Relativismus schließlich findet sich unten rechts wieder. Unter ‚komparativem Relativismus‘ wird dabei verstanden, daß versucht wird, interkulturelle Unterschiede in psychologischen Prozessen auf dem Individuallevel anhand von nationalen Unterschieden auf dem kulturellen Level zu erklären. Dies geschieht in der Regel mit Hilfe von Mehrebenenanalysen. Auf der einen Seite werden somit zwar durchaus kulturspezifische Prozesse angenommen. Auf der anderen Seite jedoch wird ein Vergleich dieser unterschiedlichen Prozesse über verschiedene Länder hinweg explizit angestrebt, und es wird versucht, genau diese Unterschiedlichkeit anhand von länderbezogenen Charakteristika zu erklären. Hinsichtlich dieser neuen vierten Form interkultureller Forschung unterscheidet Huang (2002) darüber hinaus zwischen zwei verschiedenen Arten, wie dann Erklärungen dieser Unterschiede in der Forschung oft entwickelt werden: a) durch ad hoc Erklärungsmodelle (ad hoc comparative model) und b) durch schlußfolgernde Erklärungsmodelle (inferential comparative model). Erstere Modelle versuchen, Unterschiede im Zusammenhang individual-spezifischer Variablen lediglich anhand von naheliegenden nationalen Charakteristika zu erklären, die jedoch zumeist nicht explizit gemessen wurden. Letztere Modelle hingegen messen in der Tat nationale Unterschiede z.B. hinsichtlich sozioökonomischer Variablen oder auch hinsichtlich Kulturdimensionen und versuchen diese qualitativ und quantitativ mit den Unterschieden in den individuellen Prozessen in Verbindung zu bringen. Insgesamt gesehen erfaßt somit Huangs (2002) neuer Terminus und seine neue Einordnung eine weitere wichtige Sichtweise, die aktuell in der interkulturellen Forschung zunehmend an Bedeutung gewinnt und die jedoch in den bisherigen Einteilungen noch keinen angemessenen Platz finden konnte.

Welche Sichtweise letztendlich die Richtige ist, kann so ohne weiteres nicht bestimmt werden, denn für alle Ansätze bringen die jeweiligen Anhänger gute Gründe an, warum die eigene Sichtweise die einzig korrekte ist. Die Argumente basieren dabei sowohl auf grundsätzlichen inhaltstheoretischen Unterscheidungen von qualitativer und quantitativer Forschung (Kagitcibasi & Poortinga, 2000) als auch auf zum Teil gegensätzlichen Ansichten der Naturwissenschaften auf der einen Seite und der Geisteswissenschaften auf der anderen Seite (Jahoda, 1980). Konkrete Beispiele und empirische Belege gibt es ebenso für jeden der Ansätze. So lassen sich als ein Beispiel für Absolutismus sicherlich die grundlegenden Wahrnehmungsprozesse von Menschen heranziehen, da bei diesen am ehesten davon ausgegangen werden kann, daß sie bei allen Menschen weitestgehend nach den gleichen Prinzipien funktionieren. Beispiele für Relativismus lassen sich vor allem im Bereich der „indigenous psychology“ finden (Kim, Park & Park, 2000). Dort wird vor allem versucht, kulturimmanente Phänomene aus dem jeweiligen Kulturraum selbst heraus zu verstehen und so das Verständnis vor allem über eher wenig erforschte Kulturen zu verbessern (z.B. afrikanische Kulturen, Kulturen Vorderasiens oder auch Kulturen von Ureinwohnern wie Aborigines). Ein gut geeignetes Beispiel für komparativen Relativismus ist die Studie von Huang (2002) selbst, in der er den unterschiedlichen Zusammenhang von Belohnungen hinsichtlich intrinsischer und extrinsischer Arbeitswerte einerseits und Arbeitszufriedenheit andererseits anhand der Kulturunterschiede des jeweiligen Landes erklärt. Als Ergebnis zeigt sich hier beispielsweise, daß der in der Forschung gemeinhin angenommene positive Zusammenhang von intrinsischen Belohnungen und Arbeitszufriedenheit in individualistischen Kulturen deutlich stärker ist als in kollektivistischen Kulturen (siehe auch für eine weitere ähnliche Studie: Huang & van de Vliert, 2002). In gleicher Weise relativiert sich der bis dato scheinbar gut bestätigte positive Zusammenhang von Job-Hierarchie (job level) und Arbeitszufriedenheit, wenn man eine komparativ relativistische Perspektive einnimmt. Daß die Arbeitszufriedenheit in der Regel steigen sollte, je höher jemand in der Job-Hierarchie aufsteigt, gilt so nur in individualistischen Ländern. In kollektivistischen Ländern gibt es jedoch keinen Zusammenhang zwischen den beiden Aspekten (Huang & van de Vliert, 2004). Es ist jedoch wenig überraschend, daß an vielen der bisher für bestätigt gehaltenen Zusammenhänge im Sinne des Absolutismus so lange festgehalten wurde und auch in Teilen noch festgehalten wird, denn ein Großteil der Forschung in diesem Bereich findet bekanntlich im angloamerikanischen und somit weitestgehend individualistisch geprägten Kulturraum

statt. Speziell die komparativ relativistische Sichtweise ist jedoch somit in der Lage, die bisherigen Erkenntnisse zu erweitern und sie auch für andere Kulturräume überprüfen zu können. Für die letzte der vier Sichtweisen, den Universalismus, finden sich schließlich Beispiele vor allem im Bereich von Studien, die die Konstruktvalidität verschiedener psychologischer Phänomene in verschiedenen Kulturräumen überprüfen. Borg (1986) konnte in dieser Hinsicht beispielsweise zeigen, daß die klassische Dreiteilung von Arbeitswerten in kognitive, affektive und instrumentelle Werte auch in anderen, nicht-westlichen Kulturen Bestand hat. Auch die Überprüfung der interkulturellen Gültigkeit der Big-Five-Persönlichkeitsdimensionen wäre ein weiteres prominentes Beispiel für diesen Ansatz (siehe Hofstede & McCrae, 2004). Angenommen wird dabei in der Regel, daß die bekannten zentralen Faktorenstrukturen auch in anderen Kulturräumen bestehen bleiben, selbst wenn die Faktorladungen oder auch Sub-Faktoren sich von Kultur zu Kultur leicht unterscheiden mögen.

Versucht man die hier vorliegende Arbeit in eine der Sichtweisen einzuordnen, so läßt sie sich in allererster Linie erst einmal dem universalistischen Ansatz bzw. einem *derived-etic*-Verständnis zuschreiben. Daß Arbeitszufriedenheit eine Bedeutung nicht nur in individualistischen Kulturen, sondern auch in kollektivistischen Kulturen hat, soll überhaupt nicht angezweifelt werden, denn Menschen müssen im Prinzip in allen Kulturen arbeiten, um sich einen Lebensunterhalt zu verdienen. Dabei wird es für jeden Menschen von Bedeutung sein, inwieweit er sich mit seiner Arbeit anfreunden kann; sprich, ob er mit ihr zufrieden ist. In dieser Hinsicht wird somit ebenfalls von einem universellen Konstrukt ausgegangen, bei dem sich die Bedeutung in verschiedenen Kulturen nicht grundlegend unterscheidet. Daß Arbeitszufriedenheit eine Bedeutung in den unterschiedlichsten Kulturräumen hat, konnte die Forschung ja bereits zeigen (vgl. Judge et al., 2001; Six & Felfe, 2006). Fraglich ist jedoch, ob es möglich ist, trotz dieser allgemeinen interkulturellen Gültigkeit von Arbeitszufriedenheit dieses Phänomen auch in gleicher Weise zu erfassen. Es könnte demnach durchaus möglich sein, daß beispielsweise die Art der Frage überhaupt nicht passend für das jeweilige kulturelle Denken ist. Da in der Regel eine Frage wie „Wie zufrieden sind Sie alles in allem mit Ihrem Job?“ gestellt wird und somit nach dem individuellen Erleben einer einzelnen Person ohne Bezug zu seinem Kontext gefragt wird, könnte allein schon hinter die Übertragung dieser simplen Frage von individualistischen zu kollektivistischen Kulturen ein Fragezeichen gesetzt werden. Selbst wenn aber inhaltlich gesehen die gleiche Frage in verschiedenen Kulturen grundsätzlich in der Lage wäre, die

Zufriedenheit mit der Arbeit anzusprechen, wäre immer noch nicht geklärt, ob in manchen Kulturen möglicherweise die Antwort auf diese Frage durch verschiedenen Antworttendenzen verzerrt wird und sozusagen überformt ist. Von einer weltweiten Gültigkeit im Sinne des Absolutismus könnte demnach bei der Arbeitszufriedenheitsforschung nicht ausgegangen werden, sollte sich diese Möglichkeit bestätigen. Einerseits wäre somit das Verständnis zwar durchaus universalistisch, die Messung und die Interpretation des Konstruktes würde jedoch auch emische Aspekte mit einbeziehen und somit insgesamt die Sichtweise eines *derived-etic*-Verständnisses widerspiegeln.

Da in der Arbeit aber auch zusätzlich Analysen im Sinne der Mehrebenenthematik durchgeführt werden, könnte man die Arbeit aber durchaus auch dem komparativ relativistischen Verständnis zuordnen. Wenn also beispielsweise der Zusammenhang von Antworttendenzen und Antworten zur Arbeitszufriedenheit auf individuellem Niveau untersucht wird und mögliche Unterschiede in diesem Zusammenhang auf national-unterschiedliche Faktoren zurückgeführt werden können, dann spiegelt dieses Vorgehen eher das Denken dieses vierten Ansatzes wider.

5.2 Nationale Einflußfaktoren

Folgt man nun den Überzeugungen des universalistischen oder des komparativ relativistischen Ansatzes, so gesteht man grundsätzlich zu, daß Variablen, die sich von Nation zu Nation und von Kultur zu Kultur unterscheiden, Einfluß auf innerkulturelle Phänomene nehmen können. Unmittelbar daran anschließend stellt sich jedoch die Frage, welche Variablen dies denn sein könnten und welche folglich in einer interkulturellen Forschungsstudie untersucht werden sollten. Dabei sollte jedoch zuerst einmal darauf hingewiesen werden, daß es überhaupt schon positiv zu erwähnen ist, wenn in Studien, die Daten aus verschiedenen Ländern vergleichen, tatsächlich nationale oder kulturelle Maße explizit erhoben werden, denn eine Kritik, die interkulturellen Studien oft vorgeworfen wird und bereits bei der Diskussion des Modells von Huang (2002) kurz erwähnt wurde, ist, daß bei gefundenen Unterschieden zwischen verschiedenen Länderstichproben diese zum Teil vorschnell auf kulturelle Unterschiede zurückgeführt werden, obwohl in den jeweiligen Studien überhaupt keine kulturellen Maße erfaßt wurden. Diese vorschnellen Schlüsse werden dabei meistens umso eher getroffen, je stärker die erfaßten Länder auf den ersten Blick bezüglich ihrer

kulturellen Eigenschaften unterschiedlich erscheinen. Dies trifft beispielsweise oft auf den Vergleich von klassisch westlichen Ländern (z.B. den USA) und asiatischen Ländern (z.B. Japan oder China) zu. Bei solchen Ländervergleichen wird dann häufig automatisch ein in den Daten gemessener Unterschied in einem Konstrukt durch die individualistische Kultur westlicher Staaten und umgekehrt durch die kollektivistische Kultur asiatischer Staaten erklärt. Ohne jedoch die entsprechenden Einflußgrößen explizit gemessen zu haben, sind solche Schlußfolgerungen nicht haltbar (van de Vijver & Leung, 2000). Es ist zwar generell möglich, daß Kultur tatsächlich den entscheidenden Unterschied ausmacht, jedoch muß dies immer erst methodisch und statistisch abgesichert werden. Eine explizite Messung von Maßen zur Erfassung kultureller Unterschiede muß somit immer dann erfolgen, wenn denn angenommen wird, daß Kultur als erklärende Variable einen bedeutsamen Einfluß ausüben könnte. Dies könnte einerseits durch den Einsatz der entsprechenden Meßinstrumente aus den bedeutsamsten kulturellen Vergleichsstudien geschehen, d.h. durch den VSM 94 von Hofstede, durch den SVS (Schwartz-Value-Survey) sowie den PVQ (Portrait-Value-Questionnaire) von Schwartz oder durch das Meßinstrument der GLOBE-Studie, so daß letztendlich eigene neue Werte für die in der jeweiligen Studie untersuchte Stichprobe vorliegen. Andererseits besteht ebenso die Möglichkeit, die in den einschlägigen Veröffentlichungen von Hofstede, Schwartz oder der GLOBE-Studie publizierten Gesamtwerte der jeweiligen Kulturdimensionen für jedes Land in die eigene Studie mit einfließen zu lassen und sie als stellvertretenden Kennwert für die Ausprägung der kulturellen Werte der eigenen Länderstichproben in der Untersuchung zu nutzen. Beide Vorgehensweisen haben dabei ihre spezifischen Vor- und Nachteile. Während bei der Messung eigener Kulturwerte für jede neue Stichprobe auf der einen Seite zwar eine höhere Meßgenauigkeit im Sinne der korrekten Abbildung kultureller Wertvorstellungen der Studienteilnehmer angenommen werden kann, besteht jedoch auf der anderen Seite – neben dem verständlicherweise erhöhten Aufwand – in Teilen auch das Problem der ‚single-source‘-Messung und somit die Konfundierung mehrerer Meßgrößen, die in ein und derselben Studie nacheinander erhoben werden (vgl. Podsakoff, McKenzie, Lee & Podsakoff, 2003; Söhnchen, 2007). Die genau umgekehrte Problematik tritt auf, wenn man sich auf bereits veröffentlichte kulturelle Kennwerte verläßt und diese in die eigene Studie mit aufnimmt. Während hier der ‚single-source‘-Fehler zwar ausgeschaltet wird, bleibt weitestgehend offen, inwieweit die alten, bereits früher erhobenen Kulturwerte für die eigene neue Stichprobe noch Gültigkeit besitzen.

Diese Problematik ist jedoch in der Regel eng mit der Frage nach der Art der Stichprobe und deren Rekrutierung verknüpft. Gerade wenn für das Land untypische Stichproben erfaßt werden (z.B. lediglich Studenten einer bestimmten Universität oder Arbeiter aus lediglich einer Fabrik), rückt diese Frage in den Vordergrund. Bei der Auswahl von repräsentativen Landesstichproben, die unter anderem in der hier vorliegenden Studie zum Einsatz kommen, kann hingegen davon ausgegangen werden, daß die in früheren Studien erhobenen kulturellen Kennwerte auch für die neue Studie die jeweilige Ausprägung der Kultur angemessen darstellen können, da die bereits veröffentlichten Kennwerte als Indikator für die Gesamtkultur eines Landes aufzufassen sind, welche dann wiederum in angemessener Weise durch eine repräsentative Stichprobe abgebildet wird.

Selbst wenn dann die Frage nach der angemessenen Erhebung der Kulturwerte beantwortet ist, ist dennoch eine Interpretation von Unterschieden zwischen zwei Ländern keineswegs einfach. Neben der bereits erwähnten Tatsache, daß oft überhaupt keine Kulturwerte gemessen werden und dennoch vorschnell Kultur als Ursache herangezogen wird, wird ebenso häufig kritisiert, daß auch dann teilweise zu häufig und vorschnell auf Kultur als Ursache zurückgegriffen wird, selbst wenn die Kulturwerte explizit gemessen wurden (van de Vijver & Leung, 2000). Diese Kritik beruht darauf, daß neben kulturellen Unterschieden oftmals eine Vielzahl von Variablen in Frage käme, die einen Unterschied zwischen zwei Ländern in der abhängigen Variable erklären könnten. Variablen, die diesbezüglich mögliche erklärende Faktoren sein könnten, werden zumeist im sozioökonomischen Kontext gesucht: Bruttoinlandsprodukt (BIP), Wirtschaftswachstum, Bildungsgrad, Ausgestaltung des Schulsystems, Gesundheitsvorsorge, Lebenserwartung, Einwohnerzahl, Bevölkerungsdichte und Bevölkerungswachstum, Industrialisierungsgrad, etc. (vgl. Hofstede, 1980; van de Vijver & Leung, 1997). Darüber hinaus werden jedoch auch teilweise politische Faktoren, bspw. Demokratisierungsgrad, Menschenrechte, Art der Gesetze, etc. (Basabe, et al., 2002), und neuerdings auch klimatische Faktoren wie Temperatur, Breitengrad, Luftfeuchtigkeit, etc. (van de Vliert, Huang & Levine, 2004; van de Vliert, Schwartz, Huismans, Hofstede & Daan, 1999) herangezogen, um nationale Unterschiede zu erklären. Kultur ist somit nur einer von vielen Aspekten, die in internationalen Studien als erklärender Faktor fungieren können. Viele Forschungsvorhaben im interkulturellen Bereich blenden jedoch diese sozioökonomischen, politischen oder auch klimatischen Faktoren aus und beziehen sich bei ihren Erklärungen lediglich auf Kultur.

Optimalerweise sollten jedoch eine Vielzahl von Faktoren erhoben werden, um sie auf ihre Wirksamkeit hin untersuchen zu können. Diese Forderung ist jedoch auch in Teilen problematisch, da sich insgesamt eine schier unüberschaubare Anzahl von Faktoren finden ließe, die hierbei berücksichtigt werden könnten. Die oben genannten Faktoren stellen längst nicht alle denkbaren Aspekte dar und sollten nur als Auszug aus der Gesamtzahl möglicher Einflußgrößen verstanden werden. Da nicht grundsätzlich zu bestimmen ist, welches die relevantesten Größen sind und somit in verschiedenen Studien häufig jeweils unterschiedliche Meßgrößen berücksichtigt werden sollten, wäre eine allgemeinere Kategorisierung und Taxonomie hilfreich. Diesbezüglich kann das Modell von Georgas und Berry (1995) bzw. Georgas, van de Vijver und Berry (2004) als Anhaltspunkt angesehen werden. Die Autoren unterscheiden zwischen sechs großen Faktoren, die ihrer Ansicht nach ausreichen, Nationen hinreichend zu charakterisieren: a) Ökologie, b) Wirtschaft/Wohlstand, c) Bildung, d) Massenmedien, e) Bevölkerung und f) Religion. Basierend auf anfänglichen theoretischen Überlegungen (Georgas & Berry, 1995) untermauerten Georgas et al. (2004) ihre Einteilung auch anhand von empirischen Untersuchungen. Aus einer umfassenden Anzahl von über 500 Indizes, die renommierten Quellen wie beispielsweise Datenbanken der United Nations oder statistischen Jahrbüchern der UNESCO entnommen wurden, konnten sie anhand von Clusteranalysen insgesamt 23 Indikatoren auswählen, die die zentralen sechs Faktoren ausreichen erfassen können. Gemäß ihren Analysen lassen sich die Faktoren unter anderem durch folgende Indikatoren angemessen operationalisieren:

- a) Ökologie
 - höchste monatliche Niederschlagsmenge
 - höchste und niedrigste monatliche Temperatur
- b) Wirtschaft/Wohlstand
 - Bruttonationalprodukt pro Kopf (BNP)
 - täglicher Kalorienverbrauch pro Kopf (% des Bedarfs)
 - Energieverbrauch gewerblicher und privater Verbraucher pro Kopf
 - Bevölkerungsanteil in der Landwirtschaft, in der Industrie und im Dienstleistungsbereich
- c) Bildung
 - Analphabetenrate bei Erwachsenen
 - Lehrer-Schüler-Verhältnis im ersten Ausbildungsbereich
 - Ausgaben für Forschung und Entwicklung (% des BNP)
 - Bevölkerungsanteil im 1., 2. und 3. Bildungsbereich

- d) Massenmedien
 - Anzahl Telefone, Radios, Fernseher und Tageszeitungen (pro 1000 Einwohner)
- e) Bevölkerung
 - Säuglingssterblichkeit
 - Lebenserwartung bei Geburt
 - Todes- und Geburtenrate
 - Bevölkerungswachstum
- f) Religion
 - Bevölkerungsanteil in Religionsgemeinschaften

Diese Indikatoren sollten als eine beispielhafte Auswahl für den jeweiligen Faktor angesehen werden. Bei einem anderen Ausgangspunkt und einer anderen Datenbasis wären möglicherweise auch andere Indikatoren anhand der Clusteranalysen als bedeutsam ermittelt worden. Diese Maßzahlen waren gemäß Georgas et al. (2004) jedoch in der Lage, einen Großteil der Varianz zwischen den einzelnen Ländern zu erklären. Dabei sollte jedoch bedacht werden, daß diese Faktoren unter sich ebenfalls bedeutsam korrelieren, also keinesfalls als unabhängig voneinander anzusehen sind. Dies ist zwar wiederum auch nicht überraschend, da auch aus der inhaltlichen Perspektive heraus durchaus Einflüsse der jeweiligen Variablen aufeinander angenommen werden. Methodisch gesehen bereitet es aber Probleme, wenn alle Faktoren gleichzeitig in statistischen Analysen berücksichtigt werden sollen. Bei manchen Analysen (z.B. der Regressionsanalyse) würde durch diese hohe Kollinearität beispielsweise eine der Grundvoraussetzung der Methode verletzt, so daß das Verfahren nicht mehr durchführbar wäre. Darüber hinaus verlieren einige Analysen auch an sogenannter statistischer Power, wenn zu viele Variablen gleichzeitig betrachtet werden sollen. Eindeutige signifikante Ergebnisse sind dann nicht mehr zu erwarten. Von daher ist auf der einen Seite der Forderung zwar zuzustimmen, möglichst viele Faktoren neben kulturellen Variablen in den jeweiligen Analysen zu berücksichtigen. Auf der anderen Seite kann jedoch auch nicht generell davon ausgegangen werden, daß es grundsätzlich möglich wäre, alle möglichen Alternativerklärungen in den statistischen Analysen zu prüfen. Dennoch sollte, wie in der interkulturellen Forschung oft gefordert wird, ein gefundener Effekt von Kultur immer auch auf Alternativerklärungen hin überprüft werden. Diese Form der methodischen Überprüfung wird dabei vielfach auch als ‚destruktives Testen‘ bezeichnet (vgl. beispielsweise Anderson & Anderson, 1996). Dabei wird geprüft, ob ein gemessener Zusammenhang zwischen einer unabhängigen

Variable (oftmals ‚Kultur‘) und einer abhängigen Variable auch durch andere Variablen erklärt werden kann oder ob neben dem Einfluß einer Variablen möglicherweise auch andere Variablen in gleicher Weise das abhängige Maß beeinflussen. So bedeutet beispielsweise ein einmal gefundener Zusammenhang zwischen Kultur und einer abhängigen spezifischen Variable zwar, daß Kultur scheinbar eine wesentliche Erklärung der Varianz der abhängigen Variable liefern kann. Jedoch besteht ohne weiteres die Möglichkeit, daß dieser Zusammenhang durch andere Variablen moderiert oder auch mediiert wird oder aber daß in gleichberechtigter Form andere Variablen für den länderbezogenen Unterschied verantwortlich sind, so daß eine nur auf Kultur bezogene Hypothese nicht vollständig zu halten wäre. Dennoch wird auch in weiten Teilen der internationalen und interkulturellen Forschung diese – durchaus berechtigte – Forderung nach destruktivem Testen oftmals nicht beachtet bzw. umgesetzt (van de Vijver & Leung, 2000).

Die hier vorliegende Arbeit folgt dieser Forderung in mehrfacher Hinsicht. Beispielsweise werden neben den Kulturaspekten weitere wichtige länderspezifische Faktoren, die eine Bedeutung für das Arbeitsleben haben, auf ihren Erklärungsgehalt für unterschiedlich hohe Zufriedenheitsniveaus geprüft. Zwar können nicht alle Faktoren des Rahmenmodells von Georgas und Berry (1995) aufgrund der erwähnten methodischen Probleme in die Analysen aufgenommen werden. Dennoch werden vor allem die für die Arbeitswelt wichtigsten Indikatoren berücksichtigt. Des Weiteren werden insbesondere Antworttendenzen als Mediator für die Beziehung zwischen kulturellen Unterschieden und verschiedenen hohen Zufriedenheitsniveaus geprüft. In dritter Hinsicht werden schließlich im Sinne des destruktiven Testens die ermittelten Zusammenhänge durch den gleichzeitigen Einbezug von weiteren Faktoren in die statistischen Analysen auf ihre Tragfähigkeit hin geprüft.

5.3 Forschungsdesign und Stichprobenproblematik

Die Problematik von möglichen Alternativerklärungen bekommt auch deshalb in interkulturellen Studien eine größere Bedeutsamkeit, da Studien, die in mehreren Ländern und Kulturen durchgeführt werden sollen, im Prinzip von Natur aus nur als quasi-experimentelles Design umgesetzt werden können (van de Vijver & Leung, 1997). Kultur als unabhängige Variable läßt sich nicht den einzelnen Studienteilnehmern zufällig zuordnen. Durch diese fehlende Möglichkeit der Zuordnung tritt somit auch

verstärkt das Problem der Konfundierung von Variablen auf. Eine zufällige Zuordnung der Versuchspersonen, wie sie in klassischen Experimenten zur Kontrolle von Störvariablen und somit zur Ausschaltung von Alternativerklärungen vorgenommen wird, läßt sich in interkulturellen Studien nicht umsetzen. Von daher könnten sich die jeweiligen Länderstichproben neben kulturellen Aspekten auch immer noch in anderen der bereits erwähnten Kontextvariablen unterscheiden. Da die zur Erfassung kausaler Zusammenhänge favorisierte Form von Randomisierung der Untersuchungsteilnehmer somit nicht möglich ist, könnte möglicherweise als weitere Möglichkeit auf die Parallelisierung der Stichproben zurückgegriffen werden. Diese Alternative wählte beispielsweise Hofstede (1980), als er seine Untersuchung in einem weltweit tätigen, multinationalen Konzern durchführte, der in jedem Land sehr ähnlich strukturiert war. Die Arbeitsumgebung, in der die Studienteilnehmer befragt wurden, war folglich für alle Länder relativ vergleichbar. Somit konnte er viele der möglichen Störvariablen ausschalten, indem er sie für alle Studienteilnehmer durch diese Vorgehensweise gleichschaltete. Dennoch ist auch diese Form der Datenerhebung nicht komplett in der Lage, alle potentiell möglichen Störeinflüsse und Kontextvariablen auszuschalten. Zwar wird es durchaus wahrscheinlicher, wie auch Hofstede (1980, 2002) selbst argumentierte und so seine Studie verteidigte, daß die gefundenen Unterschiede sich auf kulturelle Aspekte zurückführen lassen. Ein eindeutiger kausaler Zusammenhang zwischen Kultur und der jeweiligen abhängigen Variablen läßt sich mit der Parallelisierung von Stichproben durch die Erhebung beispielsweise in ein und demselben Unternehmen jedoch auch nicht herstellen, denn in der Regel unterscheiden sich die Mitarbeiterschaften des gleichen Unternehmens auch durchaus von einem Standort zum anderen. Es kann auch nicht zwingend gewährleistet werden, daß diese spezifische Stichprobe sich nicht doch in zentralen kulturellen Aspekten von den restlichen Menschen des jeweiligen Landes unterscheidet. Daher kommt bei der Datenerhebung in nur einem situationalen Umfeld (wie dem eines einzigen Unternehmens) auch der bereits erwähnte Nachteil zum Tragen, daß die bekannten Kulturwerte von Hofstede, Schwartz oder der GLOBE-Studie eigentlich nicht für diese Art der exklusiven Stichprobe als angemessenes Maß der Kultur herangezogen werden können. Eine erneute Erhebung der Kulturwerte würde somit notwendig.

Als letzte und dritte Art der Datenerhebung besteht schließlich die Möglichkeit, auf repräsentative Stichproben des jeweiligen Landes zurückzugreifen. Zwar ist eine Rekrutierung repräsentativer Stichproben in der Regel mit sehr großem administrativen

Aufwand verbunden (vgl. beispielsweise Inglehart & Welzel, 2004, oder aber auch House et al., 2004; MOW International Research Team, 1987) und daher nur in wenigen internationalen Kooperationen zu finden (vgl. das ISSP – International Social Survey Programme –, das Eurobarometer oder den World-Values-Survey), jedoch bilden nur repräsentative Stichproben die in dem jeweiligen Land vorherrschenden Bedingungen angemessen und umfassend ab. Nur sie sind in der Lage allgemeine Theorien über Gesellschaften und Kulturen zu prüfen und sollten daher unbedingt eingesetzt werden, wenn die Untersuchungseinheit die jeweiligen Länder sind und man sich somit bei den durchgeführten Analysen auf Nationenniveau bewegt (Inglehart & Welzel, 2004). Bei dieser Art von Stichprobe ließen sich auch dann wiederum die bekannten Kulturwerte von Hofstede, Schwartz oder der GLOBE-Studie einsetzen, da die jeweilige Kultur der Gesamtpopulation eines Landes ja auch komplett in der repräsentativen Stichprobe abgebildet wird. Der Vollständigkeit halber sollte jedoch hier nicht unerwähnt bleiben, daß auch die Ursprungsstudien von Hofstede und Schwartz sowie die GLOBE-Studie selbst nur auf nicht-repräsentative Datensätze zurückgegriffen haben. Die in diesen Studien jeweils gemessenen Länderkennwerte können dennoch aber als gute Indikatoren für die jeweilige Gesamtkultur angesehen werden, da sie sich über viele Jahre der Forschung in vielen empirischen Studien immer wieder als gute und valide Prädiktoren herauskristallisieren und gegen die durchaus auch vorhandene Kritik an der jeweiligen Meßmethode behaupten konnten. Als weiteres Argument für den Einsatz repräsentativer Stichproben kommt hinzu, daß für viele Konstrukte, die interkulturell untersucht werden sollen, zumeist Mittelwerte aus Individualdaten als Repräsentant des jeweiligen Länderkennwertes genutzt werden. Dies trifft unter anderem auch auf die Erfassung von Antworttendenzen zu, wenn diese als interkulturell unterschiedliches Kommunikationsverhalten angesehen werden, wie es in der hier vorliegenden Studie der Fall ist. Zwar wird diese Vorgehensweise auch teilweise kritisiert und über ein besseres Vorgehen zur Bildung von Maßzahlen für länderspezifische Konstrukte diskutiert (vgl. Fischer, 2006), jedoch ist in vielen Forschungsbereichen dieses Vorgehen mangels Alternativen das oftmals praktikabelste. Aufgrund dieser Mittelung von Individualwerten wird es jedoch wiederum umso bedeutsamer, daß alle Teile der Gesellschaft in der jeweiligen Stichprobe vertreten sind. Dies können jedoch nur repräsentative Stichproben ermöglichen. Insgesamt wären somit gegenüber der Parallelisierung repräsentative Stichproben vorzuziehen, bei denen dann je nach Fragestellung weitere Alternativerklärungen neben

kulturellen Variablen methodisch oder statistisch überprüft werden sollten. Dennoch finden speziell aufgrund des enormen administrativen Aufwands bei Repräsentativstichproben auch parallelisierte Stichproben durchaus ihre Berechtigung. Wollte man beide Arten von Stichproben auf ihre Meßgüte hin vergleichen, würde eine parallelisierte Stichprobe im Prinzip einer Erhöhung der internen Validität dienen, jedoch eine geringe externe Validität besitzen. Für Repräsentativstichproben gilt genau umgekehrt, daß diese zwar geringere interne Validität besäßen, jedoch aufgrund ihrer höheren externen Validität besser auf den jeweiligen kulturellen Gesamtkontext zu übertragen wären. Die geringere interne Validität im Sinne von Störvariablen ließe sich bei dieser Art von Stichproben jedoch zusätzlich auch anhand methodischer und statistischer Verfahren erhöhen (z.B. Kovarianzanalysen, Partielle Korrelationen, destruktives Testen), während dies für eine Erhöhung der externen Validität bei parallelisierten Stichproben sowohl methodisch als auch statistisch nur schwer machbar ist. Somit schneidet auch bei diesem Vergleich eine Repräsentativstichprobe letztendlich etwas besser ab.

Dieser Vergleich zwischen Randomisierung und Parallelisierung von Untersuchungsteilnehmern beziehungsweise der Repräsentativität von Stichproben bezieht sich jedoch nur auf die Rekrutierung der Studienteilnehmer innerhalb des jeweiligen Landes. In der interkulturellen Forschung besteht jedoch ebenso immer die Problematik der angemessenen Stichprobengenerierung hinsichtlich der einzelnen Länder als Untersuchungseinheit. Diesbezüglich läßt sich zwischen angefallenen, systematischen und randomisierten Stichproben unterscheiden (van de Vijver & Leung, 1997). Viele Studien, die interkulturelle Vergleiche vornehmen, beziehen lediglich 2-3 Länder mit ein, was zu Recht kritisiert wird (vgl. Harzing, 2006). Besonders die Unterschiede zwischen lediglich zwei Kulturen sind im Prinzip aufgrund der schier umfassenden Möglichkeit von Alternativerklärungen im Prinzip nicht zu interpretieren (Brett et al. 1997). Hierbei kann auch von einer Stichprobenziehung noch keine Rede sein, da oftmals die Auswahl der Länder den persönlichen Interessen der jeweiligen Forscher folgt oder auf zufällig bestehenden Forschungsk Kooperationen beruht. Sinnvoll wäre bei solch einer geringen Anzahl von Ländern möglicherweise eher eine kontrastierende Auswahl im Sinne einer systematischen Variation, d.h. eine Auswahl von Nationen, die hinsichtlich spezifischer Aspekte möglichst heterogen sind. Methodisch besser jedoch wäre es, aus der weltweit auf dem ganzen Globus vorhandenen Anzahl von Ländern eine Zufallsauswahl zu treffen, um mögliche

Selektionseffekte zu vermeiden. Dies ist jedoch zumeist in der interkulturellen Forschung als unrealistisch anzusehen. Oftmals verhindern rein pragmatische Gesichtspunkte eine Generierung solcher Zufallsstichproben. Neben der Tatsache, daß es oft sehr aufwendig ist, gerade in Nationen mit geringer Infrastruktur Daten zu erheben, scheitern solche Vorhaben oft auch an politischen oder wirtschaftlichen Hindernissen. In der Regel handelt es sich somit in der interkulturellen Forschung – selbst bei großen internationalen Studien – um lediglich angefallene Stichproben. Van de Vijver & Leung (1997) machen jedoch deutlich, daß solche angefallenen Stichproben eigentlich nur bei explorativen interkulturellen Studien von Nutzen sind und eingesetzt werden sollten. Grundsätzlich ist eine randomisierte oder systematische Auswahl von Ländern zu bevorzugen. Die Art der systematischen Auswahl sollte dabei vom speziellen Fokus der Fragestellung abhängen. Wenn nach Unterschieden in Konstrukten geforscht wird, dann mag es sinnvoll sein, sehr ähnliche Länder zu untersuchen, da dadurch die Anzahl von Alternativerklärungen minimiert wird. Wird hingegen nach der Universalität von Phänomenen gesucht, dann sollten möglichst heterogene Kulturen in der Stichprobe enthalten sein (van de Vijver & Leung, 1997).

Die Nutzung einer lediglich angefallenen Stichprobe hat oft auch zur Folge, daß manche Typen von Ländern überrepräsentiert und wiederum andere unterrepräsentiert sind. So sind westliche Industrienationen zumeist überrepräsentiert und in großen Studien fast immer vertreten, da dort die meiste internationale bekannte Forschung stattfindet. Im Gegensatz dazu sind afrikanische Länder fast immer unterrepräsentiert. Daten aus den Ländern des afrikanischen Kontinents in der Forschungsliteratur zu finden, ist oftmals schwierig. Diese Verzerrung der Länderstichprobe kann jedoch schwerwiegende Konsequenzen für die jeweiligen statistischen Analysen auf Länderebene haben. Zum einen besteht permanent aufgrund der geringen Anzahl von Nationen die Gefahr, daß Ausreißer beispielsweise in Korrelationsanalysen die gemessenen Effekte enorm stark verzerren können. Selbst bei einer Anzahl von beispielsweise 20-25 Ländern, die im Allgemeinen schon einen enormen Rekrutierungserfolg darstellen würde und auch selbst nur in wenigen publizierten Studien anzutreffen ist, könnte allein ein einziger extremer Ausreißer potentielle Zusammenhangseffekte deutlich verschieben. Darüber hinaus kann aber auch nie mit Sicherheit festgestellt werden, welchen Teil der zugrundeliegenden latenten Verteilung die angefallene Länderstichprobe eigentlich darstellt. Bei der Verteilung auf einer Variable könnte die jeweilige Stichprobe möglicherweise lediglich das obere Ende des eindimensionalen Kontinuums darstellen.

So könnten beispielsweise im Sinne der in dieser Arbeit untersuchten Thematik bei einer unzureichenden Stichprobenauswahl die untersuchten Länder nur die sehr hoch zufriedenen Länder darstellen und nur die obersten 30 Prozent der eigentlichen Verteilung der Zufriedenheitsurteile in allen Ländern erfassen. In gleicher Weise könnten die erhobenen Länder sich aber auch unter Umständen nur am unteren Ende des Kontinuums einordnen lassen. Eine genaue Schlußfolgerung diesbezüglich läßt sich aufgrund der unklaren Verteilung der restlichen Gesamtpopulation daher grundsätzlich nie ziehen. Genau dies ist der Nachteil, wenn lediglich eine angefallene Länderstichprobe untersucht wird. Betrachtet man neben diesen Verteilungsproblemen allein bei einer Variable darüber hinaus den Zusammenhang zweier Variablen und erweitert so den eindimensionalen zu einem zweidimensionalen Raum, könnte es sogar sein, daß eine für den Zusammenhangseffekt sehr untypische Stichprobe einen völlig anderen Effekt im Vergleich zum tatsächlich in der Realität vorhandenen Effekt darstellt (vgl. Abbildung 5). Während in diesem hypothetischen Fall, der zur Verdeutlichung konstruiert wurde, im linken Teil der Abbildung die roten Datenpunkte den in der Stichprobe gemessenen Zusammenhang zweier Konstrukte darstellen würden, zeigt ein Blick auf den rechten Teil der Abbildung, daß diese Stichprobe ein völlig mißverständliches Ergebnis liefern würde, wenn denn die Gesamtpopulation tatsächlich gemäß den grünen und roten Datenpunkte verteilt gewesen wäre.

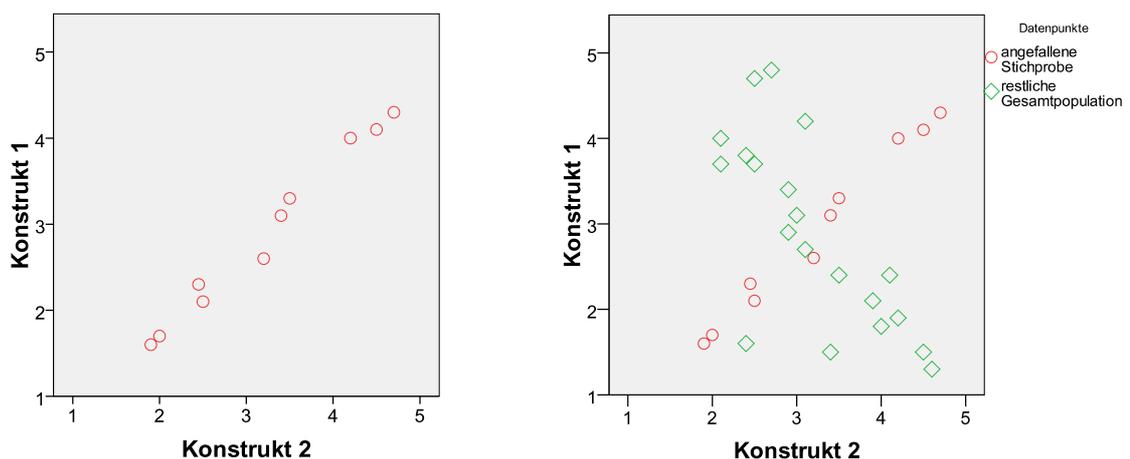


Abbildung 5: Hypothetischer Vergleich einer angefallenen Stichprobe mit einer Gesamtpopulation

Neben diesen rein methodischen Gründen, aufgrund derer allein schon eine nicht-randomisierte Auswahl von Ländern eher kritisch betrachtet werden muß, sprechen

auch inhaltliche Gründe in Teilen gegen die Nutzung von angefallenen Stichproben. Diese Art der Auswahl, bei der zumeist westliche Nationen bevorzugt werden, schneidet auch deshalb schlechter ab, da – um nur ein spezielles Beispiel zu nennen – in westlichen Ländern zumeist das Bildungsniveau höher ist. Ein höheres Bildungsniveau jedoch könnte sowohl einen Einfluß haben auf die inhaltliche Beantwortung eines Fragebogens als auch möglicherweise die Tatsache begünstigen, daß die Befragten mit Fragebögen als einer Art von Datenerhebung grundsätzlich vertrauter sind und somit diese speziellen Form von Kommunikation zwischen Forscher und Befragten eher gewöhnt sind. Gerade in Ländern, wo die Vertrautheit mit Fragebögen dann aber nicht sehr hoch ist, ist eine Verzerrung der Antworten aufgrund dessen nicht gänzlich auszuschließen (van de Vijver & Poortinga, 1997). Zumindest auf individuellem Niveau zeigte sich bereits, daß höhere Bildung mit einer geringeren Zustimmungstendenz einhergehen kann (Marin, Gamba, & Marin, 1992; Winkler, Kanouse, & Ware, 1982). Ein Einfluß auf nationalem Niveau ist somit durchaus ebenfalls möglich.

Die genannten Probleme bei der Generierung der Länderstichproben können durch verschiedene Maßnahmen zum Teil minimiert werden. Um mögliche Ausreißer erkennen zu können und somit die erste Gefahr der Ergebnisverzerrung eindämmen zu können, wäre beispielsweise zu empfehlen, die jeweiligen Zusammenhänge auch immer erst in deskriptiver Form innerhalb eines Streudiagramms zu betrachten. Länder mit extremen Werten sollten dann möglicherweise von den entsprechenden Analysen ausgeschlossen werden. Die zweite Gefahr, d.h. die nicht-repräsentative Auswahl der Länder bezüglich eines Merkmals oder eines Zusammenhangs zweier Konstrukte, läßt sich nachträglich in der statistischen Analyse leider nicht mehr kontrollieren, da ja in der Regel die Positionierung der in der Stichprobe erfaßten Länder auf dem latenten Konstrukt nicht bekannt ist. Die damit verbundene Verringerung der externen Validität muß daher vor allem bei der Interpretation von Ergebnissen grundsätzlich beachtet werden. Hinsichtlich dieser Problematik ist daher empfohlen, wann immer möglich entweder bereits bei der Auswahl der Länder besondere Vorsorge zu treffen oder aber ansonsten als Alternative immer so viele Länder wie möglich in die jeweilige Studie zu integrieren. Je mehr Länder berücksichtigt werden können, desto geringer wird folglich der Stichprobenfehler und desto höher ist die Gültigkeit und Übertragbarkeit der Ergebnisse. Ebenso gilt, daß je breiter sich die Länder auf dem zu untersuchenden kulturellen Merkmal positionieren lassen, desto valider sind die gemessenen Zusammenhänge zwischen Kultur und dem untersuchten Phänomen (Brett et al., 1997;

van de Vijver & Leung, 1997). Übertragen auf die in dieser Arbeit betrachtete Problematik bedeutet dies, daß möglichst Länder sowohl mit hohem, mit mittlerem als auch mit niedrigem Arbeitszufriedenheitsniveau in der Stichprobe erfaßt werden sollten. Nur durch eine umfassende Anzahl von Ländern steigt somit die Sicherheit, am Ende auch Ergebnisse zu erhalten, die tatsächlich eine allgemeine Theorie der jeweiligen Kulturen widerspiegeln können. Insgesamt sollte daher festgehalten werden, daß im Allgemeinen Studien mit möglichst vielen und zahlreichen kulturell heterogenen Ländern angestrebt werden sollten (Fischer et al., 2006). Falls eine umfassende Anzahl von Ländern dennoch nicht in einer Studie zu realisieren ist, so bleibt schließlich noch die Möglichkeit, die gefundenen Ergebnisse anhand einer weiteren unabhängigen Länderstichprobe zu kreuzvalidieren. Selbst wenn beide der Stichproben für sich genommen relativ klein sein sollten, würde eine Bestätigung der Ergebnisse dennoch die Gültigkeit von Ergebnissen untermauern. Solch eine Bestätigung würde auch ein weiteres Problem kleiner Länderstichproben verringern. Die zumeist zwangsweise Beschränkung auf kleine Länderstichproben bedeuten nämlich in der Regel auch, daß ein Zusammenhang zwischen den untersuchten Phänomenen nicht immer statistisch signifikant abgesichert werden kann. Entsprechende Effekte ließen sich vielleicht in einer Studie anhand der berechneten Effektgrößen zeigen, sie würden jedoch aufgrund der geringen Fallzahl möglicherweise nicht signifikant. Auch in diesem Fall kann die Gültigkeit der Ergebnisse überprüft werden, indem in weiteren Analysen mit anderen Länderstichproben die gefundenen Zusammenhänge bestätigt werden. Trotz fehlender Signifikanz könnten solche gleichgerichteten Ergebnisse dann als Indikator für das tatsächliche Vorliegen von Effekten in der Realität gedeutet werden. Da es oftmals trotz zahlreichender Anstrengungen in der interkulturellen Forschung nicht möglich ist, große Länderstichproben zu erhalten, wäre dies eine zu empfehlende Vorgehensweise. Auch wenn ein methodisch versierter Forscher immer die signifikante Absicherung von Ergebnissen fordern würde, würde ein komplettes Ignorieren solcher Effekte, die unglücklicherweise vielleicht nicht signifikant gemessen werden konnten, gerade in der interkulturellen, auf Länderniveau durchgeführten Forschung den Erkenntnisfortschritt unverhältnismäßig erschweren. Die Signifikanz von Ergebnissen sollte daher nicht als alleiniges Kriterium für die Bedeutsamkeit von interkulturellen Ergebnissen herangezogen werden. Ebenso bedeutsam wäre, daß überhaupt ein Effekt vorhanden ist und wie hoch die jeweilige Effektstärke ist. Die in dieser Arbeit durchgeführten Analysen versuchen die genannten Probleme vor allem so zu adressieren, daß einerseits

eine möglichst große Länderstichprobe genutzt wird und andererseits auch die dort gefundenen Ergebnisse an einer weiteren Länderstichprobe überprüft werden sollen.

5.4 Experimentelles Vorgehen in der interkulturellen Forschung

Kultur als unabhängige Variable läßt sich wie bereits erläutert nicht zufällig den Untersuchungsobjekten zuweisen. Interkulturelle Forschung, die Daten direkt im jeweiligen Land und der jeweiligen kulturellen Umgebung erhebt, kann somit lediglich einer quasi-experimentellen Vorgehensweise folgen. Somit ist jeder Untersuchungsteilnehmer sowohl Teil der jeweiligen allgemeinen Kultur als auch automatisch beeinflusst durch alle anderen situationalen Aspekte des untersuchten Landes. Die Konfundierung dieser verschiedenen unabhängigen Variablen läßt sich daher mit dieser Art von Erhebung nicht auflösen.

Da die beste methodische Vorgehensweise, um kausale Effekte sauber zu überprüfen, das klassische psychologische Experiment ist, wäre ein rein experimentelles Vorgehen auch durchaus in der interkulturellen Forschung wünschenswert. Ein experimentelles Design wäre in der Lage, die Problematik der Alternativerklärungen besser in den Griff zu bekommen. Wenn es möglich wäre, kulturelle Wertvorstellungen zufällig verschiedenen Probanden zuzuordnen, könnte eine solche experimentelle Variation den kausalen Einfluß von Kultur direkt und unmittelbar prüfen. Um kulturelle Effekte besser und in eindeutigerer Weise nachweisen zu können, stellt sich somit die Frage, wie Kultur sozusagen ins Labor geholt werden kann. Wie ließe sich demnach ein unterschiedliches kulturelles Denken verschiedenen Gruppen von zufällig ausgewählten Untersuchungsteilnehmern zuweisen?

In einer ersten spontanen Überlegung käme man möglicherweise auf die Idee, man könnte Probanden aus einer Kultur (beispielsweise deutsche Studenten) auffordern, sich gedanklich in eine andere Kultur hineinzusetzen, um darauf folgend entsprechende Fragen zur Messung der abhängigen Variable zu beantworten. Dieses kognitive Umdenken solle dann dazu führen, daß sich die Probanden wie ein Mitglied der jeweiligen Kultur verhielten. Solch eine spontane Idee wäre jedoch eindeutig zu kurz gedacht und nicht zur experimentellen Messung kultureller Effekte geeignet, da keineswegs davon ausgegangen werden kann, daß das sich lediglich vorgestellte Denken und Handeln anderer Kulturen überhaupt dem tatsächlichen Denken und

Handeln in der jeweiligen Kultur ähnlich sein könnte. Darüber hinaus besteht die Problematik solch einer Vorgehensweise auch darin, daß die meisten der Untersuchungsteilnehmer möglicherweise mit der sich vorzustellenden Kultur zumeist noch überhaupt keinen Kontakt gehabt hätten und daher sich nur bedingt in die andere Kultur hineinversetzen könnten. Diese Art von Messung wäre somit keinesfalls valide, da davon auszugehen ist, daß die Untersuchungsteilnehmer zumeist nur kulturelle Vorurteile aus ihrem Gedächtnis abrufen würden.

Hinweise, wie eine experimentelle Variation von Kultur in der Tat im Labor gelingen könnte, lassen sich jedoch bei den Überlegungen zum Zusammenhang von kulturellen Werten und der Ausprägung des Selbstkonzeptes der Mitglieder der jeweiligen Kultur finden (vgl. Markus und Kitayama, 1991, 1998). In weiten Teilen der Forschung und von vielen Forschern speziell dieses Bereichs wird inzwischen angenommen, daß die unterschiedliche Entwicklung und Ausprägung des Selbst – d.h. wie Menschen über sich selbst denken – einen zentralen Einfluß auf die Unterschiede im Denken und Handeln verschiedener Kulturen hat (Kühnen & Hannover, 2003). Umgekehrt beeinflußt auch die einmal vorherrschende Kultur in einem Land die Sozialisation der jeweiligen Individuen und somit die Entwicklung des Selbst. Kulturelles Umfeld und Manifestation des Selbst sind somit wechselseitig eng miteinander verbunden.

In ihren einflußreichen Studien und den daraus folgenden Überlegungen beschäftigten sich Markus und Kitayama (1991) speziell mit Unterschieden des Selbst in westlichen Nationen im Vergleich zu asiatischen Nationen, d.h. dem Unterschied zwischen eher individualistischen und eher kollektivistischen Kulturen. Gemäß ihren Überlegungen sind Menschen in individualistischen Kulturen mit der Erwartung konfrontiert, sich zu einem unabhängigen und eigenständigen Individuum zu entwickeln. Es ist in dieser Art von Kultur demnach wichtig, eigene Bedürfnisse, Motive und Einstellungen auszubilden. Das eigene Denken und die eigene Weiterentwicklung bekommen einen hohen Stellenwert zugesprochen. Auch das eigene Verhalten soll stärker von eigenen Antrieben geleitet sein als von den Erwartungen anderer. Menschen in individualistischen Kulturen werden somit insgesamt stärker dahingehend beeinflußt, sich über ihr eigenes Selbst zu definieren und ein unabhängiges Selbst im Vergleich zu den anderen Mitgliedern ihrer Kultur zu entwickeln. Markus und Kitayama (1991) sprechen hinsichtlich dieser Art des Selbst von einem independenten Selbst.

In kollektivistischen Kulturen hingegen unterliegt die Entwicklung des Selbst einem nahezu entgegengesetzten Einfluß. Bedeutsam ist hier vor allem die Verbindung des

Individuums zu den anderen Mitgliedern der Kultur. Das Selbst ist fundamental und untrennbar mit den anderen Personen verbunden. Anstatt eigene Gefühle, Bedürfnisse und Einstellungen zu entwickeln, ist in diesem Umfeld viel wichtiger, sich an andere Personen und ihre Erwartungen anzupassen. Interpersonale Beziehungen bekommen ebenso eine stärkere Bedeutung. Die eigene Person soll nicht ein eigenständiges Individuum darstellen, sondern stattdessen Teil eines sozialen Netzwerks sein. Die Individuen verspüren mehr die Motivation, sich kontingent zu den Ansprüchen ihrer Bezugspersonen zu verhalten. Menschen dieser Art von Kulturen definieren somit ihr Selbst viel stärker über die Passung mit den sie umgebenden Menschen und versuchen ihr Verhalten in Übereinstimmung mit den Gefühlen und Handlungen der anderen Personen zu bringen. Markus und Kitayama (1991) bezeichnen diese Art des Selbst als ein interdependentes Selbstkonzept.

Insgesamt ist wichtig zu betonen, daß sowohl Menschen in individualistischen als auch Menschen in kollektivistischen Kulturen immer beide Anteile des Selbst in sich vereinigt haben. Jedoch ist die Bedeutung des independenten Selbst in individualistischen Kulturen höher als die Bedeutung des interdependenten Selbst (Markus & Kitayama, 1991). Für die Menschen in kollektivistischen Kulturen gilt dies genau umgekehrt. Dort setzt sich das Selbst zu größeren Teilen aus interdependenten als aus independenten Aspekten zusammen.

Diese theoretischen Überlegungen ließen sich in einer Vielzahl von Studien belegen (z.B. Campbell, et al., 1996; Cross, Bacon & Morris, 2000; Kühnen, Hannover, Roeder, Schubert, Shah, Upmeyer & Zakaria, 2001; Rhee, Uleman, Lee & Roman, 1995; Singelis, 1994; Trafimow, Triandis & Goto, 1991). Es zeigte sich ein deutlicher Zusammenhang zwischen kulturellen Ausprägungen und der jeweiligen Form des Selbstkonzepts. Messen lassen sich die Unterschiede zwischen independentem und interdependentem Selbst dabei beispielsweise durch zwei verschiedene Meßverfahren, die eine weite Verbreitung in diesem Forschungsbereich fanden. Auf der einen Seite entwickelte Singelis (1994) eine Skala, anhand derer er die jeweilige Ausprägung des Selbstkonzeptes messen konnte. Independentes und interdependentes Selbst werden dabei jeweils durch 12 klassische Einstellungsisems und somit als voneinander unabhängige Konstrukte erhoben. Einen etwas anderen Weg wählt der TST (twenty statement test), der von Kuhn und McPartland (1954) entwickelt wurde und ebenfalls in zahlreichen Studien zur Messung der Selbstkonstrukte eingesetzt wurde (vgl. Harb & Smith, 2008). Die Studienteilnehmer werden hier gebeten, sich die Frage „wer bin

ich?“ zu beantworten und 20 Satzanfänge zu vervollständigen, die jeweils mit „Ich bin ...“ beginnen. In verschiedenen Studien zeigte sich übereinstimmend, daß individualistische Kulturmitglieder sich stärker anhand von Eigenschaftsbegriffen (z.B. ich bin intelligent, ich bin nett, ich bin schön, etc.) beschreiben, während Mitglieder aus kollektivistischen Kulturen öfter Beschreibungen wählen, die einen Gruppenbezug (z.B. ich bin ein Student, ich bin Deutscher, ich bin eine Tochter, etc.) ausdrücken (Rhee et al., 1995; Bond & Cheung, 1983). Eine verkürzte und somit ökonomischere Form von lediglich 10 Satzanfängen, die auch als TenST (ten-statement-test) bezeichnet wird, wurde ebenfalls erfolgreich in einer Reihe von Studien eingesetzt (Bochner, 1994; Khanlou, 2005; Kinket & Verkuyten, 1997; Kiss, 2004). Auch hier zeigten sich die bereits erwähnten Unterschiede in den Selbstbeschreibungen der Untersuchungsteilnehmer. Es sollte dabei jedoch betont werden, daß allen Studien gemeinsam ist, daß Studienteilnehmer aus individualistischen Kulturen und kollektivistischen Kulturen sich sowohl durch Eigenschaftsbeschreibungen als auch durch Gruppenzugehörigkeiten selbst beschreiben und charakterisieren. Sie verfügen somit – wie aus theoretischer Sicht angenommen – über independente und interdependente Anteile in ihrem Selbst und die damit verbundenen kognitiven Wissensinhalte und Korrelate. Lediglich die relative Häufigkeit der jeweiligen Anteile und daher auch der Aussagen im TST oder TenST unterscheidet sich zwischen den Kulturen.

Insgesamt zeigt sich somit, daß independente und interdependente Selbstkonzepte sich nicht gegenseitig ausschließen und nicht jeweils ausschließlich in individualistischen oder kollektivistischen Kulturen auftreten. Stattdessen besitzt jeder Mensch immer beide Anteile in seiner Konstruktion des Selbst. Entscheidend ist vielmehr die Aktivierung der jeweiligen Selbstkonzepte und das damit verbundene Handlungswissen. Von diesen Überlegungen ausgehend läßt sich in Anlehnung an die Social Cognition Forschung und das SPI-Modell (Semantic Procedural Interface Model) von Kühnen und Hannover (2003) annehmen, daß die entsprechenden Wissensinhalte auch durch spezielle kognitive Methoden aktiviert werden könnten und zu einer Art momentanen, situationspezifischen Verlagerung der Selbstkonzepte führen könnten. Kühnen und Hannover (2003) gehen dabei in ihrem SPI-Modell davon aus, daß die Selbstkonstruktion eines Menschen und daher sein Selbstkonzept sich dadurch charakterisieren läßt, welche Art von Selbstwissen aktiviert wurde (independente oder interdependente Inhalte) und wie die jeweilige Art von Selbstwissen aktiviert wurde

(situational unterschiedlich durch einen Kontextfaktor wie z.B. Priming oder aber chronisch durch das Aufwachsen in einer spezifischen Kultur). Das Selbstkonzept bildet sich somit in unterschiedlicher Art und Weise heraus, je nachdem wie hoch die Zugänglichkeit und Aktiviertheit von eher ich- oder eher wir-bezogenen kognitiven Inhalten sich gestaltet.

Verschiedene Forscher haben in Übereinstimmung mit diesen Überlegungen durch verschiedene Vorgehensweisen versucht, entsprechende Wissensinhalte durch spezielle Priming-Verfahren vorzuaktivieren und somit die entsprechenden Versuchspersonen jeweils in einen eher independenten oder aber interdependenten Zustand zu versetzen. Insgesamt lassen sich gemäß einer umfassenden Meta-Analyse von Oyserman und Lee (2008) bis zu sieben Typen von Priming-Verfahren unterscheiden, die zumeist zur Aktivierung der beiden Arten von Selbstkonzepten in der Forschung eingesetzt wurden. Sie differenzieren zwischen den folgenden Priming-Verfahren: SDFP-Priming, Pronoun-Circling-Task, Sumerian Warrior Story, Scrambled Sentences, Group instantiation, Group imagination, Language Priming. Beim SDFP-Priming (Similarities and differences with family and friends task), welches von Trafimow et al. (1991) entwickelt wurde, werden Probanden gebeten, sich zwei Minuten Gedanken zu machen, was sie von ihrer Familie und ihren Freunden unterscheidet bzw. was sie mit ihrer Familie und ihren Freunden gemeinsam haben. Erstere Variante dient dazu, die Probanden auf ihr eigenes Ich zu fokussieren und die independenten Wissensinhalte zu aktivieren, während letztere Variante den Wir-Bezug in den Vordergrund rückt und interdependente Wissensinhalte primt. In ähnlicher Form geht der Pronoun-Circling-Task (Brewer & Gardner, 1996) vor, um die unterschiedlichen Inhalte des Selbstkonzepts situational zu verstärken. Hier werden die Teilnehmer gebeten einen kurzen Text, der den Ausflug in eine Stadt beschreibt, zu lesen und alle Pronomen im Text einzukreisen. Um independentes Selbstwissen zu aktivieren, enthält der Text ausschließlich Pronomen, die einen Ich-Bezug ausdrücken (ich, mein, mir, mich, etc.), während der ansonsten identische Text zur Aktivierung von interdependentem Selbstwissen ausschließlich Pronomen mit Wir-Bezug enthält (wir, uns, unser, etc.). Bei der Sumerian Warrior Story (Trafimow et al., 1991) sollen Personen sich zu einer Entscheidung äußern, bei der ein Krieger entweder aufgrund von seinen speziellen Talenten oder aber aufgrund seiner Stammeszugehörigkeit und seiner familiären Beziehungen für eine Aufgabe in einer Schlacht ausgewählt wurde. Scrambled Sentences (Srull & Weyer, 1979; Kühnen & Hannover, 2000) schließlich versuchen

independentes und interdependentes Selbst zu primen, indem dort den Probanden ein durcheinander gewürfelter Satz vorgelegt wird, den sie in die richtige Reihenfolge bringen müssen. Zur Aktivierung des independenten Selbst enthalten die Sätze klare Worte und eindeutige Aussagen zur Einzigartigkeit von Individuen. Ein interdependentes Selbst hingegen soll mit Hilfe von Aussagen mit Gruppenbezug stärker hervorgehoben werden.

Diese vier Priming-Verfahren lassen sich gemäß Oyserman und Lee (2008) in den meisten Studien zur Untersuchung des Einflusses von independentem und interdependentem Selbst finden. Darüber hinaus versuchen einige wenige Studien Gruppenbezüge und Wettbewerb zwischen Gruppen künstlich zu erzeugen (Group instantiation: Briley & Wyer, 2002; Group imagination: Aaker & Lee, 2001), um interdependentes Denken stärker zu betonen, oder aber die Manipulation von Sprache (zum Beispiel Englisch gegenüber Chinesisch) einzusetzen, die das entsprechende Selbstkonzept stärker zum Vorschein bringen soll. Zweisprachige Probanden werden hierbei jeweils Englisch oder einer nicht-westlichen Sprachversion des Fragebogens zugewiesen; basierend auf der Annahme, daß Englisch individualistische Wertvorstellungen transportieren würde, während in nicht-westlichen Sprachen ein kollektivistisches Denken zum Vorschein käme (Bond & Yang, 1982; Kimmelmeier & Cheng, 2004; Ralston, Cunniff & Gustafson, 1995; Trafimow, Silverman, Fan & Law, 1997).

Nicht alle diese Verfahren scheinen jedoch gleich wirksam hinsichtlich des Primens von independenten oder interdependenten kognitiven Inhalten zu sein. Als erfolgreichste Verfahren, die die stärksten Effekte aufweisen konnten, resultieren in der Meta-Analyse von Oyserman und Lee (2008) eindeutig das SDFP-Priming und der Pronoun-Circling Task. Bei diesen Priming-Methoden ließen sich die stärksten Unterschiede in den jeweils abhängigen Variablen finden.

Insgesamt zeigen die zahlreichen Priming-Versuche und die Meta-Analyse von Oyserman und Lee (2008), daß tatsächlich die Möglichkeit besteht, das Selbstkonzept von Individuen in Teilen zu verändern und independente bzw. interdependente Aspekte des Selbst zumindest zeitlich begrenzt zu verändern, indem sie situativ zugänglich gemacht werden. Zusätzlich bestätigen die Ergebnisse vieler dieser Studien die Überlegungen von Markus und Kitayama (1991), daß unterschiedliche kulturelle Werte in den unterschiedlichen Arten des Selbstkonzepts verwurzelt sind. Gardner, Gabriel und Lee (1999) konnten explizit auch zeigen, daß independent und interdependent

geprimte Personen signifikante Veränderungen anhand einer Werteskala zeigten und somit eine Veränderung der Aussagen zu kulturellen Wertvorstellungen sich nur durch die kurzzeitige Veränderungen des Selbstkonzepts erreichen ließen. Dabei setzten sie zwei der klassischen Priming-Verfahren (Sumerian Warrior Story und Pronoun Circling Task) ein und erfaßten die Wertestrukturen anhand einer Werteskala von Schwartz (1992). Um die Wirksamkeit ihres Primings auch überprüfen und nachweisen zu können, setzten sie einerseits den TST (Kuhn & McPartland, 1954) als Manipulationscheck ein, der auch folgerichtig vermehrte Gruppenaussagen für die interdependent geprimte Gruppe zeigen konnte. Als weitere Überprüfung der Wirksamkeit war ebenfalls ein Teil eines Social Judgement Tasks (Miller, Bersoff & Harwood, 1990) Bestandteil ihrer Studie. In solch einer Aufgabe, die das Befolgen von zwischenmenschlichen Normen hinsichtlich helfenden Verhaltens mißt, sollen Probanden eine soziale Situation hinsichtlich ihrer Angemessenheit beurteilen, in der Amy ihre Freundin Lisa um Hilfe fragt, während diese selbst aber ein Buch lesen will und daher eigentlich nicht helfen will. Auch hier zeigte sich, daß interdependent geprimte Personen eher Lisas Verhalten als unangemessen beurteilten und ihre eher eine Verpflichtung zur Hilfeleistung zuwiesen. Die kurzzeitige Veränderung des Selbstkonzepts war somit bei Gardner et al. (1999) in der Lage, sowohl die Wertaussagen von Probanden, ihre Selbstbeschreibungsaussagen als auch die Einschätzungen in Bezug auf Hilfeleistung in sozialen Situationen zu verändern.

Die eingangs aufgeworfenen Frage, inwieweit Kultur sich ins Labor übertragen ließe, kann somit auf Basis nicht nur der Studie von Gardner et al. (1999), sondern auch der zahlreichen anderen Priming-Studien in Teilen beantwortet werden. Priming-Aufgaben können zumindest kurzzeitig unterschiedliche kognitive Inhalte aktivieren, die bedeutsam für ein eher ich-bezogenes oder aber ein wir-bezogenes Denken sind. Da dieses unterschiedliche Denken auch in entsprechenden Kulturen stärker ausgeprägt ist, kann daher davon ausgegangen werden, daß Kultur sich zumindest partiell im Labor in experimentellen Anordnungen erzeugen läßt. Es muß jedoch einschränkend hinzugefügt werden, daß durch die beschriebenen Priming-Verfahren lediglich ein Unterschied hinsichtlich Individualismus/Kollektivismus erzeugt werden kann. Nur für diese Dimension läßt sich sowohl theoretisch als auch empirisch ein Bezug zum independenten und interdependentem Selbstkonzept herstellen. Für die anderen gemeinhin bekannten Werte-Dimensionen sind keine Priming-Verfahren in der Literatur zu finden. Demnach existiert bisher kein erprobtes Verfahren, welches in der

Lage ist, eine valide und reliable experimentelle Manipulation beispielweise von Machtdistanz oder Unsicherheitsvermeidung durchzuführen. Für diese Dimensionen müßten erst ein eigenes Verfahren komplett neu konstruiert werden, welches zusätzlich auch eine eigene theoretische Einbettung benötigen würde, die so auch bisher nicht gegeben wäre. Andererseits ermöglicht jedoch die Durchführung von Primings anhand einer kulturellen Versuchspersonengruppe immerhin die Manipulation der bisher in der interkulturellen Forschung am häufigsten angewandten Kulturdimension. Somit ließen sich zumindest hinsichtlich Individualismus und Kollektivismus eindeutiger Kausalaussagen aufgrund der Möglichkeit der experimentellen Manipulation treffen. Ein Experimentaldesign, in dem eine der Probandengruppen independent und eine andere interdependent geprimt wird, vergliche somit im übertragenen Sinne individualistische Kulturen mit kollektivistischen Kulturen unter gleichzeitiger Ausschalten von Störvariablen und Alternativerklärungen. Zusätzlich zu den Analysen von repräsentativen Länderstichproben, die den besonderen Vorteil der hohen externen Validität besitzen, würde eine Laborstudie mit Hilfe von Primings die interne Validität der untersuchten Befunde erhöhen. Daher wird auch in der hier vorliegenden Arbeit eine experimentelle Studie durchgeführt, die die gefundenen Effekte möglichst replizieren soll. Auf der einen Seite ließe sich somit die Qualität der unabhängigen Variablen auf diese Weise methodisch erhöhen. Unabhängig davon besteht jedoch zusätzlich der Bedarf, auch die abhängigen Variablen in interkulturellen Studien hinsichtlich ihrer Meßgenauigkeit zu optimieren. Verbunden mit dieser Optimierung ist vor allem die Frage nach der interkulturellen Äquivalenz der abhängigen Variable.

5.5 Äquivalenz im interkulturellen Vergleich

Einer der methodischen Aspekte, der besonders in der interkulturellen Forschung an Bedeutung gewinnt, ist die Überprüfung der Äquivalenz von Studien und Datenerhebungen, die in verschiedenen Kulturen durchgeführt werden. Dies ist vor allem deswegen bedeutsam, da eine der zentralen Fragen beim Vergleich interkultureller Daten ist, ob die erfaßten Unterschiede in den gemessenen Werten tatsächlich Ausdruck von Unterschieden in dem zu untersuchenden Konstrukt sind, oder ob die Unterschiede möglicherweise auf methodische Einflüsse zurückzuführen sind und somit die Erhebungen nicht (meß-)äquivalent sind. Dieses Problem stellt sich grundsätzlich nicht erst neuerdings im Bereich der Forschung über Länder- und

Kulturgrenzen hinweg. Auch beim Vergleich von reinen Gruppenunterschieden innerhalb ein und derselben Kultur muß prinzipiell die Frage nach der Äquivalenz der gemessenen Werte beantwortet werden (Drasgow & Kanfer, 1985). Jedoch tritt der Äquivalenzaspekt insbesondere aufgrund unterschiedlicher kultureller Denk- und Verhaltensweisen in den jeweiligen Ländern im interkulturellen Vergleich in den Vordergrund. Werden beispielsweise physikalische Größen international verglichen, ist im Prinzip jeder dieser Problematik bewußt. So würde niemand auf die Idee kommen, bei Gewichten, Längen oder anderen physikalischen Maßeinheiten einen unmittelbaren Vergleich anzustreben und Pfund mit Kilogramm, Kilometer mit Meilen oder Zentimeter mit Inch direkt in Beziehung zu setzen (Drasgow & Kanfer, 1985; Poortinga, 1989). Bei der Erfassung psychologischer Aspekte in verschiedenen Kulturen wird diese Problematik jedoch oftmals übersehen oder sogar bewußt ausgeblendet und die Vergleichbarkeit als machbar angenommen und vorausgesetzt. Jedoch ist keineswegs von Grund auf als gesichert anzunehmen, daß die Erfassung eines psychologischen Konstrukts (z.B. eines Persönlichkeitsaspekts wie Extraversion) in Kultur A genauso funktioniert wie in Kultur B (Hui & Triandis, 1985a). Bevor diese Annahme getroffen werden kann, muß die Äquivalenz dieses Konstrukts in beiden Kulturen überprüft werden.

Unter Äquivalenz wird dabei grundsätzlich verstanden, daß Meßinstrumente über verschiedene Erhebungsgruppen hinweg die gleichen psychometrischen Eigenschaften besitzen müssen (Poortinga, 1989; van de Vijver & Poortinga, 1997). Dies bedeutet vor allem, daß ein Unterschied in den erfaßten Meßwerten auch immer einen Unterschied in dem zugrundeliegenden Konstrukt darstellen soll. Poortinga (1989) trennt daher auch zur Erläuterung dieser Problematik zwischen a) dem zugrundeliegenden und in der Regel latenten Konstrukt, b) einer sogenannten Vergleichsskala (comparison scale) und c) der Meßwert-Skala (measurement scale). In der Regel wird bei der Durchführung psychologischer Studien versucht, ein psychologisches und nicht direkt meßbares Konstrukt, z.B. ein Persönlichkeitsaspekt oder ein Einstellungskonstrukt, zu erfassen; für den Fall der hier vorliegenden Arbeit wäre dies ‚die Zufriedenheit mit der Arbeit‘. Dieses latente Konstrukt, welches somit nicht direkt beobachtbar und meßbar ist, wird im Zuge der Operationalisierung in die Vergleichsskala (comparison scale) überführt, d.h. es werden Aspekte gesucht (z.B. Verhaltensweisen), die das Konstrukt in den verschiedenen Kulturen gleichermaßen repräsentieren. Die konkreten Ausprägungen dieser Aspekte können dann auf der finalen Meßwert-Skala erfaßt werden. Nimmt man

einmal an, daß ein latentes Konstrukt in zwei verschiedenen Kulturen gleichermaßen existieren würde und nimmt man ebenso an, daß sich dieses Konstrukt durch eine identische Vergleichsskala operationalisieren läßt, dann läge somit nur dann Äquivalenz vor, wenn Unterschiede auf der Meßwert-Skala auch mit Unterschieden auf der Vergleichsskala einhergehen (Poortinga, 1989). Dies wäre jedoch nicht der Fall, wenn trotz der genannten beiden Voraussetzungen zusätzliche Faktoren Einfluß auf die Meßwert-Skala nehmen, jedoch hingegen keinen Einfluß auf die Vergleichsskala und das zugrunde liegenden Konstrukt haben.

Äquivalenz im interkulturellen Bereich kann jedoch noch etwas weiter gefaßt werden und bezieht sich daher auf mehr als nur diese gleichgeartete Veränderung von Vergleichs- und Meßwert-Skala. Hui und Triandis (1985a) unterscheiden diesbezüglich vier verschiedenen Arten und Ebenen von Äquivalenz:

- a) Konzeptuelle/Funktionale Äquivalenz: Ein Konstrukt muß in verschiedenen Kulturen gleichermaßen verstanden werden und wird erst durch dieses gleiche Verständnis vergleichbar. So muß beispielweise unter „Liebe“ oder auch „Loyalität von Kindern zu ihren Eltern“ in Kultur A inhaltlich das Gleiche wie in Kultur B verstanden werden.
- b) Äquivalenz in der Operationalisierung des Konstrukts: Das untersuchte Konstrukt muß sich in verschiedenen Kulturen in den gleichen Verhaltensweisen oder beispielsweise auch Einstellungen ausdrücken.
- c) Item-Äquivalenz: Ein Konstrukt mit der gleichen Bedeutung und Operationalisierung in verschiedenen Kulturen muß durch das gleiche Instrument erfaßt werden können. Nur dann können numerische Meßwerte aus den verschiedenen Kulturen direkt miteinander in Beziehung gesetzt werden.
- d) Skalare Äquivalenz: Zusätzlich zu den zuvor genannten Äquivalenzformen besitzt ein Meßwert skalare Äquivalenz, wenn seine identische Ausprägung auch der gleichen Intensität, Stärke oder Menge hinsichtlich des Konstrukts in verschiedenen Kulturen entspricht.

Anhand der Beschreibung der vier Ebenen wird deutlich, daß die jeweils zuerst genannten allgemeineren Äquivalenz-Formen die Voraussetzung für die Erreichung der spezifischeren Arten von Äquivalenz sind.

In ähnlicher Form unterscheiden van de Vijver und Leung (1997) die Konstrukt- oder auch Strukturelle Äquivalenz von der Äquivalenz der Maßeinheiten und der skalaren

Äquivalenz. Während die Konstrukt-Äquivalenz der gleichen Bedeutung wie bei Hui und Triandis (1985a) folgt, unterscheidet sich das Verständnis jedoch leicht bei den anderen Formen von Äquivalenz. Unter Äquivalenz der Maßeinheiten kann gemäß van de Vijver und Leung (1997) die Messung anhand gleicher Einheiten verstanden werden, wo jedoch der Ursprung der Skala nicht unbedingt identisch sein muß. Ein Beispiel dafür wäre die Messung von Temperatur anhand von Kelvin oder Celsius Skalen. Die höchste Form der Äquivalenz, die skalare Äquivalenz, bedeutet in ihrem Verständnis schließlich die Messung der Werte auf einer Verhältnisskala, d.h. Einheiten und Anfang der Skala sind ebenfalls identisch.

Diese verschiedenen Formen von Äquivalenz, oder auch anders ausgedrückt die sogenannten Invarianzen der zu untersuchenden Konstrukte, können dabei im interkulturellen Kontext nicht einfach angenommen werden, sondern müssen immer auch empirisch untermauert werden (van de Vijver & Poortinga, 1997). Eine Gefahr hinsichtlich der Äquivalenz stellen bei der Messung daher verschiedene Arten von Bias dar, die sich auf den eben erläuterten Leveln von Äquivalenz manifestieren können. Es läßt sich gemäß van de Vijver und Poortinga (1997) demnach ein Konstruktbias von einem Methodenbias und einem Itembias unterscheiden (vgl. auch van de Vijver & Leung, 1997).

Besteht zwischen verschiedenen Kulturen ein unterschiedliches Verständnis von einem Konstrukt, spricht man somit von einem *Konstruktbias*. Ein Vergleich über verschiedene Kulturen hinweg ist dann nicht mehr möglich und der Versuch einer Erfassung durch ein einziges gemeinsames Instrument ist damit hinfällig. Als ein prominentes Beispiel nennen die Autoren die Persönlichkeitskomponente ‚Respekt gegenüber den Eltern‘ (filial piety) (van de Vijver & Leung, 1997). Das westliche Verständnis dieses Persönlichkeitszuges ist bei diesem Konstrukt bei weitem nicht so umfangreich wie das chinesische Verständnis (Ho, 1996), was somit die Messung anhand eines gemeinsamen Instruments unmöglich macht.

Ein *Methodenbias* hingegen liegt vor, wenn insbesondere die Messung der Werte durch Faktoren beeinflusst ist, die nicht in einer Veränderung der Ausprägung des zugrunde liegenden Konstruktes anzusetzen sind. Eine Verzerrung der Werte kann im Sinne dieses Bias auch dann vorliegen, wenn das Verständnis des Konstruktes in den Kulturen identisch ist und somit ein Konstruktbias ausgeschlossen werden kann. Als klassische Formen von Methodenbias fallen beispielsweise das in dieser Arbeit behandelte, kulturspezifische Antwortverhalten und die damit verbundenen verschiedenen

Antwortstile in diesen Bereich; sowohl die Tendenz, eher Extremwerte anzukreuzen, als auch Zustimmungstendenzen und in gleicher Weise auch das Antworten im Sinne sozialer Erwünschtheit. Ebenso kann jedoch auch der Grund für eine methodische Verzerrung die unterschiedliche Gewohnheit und Vertrautheit in verschiedenen Kulturen sein, mit bestimmten Stimuli umzugehen. Beispiele dafür finden sich vor allem im Bereich der Intelligenzmessung und deren Vergleich bei verschiedenen Kulturen (vgl. sogenannte kulturfreie Intelligenztests). Gleichmaßen kann auch einen Einfluß haben, welche Art von Verhaltensreaktion bei der Anwendung spezieller Instrumente in verschiedenen Kulturen gefordert wird. Van de Vijver und Poortinga (1997) nennen Beispiele von Kindern aus England und Sambia, die bei der Messung von Coping-Stilen sich unterschieden, je nachdem ob sie dafür Bilder malen sollten oder aber Drähte verbiegen mußten. Trotz des gleichen Konstruktes, welches durch die beiden Verhaltensformen operationalisiert und erfaßt werden sollte, hatte jeweils die eine Kultur bei dem einen und die andere Kultur bei dem anderen Instrument höhere Meßwerte. Weitere klassische Einflußfaktoren im Sinne eines Methodenbias wären aber auch unterschiedliche physikalische Bedingungen in der Erhebungssituation oder auch Kommunikationsprobleme zwischen dem Testleiter und der Versuchsperson bzw. dem Interviewpartner (van de Vijver & Leung, 1997).

Die dritte Form des Bias schließlich, der *Itembias*, betrifft die Unangemessenheit einzelner Items. Während der Methodenbias sich eher auf die Veränderung der Meßqualität eines gesamten Instruments bezieht, liegt ein Itembias vor, wenn lediglich einzelne konkrete Items nicht die erforderliche Meßqualität besitzen. Im übertragenen Sinne bedeutet dies, daß bei gleicher Ausprägung des Konstrukts in verschiedenen Kulturen (zumeist gemessen über den Gesamtwert eines Instruments) die beiden Gruppen jedoch nicht den gleichen erwarteten Wert bei einem speziellen Item besitzen. Dieses spezielle Item scheint somit nicht in beiden Kulturen gleich gut zu funktionieren. Typische Gründe für einen Itembias sind gemäß van de Vijver und Leung (1997) beispielsweise eine unangemessene Übersetzung des Item-Wortlauts, zu komplexe Item-Formulierungen oder aber die Abfrage eines unangemessenen Item-Inhalts (wenn beispielsweise in einem Wissenstest Inhalte abgefragt werden, die nicht in beiden Gruppen gleichermaßen Bestandteil des Curriculums waren). Diese Aspekte führen dann dazu, daß trotz gleicher Ausprägung des zugrundeliegenden Konstrukts die verschiedenen Gruppen bei einzelnen Items nicht die gleichen Werte aufweisen. Diese Art von Items, die somit einen Bias enthalten, besitzt dann folglich auch eine

Information über die Gruppenzugehörigkeit, welche jedoch bei der Messung nicht erwünscht ist. Man könnte somit an der jeweils höheren oder niedrigeren Ausprägung erkennen, aus welcher Gruppe bzw. aus welcher Kultur die Versuchsperson stammt. Bei einem Item ohne Bias sollte der Wert einer Person jedoch generell keine Informationen über dessen Gruppenzugehörigkeit besitzen. Ein Item mit Bias weist somit jedoch diese ungewollte Information auf und sollte demnach eliminiert werden.

Ein Itembias läge auch dann vor, wenn Items zwar von der Bedeutung der einzelnen Worte her korrekt übersetzt worden sind, jedoch die Items in einem anderen kulturellen Kontext nicht mit der gleichen Bedeutung von den Befragten versehen werden. So werden in der Forschung oftmals Items in einem individualistischen Kontext entwickelt, dann übersetzt, um sie schließlich auch unter anderem in kollektivistischen Kontexten einzusetzen. Selbst wenn dabei die zumeist empfohlene und qualitativ bessere Strategie der Hin- und Rückübersetzung von Items bei der Übertragung in verschiedene Sprachen angewendet wurde (vgl. beispielsweise Werner & Campbell, 1970), ist damit jedoch keineswegs gesichert, daß Items aus einem individualistischen Kontext auch in einem kollektivistischen Kontext Sinn machen. So fanden beispielsweise Riordan und Vandenberg (1994), daß in verschiedenen Kulturen bei dem gleichen Item ein unterschiedlicher inhaltlicher Bedeutungsraum kognitiv aktiviert wurde. Beide Kulturen interpretierten die Bedeutung des Items somit unterschiedlich, obwohl eine korrekte Übersetzung vom Wortlaut her vorgenommen wurde. In ihrem konkreten Fall war die SWMSS (satisfaction with my supervisor scale), welche in einem individualistischen Kontext entwickelt wurde, nicht sinnvoll in Korea anzuwenden. In diesem Instrument wurden Items zur Beurteilung der technischen Kompetenz der jeweiligen Führungskraft eingesetzt, die für Koreaner jedoch nicht zu beantworten waren, da dies ein Aspekt ist, den sie nie bei ihrem Vorgesetzten hinterfragen würden, denn dieser wird eben genau wegen seiner höheren Kompetenz ihr Vorgesetzter. Eine Anwendung solcher Items führt somit in Korea nicht zu einer sinnvollen Messung der Zufriedenheit mit dem Vorgesetzten, obwohl eine inhaltlich korrekte Übersetzung des Iteminhalts vorgenommen wurde. Eine Zusammenfassung dieser drei bedeutsamsten Formen von Bias – d.h. Konstrukt-, Methoden- und Itembias – sowie die möglichen Gründe, die zu ihrer Entstehung führen, kann Tabelle 4 entnommen werden.

Neben der Kenntnis über verschiedenen Arten von Bias und dem Bewußtsein darüber, daß Äquivalenz ein zentraler Aspekt der interkulturellen Forschung darstellt, bleibt jedoch noch die Frage offen, welche Konsequenzen aus dem Vorhandensein von einem

Bias entstehen und wie man damit umgehen sollte. Poortinga (1989) unterscheidet diesbezüglich insgesamt vier verschiedene Herangehensweisen: a) Ausschluß jeglicher Vergleiche, b) Reduktion des Bias, c) Versuch der Interpretation des Bias und d) Ignorieren des Bias. Diese lassen sich hinsichtlich ihrer Angemessenheit und Anwendbarkeit unterscheiden.

Tabelle 4: Die drei Arten von Bias und ihre möglichen Gründe im Überblick (nach van de Vijver & Poortinga, 1997, S. 34)

Art des Bias	Gründe
Konstruktbias	<ul style="list-style-type: none"> - unvollständige Überlappung der Definitionen der Konstrukte über die Kulturen hinweg - unterschiedliche Angemessenheit des Iteminhalts (z.B. Fähigkeiten gehören nicht zum Repertoire jeder der kulturellen Gruppen) - schlechte Auswahl aller relevanten Verhaltensweisen (z.B. kurze Meßinstrumente zur Erfassung breiter Konstrukte) - unvollständige Erfassung des psychologischen Konstrukts
Methodenbias	<ul style="list-style-type: none"> - unterschiedliche Soziale Erwünschtheit - unterschiedliche Antwortstile wie beispielsweise extreme Antworten oder Zustimmungstendenzen - unterschiedliche Stimulus-Vertrautheit - mangelnde Vergleichbarkeit der Stichproben (z.B. Unterschiede hinsichtlich des Ausbildungshintergrunds, des Alters oder der Geschlechterzusammensetzung) - Unterschiede in den physikalischen Testbedingungen - unterschiedliche Vertrautheit mit der Antwortprozedur - Testleitereffekte - Kommunikationsprobleme zwischen den Untersuchungsteilnehmern und dem Testleiter in einer der beiden kulturellen Gruppen
Itembias	<ul style="list-style-type: none"> - schlechte Itemübersetzungen - unangemessene Itemformulierungen (z.B. komplizierte Wort- und Satzstrukturen) - ein oder mehrere Items erfassen auch zusätzliche Eigenschaften oder Fähigkeiten - zufällige Unterschiede in der Angemessenheit der Iteminhalte (z.B. wenn der Inhalt eines Wissensitems nicht im Curriculum in einer der beiden kulturellen Gruppen vorhanden war)

Die einfachste Form, mit einem vorhandenen Bias umzugehen, wäre somit der *Ausschluß jeglicher Vergleiche*. Jedoch ist diese Herangehensweise ebenso die konservativste Entscheidung und oftmals im interkulturellen Vergleich nicht passend,

da ganz im Gegenteil dazu ja oftmals der Vergleich der Daten in der interkulturellen Forschung geradezu erwünscht und erforderlich ist. Der Ausschluß von Vergleichen stellt somit keine wirkliche Alternative dar.

Daher schlägt Poortinga (1989) als angemessenste Form die *Reduktion des Bias* vor. Dabei bezieht sich jedoch die Reduktion des Bias oftmals nur auf auftretenden Itembias, da dieser einfacher zu beseitigen sind. Um diesen zu eliminieren wird in der Regel anhand von statistischen Item-Analysen versucht, unpassende Items aus einem Meßinstrument zu entfernen. Methoden- und Konstruktbias sind dadurch jedoch nicht zu entfernen. Item-Analysen können auch nur dann erfolgreich sein, wenn lediglich wenige Items in einem großen Itempool einen Bias aufweisen, denn um die Unbrauchbarkeit einzelner Items nachweisen zu können, wird immer ein passender Vergleichswert benötigt, der in der Regel in dem Gesamtmeßwert der entsprechenden Skala besteht. Je mehr Items jedoch mit einem Bias versehen sind, desto eher stößt diese Vorgehensweise an ihre Grenzen, denn dann würde somit auch der Gesamtwert umso mehr einen Bias enthalten. Die statistische Überprüfung eines einzelnen Items ist somit prinzipiell für sich genommen überhaupt nicht möglich, da immer ein Vergleichskriterium herangezogen werden muß, um die Brauchbarkeit eines einzelnen Items überprüfen zu können. Läßt man bei der Feststellung von Itembias sogar die Überlegung zu, daß möglicherweise der Grund des Bias doch auch in der Unvereinbarkeit der zugrundeliegenden Konstrukte liegen kann, würde folglich sogar auch diese Form der Item-Reduktion – nämlich die Herausnahme des Items aus der Skala – bei der Beseitigung des Bias nicht weiterhelfen.

Die dritte Art mit vorhandenem Bias umzugehen, *die Interpretation des Bias*, geht daher einen anderen Weg und versucht, den Bias nicht als ungewollten Einfluß auf die Meßwerte zu verstehen. Stattdessen dient der Bias als Informationsquelle für Gruppenunterschiede zwischen den verschiedenen gemessenen Kulturen. Ein Bias soll somit nicht zwangsläufig eliminiert werden, sondern es sollen vielmehr genau die Faktoren untersucht werden, die in der jeweiligen Kultur dafür sorgen, daß verschiedene Arten von Bias auftreten können. Es wird somit durchaus davon ausgegangen und auch akzeptiert, daß grundsätzlich Äquivalenzprobleme auftreten können. Das Ziel ist jedoch dann, Einflußgrößen zu finden, die diese fehlende Äquivalenz erklären können.

Schließlich ist als vierte Reaktionsmöglichkeit eine leider zu oft angewendete Form des Umgangs mit fehlender Äquivalenz das *Ignorieren des Bias* und das völlige Ausblenden solcher Probleme. Trotz zahlreicher Belege in der Forschung, die das Vorhandensein

von Bias zeigen, wird die Äquivalenz-Problematik jedoch weiterhin regelmäßig ignoriert und unterschätzt (Poortinga, 1989). Dabei wird oftmals von der Voraussetzung ausgegangen, daß die Anwendbarkeit von Instrumenten in verschiedenen Kulturen generell gegeben wäre. Dies würde jedoch auf der Annahme beruhen, daß Konstrukt- und Meßäquivalenz eines Instruments von vorneherein gegeben wären und nicht empirisch überprüft werden müßten. Diese Annahme ist im Bereich der interkulturellen Forschung jedoch sehr fragwürdig. Eine Anwendung dieser Art des Umgangs mit den verschiedenen Arten von Bias kann daher nicht empfohlen werden.

Während eine Reduktion des Bias auf Item-Ebene durchaus machbar ist und angestrebt werden sollte, stellt sich das Problem der Reduktion von Methoden- oder Konstruktbias etwas anders dar. Solche Formen von Bias können oft nicht wirklich auf einfache Weise eliminiert werden. Dennoch muß überprüft werden, inwieweit sie in den jeweiligen Studien vorhanden sind. Soll demnach ein Methodenbias überprüft und gegebenenfalls eliminiert werden, werden ähnlich wie bei der Reduktion von Itembias immer auch zusätzliche Vergleichswerte benötigt. So kann beispielsweise der Einfluß von (fehlender) Vertrautheit mit einem Stimulus durch Test-Wiederholungen relativ einfach eingeschätzt werden. Die Ergebnisse der Wiederholungsmessungen dienen dann als Vergleichsmaßstab. Sollte die Vertrautheit in der Tat eine Rolle spielen, so sollten die Werte in den Wiederholungstest folglich besser werden. Die gemessene Veränderung der Werte wäre dann ein Ausdruck eines Methodenbias, was aber natürlich dazu führen würde, daß diese Werte dann jedoch keinen direkten Rückschluß auf das zu messende Konstrukt zulassen. Diese Überlegungen treffen jedoch auch wiederum streng genommen nur für kognitive Tests oder Fähigkeitstests zu (van de Vijver & Poortinga, 1997). Nur dort kann eine höhere Vertrautheit zu ‚besseren‘ Ergebnissen führen, denn nur Probanden ohne Vorerfahrungen werden durch eine Testwiederholung profitieren. Bei Einstellungsmessungen oder aber Leistungstests mit einer natürlich oberen Grenze spielt die Testvertrautheit hingegen keine so bedeutende Rolle.

Neben einem Vertrautheitsbias kann ein Methodenbias, der in Form von speziellen Antwortstilen auftreten würde, ebenso nur mit weiteren Zusatzdaten überprüft werden. Hierbei entspricht die gängige Vorgehensweise in der interkulturellen Forschung, daß man in der Regel die Antwortstile von Befragten gesondert messen muß und diese dann mit anderen Items, die das eigentliche Konstrukt messen sollen, in Beziehung setzen kann (Harzing, 2006; Smith, 2004). Eine sehr spezielle Vorgehensweise, die versucht, eine statistische Prozedur zur Lösung der Problematik anzuwenden, haben Cheung und

Rensvold (2000) vorgeschlagen. Ihr Ansatz versucht mit Hilfe der Strukturgleichungsanalyse, die Erfassung dieser Art von Bias zu erreichen. Jedoch werden auch für ihr Vorgehen Vergleichswerte benötigt (nämlich hier in Form von mehreren Items, mit denen die statistische Analyse durchgeführt wird), so daß auch ihr Vorgehen diese Problematik der Vergleichswerte nicht lösen kann. Ein einzelnes Item ist somit auch durch dieses Vorgehen nicht zu überprüfen. Darüber hinaus ist deren Vorgehensweise wiederum auch von Little (2000) stark kritisiert worden, nämlich dahingehend, daß sie gar nicht in der Lage sei, die entsprechenden Arten von Bias zu entdecken, die sie eigentlich erfassen sollte. Über die Anwendbarkeit von Strukturgleichungsmodellen bei der Erfassung von Bias im Sinne von Antwortstilen besteht somit Uneinigkeit. Insgesamt bleibt daher festzuhalten, daß zur Erfassung von Methodenbias immer auch die Messung von weiteren Daten notwendig ist, die zusätzlich zum zu untersuchenden Konstrukt erhoben werden müssen (van de Vijver & Poortinga, 1997).

Zur Überprüfung der dritten Form von Bias, dem Konstruktbias, werden oftmals statistische Verfahren im Sinne von Faktorenanalysen, mehrdimensionalen Skalierungen und Überprüfungen von Kovarianz-Strukturen eingesetzt (vgl. zur Überprüfung der Konstruktäquivalenz des Arbeitszufriedenheitskonstrukts beispielsweise Müller, 2006). Bei der Frage nach der Vergleichbarkeit der Strukturen eines Konstruktes müssen diese Verfahren zu gleichen Ergebnissen in unterschiedlichen Kulturen kommen. Beispielsweise konnte Borg (1986) anhand einer mehrdimensionalen Skalierung zeigen, daß sich die im westlichen Kontext gemessenen Facetten von Arbeitswerten) auch in dieser Dreiteilung in anderen Kulturen replizieren ließen. Die mehrdimensionale Skalierung ergab jeweils eine klare Dreiteilung in den vom ihm untersuchten Kulturen. In gleicher Weise überprüfen (zumeist konfirmatorische) Faktorenanalysen, ob die in einem kulturellen Umfeld entwickelte Binnenstruktur eines Konstruktes sich in anderen Kulturen ebenso messen läßt. Können die gleichen Faktoren extrahiert werden, die Items in identischer Weise den gleichen Faktoren zugeordnet werden und sogar die Faktorladungen in gleicher Höhe gemessen werden, spricht dies in deutlicher Weise für die interkulturelle Übertragbarkeit eines Konstruktes. Jedoch darf solch ein Ergebnis dann auch wiederum nicht dahingehend falsch verstanden werden, daß solche statistischen Überprüfungen gleichzeitig dazu in der Lage wären, eine Aussage über Einflüsse von Methodenbias zu machen, denn dies ist nicht der Fall. Eine Bestätigung der Binnenstruktur eines Konstruktes läßt keine

Rückschlüsse auf das Vorhandensein oder Nichtvorhandensein eines Methodenbias zu. Dieser könnte ohne Probleme weiterhin vorhanden sein. So kann nämlich trotz gleicher Form der Binnenstruktur eines Konstruktes dennoch ein spezifisches Antwortverhalten, z.B. im Sinne einer Zustimmungstendenz, dazu führen, daß letztendlich die Ausprägungshöhen der einzelnen Konstruktbausteine sich in den einzelnen Kulturen unterscheiden, obwohl die Konstruktbausteine selbst interkulturell als identisch anzusehen sind. Dieser Aspekt kann nicht oft genug hervorgehoben werden, da eine Überprüfung eines eventuellen Konstruktbias anhand von geeigneten statistischen Verfahren dem Forscher die trügerische Sicherheit geben könnte, daß bei vorliegender Konstruktäquivalenz jegliche Form interkultureller Vergleichsprobleme gelöst sei. Die Bestätigung der Äquivalenz von Faktorenstrukturen ist jedoch nur eine wichtige Voraussetzung, sie löst aber nicht die komplette Problematik, denn eine fehlende Vergleichbarkeit kann trotz der Überprüfung dieser Grundvoraussetzung dennoch durch methodisch bedingte unterschiedliche Ausprägungshöhen zustande kommen, die als völlig losgelöst von dem zugrunde liegenden Konstrukt anzusehen ist und völlig unabhängig davon auftreten können. Die Replizierung der internen Faktorenstruktur eines Konstruktes in verschiedenen Kulturen reicht darüber hinaus auch deswegen nicht aus, da dieses Ergebnis nichts über die Verteilung der Antworten über die jeweilige Meßskala aussagt und somit keinen Rückschluß auf Schiefe und Kurtosis der jeweiligen Antwortverteilungen zuläßt (Byrne & Campbell, 1999). Genau diese Kennwerte von Verteilungen werden jedoch speziell durch das Vorhandensein von Methodenbias in Form von unterschiedlichem Antwortverhalten beeinflusst.

Aus den vorangegangenen Überlegungen wird deutlich, welche Bedeutung die Überprüfung von Äquivalenz in der interkulturellen Forschung besitzt. Bei der insgesamt eher kleineren Gruppe interkultureller Forscher ist dieses Bewußtsein zwar inzwischen weitestgehend vorhanden, jedoch wird die Problematik in vielen anderen Bereichen weiterhin ignoriert. Daß trotz zahlreicher Belege für die Existenz von verschiedenen Arten von Bias dennoch eine Überprüfung der Äquivalenz in vielen anderen Bereichen sowohl der Forschung als auch der Praxis oftmals ausbleibt, kann gemäß der Argumentation von Steenkamp und Baumgartner (1998) verschiedene Gründe haben. Angefangen bei zum Teil verwirrenden Terminologien und unklaren Begrifflichkeiten hinsichtlich der verschiedenen Arten von Bias, kann nach deren Ansicht ebenso die fehlende Vertrautheit der Forscher mit den anzuwendenden Verfahren oder auch deren Komplexität dazu führen, daß das Vorhandensein von Bias

nicht explizit überprüft wird. Darüber hinaus vermuten diese Autoren, daß ebenso bisher größtenteils fehlende Leitlinien zum Umgang oder zur Entfernung von Bias sowie Unklarheiten darüber, ab wann ein vorhandener Bias zur Unvergleichbarkeit von Werten führt, Forscher davor zurückschrecken lassen, die Äquivalenz empirisch zu untermauern. Wenngleich ihre Überlegungen zwar ursprünglich aus dem Bereich der interkulturellen Konsumforschung stammen, ist dennoch davon auszugehen, daß die gleichen Gründe auch für den Bereich der interkulturellen Organisationspsychologie herangezogen werden können.

Diese Aspekte können letztendlich aber auch keine Entschuldigung dafür darstellen, daß die Überprüfung von Äquivalenz oftmals nicht vorgenommen wird und leider oft der nicht optimalen Vorgehensweise des ‚Ignorieren des Bias‘ (Poortinga, 1989) gefolgt wird. Sie können lediglich als beschreibende Argumente dafür dienen, um zu erklären, warum trotz offensichtlicher Notwendigkeit dennoch die Überprüfung von Bias so selten in der Forschung und in der Praxis vorgenommen wird. Insgesamt bleibt jedoch festzuhalten, daß sich zweifelsohne dieses Forschungsfeld nicht als barrierefreie und problemfreie Zone darstellt, so daß diese Vernachlässigung zwar nicht tolerierbar aber zumindest nachvollziehbar ist:

„The analysis of equivalence, is not a highway along which cross-cultural psychology can easily proceed, it has more resemblance to irregularly placed marker sticks along a narrow path in the wilderness. But such markers are badly needed to help the researcher staying out of the swamps of groundless speculation about the psychological meaning of cross-cultural differences in psychological measurements.“ (Poortinga, 1989, S. 754)

Diese Schwierigkeiten sollte keine Art von interkulturellen Studie mehr außer Acht lassen, denn inzwischen scheint sich das Forschungsbewußtsein soweit gewandelt zu haben, daß eine Nichtberücksichtigung dieses Themas und somit eine Art von „safari type of research“ (S. 43) nicht mehr in der einschlägigen Forschungsliteratur publiziert werden kann (van de Vijver & Leung, 2000); zumindest in der interkulturellen Forschungsliteratur. Eine bestimmte Abfolge von Schritten zur Überprüfung der Äquivalenz kann es jedoch bei der Bewältigung dieser Schwierigkeiten leider nicht geben, da die notwendige Stufe der Meßäquivalenz – d.h. ob auf Konstrukt-, Methoden-

oder Itemebene – immer auch von den Zielen der jeweiligen Studie abhängt (Vandenberg & Lance; 2000). Diese Ziele lassen sich allgemein in drei Gruppen einteilen: a) es werden Unterschiede in den Beziehungen von Variablen untersucht, b) es werden Unterschiede in der Ausprägungshöhe von Variablen untersucht, c) es werden Unterschiede in der Varianz von Variablen untersucht. Speziell letzteres Vorhaben (d.h. der Vergleich von Varianzen) läßt sich jedoch eher selten in der Literatur finden. Die Ziele a) und b) hingegen sind oft der Standard in interkulturellen Studien. Während bei a) der Zusammenhang zweier Konstrukte (z.B. Intelligenz und Lernerfolg) und somit der relative Einfluß der Faktoren innerhalb jeder einzelnen Kultur geprüft wird, wird bei b) der absolute Wert des jeweiligen Konstrukts (z.B. das Intelligenzniveau) über verschiedene Kulturen hinweg verglichen. Es steht somit im Fokus, welche Kultur die größte Ausprägung beim jeweiligen Konstrukt besitzt, ungeachtet der Gründe, die dazu geführt haben. Genau dieses Ziel nehmen in der Regel Mitarbeiterbefragung für sich in Anspruch, bei denen untersucht wird, welche Niederlassung in welchem Land die größte Mitarbeiterzufriedenheit aufzuweisen hat. Da sich die vorliegende Arbeit genau mit der Problematik beschäftigt, inwieweit dieser Vergleich überhaupt zu bewerkstelligen ist, läßt sie sich somit der Gruppe von Zielen zuordnen, die unter b) erfaßt werden.

Die Bedeutung fehlender Äquivalenz und somit das Vorhandensein eines Bias unterscheidet sich letztendlich danach, um welche Art von Studien es handelt. Bei der Untersuchung der Beziehungen von Variablen untereinander ist beispielsweise das Vorhandensein eines Methodenbias weniger bedeutsam. An der Beziehung der beiden Variablen untereinander ändert sich dabei wenig, da in der Regel davon ausgegangen wird, daß ein Einfluß der Methode sich auf die gesamte Studie innerhalb einer Kultur auswirkt. Somit wären die beiden zu vergleichenden Variablen gleichermaßen davon betroffen. Ein Methodenbias, möglicherweise in Form von Antworttendenzen, hat jedoch im Gegensatz dazu gravierende Konsequenzen für den Vergleich von Absolutwerten über die Länder- und Kulturgrenzen hinweg. Dort führt er zu untypisch hohen oder niedrigen Werten, so daß ein Vergleich der Absolutwerte im Endeffekt nicht mehr zulässig und nicht mehr durchführbar ist. Dies würde nur zu ungerechtfertigten Schlußfolgerungen über die Unterschiedlichkeit von Kulturen führen, die jedoch eben nicht auf verschiedene inhaltliche Ausprägungen zurückzuführen sind, sondern auf den unterschiedlichen Einfluß eines Methodenbias. In diesem Sinne ist vor allem hinsichtlich der Anwendbarkeit von Mitarbeiterbefragungen im interkulturellen

Vergleichen die Frage nach dem Vorhandensein und dem Einfluß eines Methodenbias von entscheidender Bedeutung. Die vorliegende Arbeit konzentriert sich daher vor allem auf diese Art des Bias.

5.6 Antworttendenzen

Innerhalb der einen von Form des Bias, die eine besondere Bedeutung für die hier vorliegende Studie hat, spielen vor allem die Antworttendenzen als eine spezifische Form von Methodenbias eine zentrale Rolle und soll daher hier näher erläutert werden soll.

Hinsichtlich der Bedeutsamkeit von Antworttendenzen ist in vielen Teilen der Forschung – teilweise sogar auch in der interkulturellen Forschung – leider immer noch eine weit verbreitete, jedoch falsche Annahme, daß Personen bei der Beantwortung von Fragebögen lediglich von dem substantiellen Inhalt der Fragebogenitems beeinflusst sind. Dies ist jedoch oftmals nicht der Fall, denn ebenso wird das Kreuz auf einer Skala oft auch von Antworttendenzen beeinflusst. Antworttendenzen stellen dabei eine besondere Art des Antwortverhaltens dar, welche in der Regel nur bei Fragebögen zu beobachten sind, die mit den klassischen Formen der Likert-Skala arbeiten. Unter Antworttendenzen wird dabei verstanden, daß Befragte auf Fragen und Items eines Fragebogens eben nicht im Sinne des angesprochenen Inhalts der Frage antworten, sondern ihr Kreuz auf der Skala von nicht-inhaltlichen Aspekten beeinflusst ist (Fischer et al., 2006). Dabei werden die Items unabhängig von Inhalt, Formulierung oder auch Polung systematisch in einer bestimmten Art und Weise beantwortet.

5.6.1 Arten von Antworttendenzen

Im Allgemeinen werden bei Antworttendenzen drei Formen unterschieden, wie sich Befragte in Fragebögen äußern (Fischer et al., 2006):

- a) sie antworten auf der Skala ungewöhnlich häufig im Sinne einer Zustimmung, anstatt auch die Seite der Skala zu wählen, welche eine Ablehnung des Items implizieren würde
- b) sie nutzen tendenziell stärker die Endpunkte einer Skala als die Mitte oder die gesamte Bandbreite der Skala
- c) sie antworten so, wie sie glauben, daß es in ihrem sozialen Umfeld am Angebrachtesten sei

Die erste dieser Formen des Antwortverhaltens wird dabei als Zustimmungstendenz oder auch kurz als ARS (acquiescent response style) bezeichnet. Zeigt ein Befragter diese Form der Kommunikation in Fragebögen, dann bedeutet dies, daß er seine Kreuze überzufällig häufig auf der Seite der Skala macht, die mit einer Zustimmung betitelt ist. Bei Skalenbezeichnungen, die beispielsweise von „stimme völlig zu“ bis „stimme überhaupt nicht zu“ reichen, finden sich somit tendenziell mehr Kreuze auf der „stimme völlig zu“-Seite. Der Befragte zeigt somit generell die Tendenz, höhere oder niedrigere Antworten zu geben, unabhängig von der eigentlichen Thematik des jeweiligen Items (Cheung & Rensvold, 2000). Wichtig ist dabei zu beachten, daß Zustimmungstendenzen sozusagen per Definition als unabhängig vom Itemkontext verstanden werden. Nur wenn sich diese Tendenz über verschiedene Item-Gruppen und -inhalte zeigt, spricht man in der einschlägigen Literatur von Zustimmungstendenzen. Eine Interaktion mit dem Inhalt der Items, d.h. der Befragte antwortet nur bei bestimmten Item-Sets zustimmend, während er dies bei anderen Sets mit anderer Thematik nicht tut, wird somit ausgeklammert. Sollte dieser Fall einer Interaktion tatsächlich einmal vorliegen, würde dies gemeinhin nicht mehr als generelle Antworttendenz verstanden. Zwar könnte man sich grundsätzlich auch vorstellen, zustimmende Antworten, die nur bei einem thematischen Item-Set auftreten, als Zustimmungstendenz zu bezeichnen. Jedoch würde diese Art des Verständnisses auch methodisch in eine Art Sackgasse führen, denn aus theoretischer Sicht zeigt ein Befragter ja gerade dann eine Zustimmungstendenz, wenn er über die unterschiedlichsten – idealerweise sogar entgegengesetzten – Themenkomplexe immer die „stimme völlig zu“-Seite ankreuzt. Dieses Verhalten wäre inhaltlich gesehen – d.h. im Sinne des Item-Inhalts – nicht mehr plausibel und kann daher nur als Antworttendenz verstanden werden. Würde der Befragte stattdessen dieses Antwortverhalten nur bei manchen Item-Sets zeigen, jedoch bei anderen nicht, könnte es auch sein, daß seine hohen Antworten tatsächlich durch eine bejahende Einstellung zu dem speziellen Themenkomplex zustande gekommen sind. Der Forscher könnte somit letztendlich dann nicht zwischen Methodenbias und Inhalt trennen und könnte nicht mehr entscheiden, ob die hohen Antworten eine Antworttendenz oder die tatsächliche Meinung des Befragten ausdrücken. Daher wird somit aus theoretischer Sicht nur das zustimmende Antwortverhalten über eine Vielzahl von Themen und deren Items als Zustimmungstendenz bezeichnet. Zur Messung von Zustimmungstendenzen wird folglich dann auch in der Regel eine Vielzahl möglichst heterogener und

gegebenenfalls auch entgegengesetzter Themenkomplexe herangezogen (vgl. Harzing, 2006; Smith, 2004).

Von der Zustimmungstendenz (ARS) wird oft auch eine Art von Ablehnungstendenz (DARS = disacquiescent response style) abgegrenzt. Dabei wählt der Befragte überzufällig häufig die Seite der Skala aus, die eine Ablehnung darstellt; d.h. in Bezug zum oben genannten Beispiel finden sich überzufällig mehr Kreuze auf der „stimme überhaupt nicht zu“-Seite. Für das Verständnis der Ablehnungstendenz gilt in gleicher Weise wie bei der Zustimmungstendenz, daß diese als unabhängig vom Iteminhalt verstanden wird und in gleicher Form über möglichst heterogene Themenkomplexe erhoben wird. Um beide Seiten einer Skala gleichzeitig betrachten zu können, wird schließlich auch oftmals eine Differenz dieser beiden Tendenzen errechnet, die als Netto-Zustimmungstendenz (NARS = net acquiescent response style) bezeichnet wird (siehe bspw. Harzing, 2006). Sowohl das rein einseitige Maß der Zustimmungstendenz (ARS) als auch die Netto-Zustimmungstendenz (NARS), die beide Seiten der Skala berücksichtigen kann, werden in vielen interkulturellen Studien gleichermaßen herangezogen. Dabei scheint es in der Literatur keine eindeutige Präferenz zu geben, welches dieser beiden Maße überlegen wäre. Ein Vergleich dieser beiden Maße geschweige denn eine Diskussion über die Überlegenheit des einen oder des anderen Maßes findet nirgends statt. Es zeigt sich lediglich, daß eine Operationalisierung von Zustimmungstendenzen in der Regel nicht rein über die Messung der Ablehnungstendenz (DARS) durchgeführt wird. Ob in Studien das Maß der ARS oder der NARS angewendet wird, scheint somit mehr von der Überzeugung des jeweiligen Forschers abzuhängen.

Es ließen sich jedoch in der Tat Argumente heranziehen, die für oder gegen das jeweilige Maß sprechen würden und die interessanterweise in der Literatur an keiner Stelle erwähnt werden. Die Netto-Zustimmungstendenz (NARS) ist in der Lage ist, die Verteilung auf der gesamten Skala widerzuspiegeln, denn bei ihr wird in Beziehung gesetzt, wie oft jemand sozusagen ‚links‘ auf der Skala angekreuzt hat, in Relation dazu, wie oft er ‚rechts‘ angekreuzt hat. Dieser Vorteil basiert jedoch auch auf einer wichtigen Voraussetzung, die auch in der einschlägigen interkulturellen Literatur nicht deutlich erwähnt und diskutiert wird. Diese Berechnungsweise geht implizit davon aus, daß das Gegenteil von einer Zustimmungstendenz eine Ablehnungstendenz sei und man daher idealerweise beide in Beziehung setzen müsse, um das Antwortverhalten eines Befragten auf einer Skala optimal abbilden zu können (vgl. Cheung & Rensvold, 2000;

Harzing, 2006). Das Gegenteil von einer Zustimmungstendenz könnte aber auch durchaus eine ausgewogene Nutzung der Gesamtskala sein. Befragte, die keine Zustimmungstendenz zeigen, würden somit nicht die Items ablehnen, sondern gemäß ihrer tatsächlichen Meinung oder Einstellung die gesamte Breite der Skala nutzen. Das Gegenteil einer Zustimmungstendenz wäre daher dann nicht eine Ablehnungstendenz, sondern sozusagen eine Nicht-Zustimmungstendenz. Beide Maßzahlen, die Zustimmungstendenz (ARS) und die Netto-Zustimmungstendenz (NARS), haben somit ihre Berechtigung, gehen jedoch von anderen Annahmen aus. Da beide Maßzahlen somit sowohl Vor- als auch Nachteile haben, könnten beim konkreten empirischen Vorgehen in einer Studie auch beide Maßzahlen miteinander verglichen werden. Die Netto-Zustimmungstendenz (NARS) kann jedoch im Sinne des zugrunde liegenden Konstrukts als das bessere Maß angesehen werden.

Studien, die das Auftreten von Zustimmungstendenzen zeigen konnten, gibt es in der Literatur zahlreiche (siehe u.a. Church, 1987; Cunningham, Cunningham & Green, 1977; England & Harpaz, 1983; Johnson et al., 1997; Marin et al., 1992; Morris & Parvett, 1992; Ross & Mirowsky, 1984; Steenkamp & Baumgartner, 1998; van Herk et al., 2004). Während zwar einzelne solcher Studien speziell dem methodologischen Forschungsfeld entstammen, läßt sich der Großteil der Studien jedoch dem interkulturellen Forschungsinteresse zuordnen (vgl. auch Harzing, 2006, Smith, 2004). Gemessen werden Zustimmungstendenzen (ARS) in diesen Studien in der Regel dadurch, daß die Anzahl der Kreuze auf einer Likert-Skala, die im Sinne der Skalenbezeichnung Zustimmung ausdrücken, aufsummiert und durch die Gesamtzahl der Antworten geteilt werden (vgl. Harzing, 2006). Das heißt beispielsweise bei einer 5-Punkt-Likert-Skala, die aus 10 Items besteht, daß die Kreuze, die bei 4 oder 5 gemacht wurden (wenn 5 dabei „stimme völlig zu“ entspricht) addiert werden und dann zur Gesamtzahl ‚10‘ der Antworten in Beziehung gesetzt werden. Man erhält folglich immer einen Kennwert, der zwischen 0 und 1 schwanken kann, wobei 1 die höchste Zustimmungstendenz ausdrückt. Eine weitere Binnen-Differenzierung und -gewichtung zwischen dem Kreuz auf der 4. und dem Kreuz auf der 5. Stufe wird in der Regel nicht vorgenommen. Stattdessen werden beide Kreuze gleichgewichtet in den Index aufgenommen. Man könnte auch eine Gewichtung in der Form vorgenommen, daß dem Kreuz bei 5 der Wert 2 zugesprochen wird, während das Kreuz bei 4 den Wert 1 erhält. Das äußerste Kreuz auf der Likert-Skala ginge dann somit mit der doppelten Wichtigkeit in den Index ein. Diese Vorgehensweise wird jedoch nur sehr selten

angewendet, da ein nach diesem Vorgehen gebildeter Index einen bedeutsamen Nachteil mit sich trägt, denn er würde hoch korrelieren mit dem Maß der Extremtendenz, welches später noch ausführlicher erläutert wird. Somit ist die zu präferierende Variante, lediglich die Kreuze auf den beiden äußersten Zustimmungspunkten gleichgewichtet zu zählen. Bei breiteren Skalen mit mehr als 5 Skalenpunkten werden zur Messung der Zustimmungstendenz dann auch teilweise die 3 äußersten Kreuze addiert. Bei der Berechnung der Ablehnungstendenz (DARS) wird zumeist in gleicher Weise vorgegangen, nur daß dort im Umkehrschluß bei einer 5er-Likert-Skala die Werte 1 und 2 aufsummiert und an der Gesamtzahl der Itemantworten relativiert werden. Man erhält auch dort einen Index von 0 bis 1. Bei der Berechnung der Netto-Zustimmungstendenz (NARS) werden dann die auf diese Weise berechneten Maße von ARS und DARS voneinander subtrahiert. Somit erhält man letztendlich einen Index, der von -1 bis 1 schwanken kann. Der Wert -1 entspräche dann einer ausschließlichen Ablehnungstendenz bei den jeweils gemessenen Items, d.h. kein Item kann zustimmend beantwortet worden sein. Der Wert 1 wäre dann demgegenüber als eine reine Zustimmungstendenz zu deuten, während der Wert 0 einer Beantwortung entspricht, die sowohl Ablehnung als auch Zustimmung beinhaltet.

Als eine der wenigen Ausnahmen, bei denen Zustimmungstendenzen nicht nach dieser Vorgehensweise erfaßt werden, sei hier die Studie von Smith (2004) genannt. Dort wurden stattdessen die Mittelwerte über jeweils alle Dimensionen von verschiedenen Meßinstrumenten zur Erfassung von Wertvorstellungen (Schwartz, GLOBE-Studie, Trompenaars, Hofstede) errechnet und dienen als ein Maß von zustimmendem Antwortverhalten. Ein Unterschied in der Berechnung im Vergleich zur vorherigen Variante besteht dabei darin, daß bei dieser Art der Mittelwertsberechnung verschiedene Höhen von Zustimmungstendenz errechnet werden, je nachdem ob ein Befragter das äußerste oder das zweitäußerste Ende der Skala ankreuzt. Dies macht bei der vorherigen Variante der Auszählung der Kreuze jedoch keinen Unterschied. Der Nachteil dieser Vorgehensweise besteht daher auch darin, daß solch ein Indikator stärker mit Extremtendenzen zusammenhängen würde. Im Sinne der reinen Zustimmungstendenz wäre daher das ungewichtete Aufsummieren gemäß der oben geschilderten Vorgehensweise zu präferieren.

Ein spezielles Meßinstrument, ein spezieller Fragebogen oder eine spezielle Skala, die explizit das Ausmaß von Zustimmungstendenzen mißt, läßt sich jedoch in der Literatur bis heute nicht finden (Smith & Fischer, 2008). In der Regel werden die bereits im

Fragebogen vorgefundenen Items – möglichst aus heterogenen Themenfeldern – genutzt, um Zustimmungstendenzen zu erfassen (siehe van Herk et al., 2004).

Die zweite klassische Form des Antwortverhaltens wird als Extremtendenz oder kurz auch als ERS (extreme response style) bezeichnet. Auch diese Form des Antwortverhaltens wird zumeist beim Einsatz von Likert-Skalen beobachtet. Dabei tendieren die Befragten dazu, bei einer Likert-Skala eher die Eckpunkte zu wählen; d.h. beispielsweise bei einer 5er Skala dann die Werte 1 (stimme überhaupt nicht zu) und 5 (stimme voll und ganz zu). Auch hier existiert als ein entgegengesetztes Phänomen, welches als Mittentendenz, oder auch kurz als MRS (middle response style) bezeichnet wird (Cheung & Rensvold, 2000; Hui & Triandis, 1989). Dabei wählen die Befragten überzufällig häufig die mittlere Kategorie einer Likert-Skala aus – d.h. die 3 auf einer 5er-Skala –, oder aber die beiden mittleren Kategorien, falls es sich um eine Skala mit gerader Anzahl von Auswahlmöglichkeiten handelt. Ähnlich wie bei den Erläuterungen zu den Zustimmungstendenzen und der Gegenüberstellung von ARS und DARS läßt sich jedoch auch hier die Annahme der Gegensätzlichkeit von ERS und MRS, die u.a. von Cheung und Rensvold (2000) vertreten wird, derart kritisieren, daß das gegensätzliche Verhalten zu ERS automatisch MRS sei. Zwar ist es hier bei der Argumentation zu Extremtendenzen insgesamt plausibler, daß Befragte, die nicht extreme Antworten geben möchten, folglich gemäßigte Antworten geben und somit automatisch mehr in der Mitte der Skala landen. Dennoch kann das Gegenteil von extremem Antwortverhalten auch generell wie bei den Zustimmungstendenzen eine ausgewogene Nutzung der Skala gemäß der tatsächlichen Einstellung sein, was somit dann Nicht-Extremtendenzen als das entgegengesetzte Verhalten ausweisen würde.

Studien, die das Vorhandensein von Extremtendenzen gezeigt haben, gibt es ebenso zahlreich in der Literatur (vgl. Chun, Campbell & Yoo, 1974; Clarke III, 2001; Greenleaf, 1992a; Hui & Triandis, 1985b; Lee & Green, 1991; Schaninger & Buss, 1986; Triandis, 1994; van Herk et al., 2004). Dabei wird ähnlich wie bei den Zustimmungstendenzen zumeist die Extremtendenz in der Form gemessen, daß die Anzahl der Kreuze in den extremen Antwortoptionen, d.h. 1 und 5 bei einer 5er-Skala, gezählt werden und in Relation zu den Gesamtantworten gesetzt werden (Harzing, 2006). Das Resultat ist wiederum ein Index von 0 bis 1. Bei der Berechnung von Mittentendenzen wird in der Regel nur die Anzahl der Kreuze bei 3 oder unter Umständen bei 3 und 4 genommen, sollte beispielsweise eine 6er-Skala vorliegen.

Anders als bei den Zustimmungstendenzen wird jedoch in der Regel keine Verrechnung von ERS und MRS vorgenommen, sondern beide Indizes werden getrennt voneinander behandelt. Manche Autoren gehen sogar soweit, daß sie die Extremtendenz noch in eine positive und eine negative Extremtendenz aufteilen (s. die Studie von Harzing, 2006) und die Anzahl von Kreuzen bei 1 und bei 5 getrennt erfassen. Oftmals wird jedoch nur eine allgemeine Extremtendenz erfaßt, die dann beide Enden der Skala umfaßt.

Während auch bei dieser Antworttendenz die Auszählung der Anzahl von Kreuzen auf bestimmten Antwortstufen die häufigste Form der Berechnung darstellt, geht auch hier Smith (2004) einen anderen Weg, indem er die Standard-Abweichung als ein Maß für extremes Antwortverhalten versteht. Diese Vorgehensweise wird jedoch an anderer Stelle (Cheung & Rensvold, 2000; Greenleaf, 1992a) zu Recht dahingehend kritisiert, daß Maße für extremes Antwortverhalten und die Standard-Abweichung zwar häufig hoch korreliert sind, jedoch keineswegs identisch sind. So können beispielsweise durchaus Unterschiede in Standard-Abweichungen vorliegen, ohne daß auch nur eine einzige extreme Antwortkategorie angekreuzt wurde. Dieser Aspekt wird umso bedeutsamer, je größer die Breite der Skala ist.

Ein weiterer, jedoch eher seltener Ansatz ist die unterschiedliche Gewichtung von Antworten auf unterschiedlichen Skalenpunkten. So rekodieren Smith und Fischer (2008) ihre Daten in der Form, daß Antworten auf den äußersten Ende der Skala (z.B. 1 und 5 bei einer 5-Punkt-Likert-Skala) den Wert 2, Antworten auf den weniger extremen Werten der Skala (z.B. 2 und 4 auf einer 5-Punkt-Likert-Skala) den Wert 1 und Antworten in der Mittelkategorie den Wert 0 erhalten. Dadurch beziehen sie zwar die gesamte Skalenbreite mit ein, geben jedoch für die unterschiedlichen Antworten auch eine unterschiedliche Wertigkeit vor.

Im Gegensatz zur Messung von Zustimmungstendenzen gibt es in der Tat im Bereich der Extremtendenzen einen Versuch, ein spezielles Instrument zur Messung dieses Antwortverhaltens zu entwickeln (Greenleaf, 1992b). Dazu entwickelte der Autor eine Skala bestehend aus insgesamt 16 Items, die inhaltlich möglichst heterogen sein sollten. Dieses Instrument wurde bereits in mehreren Studien erfolgreich eingesetzt (vgl. u.a. Clarke III, 2000; 2001). Bis heute ist es auch im Prinzip überhaupt das einzige Instrument, was explizit zur Messung von Extremtendenzen angefertigt wurde (de Jong, Steenkamp, Fox & Baumgartner, 2008). Weitere Instrumente existieren nicht, so daß auch Maße von Extremtendenzen zumeist anhand der schon vorliegenden Items berechnet werden. Gegen diese Berechnung anhand vorhandener Itemblöcke ist

grundsätzlich nichts einzuwenden. Jedoch besteht zumeist die Problematik, daß in Fragebögen nur eine begrenzte Anzahl von Themenkomplexen abgefragt wird, so daß möglicherweise die vorhandenen Items nicht inhaltlich heterogen genug sind. Darüber hinaus findet selbstverständlich eine Art von Konfundierung statt, wenn das gleiche Meßinstrument auf der einen Seite zur Messung inhaltlicher Aspekte angewendet wird und auf der anderen Seite zur Messung von Antworttendenzen seinen Beitrag liefert.

Als dritte Form der klassischen Antworttendenzen läßt sich die soziale Erwünschtheit, kurz SD (social desirability), nennen. Unter sozialer Erwünschtheit wird dabei verstanden, daß Personen versuchen, ihre wahre Meinung zu verbergen und stattdessen so zu antworten, wie sie glauben, wie es in ihrem gesellschaftlichen Umfeld angemessen ist. Sie geben somit in Fragebögen weniger das an, was sie selbst wirklich denken, als vielmehr das, was sie glauben, was man von ihnen hören möchte. In dieser Form ist somit soziale Erwünschtheit ebenfalls ein verzerrender Faktor, der es erschwert, den wahren Wert eines Konstruktes, der ja eigentlich gemessen werden soll, tatsächlich festzustellen. Ein Großteil der Forschung zu sozialer Erwünschtheit fand jedoch nicht im Bereich der interkulturellen Forschung und daher nicht auf dem kulturellen Level statt. Die meiste Forschung spielte sich vielmehr auf dem individuellen Level ab. So haben sich folglich mit dieser Thematik hauptsächlich Persönlichkeitspsychologen beschäftigt. Soziale Erwünschtheit hängt gemäß einer großen Meta-Analyse von Ones, Viswesvaran und Reiss (1996) mit insgesamt 233 Studien auf dem individuellen Level vor allem mit den Big Five Persönlichkeitsdimensionen Emotionale Stabilität und Gewissenhaftigkeit zusammen. Der Fokus der Analysen zur Sozialen Erwünschtheit lag vor allem darauf, zu überprüfen, welche individuellen Aspekte ein sozial erwünschtes Antwortverhalten hervorrufen und welche individuellen Konsequenzen diese Art von Antwortverhalten hat. Ein Teil der Forschung hat sich hinsichtlich der Konsequenzen vor allem damit beschäftigt, welchen Einfluß dieses verzerrende Antwortverhalten in Auswahlentscheidungen im Personalbereich hat, wo Persönlichkeitsmeßverfahren eingesetzt werden sollten (Zerbe & Paulhus, 1987). Dort sollte diese Art von Antwortverhalten vergleichsweise häufig auftreten, da Menschen vor allem in solchen Situationen, in denen sie stark davon profitieren, ein positives Bild nach außen abzugeben – z.B. eben bei Personalauswahlentscheidungen –, sich verstärkt im Sinne von Sozialer Erwünschtheit äußern (Levin & Montag, 1987). Ziel war es daher zumeist, die Verzerrung durch sozial erwünschtes Verhalten aus den eigentlichen Meßwerten

herauszurechnen. Anstatt jedoch Soziale Erwünschtheit als reines Meßartefakt anzusehen, war vielmehr die Schlußfolgerung der umfassenden Meta-Analyse von Ones et al. (1996), daß das Antworten in Form von Sozialer Erwünschtheit ein deutlicher Ausdruck der jeweiligen Persönlichkeit ist. Beim Einsatz von Persönlichkeitsfragebögen in Personalauswahlentscheidungen würde somit ein Herausrechnen der Sozialen Erwünschtheit aus den Persönlichkeitsmeßwerten auch ein Teil der wahren Varianz vernichten.

So wie Soziale Erwünschtheit auf individueller Basis scheinbar ein Ausdruck von Unterschieden zwischen der Persönlichkeit verschiedener Menschen zu sein scheint, könnte folglich auch angenommen werden, daß es ebenfalls auf kultureller Ebene Unterschiede zwischen Nationen geben mag, inwieweit man sich in der einen Kultur eher so verhält, wie man glaubt, daß es von einem erwartet wird, und in anderen Kulturen sich nicht nach diesen Erwartungen richtet. Die zentrale Frage dabei ist dann vor allem, was die Erwartungen sind, die an die Mitglieder einer Kultur gestellt werden. In einer Befragungssituation könnte dies zum Beispiel auch sein, sich möglichst zurückhaltend in der Kommunikation mit anderen Menschen zu äußern. Diese Art von Antwortverhalten wäre dann jedoch sehr ähnlich einer Zustimmung- oder auch Mittentendenz. Vergleich man diesbezüglich die bisherigen Diskussion in der einschlägigen interkulturellen Forschungsliteratur, so findet sich in der Tat oftmals keine explizite Trennung zwischen Sozialer Erwünschtheit und anderen Formen von Antworttendenzen (z.B. Zustimmung- und Extremtendenzen). So unterscheidet beispielweise Smith (2004) überhaupt nicht zwischen diesen verschiedenen Formen, sondern nennt in einem Atemzug Zustimmungstendenzen und sozial erwünschtes Antwortverhalten. Es scheint somit so zu sein, daß der Autor beide Formen als das gleiche Phänomen betrachtet. In gleicher Weise nehmen auch Cheung und Rensvold (2000) keine Trennung vor, wenn sie als mögliche Gründe für das Auftreten von Zustimmungstendenzen angeben, daß es sozial erwünscht sei, so zu antworten.

Die Vermischung dieser beiden Konstrukte ist vor allem daher diskussionswürdig, da eigentlich aus theoretischer Perspektive diese beiden Konstrukte getrennt verstanden werden. So bezieht sich Soziale Erwünschtheit auf einzelne Aspekte und beispielsweise einzelne Themenfelder eines Fragebogens, während Zustimmungstendenzen als ein Antwortverhalten verstanden wird, welches sich themenübergreifend zeigt (Fischer et al., 2006). Soziale Erwünschtheit ist per Definition an den Iteminhalt gebunden und variiert je nach Iteminhalt (Paulhus, 1984). Zustimmungstendenzen jedoch werden per

Definition als unabhängig vom Item-Inhalt verstanden. Eine Vermischung dieser beiden Konstrukte bei der Erklärung kulturspezifischen Antwortverhaltens führt somit lediglich in eine inhaltliche Sackgasse. Zwar kann davon ausgegangen werden, daß beide Konstrukte durchaus miteinander zusammenhängen, sie sind jedoch nicht identisch. So kann es beispielsweise auch in manchen Kulturen sozial erwünscht sein, ablehnend zu antworten (Smith & Fischer, 2008). Im Gegensatz zur Forschung zu Zustimmungstendenzen sind in der langen Tradition der Forschung zur Sozialen Erwünschtheit inzwischen auch mehrere Verfahren entwickelt worden, die diese Art von Antwortverhalten messen wollen. Die bekanntesten Verfahren sind die Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960), die Lügenskala des Eysenck Personality Inventory (Eysenck & Eysenck, 1975) sowie das Balanced Inventory of Socially Desirable Responding (Paulhus, 1984). Diese Arten von Meßverfahren basieren in der Regel darauf, daß den Befragten Items vorgelegt werden, bei denen es sehr ungewöhnlich wäre, ihnen komplett zuzustimmen (z.B. „Ich habe noch nie gelogen“). Eine deutliche Bejahung vieler solcher Items wird dann als Ausdruck dessen verstanden, daß der Befragte ein möglichst gutes Bild nach außen hin darstellen und seine wahre Meinung bzw. sein wahres Verhalten verbergen will. Dabei sei zu erwähnen, daß Paulhus (1984) bei seinem Verständnis von sozial erwünschtem Verhalten eine Zweiteilung vornimmt und zwischen einerseits Impression Management und andererseits Selbst-Täuschung unterscheidet. Während ersteres das schön gefärbte Bild nach außen beschreibt, umfaßt der zweite Aspekt das schön gefärbte Bild, welches man vor sich selbst vorgibt. Die anderen beiden genannten Meßverfahren jedoch fassen Soziale Erwünschtheit als eindimensionales Konstrukt auf, welches eher dem Verständnis der Beschönigung des Bildes nach außen entspricht. Anhand dieser drei Meßverfahren wird jedoch deutlich, daß es neben dem Unterschied in der theoretischen Definition auch einen Unterschied in der methodischen Operationalisierung zwischen Sozialer Erwünschtheit einerseits und Zustimmungstendenzen andererseits gibt. Von daher sollten diese beiden Formen von Antworttendenzen nicht vermischt werden, sondern als zwei verschiedene Formen des Ausdrucksverhaltens verstanden werden, welche sich in verschiedenen Kulturen manifestieren kann.

5.6.2 Umgang mit Antworttendenzen und deren inhaltliche Interpretation

Bei der Interpretation der unterschiedlichen Antworttendenzen in Form von Zustimmung- und Extremtendenz sowie sozialer Erwünschtheit gibt es verschiedene Ansichten, wie diese zu verstehen sind und wie mit ihnen umzugehen ist.

Eine weit verbreitete Interpretation von Antworttendenzen ist, daß es sich bei dieser Art des Antwortverhaltens um Störquellen handelt, die zu beseitigen wären, um sinnvolle interkulturelle Vergleiche anstellen zu können (Leung & Bond, 1989). Antworten Personen in zustimmender, extremer oder aber sozial erwünschter Weise, dann sollte gemäß dieser Ansicht dieser Methodenbias aus den Daten entfernt werden oder aber es sollte versucht werden, ihn von vornherein zu verhindern. Eine Möglichkeit, solche Verzerrungen durch Antworttendenzen schon bei der Gestaltung des Fragebogens zu minimieren, ist beispielsweise, Antwortskalen mit zur Hälfte positiv formulierten und zur anderen Hälfte negativ formulierten Items zu generieren (Knowles & Nathan, 1997; McCrae & Costa, 1997; Ray, 1983; Ten Berge, 1999). Um bei solchen Skalen einen hohen Wert auf dem zugrunde liegenden Konstrukt zu erhalten, müßte der Befragte somit einem Teil der Items zustimmen, während er dem anderen Teil der Items widersprechen müßte. Daher ließe sich anhand der Antworten dann differenzieren, ob es sich bei den Antworten eher um rein kommunikatives Verhalten im Sinne von Antworttendenzen handelt oder aber ob tatsächlich gemäß der Ausprägung des Konstruktes geantwortet wurde. In der interkulturellen Forschung wurde diesem Vorschlag für entgegengesetzt gepolte Items aber so gut wie kaum gefolgt (Smith, 2004). Ein immer wieder zitierter Grund für dieses Versäumnis ist die Tatsache, daß es in der internationalen Forschung in manchen Sprachen schwierig ist, negative Items zu formulieren. Zusätzlich würden Befragte in manchen Kulturen durch negativ formulierte Items verwirrt oder aber wären abgeneigt, Fragebogeninhalten zu widersprechen (Segall, 1986). Hinzu kommt, daß es für manche Fragen schwer ist, logisch entgegengesetzte Formulierungen zu finden (Schuman & Presser, 1996). Oft ist ebenso in Fragebögen, die in der Praxis eingesetzt werden, aufgrund von Ressourcenbeschränkungen nicht möglich, eine beliebige Anzahl von Items und somit auch entgegengesetzt gepolte Items einzusetzen (van Herk et al., 2004).

Somit beinhalten die meisten Skalen und Meßinstrumente, die in internationalen Studien eingesetzt werden, lediglich positiv formulierte Items. Dieser Umstand führt dazu, daß anhand des Antwortmusters nun nicht mehr zwischen Antworttendenz und

inhaltlicher Aussage zu unterscheiden ist. Um dennoch unverzerrte Meßwerte zu erhalten, wurden verschiedene Verfahren der Standardisierung von Antworten diskutiert, die in der Lage sein sollen, eine ungewollte Verzerrung durch Zustimmung- oder auch Extremtendenzen zu eliminieren. Dabei lassen sich klassischerweise drei verschiedene Formen der Standardisierung unterscheiden (vgl. Fischer, 2004): a) Standardisierung innerhalb einer Person (within-subject-standardisation), b) Standardisierung innerhalb einer Gruppe (within-group-standardisation) und c) Standardisierung innerhalb einer Kultur (within-culture-standardisation). Jede dieser Standardisierungsformen hat verschiedene Vor- und Nachteile und verschiedene Implikationen. Die Standardisierung innerhalb einer Person kann dabei sicherlich als die am weitesten verbreitetste Form angesehen werden und wurde seit den Arbeiten von Hofstede (1980) oft favorisiert. Dabei wird für jede Person die Antwort zu den jeweiligen Items mit dem Gesamtmittelwert aller Items dieser Person in Bezug gesetzt. Bei einer Skala mit beispielsweise 20 Items wird somit der Mittelwert über alle Items von jedem einzelnen Item, also von dessen Rohwert, subtrahiert. Man erhält somit relative Abweichungswerte bei jedem einzelnen Item in Bezug zum Gesamtmittelwert. Das Ergebnis dieser Form der Standardisierung wird auch als ipsative Werte oder ipsative Distanzen bezeichnet (Hicks, 1970). Der Mittelwert dieser neu gebildeten ipsativen Werte ist dann insgesamt null. Dieses Vorgehen eliminiert auf der einen Seite in der Tat Tendenzen eines Befragten, beispielsweise generell zustimmend zu antworten, aus den Rohwerten. Auf der anderen Seite hat es jedoch ebenso gravierende Nachteile. So sind folglich die einzelnen Item-Werte nicht mehr unabhängig voneinander, womit eine wichtige Voraussetzung für die Anwendbarkeit zahlreicher statistischer Verfahren verloren geht (Smith, 2004). Dies betrifft sowohl korrelative als auch varianzanalytische Verfahren. Alle statistischen Verfahren, die auf Korrelationen beruhen, sind von ipsativen Werten stark beeinflusst. Ohne hier eine ausführliche Diskussion dieser Problematik darzulegen (nähere Details lassen sich nachlesen bei Fischer, 2004), läßt sich festhalten, daß ipsative Werte weder geeignet sind für faktorenanalytische Verfahren (Cheung & Rensvold, 2000; Closs, 1996; Cornwell & Dunlop, 1994), noch lassen sich Reliabilitäten oder auch Korrelationen und Regressionen mit externen Variablen problemlos interpretieren (Fischer, 2004). Hinsichtlich varianzanalytischer Verfahren bringt eine Standardisierung in Form von ipsativen Distanzen ebenfalls schwerwiegende Nachteile mit sich und verhindert im Prinzip den Vergleich über Kulturgrenzen hinweg. Fischer (2004) nutzt zur Erläuterung dieser Problematik ein sehr

anschauliches Beispiel, wobei er die Messung der Beine und Schwänze eines Elefanten und einer Maus vergleicht. Würde man hingehen und die Länge der Beine des Elefanten und der Maus vergleichen und damit auf innerhalb der Person (bzw. innerhalb des Tieres) standardisierte Werte zurückgreifen, würde man wahrscheinlich keinen großen Unterschied feststellen. Obwohl selbstverständlich die Beine des Elefanten absolut gesehen viel länger als die der Maus sind, führt diese Art der Relativierung an der Gesamtgröße des Tieres zu keinem Unterschied. Würde man nun die Schwänze der beiden Tiere vergleichen und einmal absolute Werte und einmal ipsative Werte heranziehen, würde man zu interessanten und tatsächlich widersprüchlichen Ergebnisse hinsichtlich der Frage gelangen, welcher Schwanz den nun der längere ist. Vergleicht man die absoluten Werte, gewinnt selbstverständlich der Elefant, denn sein Schwanz ist um einiges länger als der einer Maus. Zieht man jedoch die ipsativen Werte heran, gewinnt auf einmal die Maus, denn ihr Schwanz ist im Vergleich zu ihrem Gesamtkörper um Einiges länger als der Schwanz des Elefanten im Vergleich zur Gesamtgröße seines Körpers. Ob man somit absolute Werte oder ipsative Distanzen zur Messung nutzt, hat gravierende Konsequenzen für die Interpretation der Ergebnisse und sollte daher wohlüberlegt sein. Generell scheint es so zu sein, daß ipsative Werte angemessen sind für die Betrachtung intra-individueller Unterschiede, jedoch für den inter-individuellen Vergleich nicht herangezogen werden können (Chan, 2003; Fischer, 2004; Hicks, 1970). Ist man somit an der relativen Wichtigkeit einzelner Aspekte für eine Person interessiert, kann eine ipsative Standardisierung angebracht sein. Bei Betrachtung der absoluten Wichtigkeit von Aspekten in einem Vergleich von verschiedenen Personen löst sie jedoch die Problematik von Antworttendenzen nicht, denn eine der grundlegenden Schwierigkeiten der Ipsatisierung ist, daß sie die Bedeutung des Items ändert (Cheung & Rensvold, 2000). Der resultierende Wert ist dann nicht mehr die ‚personenbezogene Antwort eines Items‘ sondern nur noch die ‚personenbezogene Antwort eines Items in Relation zu anderen Items‘ (Gurwitz, 1987). Dies erschwert somit die Interpretation der jeweiligen Ergebnisse.

In dem hier in der Arbeit vorliegenden Fall des Vergleichs von Arbeitszufriedenheitswerten über Kulturgrenzen hinweg ist man in der Regel an dem Vergleich der Absolutwerte interessiert. Eine Standardisierung von Rohwerten innerhalb einer Person führt somit hier zu keinem besseren Ergebnis. Würden möglicherweise tatsächlich unterschiedliche Antworttendenzen in unterschiedlichen Kulturen vorliegen, dann würde dieses Vorgehen zwar diese Tendenzen herausrechnen,

jedoch könnte es sehr wohl auch relevante inhaltliche Unterschiede zwischen den Kulturen eliminieren. Ein Vergleich hinsichtlich der Frage, wo denn die Arbeitszufriedenheit am höchsten ist, wäre somit nicht mehr möglich. Stattdessen wären nur noch Aussagen der Art möglich, wo die Arbeitszufriedenheit am höchsten im Vergleich zu den eingesetzten Vergleichswerten oder zur Gesamt-Antworttendenz ist. Neben der Standardisierung innerhalb einer Person ließe sich auch eine Standardisierung innerhalb der Gruppe oder innerhalb der Kultur durchführen. Beim Vergleich mit der Gruppe wird der Rohwert eines Items einer Person nicht in Bezug zum Gesamtwert der Person über alle Items hinweg gesetzt, sondern stattdessen in Bezug zum Gesamtwert aller Personen auf diesem speziellen Item. Diese Form wird auch als „centering“ (Aiken & West, 1991) bezeichnet. Beim Vergleich mit der Gesamtkultur wird der Rohwert einer Person auf einem Item hingegen in Bezug zum Gesamtmittelwert über alle Personen und alle Items in einer Kultur herangezogen. Diese Art der Standardisierung wird oft als „grand-mean-centering“ (Raudenbush & Bryk, 2002) bezeichnet. Diese Arten der Standardisierung generieren nicht dieselben statistischen Schwierigkeiten wie die vorhergehende Form, sondern werden stattdessen sogar in speziellen methodischen Kontexten explizit eingesetzt, z.B. in Mehrebenenanalysen (Raudenbush & Bryk, 2002). Insgesamt gesehen ist gemäß Fischer (2004) eine Anwendung dieser Standardisierungsprozeduren durchaus zu empfehlen. Jedoch ist Vorsicht bei der Interpretation der Ergebnisse geboten, denn die Bedeutung der Ergebnisse ändert sich je nach Gruppen- oder Kultur-Standardisierung. Vor allem in Mehrebenenanalysen zeigen die ermittelten Koeffizienten andere Einflußstärken an; je nach Standardisierungsform unter Umständen den Einfluß auf dem individuellen Niveau, kontrolliert für Gruppenunterschiede, oder aber ganz im Gegenteil den Einfluß von Gruppenunterschieden auf Variablen auf individuellem Niveau, kontrolliert für individuelle Unterschiede. Der Einsatz der jeweiligen Standardisierungsprozedur hängt somit sehr stark von der Fragestellung ab, die jeweils verfolgt werden soll (Fischer, 2004).

Zusätzlich zur Standardisierung an einem Gesamtmittelwert wird oftmals auch eine Standardabweichung mit einbezogen. Während die Standardisierung an einem Gesamtmittelwert für mögliche Zustimmungstendenzen kontrollieren soll, dient die Relativierung der Rohwerte anhand der Standardabweichung zur Kontrolle von Extremtendenzen. Es können somit insgesamt Standardisierungen a) anhand von Mittelwerten, b) anhand von Standardabweichungen und c) anhand von Mittelwerten

und Standardabweichungen vorgenommen werden, gepaart mit der Möglichkeit, dies auf Individual-, Gruppen- oder Kulturlevel durchzuführen (Fischer, 2004). Am bekanntesten von diesen Möglichkeiten sind dabei sicherlich die klassische Form der Ipsatisierung (mit und ohne Einbezug der Standardabweichung) und die z-Transformation, bei der Gruppenmittelwert und -standardabweichung einbezogen werden.

Bei der Betrachtung von Unterschieden in Absolutwerten bringen diese Standardisierungsprozeduren jedoch keine Lösung des Problems. Ein Vergleich der Antworten zwischen verschiedenen Kulturen würde zu unterschiedlichen Ergebnissen führen, je nachdem ob standardisiert wurde oder nicht. Der betrachtete Antwortwert hätte nach einer Standardisierung nur noch Aussagekraft in Bezug zu seinem Vergleichswert; sei es das individuelle, das gruppenbezogene oder das kulturelle Antwortverhalten. Darüber hinaus ist ein weiterer gravierender Nachteil, daß eine Standardisierung implizit immer davon ausgeht, daß bei der Herausparsialisierung nur Anteile von Bias vernichtet werden. Es ist jedoch ebenso möglich, daß dabei inhaltlich relevante Varianz ebenfalls aus dem Rohwert herausgerechnet wird.

Eine andere Möglichkeit, um einen möglichen Bias zu entdecken, wäre der Einsatz eines sogenannten Multitrait-Multimethod-Ansatzes, bei dem verschiedene Meßinstrumenten zur Messung des gleichen Konstrukts eingesetzt werden (Bagozzi & Yi, 1991; Campbell & Fiske, 1959; Marsh & Byrne, 1993; van Herk et al., 2004). Sollte das Konstrukt anhand verschiedener Verfahren in gleicher Ausprägung gemessen werden, wäre dies ein guter Indikator dafür, daß keine Antworttendenz die Werte verzerrt hätte. Jedoch wird dieser Ansatz insgesamt aufgrund knapper Forschungsressourcen sehr selten eingesetzt.

Neben der Ansicht, daß unterschiedliches Antwortverhalten einen Bias darstellt und herausgerechnet werden sollte, besteht eine neuere Sichtweise darin, diese Antwortmuster nicht rein als methodische Störquelle anzusehen. Smith (2004), van Hemert, van de Vijver, Poortinga und Georgas (2002), Harzing (2006) und Smith und Fischer (2008) gehen in ihrer Argumentation davon aus, daß unterschiedliche Antwortmuster vielmehr ein Ausdruck kultureller Unterschiede sind und explizit untersucht werden sollten. Ein Herausrechnen in Form von Standardisierungen würde daher inhaltliche wertvolle Varianz vernichten, anhand derer Schlußfolgerungen über die Kommunikation in entsprechenden Kulturen hätten getroffen werden können.

Antwortmuster sind demnach vielmehr Manifestationen kultureller Differenzen statt lediglich ein störender Bias. Diesem Gedankengang folgend gibt Smith (2004) jedoch unmittelbar in methodischer Hinsicht zu bedenken, daß bei der Messung von kulturspezifischen Antworttendenzen die Thematik der Mehrebenenmessung eine wichtige Rolle spielt. Da die Antworten auf Fragen selbst nur von einzelnen Personen gegeben werden können, werden Maße der Antworttendenz für Individuen und nicht für Kulturen berechnet. In der Regel werden diese Maße anhand möglichst heterogenen Items errechnet, so daß am Ende jedes Individuum jeweils eine einzige Maßzahl erhält. Die einzige Möglichkeit, ein kulturbezogenes Maß zu erhalten, besteht demnach darin, diese Individual-Maße zu mitteln. Diese könnten sich aber möglicherweise je nach Stichprobe, Itemsatz oder Erhebungszeitpunkt unterscheiden. Es müßte somit überprüft werden, inwieweit sich kulturell stabile Maßzahlen überhaupt in dieser Form ermitteln ließen. Smith (2004) konnte diesbezüglich in der Tat zeigen, daß verschiedene Maße von Antworttendenzen, die an unterschiedlichen Stichproben mit unterschiedlichen Inhalten und zu unterschiedlichen Zeiten gemessen wurden, im Gegensatz zur anfänglichen Vermutung in der Tat hoch miteinander korrelieren. Zur Messung nutzte er die renommierten weltweiten Studien von Hofstede, Schwartz, der GLOBE-Gruppe, Daten einer eigenen weltweiten Studie sowie weltweite Datensätze der Eysenck-Lügenskala. Hohe Korrelationen fanden sich vor allem für Itemsätze, in denen Aspekte erfragt wurden, die persönlich relevant für die Befragten waren (d.h. Werte, Einstellungen, Verhalten). Lediglich das GLOBE-Maß, welches zur Messung von kulturellen Praktiken („as is“) verwendet wurde, zeigte geringere Zusammenhänge mit den restlichen Indizes. Die personen-bezogenen Maße zeigten jedoch hohe konvergente Validitäten, was demnach als ein klarer Indikator dafür angesehen werden kann, daß es kultur- bzw. länderspezifische Tendenzen gibt, wie auf Fragen geantwortet wird. Es sollte an dieser Stelle erneut erwähnt werden, daß dies vor allem für Likert-skalierte Fragen gilt (Smith & Fischer, 2008). Smith' (2004) Ergebnisse konnten somit zeigen, daß verschiedenes Antwortverhalten durchaus einen relativ stabilen Effekt kultureller Unterschiede darstellt. In bestimmten Kulturen scheinen demnach Vorgaben zu bestehen, wie man sich in kommunikativen Befragungssituationen zu äußern hat. Diese Situationen, in denen nach Smith (2004) kulturspezifisches Antwortverhalten auftritt, scheinen oft gekennzeichnet zu sein von schriftlicher Kommunikation (dem Ausfüllen des Fragebogens), dessen Adressat zumeist eine Art Fremder ist, der noch nicht einmal immer in der Situation präsent ist. Diese Vorgaben, wie man sich zu äußern hätte,

können gemeinhin als Norm verstanden werden. Auch Hui und Triandis (1989) sehen das Befolgen von Antwortnormen als einen möglichen Grund in ihrer Diskussion über das Auftreten von Extremtendenzen. Darüber hinaus können ihrer Meinung nach selbstverständlich auch rein individuelle psychologische Prozesse eine Rolle spielen, deren Ursachen in den generellen Prozessen menschlicher Urteilsfindung liegen. Hinsichtlich der kulturellen Perspektive wäre jedoch eine mögliche Erklärung für das Auftreten von Antworttendenzen, daß in manchen Kulturen, beispielsweise in Asien, es als wichtig erscheint, bescheiden zu sein und eher vorsichtig zu antworten. Die häufige Wahl der mittleren oder der zustimmenden Antwortkategorie würde es den Befragten ermöglichen, genau dieses Bild nach außen hin abzugeben. Mehrere Studien konnten bereits verschiedene Unterschiede im kommunikativen Verhalten insbesondere für asiatische Länder zeigen. Asiaten haben demnach insgesamt einen eher indirekten Kommunikationsstil (Holtgraves, 1997), zeigen mehr Zurückhaltung im emotionalen Ausdruck (Argyle, Henderson, Bond, Iizuka & Contarello, 1986) und besitzen einen weniger kompetitiven Verhandlungsstil (Morris et al., 1998). Im Gegensatz dazu würde in Ländern des Mittelmeerraumes es bevorzugt, wenn eine Person sich klar ausdrückt (Hui & Triandis, 1989). Die mittlere Antwortkategorie zu wählen, würde dort eher als ein Verstecken der eigenen Meinung interpretiert. Gerade die extremen Antwortkategorien würden zeigen, daß jemand aufrichtig und ehrlich antwortet. Insgesamt würden sich demnach die Befragten in Fragebögen so verhalten, wie es von ihnen in ihrer Kultur erwartet würde. Anhand des Antwortverhaltens kann somit erkannt werden, welches kommunikative Verhalten in der jeweiligen Kultur präferiert wird. Diesen neuen Überlegungen folgend, daß Antworttendenzen vielmehr ein Ausdruck der jeweiligen kulturellen Normen sind, stellt sich nun die Frage, inwieweit sich Kulturen genauer charakterisieren lassen, in denen die jeweiligen Antwortmuster auftreten.

6 Antworttendenzen und Kultur

In der Literatur zeigen sich zahlreiche Belege dafür, daß sich Antworttendenzen von Land zu Land unterscheiden lassen. Vielfach werden diese Tendenzen in rein deskriptiver Weise berichtet. Ein weiterer Teil der Forschung versucht jedoch, diese unterschiedlichen Antworttendenzen mit kulturellen Variablen in Verbindung zu bringen und somit zu erklären, warum in den unterschiedlichen Kulturen jeweils ein unterschiedliches Kommunikationsverhalten bei der Beantwortung von Fragebögen auftritt.

6.1 Deskriptive Ergebnisse der bisherigen Forschung

Anhand der deskriptiven Vergleiche der entsprechenden Länder werden erste Anhaltspunkte deutlich, wie sich die Länder, in denen vermehrt Antworttendenzen auftreten, von solchen unterscheiden lassen, in denen diese Art von Antwortverhalten nicht zu beobachten ist. Dabei wurden in den Studien oft Vergleiche zwischen (nord-) amerikanischen und asiatischen Stichproben vorgenommen. Hinsichtlich dieser Vergleiche konnten mehrere Studien zeigen, daß amerikanische oder kanadische Befragte höhere Extremtendenzen zeigten als japanische Befragte, die im Gegenzug dafür stärkere Mittentendenzen zeigten (Chen, Lee & Stephenson, 1995; Shiomi & Loo, 1999; Takahashi, Ohara, Antonucci & Akiyama, 2002; Zax & Takahashi, 1967). In gleicher Weise unterschieden sich koreanische Befragte von amerikanischen Studienteilnehmern (Chun et al., 1974; Lee & Green, 1991). Auch philippinische Studenten zeigten höhere Antworttendenzen, hier in Form von Zustimmungstendenzen, als amerikanische Studenten (Grimm & Church, 1999). In gleicher Weise konnte bereits Church (1987) zeigen, daß philippinische Befragte offenbar generell hohe Zustimmungstendenzen zeigen. In einer einzigen Studie hingegen zeigten im Gegensatz zu den anderen Studien indonesische und malaysische Befragte stärkere Extremtendenzen als amerikanische Befragte (Stening & Everett, 1984).

Bei einer detaillierten Betrachtung von europäischen Ländern konnten van Herk et al. (2004) zeigen, daß südeuropäische Länder (Griechenland, Spanien, Italien) höhere Extremtendenzen zeigten als Frankreich, Deutschland oder auch England. Dieses Resultat steht in Einklang mit den früheren Überlegungen von Hui und Triandis (1989), wonach gerade ehrliche und deutliche Antworten in Ländern des Mittelmeerraums

erwartet würden. Jedoch zeigten in der Studie von van Herk et al. (2004) die genannten Länder auch in gleicher Reihenfolge ein unterschiedliches Ausmaß an Zustimmungstendenzen. Auch hier war dies am deutlichsten in Griechenland und am schwächsten in England. Dies wiederum wäre somit weniger mit den Erklärungen von Hui und Triandis (1989) in Übereinstimmung zu bringen. Eine genaue Erklärung hinsichtlich des kulturellen Zusammenhangs kann aus dieser Studie somit nicht geschlossen werden. Trotz dieser Unklarheiten spricht für die generelle Bedeutung des Einflusses von Antworttendenzen auf die Beantwortung von Fragebögen und deren Verständnis als eine Form von kommunikativem Verhalten jedoch in der Studie von van Herk et al. (2004) vor allem, daß sich interessanterweise überhaupt kein Zusammenhang zwischen den Antworten zu den Einstellungsfragen und zusätzlichen Verhaltensdaten zeigte. Dies bedeutet, daß die Befragten zwar die in dieser Studie zu beurteilenden Produkte positiver bewerteten, jedoch keineswegs ein anderes Verhalten bezüglich der Produkte zeigten. Weder wurden die Produkte in den Ländern mit positiver Bewertung häufiger gekauft noch wurden sie häufiger genutzt. Dieser Widerspruch läßt mit großer Wahrscheinlichkeit die Schlußfolgerung zu, daß die Bewertung der Produkte aufgrund von Zustimmung- und Extremtendenzen verzerrt gewesen sein muß. Die Befragten sagen halt nur, daß die Produkte toll sind ohne jedoch eine Entsprechung im Verhalten vorweisen zu können. Hinzu kommt noch, daß in den entsprechenden Ländern nicht nur ein Produkt, sondern alle Produkte im Durchschnitt höher bewertet wurden als in anderen Ländern. Auch dies kann als ein klarer Indikator für Antworttendenzen verstanden werden. Insgesamt spricht somit diese deutliche Unabhängigkeit von Verhaltensmaßen und Antwortmaßen dafür, daß hohe Skalenwerte eher in Antworttendenzen und somit dem jeweiligen kulturellen Kommunikationsverständnis begründet waren statt auf tatsächlichen Einstellungen gegenüber den Produkten (van Herk et al., 2004). Betrachtet man neben dieser interessanten Studie weitere Untersuchungen, zeigt sich beim Vergleich europäischer Länder ebenfalls bei Steenkamp und Baumgartner (1998), daß griechische Befragte höhere Zustimmungstendenzen als belgische oder englische Teilnehmer zeigten. Auch bei Baumgartner und Steenkamp (2001) zeigten südeuropäische Länder, hier Spanien und Portugal, höhere Zustimmungstendenzen als andere europäische Länder. Betrachtet man neben Asien und Europa weitere Regionen, zeigte sich in einer Studie von Clarke III (2000), daß bei mexikanischen Studenten stärkere Extremtendenzen hinsichtlich der Skalenendpunkte auftraten als bei amerikanischen Studenten. Dies

konnte Clarke III (2001) in einer weiteren Studie bestätigen. Auch hier zeigten sogar über verschiedene Likert-Skalen hinweg (von 3- bis 10-Punkt-Skalen) die Befragten in Mexiko leicht höhere Extremtendenzen als in den USA. Am stärksten waren die Extremtendenzen hier jedoch bei Franzosen ausgeprägt, während in Australien die Extremtendenzen über alle Skalentypen hinweg am niedrigsten waren. Bei einer Studie an Kindern hingegen zeigten hingegen Watkins und Cheung (1995) wiederum, daß hier die australischen Kinder stärker extreme Skalenpunkte nutzten als Kinder aus China, Nepal, Nigeria oder den Philippinen.

Während die bisherigen Studien das Kommunikationsverhalten zwischen Ländern verglichen, zeigten sich auch innerhalb eines einzigen Landes, vornehmlich den USA, ebenso Unterschiede zwischen verschiedenen kulturellen Gruppen. So fanden sich höhere Extremtendenzen sowohl bei lateinamerikanischen Befragten (Clarke III, 2000; Hui & Triandis, 1989; Johnson et al., 1997; Marin et al., 1992; Ross and Mirowsky, 1984) als auch bei afroamerikanischen Befragten (Bachman & O'Malley, 1984; Clarke III, 2000; Johnson et al., 1997), jeweils im Vergleich mit weißen amerikanischen Befragten. In gleicher Weise unterschieden sich die kulturellen Gruppen der Befragten hinsichtlich der Zustimmungstendenzen.

Festzuhalten bleibt somit insgesamt, daß sich verschiedene Länder hinsichtlich ihres kommunikativen Verhaltens in Befragungssituationen unterscheiden lassen. Jedoch scheinen auf den ersten Blick diese Resultate relativ gemischt zu sein und keine klaren Tendenzen zu zeigen. Bei genauerem Hinsehen wird jedoch deutlich, daß sich dennoch eine gewisse Struktur in die Resultate bringen läßt. Demnach scheint es so zu sein, daß in angloamerikanischen Ländern Antworttendenzen weniger stark ausgeprägt sind als in asiatischen, südeuropäischen oder süd-/mittel-amerikanischen Ländern. Jedoch bleiben diese Resultate in Teilen bruchstückhaft, da in vielen dieser Studien nur wenige Länder miteinander verglichen wurden; ein Problem, welches sich in weiten Teilen der interkulturellen Forschung finden läßt. Aus einer kontrastierenden Gegenüberstellung beispielsweise nur eines angloamerikanischen und eines asiatischen Landes läßt sich somit bei weitem kein genereller Rückschluß über Unterschiede ganzer Regionen ziehen. Ebenso lassen sich anhand dieser deskriptiven Vergleiche weniger Länder zwar Vermutungen darüber anstellen, welche Kriterien dieser Länder für die Unterschiede in den Antworttendenzen zuständig sein könnten und daß die jeweilige Kultur möglicherweise dabei eine große Rolle spielt. Eine eindeutige Antwort können diese

Studien auf diese Frage jedoch nicht geben, denn Kultur-Variablen wurden in diesen Studien in der Regel nicht miterhoben.

6.2 Kulturelle Werte und Antworttendenzen

Einige wenige Studien (Harzing, 2006; Johnson, Kulesa, Cho & Shavitt, 2005; Smith, 2004; Smith & Fischer, 2008) haben jedoch versucht, diese Frage genauer zu beleuchten, inwieweit der kulturelle Unterschied zwischen verschiedenen Nationen dafür verantwortlich sein kann, daß Mitglieder der jeweiligen Kulturen sich in Fragebögen unterschiedlich äußern. Smith (2004) korrelierte daher in seiner bereits zuvor an anderer Stelle erwähnten Studie Zustimmungsindizes einer Vielzahl von Ländern mit den entsprechenden Kulturwerten dieser Länder. Zur Interpretation seiner Ergebnisse ist wichtig zu wissen, wie genau Zustimmungsindizes erfaßt wurden. Daher soll die genaue Berechnung hier noch einmal detaillierter beschrieben werden. Zur Berechnung verwendete er dabei insgesamt sieben verschiedene Maße aus den Studien von Hofstede, Schwartz und der GLOBE-Gruppe sowie Daten einer eigenen weltweiten Studie, weltweite Datensätze der Eysenck-Lügenskala und einer Studie zu sozialen Axiomen. Das erste Maß von Hofstede (2001) errechnete sich aus der gemittelten Wichtigkeit von 14 in dessen Studie erhobenen Arbeitszielen. Aus den Studien von Schwartz (1994), in denen anhand von 56 Items kulturelle Werte erhoben wurden, nutzte er als zweites Maß das mittlere Urteil über alle 56 Items, welches Schwartz selbst zur Standardisierung und zur Kontrolle von Antworttendenzen nutzte. Aus dem GLOBE-Projekt (House et al., 2004) nutzte er den Mittelwert je Land über alle Kulturdimensionen sowohl aus den Messungen der kulturellen Praktiken („as is“) als auch aus den Messungen der kulturellen Werte („should be“). Ein fünftes Zustimmungsmaß generierte Smith (2004) aus einer eigenen Studie, in der Smith et al. (2002) Manager aus 53 Nationen danach fragten, aus welcher Art von Informationsquelle sie sich in verschiedenen Situationen jeweils Informationen zur Lösung eines Problems heranholen würden. Auch hier wurden die über alle Situationen und alle Informationsquellen gemittelten Antworten als Zustimmungsinde herangezogen. Als sechstes und siebtes Maß dienten schließlich gemittelte Werte auf Länderebene der Eysenck-Lügenskala aus einer Studie von van Hemert et al. (2002) sowie Länderkennwerte aus einer Studie von Leung et al. (2002), in der die Akzeptanz sozialer Axiome in insgesamt 34 Ländern gemessen wurde. Alle genutzten

Zustimmungsindizes lagen somit auf Länderebene vor. Um den Zusammenhang mit kulturellen Werten festzustellen nutzte Smith (2004) nun die einzelnen Werte der Kulturdimensionen je Land aus den bekannten Studien von Hofstede, Schwartz und dem GLOBE-Projekt, wobei zur Vermeidung von Autokorrelationen selbstverständlich bei der Korrelationen eines Zustimmungsmaßes aus einer Wertestudie (Bsp. Hofstedes Zustimmungsindex) nur die jeweiligen Kulturwerte der anderen Studien herangezogen wurden (d.h. Hofstedes Kulturwerte wurden bei der Berechnung ausgeklammert). Wie von ihm erwartet, zeigte sich in der Tat bei seinen Berechnungen eine hohe Konsistenz in den ermittelten Korrelationen; sowohl bei der Betrachtung aller Zustimmungsindizes hinsichtlich eines einzelnen Kulturwertes (z.B. Kollektivismus) als auch bei dem Vergleich mehrerer Kulturwerte bezüglich der Zustimmungsindizes (z.B. Kollektivismus bei Hofstede und Konservatismus bei Schwartz). So korrelierte unter anderem Hofstedes Kollektivismusmaß mit 6 der 7 Zustimmungsindizes im hoch positiven Bereich ($r = .46$ bis $r = .65$). Lediglich der Zustimmungsindex der kulturellen Praktiken aus der GLOBE-Studie zeigte keinen Zusammenhang. Dieser Meßwert zeigte jedoch insgesamt keine klaren Zusammenhänge und wurde daher von Smith (2004) gesondert behandelt. In gleicher Weise wie bei Kollektivismus korrelierte auch Hofstedes Maß der Machtdistanz mit den 6 Zustimmungsindizes im hoch positiven Bereich ($r = .52$ bis $r = .67$). Zustimmungstendenzen tauchen somit vor allem in den Ländern auf, die gemäß Hofstede ein hohes Maß von Kollektivismus und ein hohes Maß von Machtdistanz besitzen. Dieser Zusammenhang bestätigte sich auch bei der Betrachtung der Kulturwerte von Schwartz, jedoch in etwas schwächerer Form. Zustimmungstendenzen sind hinsichtlich seiner Dimensionen vor allem in Ländern mit hohen Hierarchiewerten ($r = .12$ bis $r = .70$) und mit hohen Werten von Konservatismus/Eingebettetsein ($r = .35$ bis $r = .67$) zu finden. Trotz der unterschiedlichen Bezeichnung von Schwartz' Dimensionen im Vergleich zu Hofstede wird jedoch deutlich, daß sich auch hier der inhaltlich gleiche Zusammenhang mit kulturellen Werten wiederfinden läßt. Bei dem Vergleich mit den GLOBE-Werten konnte dieser Zusammenhang ebenfalls bestätigt werden, da alle Zustimmungsindizes hoch mit den „As is“-Werten der jeweiligen Ländern hinsichtlich InGroup Kollektivismus ($r = .45$ bis $r = .70$) und Machtdistanz ($r = .30$ bis $r = .51$) korrelierten. Darüber hinaus zeigte sich ein deutlicher Zusammenhang der Zustimmungsindizes mit den GLOBE „Should Be“-Maßen von Unsicherheitsvermeidung ($r = .55$ bis $r = .69$). Betrachtet man somit die Maße der GLOBE-Studie insgesamt, tritt zustimmendes

Antwortverhalten vor allem in Ländern auf, in denen derzeit ein hohes Maß an Praktiken in Form von Kollektivismus und Machtdistanz vorherrschen. Zusätzlich antworten vor allem die Kulturen zustimmender, bei denen der Wunsch nach stärkerer Unsicherheitsvermeidung hoch ausgeprägt ist.

Die Zusammenhänge mit den kulturellen Wertedimensionen aus drei völlig unabhängigen Wertestudien deuten somit alle in die gleiche Richtung. Bedeutsam beim Auftreten von Zustimmungstendenzen scheinen somit vor allem Werte von Kollektivismus, von Machtdistanz oder Hierarchie und möglicherweise auch von Unsicherheitsvermeidung zu sein. In einer weiteren regressionsanalytischen Berechnung, die er mit den Kulturdimensionen des GLOBE-Projekts durchführte, prüfte Smith (2004), welcher dieser kulturellen Werte denn insgesamt die stärkste Bedeutung besaß. Es zeigte sich, daß vor allem das GLOBE-„As is“-Maß von InGroup Kollektivismus und das GLOBE-„Should be“-Maß von Unsicherheitsvermeidung die meisten Varianz zwischen den Ländern hinsichtlich der Zustimmungsindizes aufklären konnten. Man hätte vielleicht aufgrund der zuvor erwähnten Zusammenhänge mit den Hofstede-Dimensionen vermuten, daß doch eigentlich die Kulturwerte der GLOBE-Studie („should be“) noch stärker mit den Zustimmungsindizes hätten korrelieren müssen und die reinen Praktiken („as is“) in den jeweiligen Ländern von geringer Bedeutung gewesen sein müßten. Jedoch wurde bereits bei der Darstellung der Zusammenhänge der Kulturdimensionen untereinander deutlich (vgl. Kapitel 3.1), daß vor allem die Praktiken der GLOBE-Studie und nicht deren Kulturwerte mit den Kulturdimensionen von Hofstede zusammenhingen und daß darüber hinaus die beiden Maße (Praktiken und Kulturwerte) allein innerhalb der GLOBE-Studie mehrheitlich negativ korreliert waren (Heine et al., 2002). Beachtet man somit zusätzlich diese bereits bekannten Resultate über den Zusammenhang der verschiedenen Kulturwerte untereinander, so sind die Ergebnisse von Smith (2004) keineswegs verwunderlich, sondern geradezu logisch vereinbar und so zu erwarten gewesen.

Daß diese Ergebnisse von Smith (2004), die erstmalig das Auftreten von speziellem Antworttendenzen in Verbindung mit kulturellen Werten brachte, kein rein zufälliges Ergebnis waren, zeigte sich durch eine weitere große multinationale Studie von Harzing (2006), die insgesamt 26 Länder beinhaltete. Im Großen und Ganzen zeigten sich dort nämlich fast identische Zusammenhänge zwischen Kulturdimensionen und Antworttendenzen. Zur Messung von Antworttendenzen diente in dieser Studie ein weltweit eingesetzter Fragebogen aus einer groß angelegten Studie zur Messung des

Einflusses der Sprache des Fragebogens auf das Antwortverhalten von Studenten. Anhand vier verschiedener inhaltlicher Themenkomplexe (Normen hinsichtlich genereller Aktivität, Normen und Werte hinsichtlich Beziehungen, Gründe für die Auswahl von Wahlkursen an der Universität sowie Beschreibungen der idealen Tätigkeitsbereiche nach der Erlangung des Universitätsabschlusses) wurden Antworttendenz-Indizes auf Länderebene gemäß den in Kapitel 5.6 beschriebenen Vorgehensweisen errechnet. Die Zusammenhänge in dieser Studie zwischen Zustimmungswerten und Kulturwerten konnten die Ergebnisse von Smith (2004) eindrucksvoll bestätigen. Auch hier zeigte sich in den regressionsanalytischen Berechnungen hinsichtlich des Zustimmungswertes ein deutlicher Einfluß von Kollektivismus ($\beta = .45$) und von Machtdistanz ($\beta = .37$) nach Hofstede sowie von InGroup-Kollektivismus-Praktiken ($\beta = 1.03$) und von Unsicherheitsvermeidungswerten ($\beta = .53$) der GLOBE-Studie. Diese Ergebnisse sind darüber hinaus auch daher bedeutsam, da sie die zeitliche Stabilität von Antworttendenzen innerhalb eines Landes zeigen (Harzing, 2006), denn die in dieser Studie zugrunde liegenden Daten wurden deutlich später als die Daten von Smith (2004) erhoben.

Neben der Erfassung von Zustimmungswerten wurden in der Studie von Harzing (2006) jedoch ebenso Extrem- und Mitteltendenzen untersucht und mögliche Erklärungsmuster für deren Auftreten diskutiert. Dazu wurden Indizes für jedes Land anhand desselben Fragebogens ebenfalls gemäß den in Kapitel 5.6 beschriebenen Vorgehensweisen für diese Art von Antwortverhalten errechnet. Harzing (2006) trennt dabei zwischen positiven Extremtendenzen, die nur die eine Seite der Skala berücksichtigen, und negativen Extremtendenzen, die sich auf das entgegengesetzte Ende der Skala beziehen. Ihre Argumentation stützt sich wie auch die von Smith (2004) darauf, daß Antworttendenzen als eine Art von Kommunikationsstil, der in der jeweiligen Kultur tradiert wird, angesehen werden sollten. Dabei wird dieser Kommunikationsstil nicht nur von dem kulturellen Umgang mit Beziehungssituationen bestimmt (mit denen sich die Dimensionen von Machtdistanz, Kollektivismus und Unsicherheitsvermeidung beschäftigen), sondern ihrer Meinung nach auch von der Tatsache, wie stark Menschen generell nach außen hin agieren und expressives Verhalten zeigen. Um den Einfluß dieser Expressivität zu messen, bezog sie Werte der Extraversionsskala aus dem Eysenck-Persönlichkeitsfragebogen mit in ihre Studie ein. Die entsprechenden Länderkennwerte für Extraversion stammten dabei aus der Studie von van Hemert et al. (2002) sowie aus einer Studie von Lynn und Martin (1995) und

stellten das gemittelte Ausmaß der auf Individualniveau gemessenen Extraversionswerten dar. Extrovertiertes Verhalten sollte gemäß Harzing (2006) vor allem die Tendenz zu Extremwerten auf einer Likert-Skala vorhersagen können, da in den entsprechenden Kulturen, wo Menschen sich stark expressiv verhalten, auch verlangt wird, daß man sich klar und deutlich äußern sollte. Die Ergebnisse ihrer Studie konnten diese Überlegungen jedoch überraschenderweise nur zum Teil bestätigen. Ein Zusammenhang von Extraversion und Extremtendenzen zeigte sich nur bezüglich der positiven Extremtendenz. Hier war in der Tat Extraversion der stärkste Prädiktor in den jeweiligen Regressionsanalyse ($\beta = .54$ bis $\beta = .72$), egal ob zusätzlich die Kulturdimensionen von Hofstede oder der GLOBE-Studie mit in die Analyse einbezogen wurden. Extremtendenzen, bei denen das äußerst positive Ende der Likert-Skala überzufällig häufig ausgewählt wird, zeigen sich demnach vor allem in Ländern mit hohen Extraversionskennwerten. Bei negativen Extremtendenzen, d.h. dem extremen Ablehnen von Items, konnte dieser Zusammenhang jedoch nicht beobachtet werden. Betrachtet man neben dem Extraversionsmaß auch die kulturellen Maße, zeigte sich, daß kulturelle Unterschiede gemessen an den Dimensionen von Hofstede das Auftreten von Extremtendenzen interessanterweise so gut wie gar nicht vorhersagen konnten, weder bezüglich positiver noch bezüglich negativer Tendenzen. Nur hinsichtlich weniger Kulturdimensionen des GLOBE-Modells zeigten sich noch signifikante Zusammenhänge. So waren neben Extraversion noch die InGroup-Kollektivismus-Praktiken ein signifikanter Prädiktor für positive Extremtendenzen in einer der Regressionsanalysen sowie die Unsicherheitsvermeidungswerte in einer der anderen Regressionsanalysen. Für negative Extremtendenzen zeigten sich lediglich negative Zusammenhänge mit InGroup-Kollektivismus-Praktiken und Machtdistanzwerten. Die Zusammenhänge zwischen den beiden Maßen für Extremtendenzen und den Kulturdimensionen scheinen somit in Teilen das Muster zu replizieren, welches auch zwischen Zustimmungstendenzen und Kollektivismus, Unsicherheitsvermeidung oder auch Machtdistanz gefunden wurde. Dies ist auf den ersten Blick inhaltlich nicht so einfach nachzuvollziehen. Bei genauerer Betrachtung der methodischen Vorgehensweise wird aber deutlich, daß der Grund für diese Ergebnisse weitestgehend in der in dieser Studie vorhandenen Kovariation von Zustimmungstendenzen und positiven Extremtendenzen ($r = .66$) sowie von Zustimmungstendenzen und negativen Extremtendenzen ($r = -.61$) zu finden ist. Extremtendenzen und Zustimmungstendenzen sind nämlich gerade dann oft korreliert,

wenn – wie in dieser Studie augenscheinlich der Fall – die entsprechenden Items lediglich in eine Richtung formuliert sind. Besser wäre es diesbezüglich gewesen, einen Gesamtindex auch für die Extremtendenzen zu berechnen, welcher möglicherweise andere Zusammenhänge gezeigt hätte. Eine eindeutige Aussage hinsichtlich der Beeinflussung von Extremtendenzen durch kulturelle Werte läßt sich somit nicht treffen. Als letzten Indikator untersuchte Harzing (2006) ebenfalls den möglichen Gegenpol von Extremtendenzen, d.h. die Mittentendenzen. Hier zeigten sich jedoch überhaupt keine eindeutigen Ergebnisse. Zwar hing das Auftreten von Mittentendenzen in einer der Analysen stark von Hofstedes Machtdistanz ($\beta = -.80$) und Kollektivismus ($\beta = .57$) ab, jedoch zeigten sich überhaupt keine ähnlichen Zusammenhänge bei der Betrachtung der GLOBE-Dimensionen. Insgesamt ergab sich somit kein klares Bild bei der Entstehung von Mittentendenzen. Ebenso zeigte sich somit auch nicht das entgegengesetzte Beziehungsmuster von Extrem- und Mittentendenzen, welches eigentlich zu erwarten gewesen wäre. Dieses Ergebnis deutet somit darauf hin, daß Mittentendenzen keineswegs das Gegenteil von Extremtendenzen sein müssen, wie teilweise vorschnell angenommen wird (vgl. die eigene Argumentation in Kapitel 5.6.1). Der Vollständigkeit halber sei erwähnt, daß Harzing (2006) in einer zweiten Studie versucht zu zeigen, daß speziell negative Extremtendenzen sowie das Auftreten von Mittentendenzen stärker von der Sprache, in der der Fragebogen vorgegeben wird, abhängen als von kulturellen Aspekten. Höhere Extremtendenzen und geringere Mittentendenzen treten gemäß ihren Analysen beispielsweise auf, wenn der Fragebogen in der Muttersprache statt in Englisch vorgegeben wird.

Insgesamt konnten somit die Ergebnisse von Harzing (2006) die Resultate von Smith (2004) dahingehend bestätigen, daß Zustimmungstendenzen vor allem in kollektivistisch geprägten Kulturen, in denen auch ein hohes Maß von Unsicherheitsvermeidung herrscht, auftreten. Die in ihrer Analyse neu hinzugefügten Analysen bezüglich Extremtendenzen und Mittentendenzen zeigen zwar kein eindeutiges Bild. Sie deuten jedoch darauf hin, daß diesbezüglich vor allem Machtdistanz und Extraversion eine Rolle spielen könnten.

Betrachtet man abschließend neben den Zustimmungs- und Extremtendenzen die kulturellen Unterschiede hinsichtlich der Sozialen Erwünschtheit, scheinen bisherige Ergebnisse darauf hinzudeuten, daß diese Form des Antwortverhaltens sich ebenfalls stärker in kollektivistischen Kulturen wiederfinden läßt. So konnte bereits Hui (1988) zeigen, daß in seiner chinesischen Stichprobe ein Zusammenhang zwischen

Kollektivismus und sozial erwünschtem Antworten bestand. Weitere Studien, die zwar nicht immer explizit individualistische/kollektivistische Werte gemessen haben und soziale Erwünschtheitsskalen eingesetzt haben, deuten ebenfalls in die Richtung, daß Soziale Erwünschtheit eher in dieser Art von Kultur aufzutreten scheint. So konnten Triandis et al. (2001) zeigen, daß der Hang dazu, in geschäftlichen Verhandlungen zu lügen und zu betrügen, in kollektivistischen Ländern höher ausgeprägt war. Triandis und Suh (2002) argumentieren in ähnlicher Weise, daß Lügen eher in kollektivistischen Kulturen als in individualistischen Kulturen akzeptiert wird; vor allem dann, wenn es dadurch möglich wird, sein Gesicht zu wahren oder es für Vorteile für die InGroup sorgen kann. Mehrere Forscher haben diesbezüglich die Bedeutung des Gesicht-Wahrens in kollektivistischen Kulturen betont (Ho, 1976; Hu, 1944), was ja übrigens für sich genommen ebenso eine Art von sozial erwünschtem Verhalten darstellt. Analog argumentieren Trilling (1972) und Triandis (1995), daß Aufrichtigkeit und Ehrlichkeit vor allem in individualistischen Kulturen von Bedeutung sind, während diese in kollektivistischen Kulturen von nicht so hohem Wert sind. Neuere empirische Studien untermauern diese Argumentation. Van Hemert et al. (2002) berichten demnach auf Basis von 23 Ländern eine hohe negative Korrelation ($r = -.68$) zwischen dem länderspezifischen Individualismuswert und dem entsprechenden mittleren Wert der Eysenck-Lügenskala. Auch Bernardi (2006) kann diese Ergebnisse bestätigen. In seiner 12 Länder umfassenden Studie korrelierte ebenfalls Hofstede's Individualismuswert hoch negativ ($r = -.54$) mit sozial erwünschtem Verhalten, gemessen sowohl durch die bekannte Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960) als auch das Balanced Inventory of Socially Desirable Responding (Paulhus, 1984). Darüber hinaus zeigte sich in dieser Studie, daß auch Unsicherheitsvermeidung ($r = .54$) und Machtdistanz ($r = .48$) mit der länderspezifischen Tendenz zu Sozialer Erwünschtheit zusammenhängen. Zusätzlich kann man noch die spezifischen kulturellen Zusammenhänge gemäß der Unterscheidung von Paulhus (1984) in einerseits Selbsttäuschung und andererseits Fremdtäuschung betrachten. Diesbezüglich zeigten die Untersuchungen von Lalwani, Shavitt und Johnson (2006) ein etwas differenzierteres Bild. Ihre Studien können bestätigen, daß Kollektivisten höhere Werte bei Fremdtäuschung erhalten und sich bei Verhaltensweisen, die nach außen gerichtet sind, stärker sozial erwünscht verhalten als Individualisten. Hinsichtlich der Selbsttäuschung erzielen jedoch Individualisten höhere Werte als Kollektivisten und

besitzen demnach eher ein übersteigertes Bild von ihren persönlichen Fähigkeiten gegenüber sich selbst.

Insgesamt kann gemäß dieser unterschiedlichen Studien somit davon ausgegangen werden, daß Soziale Erwünschtheit, mit höherer Wahrscheinlichkeit in kollektivistischen als in individualistischeren Kulturen auftreten wird; zumindest insoweit darunter ein verschleiern des Verhalten einer anderen Person gegenüber verstanden wird.

6.3 Kulturelle Werte, individuelle Unterschiede, Antworttendenzen und deren Wechselwirkung

Die im vorherigen Kapitel dargestellten Studien befaßten sich mit dem Zusammenhang von Kultur und Antworttendenzen auf dem kulturellen Level. Die für die Analysen notwendigen kulturspezifischen Kennwerte, die das jeweilige Länderniveau abbilden sollen, wurden dabei in aller Regel als Mittelwert von personenspezifischen Daten gebildet, die auf Individualniveau gemessen worden sind. Zwar ist in der interkulturellen Forschung auch die Frage gestellt worden, ob diese Vorgehensweise immer der beste Weg ist, um geeignete Maße für die länderspezifischen Dimensionen zu erhalten (Fischer, 2006). Jedoch ist dies oftmals die praktikabelste Lösung. Wichtig ist dann nur, genau zu betrachten, was durch die einzelnen Kennwerte eigentlich ausgedrückt wird. Im Allgemeinen gilt dabei, daß die Kennwerte auf dem Individualniveau rein personenspezifisches Verhalten abbilden, während die Kennwerte auf dem Länderniveau das durchschnittliche Verhalten aller Menschen des jeweiligen Landes darstellen. Dieser länderbezogenen Kennwerte könnten daher beispielsweise als Indikatoren für eine Art Norm aufgefaßt werden. Die Problematik der Unterscheidung von Individual- und Länderkennwerten betrifft dabei grundsätzlich jede Fragestellung, bei der Vergleiche zwischen Ländern durchgeführt werden. Somit gilt auch für die spezifische Thematik von Antworttendenzen in dieser Arbeit im Prinzip genau dasselbe. Auch hier sollte zwischen individualspezifischem Verhalten und der kulturellen Ausprägung unterschieden werden.

Im neueren Verständnis der Forschung wird diese Unterscheidung getroffen und Antworttendenzen auf Länderniveau werden als kommunikative Norm verstanden, die vorgibt, wie man sich in einem Fragebogen (speziell bei Likert-Skalen) zu verhalten hat und was dabei jeweils die ‚angemessene‘ Antwort auf eine Frage ist (Smith, 2004).

Darüber hinaus wird das Auftreten von Antworttendenzen aber auch von personenspezifischen Aspekten beeinflusst. So konnte die bisherige Forschung ebenfalls zeigen, daß Zustimmungstendenzen unter anderem eher bei Personen mit geringerem Status (Carr & Krause, 1978; Greenleaf, 1992a; Phillips & Clancy, 1970; Ross & Mirowsky, 1984), bei Personen mit geringerer Bildung (Marin et al., 1992; Winkler et al., 1982) bei älteren Personen (Greenleaf, 1992a; Ross & Mirowsky, 1984; Winkler et al., 1982) und teilweise auch eher bei Männern (Greenleaf, 1992a) auftreten. Von daher sollte neben der Untersuchung des Einflusses von Kultur auf das Antwortverhalten ebenso berücksichtigt werden, welche personenspezifischen Einflußmerkmale es geben kann. Die methodische Vorgehensweise, die in der Lage ist, sowohl Einflußgrößen auf Länderniveau als auch auf Individualniveau zu berücksichtigen, ist dabei die Mehrebenenanalyse (Bryk & Raudenbush, 1992; Hox, 1995; Engel, 1998; Nezlek, Schröder-Abé & Schütz, 2006). Eine Mehrebenenanalyse, oft auch als Multilevel-Analysis, Hierarchical Linear Modeling (HLM), Mixed Models, Random Coefficient Models oder auch Hierarchical Linear Regression bezeichnet, ist in der Lage, die Varianz einer abhängigen Variable auf Länderniveau aufzuteilen in Varianzanteile, die von Länderkennwerten beeinflusst werden, und in Varianzanteile, deren Ursprung in individuellen Unterschieden liegt.

In einer eben solchen Analyse untersuchten Smith und Fischer (2008), welche weiteren individuellen Einflußgrößen es neben Kultur auf länderspezifisches Antwortverhalten gibt und inwieweit der kulturelle oder der individuellere Einfluß der jeweils stärkere ist. Dabei konnten sie im erster Teil ihrer umfangreichen Analysen anhand von Korrelationsstudien auf Länderniveau die bisherigen Ergebnisse anderer Studien bestätigen und ebenfalls zeigen, daß bei der Betrachtung des reinen Länderniveaus Zustimmungstendenzen eher auftreten in Nationen mit hohem Kollektivismus. Dies galt sowohl für Hofstedes Kennwert, die GLOBE-Kollektivismus-Praktiken, einem eigens neu kreierten Kollektivismusmaß als auch für die Schwartz-Dimensionen Eingebettetsein. Einerseits sei jedoch angemerkt, daß diese Ergebnisse durchaus zu erwarten waren, da auch in den Analysen von Smith und Fischer (2008) als Basis die Daten aus der gleichen Studie von Smith et al. (2002) genutzt wurden, die auch den Analysen von Smith (2004) zu Grunde lag. Andererseits wurden jedoch von den Autoren in diesen neueren Analysen weitere Nationen ergänzt, so daß die Ergebnisse von Smith und Fischer (2008) als wichtige Bestätigung der bereits zuvor gefundenen Ergebnisse verstanden werden sollten. Darüber hinaus blieben die korrelativen

Zusammenhänge der Konstrukte auf dem Länderniveau auch in ihren weiteren Analysen anhand der Mehrebenenmodelle trotz des Einbezugs der Individualdaten bestehen.

Hinsichtlich der individuellen Unterschiede spielte in den Analysen von Smith und Fischer (2008) das Alter, das Geschlecht und die Ausprägung des individuellen Selbstkonzepts eine bedeutende Rolle. Bezüglich der soziodemographischen Aspekte fanden die Autoren, daß Zustimmungstendenzen höher bei älteren Befragten waren und stärker bei Frauen auftraten. Die altersbezogenen Ergebnisse sind somit in Einklang mit den früheren Ergebnissen der zuvor erwähnten Studien, während die geschlechtsbezogenen Resultate jedoch entgegengesetzt zu früheren Studien sind (vgl. beispielsweise Greenleaf, 1992a). Eine genauere Erklärung für diese Zusammenhänge gaben die Autoren jedoch nicht, da die Analyse der soziodemographischen Variablen nicht im Fokus ihrer Studie stand und für etwaige Unterschiede diesbezüglich in den statistischen Mehrebenenmodellen nur kontrolliert werden sollte. Ihre Studie hatte vielmehr zum Ziel, den Einfluß von interdependentem bzw. independentem Selbstkonzept und dessen Wechselwirkung mit der jeweils kulturellen Umgebung – vor allem in Bezug auf Individualismus/Kollektivismus – zu untersuchen. An dieser Stelle sei erneut kurz darauf hingewiesen, daß Interdependenz und Kollektivismus bzw. Independenz und Individualismus inhaltlich im Prinzip jeweils das gleiche Konstrukt beschreiben, jedoch Interdependenz/Independenz die Ausprägung auf dem Individuallevel beschreibt, wohingegen Kollektivismus/Individualismus die kulturelle Ausprägung bezeichnet.

In den Analysen von Smith und Fischer (2008) hing die Ausprägung von Zustimmungstendenzen in der Tat deutlich mit dem jeweiligen individuellen Selbstkonzept der Befragten zusammen. In allen ihrer Analysen zeigte sich durchgängig, daß je stärker das interdependente Selbstkonzept ausgeprägt war, desto stärker zeigten die jeweiligen Personen zustimmendes Antwortverhalten. Dieser Zusammenhang bestand in allen Mehrebenenmodellen, die die Autoren jeweils getrennt mit Hofstede's Kulturwerten, den Werten der GLOBE-Studie und den Kulturwerten von Schwartz berechneten. Obwohl somit jeweils eine unterschiedliche Anzahl von Ländern und jeweils unterschiedliche Kulturwerte in die Analysen eingingen, zeigte sich dennoch dieses durchgängige Bild: Menschen, die sich stärker über den Gruppenbezug definieren, stimmen Items in Fragebögen unabhängig von deren Inhalt generell stärker zu.

Besonders interessant und in keiner der vorhergehenden Studien jemals analysiert worden ist vor allem die Wechselwirkung von independentem/interdependentem Selbstkonzept und der Kultur, in der die jeweiligen Personen leben. Die Frage, die sich dabei stellt, lautet, ob möglicherweise eine eher interdependente Person in einer kollektivistischen Kultur, die ihrem Selbstkonzept weitestgehend entspricht, unterschiedlich antwortet im Vergleich dazu, wenn sie in einer individualistischen Kultur leben würde, wo die Mehrheit der sie umgebenden Menschen ein eher entgegengesetztes Selbstkonzept hat.

Die Ergebnisse von Smith und Fischer (2008) zeigen diesbezüglich in der Tat interessante Wechselwirkungen. Diese sind jedoch auch keineswegs völlig eindeutig und unterscheiden sich zum Teil, je nachdem welche Kulturdimension betrachtet wird. Hinsichtlich Zustimmungstendenzen konnten sie beispielsweise feststellen, daß in Kulturen, die geringe Ausprägungen hinsichtlich Institutioneller Kollektivismuspraktiken (GLOBE-„as-is“-Maß) haben, der Zusammenhang zwischen interdependentem Selbstkonzept und Zustimmungstendenzen höher ist als in Kulturen, die hinsichtlich dieser Dimension hoch ausgeprägt sind (siehe Abbildung 6).

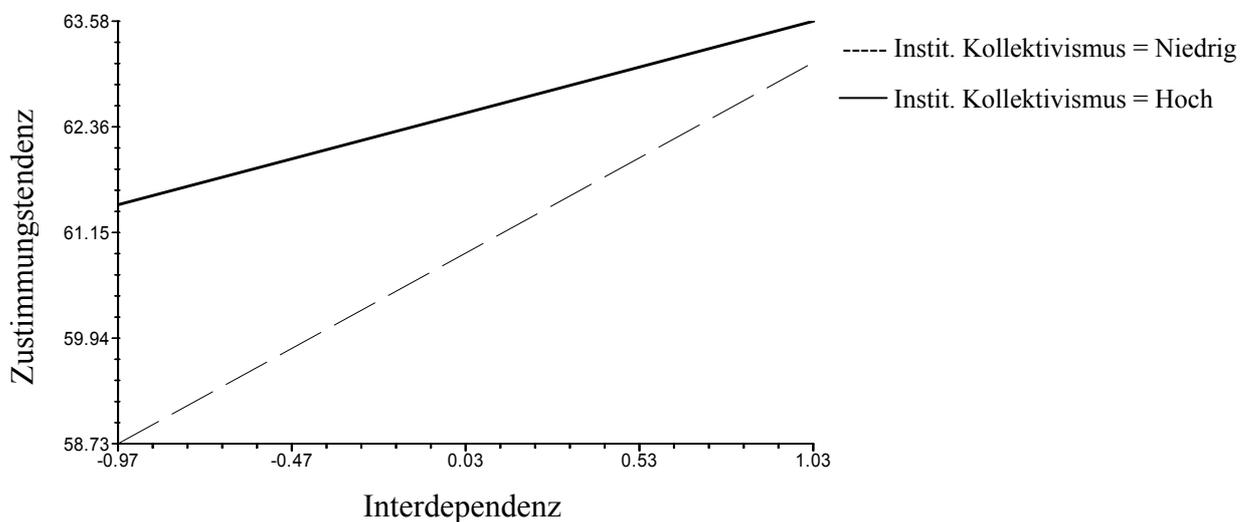


Abbildung 6: Zusammenhang von interdependentem Selbstkonzept und Zustimmungstendenz in Kulturen mit hohen und niedrigen Kollektivismuspraktiken (aus Smith & Fischer, 2008)

Wie aus dieser Abbildung hervor geht, ist zwar in Ländern, in denen Institutionelle Kollektivismuspraktiken hoch ausgeprägt sind und somit das Befolgen von

gesellschaftlichen Normen stark praktiziert wird, das Niveau der Zustimmungstendenz insgesamt höher, der Zusammenhang mit dem Selbstkonzept ist jedoch dort niedriger. Eine ähnliche, jedoch nur marginal signifikante Interaktion zeigte sich in den Berechnungen von Smith und Fischer (2008) auch bei den InGroup-Kollektivismus-Praktiken. Anhand der GLOBE-Dimensionen zeigen sich somit durchaus die erwarteten Effekte.

Die Analysen mit den Kulturdimensionen von Schwartz, die individualistische und kollektivistische Werte erfassen, resultierten hingegen in nicht so leicht nachvollziehbaren Ergebnissen. So zeigte sich in Kulturen mit geringer Intellektueller Autonomie ein starker Zusammenhang zwischen interdependentem Selbstkonzept und Zustimmungstendenzen, während er in Ländern mit hoher Intellektueller Autonomie so gut wie gar nicht vorhanden war (siehe Abbildung 7, obere Grafik).

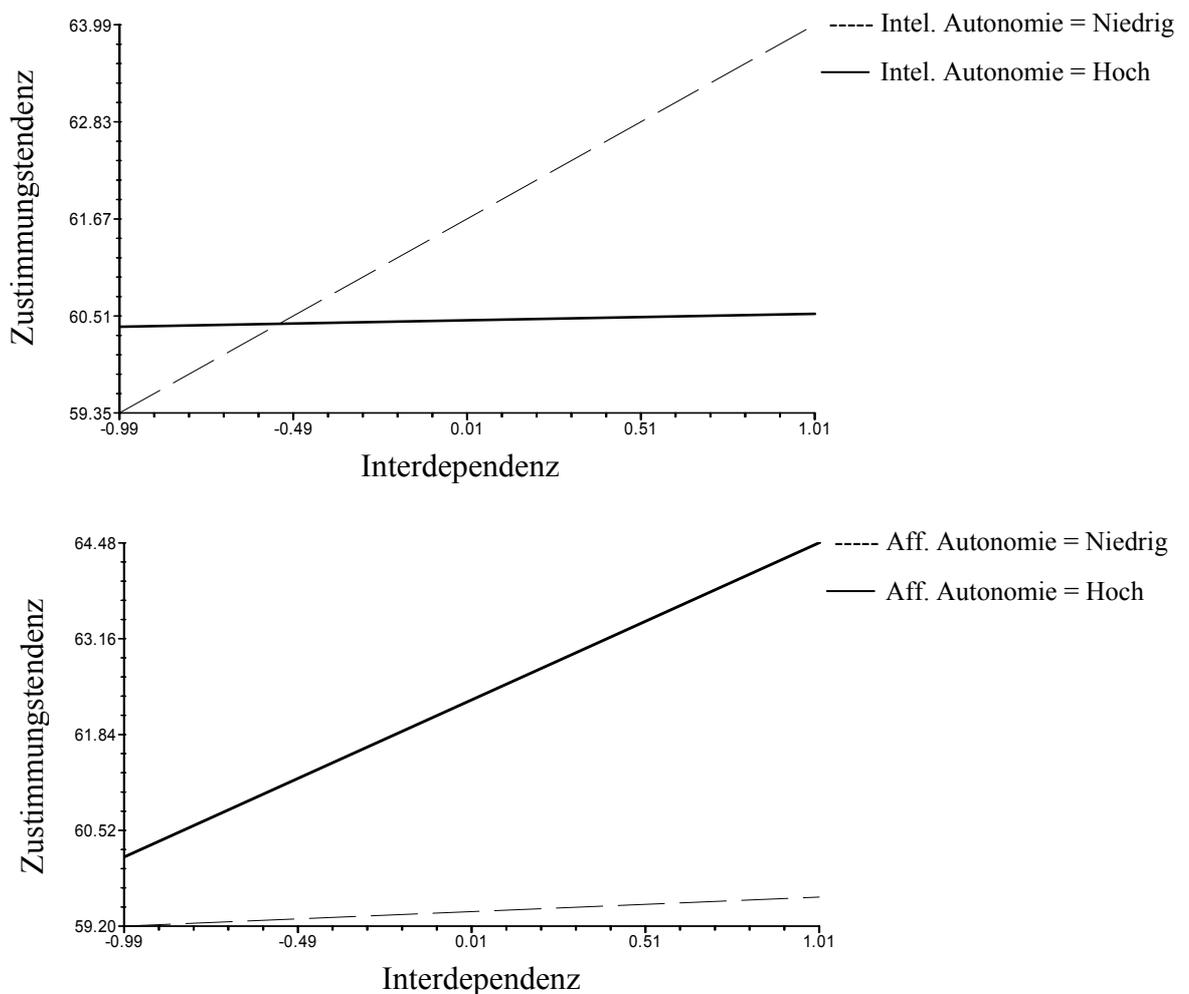


Abbildung 7: Zusammenhang von interdependentem Selbstkonzept und Zustimmungstendenz in Ländern mit hoher und niedriger Intellektueller sowie Affektiver Autonomie (aus Smith & Fischer, 2008)

Hinsichtlich der Ausprägung von Affektiver Autonomie zeigte sich erstaunlicherweise sogar ein genau entgegengesetzter Zusammenhang (siehe Abbildung 7, untere Grafik). Obwohl somit beide Dimensionen, Intellektuelle und Affektive Autonomie gemäß Schwartz (1994) Kulturen charakterisieren, in denen das Individuum mit seinen Ideen und Wünschen stärker im Vordergrund steht, zeigen sich hier scheinbar unterschiedliche Ausgangsbedingungen für das Auftreten von Zustimmungstendenzen (Smith & Fischer, 2008). Bei weiteren Analysen mit Schwartz' Dimension des Konservatismus/Eingebettetsein sowie Hofstede's Kollektivismus-Dimension konnten schließlich bezüglich der Zustimmungstendenzen überhaupt keine signifikanten Interaktionen festgestellt werden.

Die Analysen von Smith und Fischer (2008) zu den Wechselwirkungen zwischen individuellem Selbstkonzept und der kulturellen Umgebung im Hinblick auf Zustimmungstendenzen zeigten somit eher gemischte Ergebnisse. Trotz dieser nicht so durchgängig klaren Interaktionseffekte bleibt dennoch festzuhalten, daß sich insgesamt im Sinne eines Haupteffekts ein positiver Zusammenhang von interdependentem Selbstkonzept und Zustimmungstendenzen feststellen ließ, wenngleich dieser Zusammenhang je nach kultureller Umgebung zum Teil unterschiedlich stark ausgeprägt ist. Ebenso zeigte sich im Sinne eines Haupteffekts, daß Zustimmungstendenzen vor allem in kollektivistischen Kulturen auftreten. Alles in allem ergaben sich somit bei den Haupteffekten durchaus Zusammenhänge, wie sie zu erwarten waren.

Smith und Fischer (2008) beschränkten sich bei ihren Analysen jedoch nicht nur auf Zustimmungstendenzen, sondern untersuchten ebenfalls den Einfluß von individuellen und länderbezogenen Größen bei Extremtendenzen. Hierbei zeigte sich im Gegensatz zu den vorherigen Ergebnissen insgesamt auf individuellem Niveau ein negativer Zusammenhang von interdependentem Selbstkonzept und Extremtendenzen. Demnach zeigen vor allem Personen, die sich über ihr eigenes Ich definieren, extremes Antwortverhalten und äußern somit ihre Ansichten sehr deutlich.

Auf dem Länderniveau hingegen existierten eher unklare Zusammenhänge. Bezüglich der Kollektivismus-Dimension von Hofstede und der von den Autoren selbst erzeugten Kollektivismus-Dimension ergaben sich keinerlei Zusammenhänge, was somit die Ergebnisse von van Herk et al. (2004) und Harzing (2006) bestätigt. Nur bei der Analyse der GLOBE- und Schwartz-Dimensionen deutete sich an, daß extremes Antwortverhalten eher in Kulturen mit geringen Institutionellen

Kollektivismuspraktiken, mit geringer Intellektueller Autonomie oder mit hoher Affektiver Autonomie aufzutreten scheint.

Die Analyse der Interaktionen ergeben jedoch wieder ein etwas klareres Bild und deuten hier darauf hin, daß vor allem in Kulturen, die hoch individualistisch sind, die geringe Institutionelle Kollektivismuspraktiken besitzen oder aber bei denen die Werte von Konservatismus/Eingebettetsein weniger wichtig sind, Menschen mit hohem independentem Selbstkonzept stark extremes Antwortverhalten zeigen (vgl. beispielhaft den Interaktionseffekt bzgl. der eigenen Kollektivismusdimension von Smith & Fischer, 2008, in Abbildung 8). Dies kann daher so interpretiert werden, daß in Kulturen, in denen der Gruppenzusammenhalt keine wesentliche Prämisse darstellt, es scheinbar problemlos möglich ist, seine Meinung deutlich zu äußern. Genauso gut könnte man das Ergebnis in ähnlicher Form so interpretieren, daß in Kulturen, in denen das einzelne Individuum im Vordergrund steht, auch von jedem Mitglied erwartet wird, sich klar und deutlich zu äußern und seine Meinung nicht zu verbergen. Betrachtet man die weiteren Individualismus/Kollektivismus-Dimensionen zeigten die Ergebnisse von Smith und Fischer (2008) bezüglich der Autonomie-Dimensionen von Schwartz erstaunlicherweise jedoch erneut keine klar zu interpretierenden Interaktionseffekte.

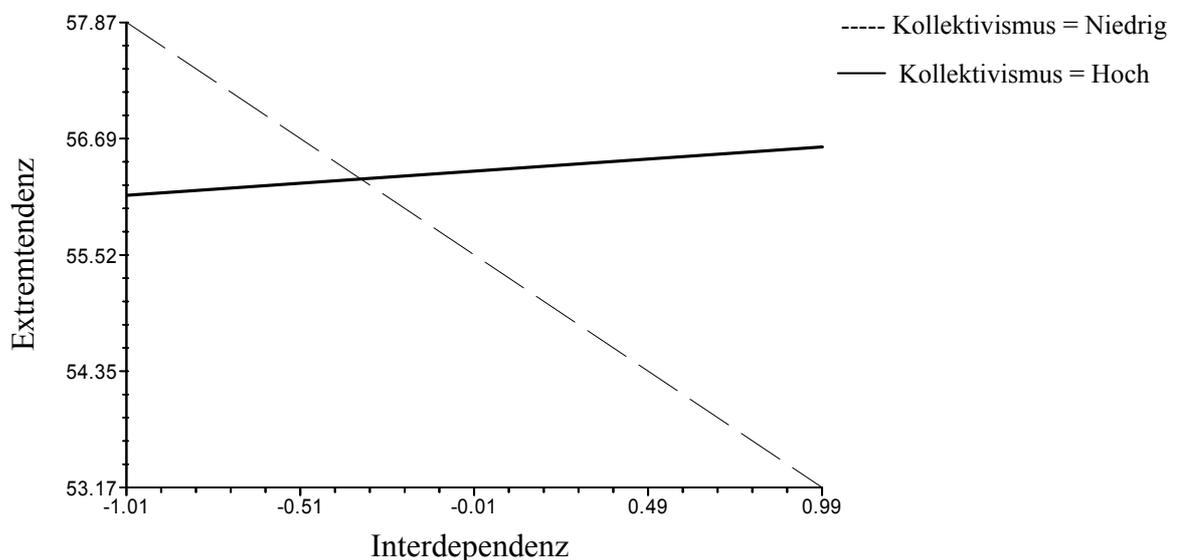


Abbildung 8: Zusammenhang von Interdependentem Selbstkonzept und Extremtendenzen in Ländern mit hohem und niedrigem Kollektivismus (aus Smith & Fischer, 2008)

Fast man die Ergebnisse zu beiden Formen von Antworttendenzen, d.h. den Zustimmung- und den Extremtendenzen, zusammen, läßt sich festhalten, daß zwar die Haupteffekte auf dem Länderniveau sowie die Haupteffekte auf dem individuellen Niveau weitestgehend nachvollziehbar sind. Jedoch lassen sich nicht alle Interaktionen zum jetzigen Zeitpunkt eindeutig interpretieren. Auch Smith und Fischer (2008) selbst lassen einzelne der in Teilen nicht erwarteten Interaktionseffekte gänzlich unkommentiert.

Alles in allem können diese Ergebnisse aber inhaltlich zeigen, daß erstens kollektivistische Wertvorstellungen in Kulturen Menschen im Mittel eher zustimmend antworten lassen, während individualistische Wertvorstellungen Menschen eher dazu drängen, extreme Antworten zu geben. Zweitens scheint es mehrheitlich so zu sein, daß Menschen mit einem interdependentem Selbstkonzept in einem kollektivistischen Kontext umso mehr zustimmend antworten, während Menschen mit einem independentem Selbstkonzept hingegen in einem individualistischen Kontext umso extremer antworten (Smith & Fischer, 2008). Drittens gibt es feine Unterschiede in den Zusammenhängen, je nachdem wie genau individualistische bzw. kollektivistische Wertvorstellungen erhoben wurde. Diesbezüglich unterscheiden sich die Einteilungen von Hofstede, Schwartz und der GLOBE-Studie trotz konzeptueller Ähnlichkeiten nämlich zum Teil deutlich. Schließlich betonen die Ergebnisse von Smith und Fischer (2008) aber vor allem, von welcher großer Bedeutung es ist, die verschiedenen Ebenen von Individualeinflüssen und nationalen Einflüssen getrennt zu betrachten und nicht einfach von einem Ergebnis auf der einen Ebene auf ein Ergebnis auf der anderen Ebene zu schließen.

7 Arbeitszufriedenheit, Kultur und Antworttendenzen

Die bisherigen Ausführungen konnten zeigen, daß in zahlreichen Studien immer wieder deutliche Unterschiede zwischen den Arbeitszufriedenheitsniveaus einzelner Länder auftreten. Diese scheinen in Teilen auch über die Studien und die Zeit hinweg stabil zu sein. Des weiteren gibt es Hinweise darauf, daß die Zufriedenheit möglicherweise mit kulturellen Wertvorstellungen kovariert und in bestimmten Kulturen grundsätzlich höher als in anderen ist. Schließlich gibt es dann noch klare Hinweise darauf, daß Kulturen sich in ihrer Art und Weise, auf Fragebögen zu antworten, sehr stark unterscheiden. Die Frage, die sich demnach nun anschließend stellt, ist, wie sich diese drei Aspekte miteinander verbinden lassen.

Zur Beantwortung dieser Frage ist von Bedeutung, daß diese Aspekte in einen theoretischen Rahmen eingebettet werden sollten, da oftmals eine Vielzahl von interkulturellen Studien einem rein deskriptiven Vorgehen ohne jegliche theoretische Fundierung folgen (Aycan, 2000; Aycan & Kanungo, 2001). Dies gilt auch für wissenschaftliche Vergleichsstudien zur Arbeitszufriedenheit. Bei Untersuchungen im praktischen Unternehmens- und Managementkontext ist es sogar fast die Regel, daß ohne Rückgriff auf theoretische Überlegungen und bisherige empirische Ergebnisse die Unterschiede in den Zufriedenheitsniveaus interpretiert werden und Konsequenzen daraus abgeleitet werden. In der hier vorliegenden Studie wurde daher Wert darauf gelegt, jedem dieser drei Aspekte eine klare theoretische Basis zu schaffen. Erst diese theoretische Einbettung ermöglicht es auch, die Vielzahl von potentiellen Alternativerklärungen einzudämmen. Für die inhaltliche Fundierung von Kulturunterschieden werden daher die kulturellen Rahmenmodelle von Hofstede, Schwarz und der GLOBE-Studie herangezogen, die in einem der vorherigen Kapitel ausführlich vorgestellt wurden. Dabei soll ein besonderer Fokus auf der Unterscheidung von individualistischen und kollektivistischen Kulturen liegen, da sich anhand dieser Dimension verschiedene Nationen am deutlichsten in der bisherigen Forschung unterscheiden ließen. Die Bedeutung des Antwortverhaltens und seine Beziehung zu kulturellen Werten und kulturellem Denken stützt sich auf die inzwischen oft replizierten Befunde der interkulturellen Forschung, wonach beispielsweise Zustimmungstendenzen vermehrt in kollektivistischen Kulturen auftreten (vgl. Smith, 2004; Smith & Fischer, 2008). Die Problematik der Arbeitszufriedenheitsmessung wird

hinreichend deutlich, wenn man sich das Modell der semantischen Prozesse der Urteilsbildung vor Augen führt (vgl. Fischer, 1991; Fischer & Belschak, 2006). Eine Verbindung dieser drei Bereiche ist jedoch bisher in der Forschung nicht vorgenommen worden und daher zentraler Bestandteil der hier durchgeführten Untersuchung. Neben dem Versuch einer Integration dieser drei Ansätze werden jedoch ebenso die bilateralen Beziehungen von jeweils zwei dieser Aspekte betrachtet. Dabei steht vor allem der Zusammenhang von Antworttendenzen und dem Arbeitszufriedenheitsniveau im Fokus des Interesses. Die Zusammenhänge von Kultur und Antworttendenz auf der einen Seite sowie Kultur und Arbeitszufriedenheit auf der anderen Seite müssen jedoch ebenso überprüft werden und sollten möglichst die bisherigen Ergebnisse empirischer Studien replizieren, um darauf aufbauend eine Gesamtintegration der drei Aspekte Kultur, Antworttendenz und Arbeitszufriedenheit zu ermöglichen.

7.1 Die Beziehung von Kultur und Arbeitszufriedenheit

Die bisher in der Forschung veröffentlichten deskriptiven Ergebnisse zum Arbeitszufriedenheitsniveau in verschiedenen Ländern legen die Vermutung nahe, daß die kulturelle Umgebung einen Einfluß auf das Niveau besitzt (siehe Judge et al., 2001). Dabei stand vor allem die kulturelle Dimension von Individualismus und Kollektivismus im Fokus des Forschungsinteresses. Dies ist sicherlich unter anderem darauf zurückzuführen, daß diese Dimension den prägnantesten und eingängigsten Kulturwert darstellt, anhand dessen sich Nationen unterscheiden lassen. Die Bedeutung gerade dieser Kulturdimension wird vor allem auch durch den oftmals in der Forschung stattfindenden Vergleich von westlichen und asiatischen Ländern deutlich. Dennoch sei nebenbei erwähnt, daß die asiatischen Länder nicht grundsätzlich die kollektivistischsten Länder darstellen (vgl. Hofstede, 1980, 2001 & 2005). Ein Großteil südamerikanischer Länder läßt sich beispielweise ebenfalls sehr hoch auf der Kollektivismusskala einordnen. Die Ergebnisse zum sich andeutenden Zusammenhang von Individualismus/Kollektivismus und Arbeitszufriedenheit sind jedoch bisher nicht häufig durch die explizite Messung des Kulturwerts untermauert worden, sondern begründen sich vielfach lediglich auf deskriptiven Unterschieden von Ländern. Diese deskriptiven Ergebnisse zeigen zwar eine scheinbar deutliche Tendenz, sind jedoch auch keineswegs immer eindeutig. Daher soll hier in einem ersten Schritt überprüft werden, inwieweit sich die bisherigen Ergebnisse und Vermutungen hinsichtlich des

Zusammenhangs von Individualismus/Kollektivismus und Arbeitszufriedenheitsurteilen auch in dieser Arbeit bestätigen lassen. Eine erste untersuchungsleitende Hypothese läßt sich daher wie folgt formulieren:

Hypothese 1: Das Arbeitszufriedenheitsniveau ist höher in individualistischen Kulturen als in kollektivistischen Kulturen.

Zwar hat der Individualismus/Kollektivismus-Vergleich bisher das größte Interesse in der interkulturellen Forschung auf sich gezogen, jedoch unterscheiden sich Kulturen darüber hinaus auch anhand zahlreicher weiterer Dimensionen. Eine genaue Vorhersage der Zusammenhänge zwischen dem Arbeitszufriedenheitsniveau und den weiteren Kulturdimensionen (bspw. Machtdistanz, Unsicherheitsvermeidung oder auch Future Orientation) läßt sich jedoch aufgrund der bisherigen Forschungsergebnisse und der deskriptiven Ländervergleiche nicht treffen. Die Kulturdimensionen sind zwar untereinander keineswegs unabhängig, so daß beispielsweise Kulturen mit einer hohen Bedeutung des Individuums auch tendenziell durch eine geringe Machtdistanz und eine geringere Unsicherheitsvermeidung charakterisiert sind (Hofstede, 2001). Daraus ließe sich schlußfolgern, daß auch möglicherweise das Arbeitszufriedenheitsniveau in Nationen mit hoher Machtdistanz und hoher Unsicherheitsvermeidung geringer sein müßte als in Nationen, deren Kultur auf diesen Dimensionen eher niedrig ausgeprägt ist. Die bisherigen Ländervergleiche lassen diesbezüglich jedoch keine eindeutige Argumentation zu, so daß die Betrachtung dieser Zusammenhänge als weitestgehend explorativ anzusehen ist. Um abschließend aber eine eindeutigere Aussage über den Bedeutungsgehalt von Kultur treffen zu können, werden daher auch die Zusammenhänge mit den anderen kulturellen Dimensionen genauer analysiert. Eine präzise Hypothese läßt sich folglich jedoch nicht apriori aufstellen. Stattdessen wird lediglich davon ausgegangen, daß auch weitere Kulturdimensionen einen bedeutenden Erklärungsgehalt für die Niveauunterschiede hinsichtlich der Arbeitszufriedenheit besitzen:

Hypothese 2: Unterschiede im Arbeitszufriedenheitsniveau lassen sich durch weitere kulturelle Dimensionen, vor allem Machtdistanz und Unsicherheitsvermeidung, erklären.

Um zudem eine Aussage über mögliche Alternativerklärungen treffen zu können, sollen neben der spezifischen Betrachtung der Zusammenhänge von Kultur und

Arbeitszufriedenheit auch ausgewählte Korrelate mit anderen sozioökonomischen Daten betrachtet werden. Dabei werden in der hier vorliegenden Studie vor allem die nationalen Indikatoren mit einbezogen, von denen anzunehmen ist, daß sie in einem engen Bezug zum Arbeitsleben stehen und daher ein direkter Einfluß auf das Arbeitszufriedenheitsniveau anzunehmen ist. Zu nennen sind hier beispielsweise das Bruttoinlandsprodukt als Indikator für die Ausprägung der Arbeitsqualität oder auch die Arbeitslosenquote, die in einem Land anzutreffen ist. Diese Vorgehensweise folgt somit der Forderung nach destruktivem Testen, um den Erklärungsgehalt von Kultur gegenüber anderen Faktoren absichern zu können.

7.2 Die Beziehung von Kultur und Antworttendenz

Die Ergebnisse der interkulturellen Forschung zeigen in übereinstimmender Weise, daß es zwischen der Kultur eines Landes und dem kommunikativen Verhalten der jeweiligen Menschen eine deutliche Beziehung gibt. Antworttendenzen treten somit verstärkt nur in einem für sie günstigen kulturellen Umfeld auf. Sich zu Items eines Fragebogens überwiegend zustimmend zu äußern, ist somit ein prototypisches Verhalten vor allem in kollektivistischen Kulturen (Harzing, 2006; Smith, 2004). Da auch in dieser Studie ein besonderes Augenmerk auf den Unterschieden von Ländern hinsichtlich Individualismus/Kollektivismus liegt, soll daher vor allem der Zusammenhang zwischen dieser Kulturdimension und kommunikativem Verhalten untersucht werden. Anhand der in dieser Studie untersuchten Daten sollte sich daher dieser bisher in Teilen nachgewiesene Zusammenhang für Zustimmungstendenzen replizieren und folgende Hypothese bestätigen lassen:

Hypothese 3: Zustimmungstendenzen sind im Mittel häufiger in kollektivistischen als in individualistischen Kulturen anzutreffen.

Ebenso zeigten bisherige Ergebnisse aber auch zum Teil Zusammenhänge zwischen Zustimmungstendenzen und zwei der weiteren Kulturdimensionen. Deutlich wurde demnach, daß diese Art des Antwortverhaltens vor allem auch in Nationen mit hoher Machtdistanz und dem Wunsch nach mehr Unsicherheitsvermeidung aufzutreten scheinen (Smith, 2004). Auch diese Hypothese sollte sich daher in dieser Studie bestätigen lassen und die bisherigen Erkenntnisse replizieren:

Hypothese 4: Zustimmungstendenzen sind im Mittel häufiger in Kulturen anzutreffen, die von hoher Machtdistanz zwischen den Menschen und von dem Streben nach mehr Unsicherheitsvermeidung charakterisiert sind.

Das Auftreten der zweiten bedeutsamen Art des Antwortverhaltens, d.h. die Auswahl von extremeren Ausprägungen einer Skala beim Beantworten von Fragebögen, konnte in empirischen Studien zwar ebenfalls häufig nachgewiesen werden, die Zusammenhänge mit spezifischen Kulturwerten sind jedoch nicht eindeutig. Es gibt lediglich erste Hinweise darauf, daß Extremtendenzen eher dort auftreten, wo das Individuum im Zentrum der Aufmerksamkeit steht und Gruppenbezüge nicht vorherrschend sind (Smith & Fischer, 2008). Von daher soll überprüft werden, inwieweit sich folgende Hypothese bestätigen läßt:

Hypothese 5: Extremtendenzen sind stärker in einem individualistischen als in einem kollektivistischen Kontext anzutreffen.

Die Aussage dieser Hypothese ist jedoch aufgrund der bisherigen mehrdeutigen Forschungsergebnisse in Teilen als explorativ anzusehen. Daher soll der Zusammenhang von Extremtendenzen insgesamt mit allen Kulturdimensionen geprüft werden, um genauere Angaben darüber machen zu können, in welchem kulturellen Umfeld Aussagen stärker auf den Endpunkten einer Likert-Skala getroffen werden.

Gleiches gilt für die Überprüfung des Zusammenhangs von Mittendenzen und kulturellen Werten. Auch wenn in anekdotischer Weise in einzelnen Studien berichtet wird, daß beispielsweise Japaner überzufällig häufig die Mitte einer Skala auswählen (Shiomi and Loo, 1999; Takahashi et al., 2002), läßt sich bisher kein eindeutiger Zusammenhang mit irgendeiner der Kulturdimensionen herstellen. Die Überprüfung des Zusammenhangs in der hier vorliegenden Untersuchung ist somit vollständig explorativ. Zwar wäre durchaus anzunehmen, daß die Auswahl der mittleren Antwortkategorie gerade in solchen Kulturen häufiger anzutreffen sein sollte, in denen Unsicherheiten vermieden werden oder aber eine hohe Machtdistanz zwischen den Menschen herrscht. Der Grund dafür wäre diesbezüglich, daß es in dieser Art von Kultur für den einzelnen Befragten eher von Nachteil sein sollte, wenn er sich klar und deutlich über seine Meinung äußert. Erstaunlicherweise konnten jedoch bisherige Studien diesen durchaus plausiblen Zusammenhang nicht durchgängig nachweisen. Eine genaue Aussage über

die vermuteten Zusammenhänge von kulturellem Denken und Antworttendenzen läßt sich somit vorab nicht treffen.

Für die dritte klassische Form des verzerrten Antwortverhaltens, der sozialen Erwünschtheit, lassen sich einzelne empirische Hinweise als auch theoretische Überlegungen finden, die dafür sprechen, daß dieses Antwortverhaltens stärker bei Kollektivistinnen ausgeprägt sein sollte. Basierend auf den Studien von van Hemert et al. (2002) sowie Bernardi (2006) wird daher angenommen, daß dieser Zusammenhang sich auch in dieser hier vorliegenden Studie nachweisen lassen sollte:

Hypothese 6: Antworten im Sinne Sozialer Erwünschtheit treten häufiger bei kollektivistischen Kulturen auf.

Für den Zusammenhang mit den weiteren Kulturdimensionen bestehen auch hier bisher keine tiefreichenden Erkenntnisse, so daß sich diesbezüglich keine genauen Hypothesen formulieren lassen.

7.3 Die Beziehung von Antworttendenz und Arbeitszufriedenheit

Während über die Zusammenhänge zwischen Kultur und Antworttendenz einerseits und Kultur und Arbeitszufriedenheit andererseits schon in Teilen Erkenntnisse in der Forschung vorliegen, aus denen sich die bisherigen Hypothesen ableiten ließen, gibt es für den Zusammenhang zwischen Antworttendenzen und dem Arbeitszufriedenheitsniveau bisher überhaupt keine veröffentlichten Ergebnisse. Hier betritt die vorliegende Studie empirisches Neuland. Somit lassen sich auch keine Überlegungen in der Literatur finden, die sich explizit mit dieser Thematik beschäftigen und anhand derer eine spezifische Hypothese abgeleitet werden könnte. Dies ist umso überraschender, da aufgrund der zahlreichen Erkenntnisse zum Auftreten von kulturspezifischen Antworttendenzen es sehr plausibel ist, daß auch zwischen einem Urteil zu einer Arbeitszufriedenheitsfrage und der generellen Art und Weise, wie in Fragebögen kommuniziert wird, ein Zusammenhang bestehen sollte. Von einem Zusammenhang zwischen der Antworttendenz, die in einer Kultur vorherrscht, und dem Arbeitszufriedenheitsniveau, das in dem jeweiligen Land angegeben wird, sollte somit grundsätzlich auszugehen sein.

Dies ist vor allem deshalb anzunehmen, da Arbeitszufriedenheitsfragen in der Regel nach genau dem gleichen Muster erhoben werden, anhand dessen sich auch kulturell

unterschiedliche Antworttendenzen nachweisen ließen. Wenn also Befragte auf Likert-skalierte Fragen beispielsweise zu politischen Einstellungen, zu studienbezogenen Aspekten, zu der Wichtigkeit von Arbeitsaspekten, etc. zustimmender oder auch extremer antworten, ist anzunehmen, daß sie gleiches Verhalten auch bei Fragen zur Arbeitszufriedenheit zeigen sollten. Auch Arbeitszufriedenheitsfragen werden nämlich zumeist in der gleichen Art von Erhebungssituation (d.h. zumeist Fragebögen) mit der gleichen Methode (d.h. Likert-skaliert) erhoben, wie dies in den Studien zu Antworttendenzen der Fall war. Es wäre somit nur schwer nachvollziehbar, warum gerade ein Urteil der Arbeitszufriedenheit nicht verzerrt sein sollte. Ganz im Gegenteil ist davon auszugehen, daß Befragte sich auch zu ihrer Zufriedenheit mit der Arbeit aufgrund der kommunikativen Norm, der sie in ihrer Kultur folgen, und aufgrund der daraus resultierenden allgemeinen und inhaltsunspezifischen Antworttendenz etwas anders äußern, als sie sich tatsächlich fühlen. Wenn somit im Mittel eine Kultur sich tendenziell sehr stark in Form von Antworttendenzen äußert, dann sollte auch das mittlere Urteil der individuellen Arbeitszufriedenheitsaussagen, d.h. das Arbeitszufriedenheitsniveau, davon beeinflusst sein. Daher ergibt sich auf Grundlage bisheriger Studien zur Antworttendenz-Thematik und dieser hier angestellten Überlegungen bezüglich dieses Zusammenhangs folgende erste, ungerichtete Hypothese:

Hypothese 7: Unterschiede im Arbeitszufriedenheitsniveau lassen sich durch die Höhe der Antworttendenz einer Kultur erklären.

Über diese generelle Tendenz hinaus, von der nun grundsätzlich auszugehen sein sollte, stellt sich jedoch die spezifischere Frage, in welche Richtung eine kulturell bedingte Antworttendenz ein Arbeitszufriedenheitsniveau verschieben würde. Die Fragen, die sich hier stellen, lauten daher: Würde gelten, je höher eine Zustimmungsnorm ist, desto höher fällt auch das Arbeitszufriedenheitsniveau aus, oder gilt genau der umgekehrte Zusammenhang, daß das Arbeitszufriedenheitsniveau durchschnittlich geringer ist, je stärker die Kommunikationsnorm „man sollte zustimmend antworten“ ausgeprägt ist? Gibt es unterschiedliche Zusammenhänge auf dem kulturellen Level zwischen Zustimmungsnormen und dem Zufriedenheitsniveau, verglichen mit dem Zusammenhang auf dem individuellen Level zwischen individuellen Zustimmungstendenzen und dem persönlichen Arbeitszufriedenheitsurteil? Welcher Zusammenhang existiert darüber hinaus mit den anderen Antworttendenzen in Form

von Extrem- und Mittentendenzen? Da diesbezüglich bisher keine empirischen Erkenntnisse vorliegen, lassen sich nur theoretische Überlegungen über den möglichen Zusammenhang anstellen.

Betrachtet man zuerst das individuelle Level, wäre sicherlich intuitiv naheliegend anzunehmen, daß bei einer hohen Zustimmungstendenz auch ein Arbeitszufriedenheitsurteil im Mittel höher ausfallen sollte. Der Grund dafür wäre, daß eine Zustimmungstendenz gleichbedeutend ist mit dem Verschieben der Antworten hin zu einem Skalenende ist, womit unter Annahme des gleichen Effekts bei Zufriedenheitsfragen die Antworten dort grundsätzlich ebenfalls zu einem Ende der Skala tendieren sollten. Weiterhin könnte man vielleicht annehmen, daß die zustimmende Seite der „stimme zu – stimme nicht zu“-Likert-Skala eine gleiche inhaltliche Qualität mit der Zufriedenheitsseite der „äußerst zufrieden – überhaupt nicht zufrieden“-Skala besitzt, da beide eher die bejahende Seite der Skala darstellen. Daher würde eine zustimmende Auswahl der Antwortalternativen gleichbedeutend mit einer „ich bin zufrieden“-Äußerung bei einer Zufriedenheitsfrage sein.

Hinzu käme, daß, wenn Personen Zustimmungstendenzen zeigen und sich daher unabhängig vom Inhalt der Frage zustimmend äußern, sie die Aussage des Fragenden und deren implizite Botschaft nicht hinterfragen möchten und dieser Aussage nicht widersprechen möchten. Diese Form des Widersprechens wird in manchen Kulturen nicht gerne gesehen, wie die bisherige Forschung bereits zeigen konnte. Wenn nun aber jemand persönlich gefragt werden würde, wie es ihm geht, wird sicherlich oftmals eine Aussage im Sinne eines „gut, alles in Ordnung“ erwartet. Dieser Art und Weise, auf die Frage nach dem Wohlbefinden zu antworten, würde sicherlich erst recht erwartet, wenn es sich bei dem Fragenden um einen Fremden handeln würde. Auch unter der Annahme dieser kommunikativen Erwartung wäre dann somit eine zustimmende Aussage auf eine Zufriedenheitsfrage die Aussage „mir geht es gut, ich bin zufrieden“. Jemand, der in seinem persönlichen kommunikativen Verhalten sich sehr zurückhaltend äußert und oft dem Gegenüber zustimmt, sollte folglich auch ein eher positiv gefärbtes Zufriedenheitsurteil nach außen hin abgeben.

Aufgrund dieser Überlegungen wäre daher anzunehmen, daß sich Personen, die insgesamt eine Ja-Sage-Tendenz aufweisen (also eine höhere Zustimmungstendenz zeigen), auch bei Arbeitszufriedenheitsfragen tendenziell zustimmender und somit eher positiv äußern. Auf dem individuellen Level sollte sich daher folgende Hypothese bestätigen lassen:

Hypothese 8: Personen mit hoher Zustimmungstendenz äußern sich auch bei Fragen zur Arbeitszufriedenheit positiver.

Trotz dieser grundsätzlichen plausiblen Annahme soll jedoch auch in kritischer Hinsicht ein weiterer Gedanke kurz erläutert werden. Es könnte nämlich durchaus auch die Frage gestellt werden, wie eine Person überhaupt auf eine „Wie“-Frage (und Arbeitszufriedenheitsfragen werden in aller Regel nach dem Prinzip „Wie zufrieden sind Sie mit Ihrem Job?“ gestellt) zustimmend antworten kann. In einer „Wie“-Frage wird semantisch eigentlich gar keine Richtung vorgegeben und keine Aussage des Fragenden transportiert, der man somit dann auch im Prinzip überhaupt nicht zustimmen kann. Anders verhält es sich bekanntermaßen bei den Item-Batterien, die zur Messung von Zustimmungstendenzen eingesetzt werden. Dort wird durch ein Item bereits eine Aussage in den Raum gestellt (z.B. „Deutschland ist ein schönes Land“ oder auch „politische Sanktionen sind sinnvolle Maßnahmen in einer Demokratie“), der erst einmal widersprochen werden müßte. Bei Zufriedenheitsfragen jedoch, die im „Wie“-Stil gefragt werden, gibt es eigentlich keine Aussage, der somit auch nicht widersprochen werden kann. Es muß unter Umständen also gar nicht gesichert sein, daß ein zustimmendes Antwortverhalten zwingend gleichbedeutend mit einem höheren Zufriedenheitsurteil sein müßte. Des weiteren ist offensichtlich ein „stimme zu – stimme nicht zu“-Skalenlabel nicht identisch mit einem „zufrieden – nicht zufrieden“-Label. Automatisch von einem gleichgerichteten Einfluß auszugehen, obwohl die Skalenbezeichnungen unterschiedlich sind, könnte daher fragwürdig erscheinen. Erstaunlicherweise hat die bisherige Forschung diesem Umstand überhaupt keine Rechnung getragen, wenn – wie beispielsweise in der Studie von Harzing (2006) geschehen – verschiedene Skalenlabel einfach zu einem gemeinsamen Zustimmungsindex verrechnet werden. Stattdessen sollte aber eher eine Trennung in die verschiedenen Skalenlabel durchgeführt werden, denn es gibt keinen zwingenden Grund anzunehmen, daß nicht auch unterschiedliche Skalen unterschiedliche Antworteffekte hervorrufen könnten. Vor diesem Hintergrund wäre es daher auch grundsätzlich möglich, daß ein zustimmendes Antwortverhalten eher zu geringeren Arbeitszufriedenheitswerten führen würde.

Obwohl gemäß diesen letzten Überlegungen somit unter Umständen sogar ein negativer Zusammenhang auf individuellem Level zwischen dem generellen kommunikativen Verhalten einer Person und der entsprechenden Aussage zu einer Zufriedenheitsfrage

bestehen könnte, erscheint jedoch die vorangegangene Argumentation insgesamt plausibler. Es wäre demnach – wie bereits in Hypothese 8 ausgedrückt – insgesamt wahrscheinlicher, davon auszugehen, daß Zustimmungstendenzen auf dem individuellen Level ein höheres Zufriedenheitsurteil erzeugen würden. Dies gilt vor allem auch wegen der semantisch ähnlichen Konnotation der ‚stimme-zu‘-Seite einer Likert-Skala und der ‚zufrieden‘-Seite der entsprechenden Zufriedenheitsskala. Da bei beiden etwas Positives mitschwingt und bei beiden Aussagen der Antwortende die individuelle Erwartung des Anderen bestätigt, sollte sich folglich doch ein positiver Zusammenhang zeigen.

Wenngleich diese Überlegungen auf dem individuellen Niveau hinsichtlich der Aussagen jeder einzelnen Person eine hohe Plausibilität besitzen, ergibt sich ein möglicherweise anderes Bild, wenn man sich die bisher schon zusammengetragenen Erkenntnisse auf dem gesellschaftlichen, kulturellen Level anschaut.

Triandis, McCusker und Hui (1990) stellen als Fazit ihrer Forschungsergebnisse fest, daß Personen aus westlichen Kulturen sich stärker darüber definieren, unterschiedlich zu sein, während Personen aus östlichen (d.h. asiatischen) Kulturen darauf achten, gleich zu sein. Diesbezüglich verweisen sie auf ein prototypisches japanisches Sprichwort, welches besagt, „a nail sticking out invites hammering down“ (Triandis et al., 1990, S. 1018). In der japanischen Kultur, die deutlich kollektivistischer ist als die deutsche oder auch die amerikanische Kultur, wird es somit nicht gern gesehen und ist nicht angenehm für den Einzelnen, aus der Gruppe hervorzustechen. Stattdessen versucht man sich eher in Reih und Glied mit den anderen Nägeln zu präsentieren, wenn man das Beispiel des japanischen Sprichworts aufnehmen wollte. Dies deckt sich dann auch mit der Beobachtung der stärkeren Zustimmungstendenzen in den kollektivistischen Kulturen. Darüber hinaus zeigt sich, daß asiatische und somit kollektivistische Kulturen sich mit stärkerer Zurückhaltung präsentieren und versuchen, bescheiden zu wirken (Triandis et al., 1990). Ebenso können Kitayama, Markus, Matsumoto und Norasakkunit (1997) zeigen, daß japanische Kulturen eher eine maßvolle Selbst-Präsentation an den Tag legen. In die gleiche Richtung deuten die Schlußfolgerungen von Markus und Kitayama (1991), die feststellen, daß kollektivistischere Kulturen und Personen mit einem stärker interdependenten Selbst deutlich dazu tendieren, ihre Gefühle zurückzuhalten, sie nicht offen zu zeigen und sich sehr zurückhaltend über ihre privaten Gedanken zu äußern. Dies sollte vor allem dann gelten, wenn es sich bei den Kommunikationspartner um Mitglieder der Outgroup

handelt und nicht um Personen, die zur Ingroup gezählt werden, denn gerade in kollektivistischen Kulturen haben die Menschen einen sehr engen Bezug zur Ingroup, jedoch eine größere Distanz zur Outgroup (Nakane, 1970). In ihren oft zitierten Überlegungen können Markus und Kitayama (1991) ebenfalls zeigen, daß speziell Emotionen, die sehr eng mit der eigenen Person verbunden sind, in kollektivistischen Kulturen nicht gezeigt werden (dürfen). So gilt gemäß ihrer Argumentation hinsichtlich des Zeigens von Stolz beispielsweise, daß „*pride may mean hubris*“ (S. 237). Um diese theoretische Annahme auch empirisch zu untermauern, verweisen sie unter anderem auf eine Studie von Stipek, Weiner und Li (1989), in der in der Tat Chinesen weniger Stolz hinsichtlich ihrer persönlichen Leistungen zeigten als Amerikaner. Insgesamt besteht ein zentrales Fazit der umfangreichen Überlegungen von Markus und Kitayama (1991) über die Unterschiedlichkeit von individualistischen und kollektivistischen Kulturen darin, daß letztere sich vor allem durch ein deutliches Maß an Bescheidenheit und Zurückhaltung auszeichnen:

„The nature of modesty. [...] Perhaps those from interdependent cultures have simply learned that humility is the desired response, or the culturally appropriate response, and that it is wise not to gloat over their performance or to express confidence in their ability. This interpretation implies that the modesty biases [...] are primarily the result of impression management and that the subjects involved actually could have held different, perhaps opposite, beliefs about themselves and their ability. [...] given the press not to stand out and to fit in, people in interdependent cultures may acquire through socialization a habitual modest-response tendency. [...] Consequently, for those with interdependent selves, whenever certain aspects of self need to be appraised in public, a modest, self-effacing pattern of responses may occur spontaneously. Furthermore, this modesty can be motivated by many other-focused emotions that are central to the construal of self as an interdependent entity.” (S. 244)

In individualistischen Kulturen hingegen definieren sich Menschen nicht über einen starken Gruppenzusammenhalt, sondern eher über die eigenen, persönlichen positiven Erlebnisse. Hedonistisches Denken ist daher in individualistischen Kulturen viel verbreiteter (Triandis et al., 1990). Und zur Erreichung von persönlicher Lust und der

damit verbundenen Freiheit im Handeln können Gruppen fast eher eine Last sein, wie Triandis et al. (1990) es formulieren. Für Menschen in individualistischen Kulturen ist es insgesamt viel wichtiger, daß es ihnen selbst gut geht; sowohl im Leben generell als auch beispielsweise in ihrem Job. Man könnte daher annehmen, daß Menschen in individualistischen Kulturen bei sich selbst einen Widerspruch auslösen würden, wenn sie über ihren eigenen Job sagen müßten, daß er sie nicht zufriedenstellt. Wenn darüber hinaus noch gilt, daß gemäß den Erkenntnissen von Markus und Kitayama (1991) gerade individualistische Kulturen versuchen, kognitive Dissonanzen zu verringern, wird deutlich, daß ein hohes Zufriedenheitsurteil eine umso größere Bedeutung in individualistischeren Kulturen erhält. Zusätzlich betonen Markus und Kitayama (1991) in ihren Überlegungen die hohe Bedeutung des Herausstellens eigener Fähigkeit und Leistungen in individualistischen Kulturen und machen deutlich, daß „from an independent viewpoint, [...] modesty seems false and the result of suppressing a „natural“ pride in one’s attributes.“ (S. 244). Ruft man sich ebenfalls noch einmal die Ergebnisse der Forschung zu Antworttendenzen ins Gedächtnis, wird deutlich, daß diese sich mit dieser Argumentation von Markus und Kitayama (1991) decken und im Prinzip die andere Seite der gleichen Medaille darstellen. Kollektivistische Kulturen sind geprägt von Bescheidenheit und Zurückhaltung, weshalb sie folglich höhere Zustimmungstendenzen zeigen, während individualistische Kulturen im Gegensatz dazu eine klare Meinungsäußerung erwarten. Somit würde in einem individualistischen Kontext eine Kommunikationsnorm vielmehr vorschreiben, die Gefühle und Gedanken frei zu äußern.

Überträgt man diese Erkenntnisse über die interkulturellen Unterschiede zwischen individualistischen und kollektivistischen Nationen aber nun auf die Frage nach der Zufriedenheit mit der Arbeit, dann wird deutlich, daß es auf dem kulturellen Level möglicherweise auch einen anderen Zusammenhang zwischen Antworttendenznormen und dem Arbeitszufriedenheitsniveau geben kann als dies auf dem individuellen Level der Fall wäre. Betrachtet man die Zustimmungsnormen, dann könnte es umso plausibler sein, daß die Art von zurückhaltender Kommunikation, die durch die Zustimmungsnorm ausgedrückt wird, eher dazu führen müßte, daß das Urteil zur Arbeitszufriedenheit im Mittel tendenziell niedriger sein müßte und daher vielmehr auf dem kulturellen Level ein negativer Zusammenhang von Zustimmungsnorm und Arbeitszufriedenheitsniveau vorläge. Vergegenwärtigt man sich noch einmal, was genau das Auftreten von kulturell bedingten Zustimmungsnormen bedeutet, nämlich das

Zurückstellen der eigenen Meinung und eine generelle Zurückhaltung in der interpersonellen Kommunikation, dann sollte sich diese Art von Zurückhaltung bei Fragen, die die eigene Leistung betreffen, eher in einem gemäßigeren Urteil äußern (siehe Markus & Kitayama, 1991). Aussagen zur Beurteilung des eigenen Arbeitsplatzes können dabei durchaus als eine Art eigener Leistungsbeurteilung angesehen werden. Wenn daher auf der einen Seite hohe Zustimmungsnormen in einer Kultur vorliegen, sollte demnach ein Arbeitszufriedenheitsurteil in dieser Kultur im Mittel umso geringer ausfallen, bzw. umgekehrt würde gelten, daß äußerst hohe Zufriedenheitsurteile nur in solche Kulturen vorkommen dürften, wo dieses Herausstellen der eigenen Zufriedenheit einerseits wichtig ist (d.h. in individualistischen Kulturen) und andererseits auch ‚erlaubt‘ ist (d.h. bei geringen Kommunikationsnormen zu zustimmendem Verhalten). Ein Mensch in einer individualistischen Kultur müßte demnach daran interessiert sein, sich eher sehr zufrieden über die eigene Arbeit zu äußern; sowohl aufgrund des hedonistischen Bedeutungsgrads, aufgrund des Strebens, die eigene Meinung nicht zu verbergen als auch aufgrund der Intention, sich von anderen abzuheben. In kollektivistischen Gesellschaften jedoch müßte es hochgradig unerwünscht sein, sich als extrem zufrieden zu bezeichnen, da man damit mit großer Wahrscheinlichkeit aus der Gruppe hervorsticht und sich anmaßend über die eigenen Gefühle äußert. Eine passende Verhaltensweise, um dieses Risiko zu minimieren, wäre folglich, daß man sich eher gemäßigt über die Zufriedenheit äußert.

Betrachtet man noch einmal kurz die deskriptiven Länderunterschiede im Arbeitszufriedenheitsniveau (vgl. Kapitel 4.1), zeigen sich Ergebnisse, die durchaus in Einklang mit diesen Überlegungen zu bringen sein könnten. Asiatische (d.h. mehrheitlich kollektivistische) Länder waren zumeist eher im unteren Bereich der Rangfolgen zu finden. Diese äußerten sich zwar auch insgesamt zufrieden mit der Arbeit, da auch deren Mittelwerte im positiven Bereich der Zufriedenheitsskala lagen, jedoch insgesamt scheinbar gemäßiger als die individualistischen Länder. Folgt man nun aber der gerade dargelegten Argumentationsperspektive, wäre demnach eine mögliche Erklärung für das niedrige Zufriedenheitsniveau die unterschiedliche Kommunikationsnorm in den jeweiligen Kulturen. Je stärker man sich im Mittel zustimmend zu äußern hat und somit die eigene Meinung zurückzustellen hat, desto gemäßiger würde man sich im Mittel über die Arbeitszufriedenheit äußern. Dabei ist zu beachten, daß bei der Erfassung der Zustimmungsnormen in der Regel kognitive Einstellungsitems erfaßt werden, die keinen direkten persönlichen Bezug zum Befragten

haben, während die Frage nach der Arbeitszufriedenheit ein eher emotionales und motivationales Urteil darstellt, bei dem der Befragte seine eigenen persönlichen Gefühle äußert. Somit ließe sich auch der dann auftretende, möglicherweise überraschende Zusammenhang erklären, daß die eine Maßzahl einem höheren Wert auf der Likert-Skala entsprechend würde, während das Zufriedenheitsniveau einen geringeren Wert annehmen würde. Die gleiche Art kommunikativen Verhaltens, nämlich Bescheidenheit und Zurückhaltung, würde sich folglich bei unterschiedlichen Fragetypen unterschiedlich manifestieren. Damit diese Argumentation Bestand haben könnte, sollte sich somit folgende Hypothese empirisch bestätigen lassen:

Hypothese 9: Je höher die nationale Zustimmungsnorm ist, desto geringer ist das Arbeitszufriedenheitsniveau.

Im Gegensatz zum vermeintlich naheliegenden positiven Zusammenhang wäre es somit eher im Einklang mit früheren interkulturellen Befunden, wenn sich auf kulturellem Niveau eher ein negativer Zusammenhang zwischen diesen beiden Aspekten zeigen würde. Es wird folglich ein entgegengesetzter Zusammenhang angenommen, je nachdem ob man die individual-psychologische Fragestellung betrachtet oder aber den interkulturellen Zusammenhang.

Insgesamt würde somit gelten, daß individualpsychologisch gesehen zustimmendes Antwortverhalten tendenziell zu höheren Arbeitszufriedenheitswerten führt. Dies spielt sich jedoch in individualistischen Kulturen auf einem höheren Niveau ab, während es in kollektivistischen Kulturen eher auf einem gemäßigeren Zufriedenheitsniveau stattfindet, so daß auf dem Länderlevel insgesamt ein negativer Zusammenhang zwischen Zustimmungsnorm und Arbeitszufriedenheitsniveau resultiert. Folgende Hypothese macht diesen angenommen Zusammenhang deutlich:

Hypothese 10: Betrachtet man beide Ebenen gemeinsam, so besteht auf dem individuellen Niveau ein positiver Zusammenhang zwischen Zustimmungstendenz und Arbeitszufriedenheitsurteilen. Auf dem kulturellen Niveau ist jedoch das Arbeitszufriedenheitsniveau insgesamt niedriger je höher die Zustimmungsnorm in einer Kultur ausgeprägt ist.

Während Bescheidenheit in der Kommunikation sich gemäß oben genannten Überlegungen bei kognitiven Einstellungsitems in zustimmenden Antworten äußern

könnte, könnte man gegebenenfalls auch annehmen, daß sich diese Bescheidenheit vielleicht auch eher in mittleren Kreuzen auf einer Skala ausdrücken würde. Für den Zusammenhang mit dem Arbeitszufriedenheitsniveau würde dies dann bedeuten, daß bezüglich Mittentendenzen ebenfalls ein negativer Zusammenhang auf dem kulturellen Level existieren müßte. Umgekehrt sollte eine Kommunikationsnorm im Sinne der Extremtendenz dann positiv mit dem Arbeitszufriedenheitsniveau zusammenhängen. Erstaunlicherweise zeigten sich jedoch – wie bereits dargestellt – nicht immer eindeutige und klare Ergebnisse für das Auftreten von Extrem- und Mittentendenzen in individualistischen oder kollektivistischen Kulturen, so daß bezüglich der Zusammenhänge dieser Tendenzen mit dem Arbeitszufriedenheitsniveau sich auf Basis der bisherigen Forschungserkenntnisse keine ausreichend fundierten Hypothesen aufstellen lassen.

7.4 Die Beziehung von Kultur, Antworttendenz und Arbeitszufriedenheit auf nationalem Level

Die bisherigen Hypothesen beschäftigten sich ausschließlich mit dem Zusammenhang von lediglich jeweils zwei der insgesamt drei bedeutsamen Aspekte Kultur, Antworttendenz und Arbeitszufriedenheit. Zusammengefaßt lassen sich die bisherigen angenommenen Beziehungen, die sich auf das kulturelle Level beziehen, folgendermaßen darstellen:

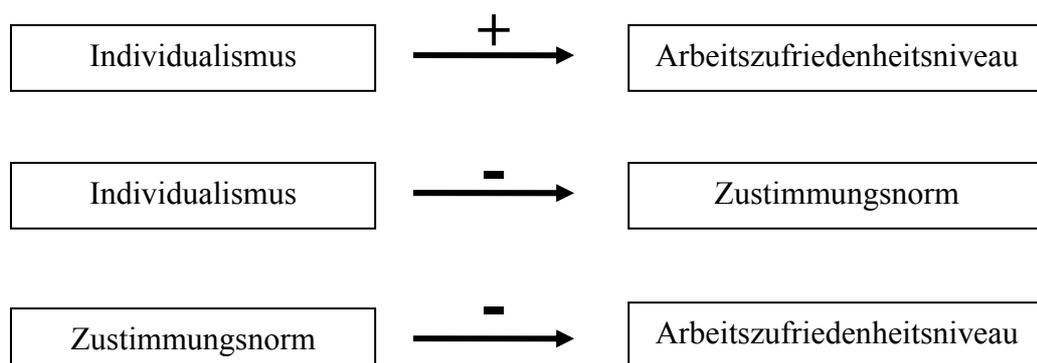


Abbildung 9: Angenommener bivariater Zusammenhang von Kultur, Antworttendenz und Arbeitszufriedenheit

Da diese drei Aspekte jeweils wechselseitig miteinander verbunden sind, liegt es nahe, sie in einen gemeinsamen Zusammenhang zu bringen. Ordnet man die Aspekte gegenseitig zu, ergibt sich folglich folgendes Bild (vgl. Abbildung 10): Individualismus steht in negativer Beziehung zu Zustimmungsnormen und in positiver Beziehung zum Arbeitszufriedenheitsniveau. Zustimmungsnormen werden ihrerseits negativ von Individualismus beeinflusst und beeinflussen selbst das Arbeitszufriedenheitsniveau in negativer Hinsicht. Das Arbeitszufriedenheitsniveau schließlich steht in positivem Zusammenhang zum Kulturwert Individualismus und in negativer Beziehung zu Zustimmungsnormen.

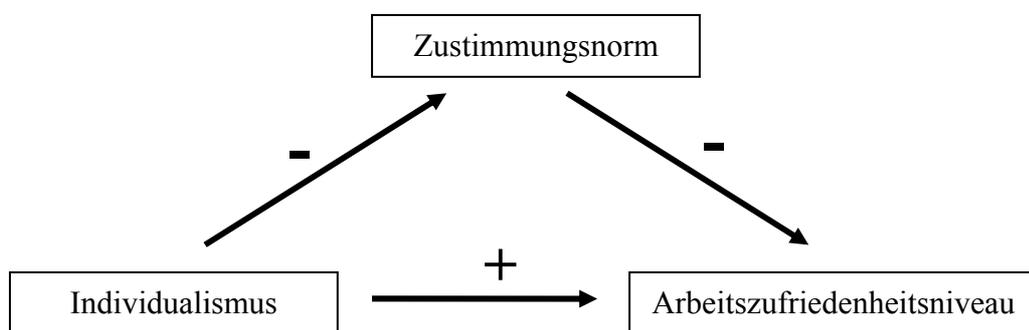


Abbildung 10: Angenommener Gesamtzusammenhang von Kultur, Antworttendenz und Arbeitszufriedenheit

In Abbildung 10 sind explizit nur Individualismus als ein bedeutsamer Kulturwert und Zustimmungstendenz als eine wichtige Antworttendenz dargestellt, da sich nur für diese Aspekte fundierte, begründete Hypothesen aufstellen lassen. Die Zusammenhänge von weiteren Kulturdimensionen mit weiteren Antworttendenzen sind bisher in der Literatur nicht so eindeutig, so daß diese sich auch nicht in solch einer Dreiecksbeziehung klar zuordnen lassen.

Eine weitere Frage, die sich bei solch einer wechselseitigen Beziehung fast aufdrängt, wäre dann, ob nicht einer der Aspekte in der Lage ist, die Beziehung der jeweils anderen Faktoren zu medieren. Dabei wäre es zwar grundsätzlich methodisch möglich, daß jeder der drei Aspekte einen medieren Faktor darstellen könnte. Aus inhaltlicher Sicht und aufgrund theoretischer Begründungen kommt jedoch im Prinzip nur einer der drei Faktoren als Mediator in Frage. Aufgrund der Tatsache, daß Kultur als der grundlegendste und basalste Faktor anzusehen ist, ist daher aus theoretischer Sicht davon auszugehen, daß sich folgender Zusammenhang ergibt: Wenn Menschen in eine

Kultur geboren werden, werden sie mit deren Denkmuster konfrontiert. Sie werden sozialisiert, bestimmte Werte und Denkweisen zu übernehmen und sich in der für die Kultur akzeptablen Weise zu verhalten. Ein wichtiger Bestandteil des Verhaltens den Mitmenschen gegenüber ist folglich auch die Art und Weise, wie Menschen in dieser Kultur miteinander kommunizieren bzw. kommunizieren sollen. Die Kultur und ihre Denkmuster führen demnach zu einer bestimmten Kommunikationsform und daher auch zu mehr oder weniger starken Zustimmungsnormen in der jeweiligen Kultur (vgl. linker Pfeil in Abbildung 10).

Das in einer Kultur akzeptierte und somit vorgegebene Kommunikationsverhalten sollte dann folglich auch beeinflussen, wie Menschen sich in einem Fragebogen äußern. Die von den Befragten geschilderte Arbeitszufriedenheit, die beispielsweise in solch einem Fragebogen abgefragt würde, wäre somit auch durch Antworttendenz-Normen mitbestimmt (vgl. rechter Pfeil in Abbildung 10).

Das Kommunikationsverhalten im Allgemeinen bzw. die Antworttendenz-Normen im Spezifischen wären somit eine mögliche Erklärung dafür, warum das Arbeitszufriedenheitsniveau sich in verschiedenen Kulturen unterscheiden könnte (vgl. unterer Pfeil in Abbildung 10). Zustimmungsnormen könnten somit einen Faktor darstellen, der in der Lage wäre, die bisherige kulturelle Blackbox genauer zu erklären, und würden demnach den mediiierenden Faktor darstellen, der zwischen Kultur auf der einen Seite und dem Arbeitszufriedenheitsniveau auf der anderen Seite vermitteln würden. Statt der Annahme, daß Kultur an sich zu einem höheren Arbeitszufriedenheitsniveau führen würde (in jedoch bisher ungeklärter Weise), würde vielmehr gelten, daß Kultur zu einer speziellen Art von Kommunikation führen würde, die wiederum dann einen Einfluß auf Arbeitszufriedenheitsurteil ausübt und somit deren Ausprägung in einem Fragebogen überformt (vgl. Abbildung 11).

Der Zusammenhang zwischen Kultur und dem Arbeitszufriedenheitsniveau müßte dann als eine Art Scheinkorrelation aufgefaßt werden, sollte sich diese Annahme bestätigen lassen. Zumindest für die Kombination aus der Kulturdimension Individualismus/Kollektivismus, aus Zustimmungsnormen und dem Arbeitszufriedenheitsniveau läßt sich somit folgende Hypothese aufstellen:

Hypothese 11: Zustimmungsnormen sind in der Lage, den Zusammenhang von Individualismus und dem Arbeitszufriedenheitsniveau zu mediiieren.

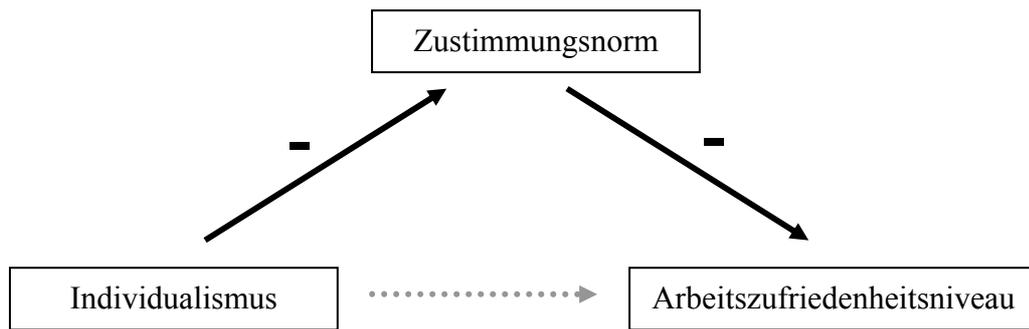


Abbildung 11: Antworttendenz als möglicher medierender Faktor zwischen Kultur und Arbeitszufriedenheit

Auch für andere Kombinationen aus z.B. Machtdistanz, Zustimmungsnorm und Arbeitszufriedenheit oder aber Individualismus, Extremtendenz und Arbeitszufriedenheit sollen die jeweiligen Zusammenhänge zwar statistisch überprüft werden. Da sich hinsichtlich dieser Zusammenhänge aber keine klaren Vorhersagen aufgrund der bisherigen empirischen Ergebnisse treffen lassen, sind diese Analysen als explorativ anzusehen.

Empirischer Teil

8 Methodisches Vorgehen

Ein zentraler Fokus der hier vorliegenden Arbeit ist die Untersuchung der Einflußgrößen auf das Arbeitszufriedenheitsniveau in verschiedenen Kulturen. Dabei stehen vor allem die Kultur und Antworttendenz-Normen im Blickpunkt. Zusätzlich sollen aber auch – wie im vorherigen Kapitel erläutert –, bereits in vorherigen Studien gefundene Ergebnisse zum Zusammenhang von Antworttendenzen und Kultur repliziert werden. Um die entsprechenden Hypothesen diesbezüglich prüfen zu können, müssen daher Analysen herangezogen werden, die die entsprechenden Kennwerte der Länder miteinander in Beziehung setzen. Hinsichtlich dieser Frage handelt es sich somit um Berechnungen auf dem Länderniveau. Meßgrößen sind hierbei die entsprechenden Indikatoren für die länderspezifischen Ausprägungen der Konstrukte. Die Untersuchungseinheit ist folglich das Land und nicht das Individuum. Um aussagekräftige Ergebnisse zu erhalten, ist es dabei hilfreich, eine möglichst umfassende Anzahl von Ländern in den Analysen berücksichtigen zu können und daher auf eine möglichst große Stichprobe von Ländern zurückgreifen zu können. Der Umfang der Daten innerhalb eines Landes ist dabei weit weniger wichtig als die reine Anzahl von Kennwerten aus möglichst vielen verschiedenen Ländern. Im Rahmen der hier vorliegenden Studie wäre es jedoch aus Kapazitätsgründen unmöglich gewesen, Daten in einer Vielzahl von Ländern neu und eigenständig zu erheben. Daher soll zur Prüfung dieser ersten Hypothesen auf bereits veröffentlichte Daten zurückgegriffen werden und anhand von großen internationalen Untersuchungen beziehungsweise bereits in der Forschungsliteratur veröffentlichten Daten der entsprechende Zusammenhang untersucht werden. Bei diesem Teil der Analysen handelt es sich somit um länderbezogene Sekundäranalysen. Anhand solcher internationaler Datensätze lassen sich sekundäranalytisch dann mehrere der in den Hypothesen formulierten Zusammenhänge prüfen. Ein großer Teil der Analysen bezieht sich dabei auf die Messung der jeweils bivariaten korrelativen Zusammenhänge der entsprechenden Konstrukte. Darüber hinaus lassen sich jedoch ebenso anhand von Mediationsanalysen die entsprechenden Dreiecksbeziehungen der hier untersuchten Konstrukte überprüfen.

Diese Überprüfung der mediierenden Effekte stellt somit den zweiten wichtigen Teil der länderbezogenen Analysen dar.

Während diese Analysen auf dem Länderlevel speziell die Zusammenhänge der Konstrukte über die Länder hinweg prüfen, sollen in einem zweiten Schritt zusätzlich die Zusammenhänge individualpsychologischer Natur auf dem individuellen Level untersucht werden. Da es sich bei der Messung von länderbezogenen Arbeitszufriedenheits- und Antworttendenz-Daten nicht umgehen läßt, diese Daten durch die Befragung von Individuen zu erhalten, ist es daher ratsam, beide Level (d.h. Länderebene und Individualniveau) in den Analysen gemeinsam zu berücksichtigen. Da es sich bei der Analyse der Daten aus den verschiedenen Ländern zusätzlich um das methodische Phänomen einer hierarchischen Datenstruktur mit genisteten Daten handelt, bietet sich der Einsatz von Mehrebenenverfahren an, um die Einflußgrößen auf dem Länderebene von jenen auf dem Individualniveau trennen zu können. Unter genisteten Daten wird dabei verstanden, daß Individuen innerhalb einer Gruppe sich generell ähnlicher sind als Individuen zwischen den Gruppen (Nezlek, Schröder-Abé & Schütz, 2006). Dies ist beispielweise der Fall bei Mitgliedern eines Teams in einer Organisation, bei Schülern in Schulklassen oder eben bei Einwohnern eines Landes. Wenn dieser Fall einer Datenstruktur vorliegt, sind wichtige statistische Voraussetzungen (u.a. die Unabhängigkeit der Meßpunkte untereinander) verletzt, so daß herkömmliche statistische Verfahren (z.B. Varianz- oder Regressionsanalysen des Gesamtdatensatzes) in solch einem Fall nicht problemlos eingesetzt werden können (Hox, 1995). Daher wird in der hier vorliegenden Studie eine Mehrebenenanalyse eingesetzt, die nicht nur den länderspezifischen Einfluß von Antworttendenz-Normen sondern auch die Erklärungskraft von individuellem Kommunikationsverhalten im Sinne von Antworttendenzen mitberücksichtigt.

Den dritten Teil der methodischen Vorgehensweise stellt schließlich der Einsatz von experimentellen Verfahren dar, die die im Rahmen der länderbezogenen Sekundäranalysen sowie der Mehrebenenanalysen gefundenen Zusammenhänge empirisch bestätigen sollen. Dies ist vor allem begründet in den Nachteilen der sekundäranalytischen Verfahren hinsichtlich der internen Validität. Ein Einbezug aller theoretisch möglichen erklärenden Variablen ist weder methodisch noch statistisch bei diesen Verfahren möglich. Um dennoch eine bessere Kontrolle von möglichen Alternativ- und Störvariablen zu erreichen, soll der vermutete Zusammenhang daher ebenfalls experimentalpsychologisch abgesichert werden.

9 Länderbezogene Sekundäranalysen

Der erste und umfassendste Teil der empirischen Analysen beschäftigt sich mit der Frage, wie sich der Zusammenhang der in dieser Studie untersuchten Konstrukte auf Länderniveau darstellt. Dabei werden nacheinander die statistischen Zusammenhänge der drei zentralen Konstrukte, die im Fokus dieser Studie liegen – d.h. Arbeitszufriedenheitsniveau, Kultur und Antworttendenz-Norm –, überprüft. In einem ersten Schritt werden dabei die Zusammenhänge von Kultur und Antworttendenzen dargestellt und es wird somit versucht, gemäß den Überlegungen und Hypothesen in Kapitel 7.2 die bisherigen Ergebnisse der Forschung zu replizieren. Die in dieser Arbeit eigentlich zentrale Frage nach Erklärungsmöglichkeiten für unterschiedlich hohe Arbeitszufriedenheitsniveaus in unterschiedlichen Ländern rückt dann im zweiten Schritt ins Zentrum der Betrachtung, wenn einerseits die bisherigen Erkenntnisse zum Zusammenhang von Kultur und Arbeitszufriedenheitsniveaus anhand neuerer Daten überprüft werden und andererseits die explizit neuartige Fragestellung nach dem Einfluß einer Antworttendenz-Norm auf das Zufriedenheitsniveau beleuchtet wird. Den Abschluß der sekundäranalytischen Berechnungen bildet schließlich der Versuch, die drei verschiedenen Konstrukte in einen gemeinsamen Bezug zu stellen und die Kausalkette, in der die kulturelle Umgebung zu einem bestimmten Kommunikationsverhalten und dieses wiederum zu einem spezifischen Arbeitszufriedenheitsniveau führt, anhand von Mediationsanalysen auf ihre Belastbarkeit zu prüfen.

9.1 Die Stichproben

Um den Anforderungen an eine angemessene Länderstichprobe für interkulturelle Vergleiche gerecht zu werden, waren mehrere Kriterien bedeutsam, die bei der Suche nach einer geeigneten Datenquelle beachtet werden mußten. Grundsätzlich galt die Maßgabe, daß je mehr Länder ein Datensatz enthält, umso geeigneter ist er für die hier angedachten Untersuchungen. Eine hohe Anzahl von Ländern würde auch die Risiken einer nicht-zufälligen Auswahl von Ländern und somit die klassische Verzerrung vieler internationaler Studien zu westlichen Kulturen hin vermeiden. Des weiteren mußte die Frage beantwortet werden, inwieweit die bei interkulturellen Studien nicht vorhandene Möglichkeit der Randomisierung von Probanden innerhalb eines Landes entweder

durch eine Parallelisierung der Stichprobe oder durch Repräsentativität ausgeglichen werden sollte. Eine Parallelisierung ist jedoch zumeist nur in Stichproben möglich, die innerhalb eines einzigen spezifischen Kontextes, z.B. innerhalb eines Unternehmens, rekrutiert wurden. Da diese Form der Datenerhebung in der hier vorliegenden Studie aus organisatorischen und ressourcenbezogenen Gründen nicht realisiert werden konnte, war daher das Ziel, daß der optimale Datensatz aus repräsentativen Stichproben innerhalb der einzelnen Länder bestehen sollte. Als drittes und insgesamt wichtigstes Merkmal mußte der Datensatz schließlich zumindest eine, wenn nicht mehrere Fragen zur Arbeitszufriedenheit beinhalten. Ohne solch eine Frage wäre verständlicherweise die hier angestrebte Analyse überhaupt nicht möglich gewesen. Der Forscher ist dabei – wie es bei sekundäranalytischen Untersuchungen grundsätzlich immer der Fall ist – auf das angewiesen, was er überhaupt an Daten bekommen kann und auf welche Datensätze er schließlich Zugriff erhält.

Eine Recherche zur Verfügbarkeit und Zugänglichkeit großer internationaler Datensätze hatte als Ergebnis, daß die Datensätze des World-Values-Surveys (WVS) und des Eurobarometers (EB) gemäß den oben genannten Kriterien als geeignet ausgewählt wurden. Beide enthalten mindestens ein Frage zur Arbeitszufriedenheit, umfassen eine Vielzahl von weltweiten (im Falle vom WVS) bzw. europäischen (beim EB) Ländern und bestehen aus repräsentativen Stichproben innerhalb der jeweiligen Länder.

Der World-Values-Survey (WVS)

Der World-Values-Survey (WVS) ist ein internationales Befragungsprojekt, welches durch die World Values Survey Association (WVSA) mit Sitz in Stockholm, Schweden, betreut und organisiert wird. In insgesamt fünf Wellen wurden seit 1981 inzwischen Daten aus 80 Ländern zusammengetragen und insgesamt Antworten von ca. 257.000 Befragten generiert (WVSA, n.d.). Basierend auf den anfänglichen Erhebungen des European Values Surveys (EVS) wurden die Befragungen kontinuierlich auf die gesamte Welt ausgedehnt, so daß inzwischen laut eigenen Angaben der Organisation fast 88% der weltweiten Bevölkerung erfaßt ist (WVSA, n.d.). Alle Erhebungen werden persönlich im face-to-face Kontakt durch fragebogenbasierte Interviews durchgeführt, wobei die jeweiligen Stichproben innerhalb der Länder zufällig ausgewählt werden und somit ein repräsentatives Abbild der jeweiligen Bevölkerung ergeben. In jedem Land beträgt der Stichprobenumfang dabei ca. 1000-2000 Befragte. Die Durchführung der Interviews folgt strengen Richtlinien, die von der WVSA vorgegeben werden und unter

anderem vorschreiben, daß jede Befragung in der nationalen Sprache durchgeführt wird. Erhoben werden die Daten schließlich von renommierten Forschungsinstitutionen oder Universitäten in dem jeweiligen Land. Bevor die gesammelten Daten in den Gesamtdatensatz integriert werden, wird anhand strenger Datensäuberungsprozeduren und interner Konsistenzchecks sichergestellt, daß die Daten die notwendige Qualität besitzen. Insgesamt ist daher von einer hohen Datenqualität auszugehen, die notwendig für vertrauenswürdige Ergebnisse ist.

In der hier vorliegenden Studie wurde vom WVS der integrierte Datensatz aller Länder über alle Wellen hinweg in der Version 20060423 genutzt, der über die entsprechende Webseite der Organisation (<http://www.worldvaluessurvey.org/>) bezogen werden kann. Der standardisierte Fragebogen, der in allen Ländern in gleicher Weise eingesetzt wurde, enthält neben verschiedenen allgemeinen Fragen zum Arbeitsumfeld eine konkrete Frage zur Arbeitszufriedenheit („Alles in allem, wie zufrieden oder unzufrieden sind Sie mit Ihrer Arbeit?“), die auf einer Skala von 1 = unzufrieden bis 10 = zufrieden erfaßt wurde. Für die zentralen Analysen zum Einfluß auf das Arbeitszufriedenheitsniveau wurden alle Länder aufgenommen, in denen diese Frage gestellt wurde. Insgesamt konnten somit für diese Analysen Daten aus 48 Ländern genutzt werden. Für die Berechnungen, bei denen der Zusammenhang mit Zustimmungstendenzen geprüft wurde, wurden alle Länder aufgenommen, in denen sich ein entsprechender Index überhaupt erstellen ließ, so daß bezüglich dieser Fragestellungen insgesamt bis zu 82 Länder in die Berechnungen mit eingehen konnten.

Das Eurobarometer (EB)

Das Eurobarometer (EB) ist ein Befragungsprojekt, welches seit den 70er Jahren regelmäßig in den Mitgliedsstaaten der Europäischen Union durchgeführt wird. Neben einem festen Fragenkomplex, der sich regelmäßig mit den Einstellungen zur Europäischen Union beschäftigt, werden jeweils im Wechsel verschiedene aktuelle politische, wirtschaftliche oder soziale Themenstellungen neu aufgenommen und ergänzen die Standardfragen. Mindestens zweimal im Jahr werden jeweils neue Stichproben in den einzelnen Ländern rekrutiert und in persönlichen face-to-face-Interviews nach ihrer Meinung zu den genannten Themen gefragt. Die Stichproben werden jeweils nach einem Zufallsverfahren zusammengestellt, so daß auch hier auf repräsentative Stichproben in den einzelnen Ländern zurückgegriffen wird. Die Durchführung der Interviews ist durch klare Vorgaben geregelt und wird verantwortlich

von renommierten Forschungsinstituten geleitet. Von einer hohen Datenqualität ist daher auch beim Eurobarometer auszugehen.

Der in dieser Studie eingesetzte Datensatz stammt aus dem Eurobarometer 56.1, welches im Jahr 2001 durchgeführt wurde. In dieser Welle des Eurobarometers wurden auch Fragen zum Arbeitsleben mit aufgenommen, so daß im Fragebogen 56.1 ebenfalls eine Frage zur Gesamtzufriedenheit mit der Arbeit enthalten war: „Alles in allem gesehen, wie zufrieden sind Sie mit Ihrer Arbeit?“. Da diese Frage in allen in dieser Welle teilnehmenden Nationen gestellt wurde, konnten insgesamt Daten aus 18 Ländern genutzt werden.

9.2 Die Meßgrößen

Die Werte der Meßgrößen, die in der hier vorliegenden Studie in den Analysen eingesetzt wurden, sind teilweise der bereits publizierten, empirischen Forschungsliteratur entnommen, wurden aber auch in Teilen anhand vorhandener und veröffentlichter Daten eigenständig neu berechnet.

Das Arbeitszufriedenheitsniveau

In der Befragung des World-Values-Survey wurde innerhalb eines thematischen Blocks zum Arbeitsumfeld eine spezielle Frage nach der Gesamtzufriedenheit mit der Arbeit gestellt. Diese Frage, die im englischen Original-Wortlaut „Overall, how satisfied or dissatisfied are you with your job?“ lautete und in die Sprachen der jeweiligen Länder übersetzt wurde, diente als Meßwert zur Ermittlung des Arbeitszufriedenheitsniveau. Diese Frage wurde im WVS nur denjenigen Personen gestellt, die zum Zeitpunkt der Befragung über eine Arbeitsstelle verfügten. Da innerhalb der Befragung des WVS keine Item-Batterie zur Erfassung der Arbeitszufriedenheit verwendet wurde, die mehrere Facetten der Arbeitszufriedenheit erfaßt hätte, diente dieses Einzelitem, welches das globale Gesamturteil der Befragten erfaßt, zur Operationalisierung des Arbeitszufriedenheitsniveaus. Dieses Globalurteil ist jedoch keineswegs einer Erfassung anhand einer Item-Batterie unterlegen (vgl. die Diskussion in Kapitel 2.2) und daher ein ebenso gut geeigneter Indikator für das Arbeitszufriedenheitsniveau. Um das jeweilige Niveau eines Landes abzubilden, wurden die Einzelaussagen zur persönlichen Arbeitszufriedenheit der Befragten auf Länderebene gemittelt. Die ursprüngliche Gewichtung der Daten innerhalb der Länder, so wie sie im WVS zum Ausgleich

unterschiedlicher Rekrutierungsprozesse in den einzelnen Ländern angelegt war, wurde dabei beibehalten. Das Arbeitszufriedenheitsniveau stellte somit die mittlere Zufriedenheit aller in einem Land Befragten dar, die zum Zeitpunkt der Befragung zur arbeitenden Bevölkerung zählten. Dabei ist wichtig zu erwähnen, daß in zahlreichen Ländern diese Frage in mehreren Wellen gestellt wurde (z.B. in Deutschland, Spanien, Großbritannien, Belgien, Dänemark, Island, Irland, Italien, Niederlande, etc.: 1981, 1990, 1999; in den USA, in Kanada, Norwegen, etc.: 1982; 1990; in Lettland, Litauen, Portugal, etc.: 1990, 1999). Zur Erfassung des Arbeitszufriedenheitsniveaus wurde letztendlich der Gesamtmittelwert aus allen Wellen gebildet. Insgesamt handelte es sich somit bei dem länderbezogenen Arbeitszufriedenheitswert um das durchschnittliche Zufriedenheitsniveau, welches in der Zeit von 1981 bis 2007 anhand des WVS ermittelt wurde. Der Vorteil dieses durchschnittlichen Arbeitszufriedenheitsniveaus gegenüber einer einmaligen Erfassung in nur einem Befragungsjahr besteht vor allem darin, daß diese Maßzahl unabhängiger von zyklischen Schwankungen des wirtschaftlichen Umfeldes und von situativen Bedingungen in dem jeweiligen Land zum jeweiligen Befragungszeitpunkt ist.

Im Eurobarometer 56.1 wurde ebenfalls eine allgemeine Frage zur Arbeitszufriedenheit gestellt, deren englischer Originalwortlaut „All in all, how satisfied would you say you are with your job?“ lautete und anhand einer Skala von „1 = completely dissatisfied“ bis „7 = completely satisfied“ beantwortet werden sollte. Diese Frage wurde im Befragungsprojekt anhand der klassischen Standardprozedur des Hin- und Rückübersetzens in die jeweilige Landessprache transferiert und ebenfalls nur Personen gestellt, die zum Zeitpunkt der Befragung berufstätig waren. Sie diente der Erfassung des Arbeitszufriedenheitsniveaus in dieser Befragung und bildete somit hier den entsprechenden Wert in den Staaten der Europäischen Union ab. Auch bei dieser Datenquelle wurden die Einzelaussagen zur persönlichen Arbeitszufriedenheit anhand des gewichteten Datensatzes auf Länderebene gemittelt. Im Gegensatz zum WVS wurde hier jedoch nur ein Meßzeitpunkt im Herbst des Jahres 2001 erfaßt, so daß der gemittelte Wert das in den Ländern berichtete Zufriedenheitsniveau in jenem Jahr wiedergibt.

Die Kulturwerte der einzelnen Länder

Um angemessene Indikatoren für die Ausprägung der Kultur in den einzelnen Ländern zu erhalten, wurde auf bereits veröffentlichte Daten aus renommierten Kulturstudien

zurückgegriffen. So wurden jeweils die Kennwerte von Hofstede (2001, 2005), von Schwartz (1994) und aus der GLOBE-Studie (House et al., 2004) in die jeweiligen Analysen mit einbezogen und dienten somit der Operationalisierung der Kulturdimensionen. Da ein besonderer Fokus auf der Dimensionen Individualismus/Kollektivismus in der hier vorliegenden Studie liegt, wurde zusätzlich ein weiteres Maß für diese Dimensionen verwendet, welches ein Rating der Kultur von Ländern von Harry Triandis – einem der erfahrensten Kulturforschern und Experten für diese Dimension – darstellte und einer renommierten Studie von Diener, Diener und Diener (1995) entnommen wurde.

Die Antworttendenz-Normen

Um die Antworttendenz-Normen zu erfassen, wurde eine zweigeteilte Vorgehensweise gewählt. Auf der einen Seite wurden auch hier bereits in der Literatur veröffentlichte Kennwerte mit einbezogen, die die Höhe einer Antworttendenz in den jeweiligen Ländern repräsentieren. So berichtete auch Hofstede (2001) bereits ein Maß für eine Zustimmungstendenz in den von ihm untersuchten Ländern, welches er anhand von Fragen zur Wichtigkeit von Arbeitszielen in seiner Studie erzeugt hat. Dazu errechnete er für jedes Land den Mittelwert über 14 insgesamt sehr heterogene Arbeitsziele und über alle Befragten und interpretierte ihn als Maß von zustimmendem Antwortverhalten, da es gemäß seiner Argumentation ungewöhnlich wäre, wenn alle Arbeitsziele als gleich wichtig bewertet werden würden. Auch Harzing (2006) berichtet in ihrer Studie die präzisen Kennwerte zu Antworttendenz-Normen, die jedes Land in dem bei ihr eingesetzten Fragebogen zeigte. Dabei berichtet sie sowohl ein Maß für die reine Zustimmungstendenz (ARS) als auch ein Maß für die Netto-Zustimmungstendenz (NARS), wovon letzteres in die hier durchgeführten Analysen einbezogen wurde. Darüber hinaus werden ebenfalls die von ihr berichteten Kennwerte für positive und negative Extremtendenzen sowie die Werte für Mittentendenzen in die Analysen mit einbezogen. Um aus diesen Angaben ein Gesamtmaß für Extremtendenzen der einzelnen Länder zu erhalten, wurden die Kennwerte von Harzing (2006) für positive und negative Extremtendenzen in dieser Studie zu einem einzigen Gesamtwert für Extremtendenzen in jedem Land gemittelt.

Als zweite Vorgehensweise wurden zusätzlich neue Maße der Antworttendenz anhand der Fragen des World-Values-Surveys errechnet. Diese Alternative wurde zusätzlich gewählt, um jeweils einzelne Indikatoren für verschiedene Skalenlabel zu erhalten; eine

Möglichkeit, die anhand der Daten von Harzing (2006) nicht gegeben war und kritisiert werden konnte. Es wurden daher getrennte Indikatoren für Skalen mit einem „wichtig-unwichtig“-Label und mit einem „stimme zu-stimme nicht zu“-Label erstellt. Bei letzteren wurde zusätzlich noch zwischen 4er- und 5er-Likert-Skalen unterschieden. Zur Berechnung dieser drei Indikatoren wurden dann der gesamte Fragebogen des World-Values-Surveys herangezogen und alle Fragen jeweils innerhalb dieser drei Ausprägungsformen getrennt zusammengefaßt. Dies bedeutet, daß jeweils Fragen mit einer 4er-Likertskala mit Zustimmungslabel, einer 5er-Likertskala mit Zustimmungslabel und einer 5er-Likert-Skala mit Wichtigkeitslabel zusammengefaßt wurden. 4er-Likert-Skalen mit Wichtigkeitslabel sowie andere Likert-Skalen mit Zustimmungslabel mit mehr oder auch weniger Stufen sind im WVS nur selten vorhanden, so daß dafür keine gesonderten Indizes errechnet wurden. Generell wurden alle Fragen bzw. Items der gleichen Kategorie zusammengefaßt, völlig gleichgültig, welche thematischen Inhalte in den Items angesprochen wurden, denn je heterogener die Iteminhalte sind, umso valider läßt sich der gebildete Indikator als Index einer Antworttendenz verstehen. In den jeweiligen Indikatoren sind somit Fragen zur Familie, zu Beziehungsaspekten, zum ökonomischen und politischen System, zum Umgang mit Gerechtigkeit, zur Religiösität, zur Bestrafung von Delikten, etc. zusammengefaßt; insgesamt somit verschiedene Fragenkomplexe, die mehrheitlich als unabhängig voneinander anzusehen sind.

Die einzelnen Indizes wurden gemäß einer der klassischen Vorgehensweisen in der interkulturellen Forschung, die auch Harzing (2006) wählte, gebildet. So wurde für den ZustimmungsindeX (ARS) der 5er-Likert-Skalen jeweils die Anzahl der Werte 1 und 2, die die Zustimmungseite oder die Wichtigkeitsseite der Skala ausdrückten, für jede einzelne Person aufsummiert und durch die entsprechende Itemanzahl geteilt. Beim AblehnungsindeX (DARS) wurden dementsprechend die Anzahl der Werte 4 und 5 errechnet und anhand der Gesamtitemzahl relativiert. Beide Indizes können somit theoretisch zwischen 0 und 1 schwanken. Die entsprechenden Indizes dieser Art des Antwortverhaltens bei 4er-Likert-Skalen wurden analog errechnet, nur daß hier für die Ablehnungstendenz die Anzahl der Werte 3 und 4 erfaßt wurde.

Insgesamt gingen 39 Items in den Index ‚Zustimmung 5er-Skala‘, 39 Items in den Index ‚Zustimmung 4er Skala‘ und 29 Items in den Index ‚Wichtigkeit 5er-Skala‘ ein. Die letztendlich benötigten länderbezogenen Kennwerte der Zustimmungs-/Ablehnungstendenz-Norm wurden dann durch den Mittelwert der Individualwerte

repräsentiert. Die Netto-Zustimmungstendenz (NARS) wurde schließlich auf Länderebene durch das Verhältnis von ARS und DARS bestimmt, indem der DARS-Index vom ARS-Index subtrahiert wurde. Die Netto-Zustimmungstendenz-Norm in den einzelnen Ländern kann somit zwischen -1 und 1 variieren.

Die Extremtendenz wurde nur für die 5er-Skala mit Zustimmungslabel errechnet und durch die Anzahl der Werte in den beiden äußeren Antwortkategorien (d.h. 1 und 5) gebildet, relativiert an der Gesamtzahl der Items. Mittentendenzen wurden ebenfalls nur für diese 5er-Skala errechnet, indem die Anzahl des mittleren Werts 3 anhand der Gesamtanzahl aller Items relativiert wurde.

Ein Maß der Sozialen Erwünschtheit konnte nicht neu anhand des WVS gebildet werden, da keine dafür geeigneten Items im Datensatz vorhanden waren. Stattdessen wurde hier auf bereits veröffentlichte Länder-Kennwerte von Bernardi (2006) zurückgegriffen.

Sozioökonomische Indikatoren

Die Indikatoren, die die ökonomische Situation in einem Land abbilden sollten und als Kontrollfaktoren mit in die Analyse gingen, wurden renommierten internationalen Datenquellen entnommen. Betrachtet wurden dabei einerseits der Human-Development-Index aus dem Jahre 2000 (United Nations, 2008) und das Bruttoinlandsprodukt (BIP) aus dem Jahre 2004 (United Nations, 2006a), welche zum Zeitpunkt der Untersuchung auf den Webseiten der United Nations verfügbar waren. Zwar liegen diese BIP-Daten z.T. zeitlich gesehen später als die Befragungen zum Arbeitszufriedenheitsniveau. Dies ist jedoch statistisch gesehen wenig problematisch, da in der Regel die ökonomischen Kennwerte aus verschiedenen Jahren sehr hoch miteinander korreliert sind. Dies zeigen auch eigene Analysen, die mit weiteren BIP-Daten aus dem Jahre 2007 (United Nations, 2009) sowie mit Daten des Bruttosozialprodukts (BSP) aus dem Jahre 1997 durchgeführt wurden, die einer Untersuchung von van de Vliert (2006) entnommen wurden. Das BIP aus dem Jahr 2004 ist sowohl mit den Daten aus 2007 ($r = .98$, $p < .001$) als auch mit den Daten aus 1997 ($r = .95$, $p < .001$) enorm hoch korreliert.

Zusätzlich wurde ebenfalls die Arbeitslosenquote der Länder (United Nations, 2006b) einbezogen, die in der Quelle der Vereinten Nationen die zum Zeitpunkt der Untersuchung aktuellste Zusammenstellung dieser Daten darstellte und insgesamt aus den jeweils neuesten Werten aus den jeweiligen Ländern bestand. Je Land stammten diese Kennwerte jedoch insgesamt aus einem der Jahre von Anfang der 90er Jahre bis

2004. Des Weiteren wurde noch ein Quality-of-Life-Indikator aufgenommen, der einer Studie von Diener (1995) entnommen wurde und das Niveau der Lebensqualität der Länder erfassen sollte.

In den Mehrebenenanalysen wurden schließlich noch Alters- und Geschlechtsangaben sowie Indikatoren über den individuellen Bildungsstand und die persönliche Einkommenskategorie aus dem WVS mit einbezogen.

9.3 Ergebnisse der bivariaten Analysen

9.3.1 Deskriptive Ergebnisse zum Arbeitszufriedenheitsniveau

Die Ergebnisse des World-Values-Surveys zeigen zum Teil deutliche Unterschiede zwischen den einzelnen Zufriedenheitsniveaus in den jeweiligen Ländern (siehe Abbildung 13). Das höchste Niveau der Arbeitszufriedenheit wurde demnach in der Schweiz berichtet, gefolgt von Malta und Dänemark. Die niedrigsten Niveaus existierten gemäß diesen Daten in Weißrußland, der Ukraine und der Türkei. Die Werte von Deutschland-West und -Ost befanden sich ungefähr in der Mitte der Rangfolge. Insgesamt zeigte sich, daß vor allem ost-europäische Länder sich am unteren Ende der Rangfolge wiederfinden ließen, während nord-europäische Länder sich eher im oberen Drittel befanden. Eine Varianzanalyse der Mittelwertsunterschiede auf Basis der Individualdaten des gesamten WVS ergab ein hoch signifikantes Ergebnis mit einer mittleren Effektstärke ($\eta^2 = .064$; $p < .001$). Die Arbeitszufriedenheitsniveaus unterschieden sich somit deutlich voneinander. Die Zugehörigkeit zu einem bestimmten Land erklärte insgesamt 6,4% der Varianz der Zufriedenheitsurteile.

Wichtig ist, darauf hinzuweisen, daß bei allen Ländern der Gesamtmittelwert sich immer über dem Mittelwert der Likert-Skala befindet, anhand derer die Zufriedenheitswerte ermittelt wurden (Skala von 1 = unzufrieden bis 10 = zufrieden). Insgesamt wird somit in allen erfaßten Ländern im Mittel durchaus eine Zufriedenheit mit der Arbeit berichtet, jedoch ergeben die Befragungsergebnisse in manchen Ländern ein extrem hohes Zufriedenheitsniveau, während in anderen Ländern dieses Niveau eher einen gemäßigten Zufriedenheitswert auf der Skala darstellt. Die bisher schon bekannte Verschiebung von individuellen Arbeitszufriedenheitswerten hin zum positiven Ende der Skala zeigt sich folglich auch anhand der Ländermittelwerte der WVS-Daten.

Die Ergebnisse des Eurobarometers (vgl. Abbildung 14) zeigen, daß auch hier Dänemark im vorderen Bereich lag und sogar das höchste Zufriedenheitsniveau besaß. Am unteren Ende der Rangfolge fanden sich hier eher süd-europäische Länder. Die Zufriedenheitsniveaus von Deutschland-West und -Ost lagen auch hier im mittleren Bereich der Rangfolge. Eine Varianzanalyse der Mittelwertsunterschiede auf Basis der Individualdaten des gesamten Eurobarometerdatensatzes ergab auch hier hoch signifikante Ergebnisse mit einer mittleren Effektstärke ($\eta^2 = .046$; $p < .001$). Auch die Zufriedenheitsniveaus im Eurobarometer unterschieden sich somit deutlich voneinander. Hier erklärte die Zugehörigkeit zu einem Land insgesamt 4,6% der individuellen Zufriedenheitsunterschiede.

Betrachtet man beide Studien zusammen, zeigt sich, daß trotz unterschiedlicher Erhebungen die Rangfolge der Zufriedenheitsniveaus der Länder relativ stabil ist. Insgesamt 17 Länder ließen sich in beiden Erhebungen finden. Deren Zufriedenheitsniveaus korrelierten mit Spearman-Rho = .67 ($p < .01$) (vgl. Abbildung 12).

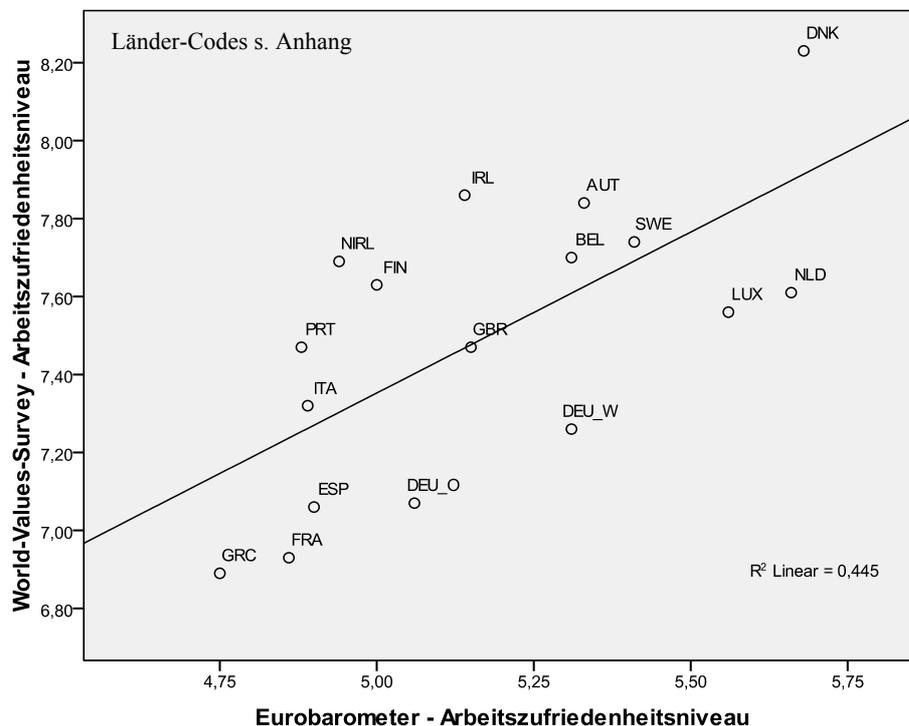


Abbildung 12: Die Arbeitszufriedenheitsniveaus vom WVS und EB im Vergleich

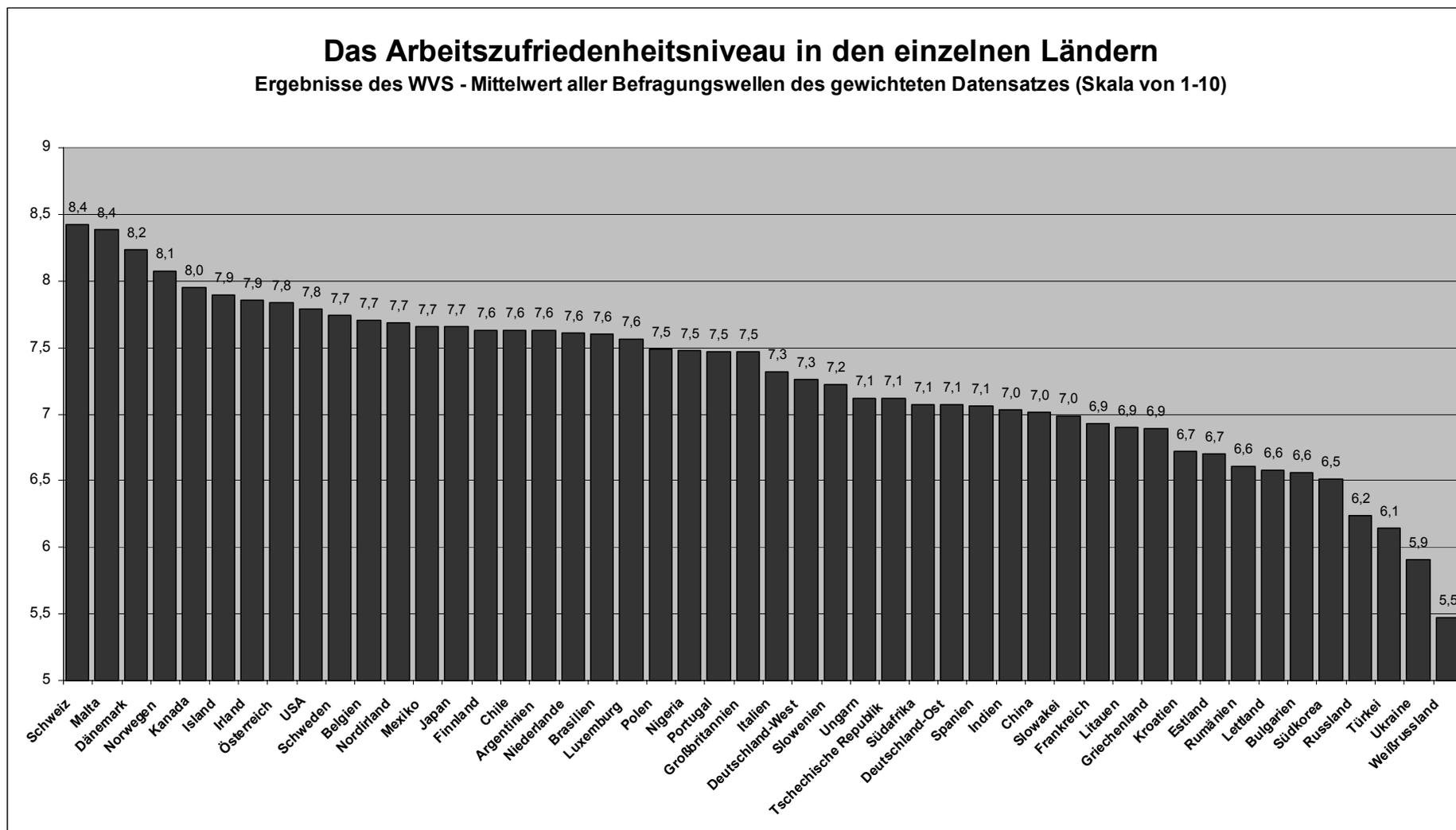


Abbildung 13: Das Arbeitszufriedenheitsniveau gemäß den Daten des World-Values-Surveys

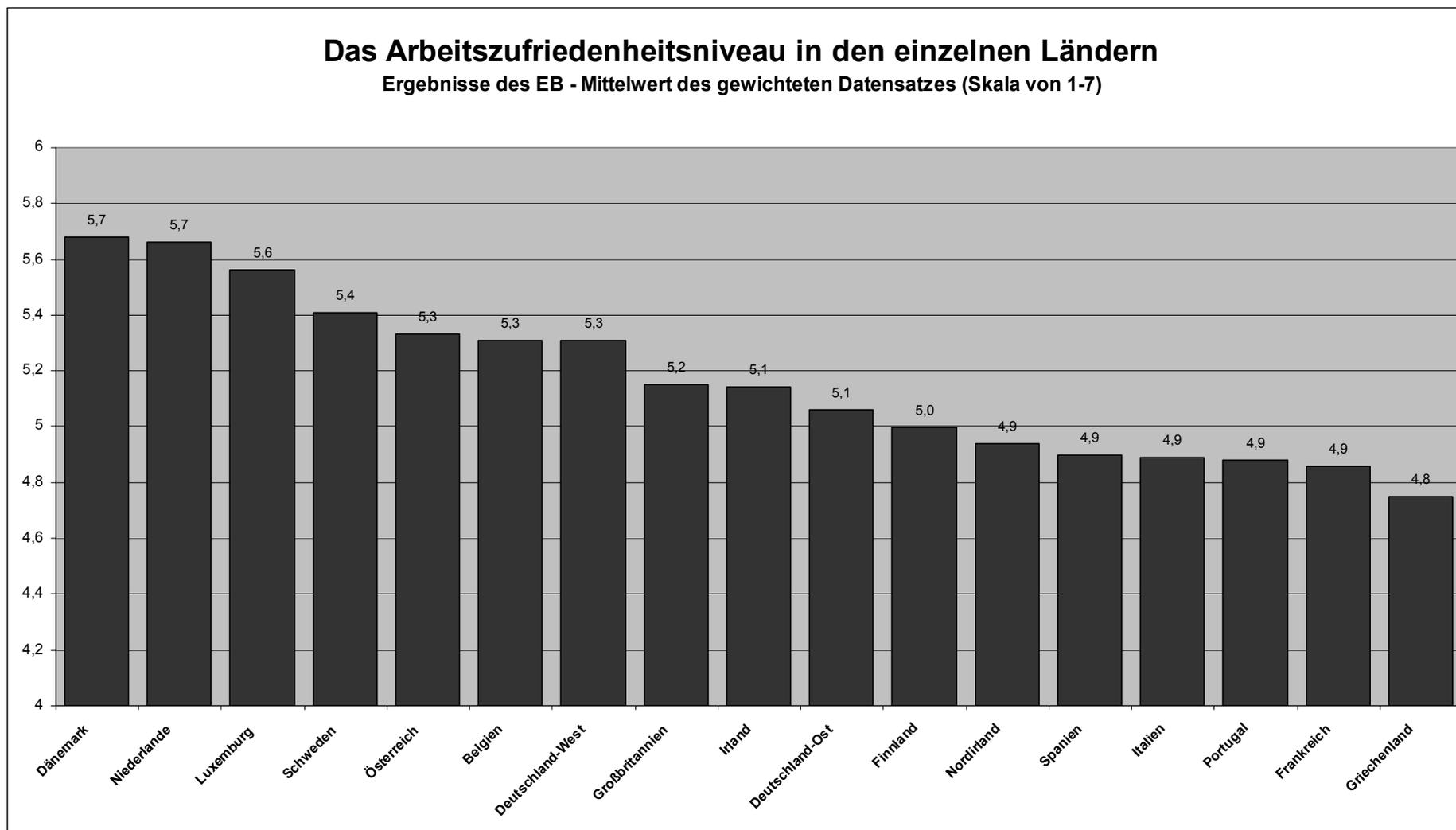


Abbildung 14: Das Arbeitszufriedenheitsniveau gemäß den Daten des Eurobarometers

9.3.2 Der Zusammenhang von Kultur und Arbeitszufriedenheitsniveau

Die Darstellung der deskriptiven Ergebnisse zum Arbeitszufriedenheitsniveau im WVS und EB zeigt, daß die Niveaus sich zwischen den Ländern signifikant unterscheiden. Sie bietet jedoch noch keine Erklärung für die Unterschiede an. Im nächsten Schritt wurde daher der Zusammenhang von kulturellen Werten in einem Land und dem jeweiligen Arbeitszufriedenheitsniveau untersucht. Gemäß den Hypothesen 1 und 2 lag dabei ein besonderer Fokus auf dem Zusammenhang mit den Dimensionen Individualismus/Kollektivismus, Machtdistanz und Unsicherheitsvermeidung.

Die Ergebnisse der Korrelationsanalysen zeigen in der Tat zahlreiche Zusammenhänge mit der Kultur eines Landes (vgl. Tabelle 5), jedoch unterschieden sich die Zusammenhänge teilweise je nach kulturellem Rahmenmodell. So resultierten hinsichtlich des Modells von Schwartz mehrheitlich nur Zusammenhänge mit dem Arbeitszufriedenheitsniveau im WVS. Bezüglich der Modelle von Hofstede und GLOBE sowie der Werte von Triandis ergaben sich jedoch deutliche Zusammenhänge mit beiden Datensätzen.

Bei der Betrachtung der Ergebnisse in Tabelle 5 sollten die Zusammenhänge, die mit dem WVS ermittelt wurden, als bedeutsamer angesehen werden, da sie auf einer größeren Länderstichprobe beruhen. Die Ergebnisse hinsichtlich des EB sollten hingegen nur als Bestätigung der mit dem WVS ermittelten Ergebnisse interpretiert werden und nicht zu eigenen Schlußfolgerungen führen, da die Länderstichprobe sich nur auf europäische Länder bezog und insgesamt sehr viel kleiner war.

Vergleicht man die einzelnen kulturellen Modelle nacheinander, zeigt sich zuerst hinsichtlich der Dimensionen von Hofstede, daß das Arbeitszufriedenheitsniveau signifikant höher ist in Ländern mit hohem Individualismus ($r = .46$, $p < .01$ und $r = .48$, $p < .10$), jedoch insgesamt geringer ist in Ländern mit hoher Machtdistanz ($r = -.58$, $p < .01$ und $r = -.61$, $p < .05$) und hoher Unsicherheitsvermeidung ($r = -.32$, $p < .05$ und $r = -.62$, $p < .05$). Diese Korrelationen stellen somit ein erstes Ergebnis dar, welches im Einklang mit den Hypothesen 1 und 2 steht und die Ergebnisse früherer Forschung repliziert.

Die Ergebnisse bezüglich des Ratings von Triandis hinsichtlich der Ausprägung von Individualismus/Kollektivismus bestätigten ebenfalls die Ergebnisse, die sich mit der

entsprechenden Hofstede-Dimension zeigten. Auch die Werte von Triandis korrelierten hoch positiv mit dem Arbeitszufriedenheitsniveau ($r = .51, p < .01$ und $r = .63, p < .01$). Somit gilt auch für diesen Zusammenhang, daß das Arbeitszufriedenheitsniveau höher in Ländern war, die Triandis als hoch individualistisch einstuft.

Tabelle 5: Zusammenhang des Arbeitszufriedenheitsniveaus mit kulturellen Dimensionen auf Länderebene

	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem Eurobarometer - EB)
Kulturdimensionen nach Hofstede		
Individualismus	.46*** (41)	.48* (15)
Machtdistanz	-.58*** (41)	-.61** (15)
Unsicherheitsvermeidung	-.32** (41)	-.62** (15)
Maskulinität	-.01 (41)	-.36 (15)
Langzeitorientierung	-.23 (27)	.49* (13)
Kulturdimension nach Triandis		
Individualismus	.51*** (34)	.63*** (17)
Kulturdimensionen nach Schwartz		
Konservatismus	-.43** (23)	-.01 (10)
Affektive Autonomie	.29 (23)	.02 (10)
Intellektuelle Autonomie	.42** (23)	.03 (10)
Hierarchie	-.36* (23)	.11 (10)
Beherrschung	.12 (23)	-.33 (10)
Gleichheitsdenken	.14 (23)	-.09 (10)
Harmonie	-.12 (23)	-.59* (10)
Kulturdimensionen nach GLOBE: Praktiken		
Machtdistanz	-.63*** (31)	-.84*** (14)
Unsicherheitsvermeidung	.55*** (31)	.70*** (14)
Institut. Kollektivismus	.15 (31)	.69*** (14)
InGroup Kollektivismus	-.59*** (31)	-.77*** (14)
Gleichheitsdenken	.18 (31)	.29 (14)
Bestimmtheit	.09 (31)	.50* (14)
Humanitäre Orientierung	.37** (31)	.45 (14)
Zukunftsorientierung	.57*** (31)	.87*** (14)
Leistungsorientierung	.42** (31)	.57** (14)
Kulturdimensionen nach GLOBE: Werte		
Machtdistanz	.11 (31)	.30 (14)
Unsicherheitsvermeidung	-.56*** (31)	-.82*** (14)
Institut. Kollektivismus	-.06 (31)	-.62*** (14)
InGroup Kollektivismus	-.01 (31)	-.15 (14)
Gleichheitsdenken	.53*** (31)	.55** (14)
Bestimmtheit	.14 (31)	-.31 (14)
Humanitäre Orientierung	.10 (31)	.02 (14)
Zukunftsorientierung	-.45** (31)	-.68*** (14)
Leistungsorientierung	.40** (31)	-.35 (14)

*** $p < .01$; ** $p < .05$; * $p < .10$; in Klammern jeweils die Anzahl der Fälle/Länder

Eine weitere Bestätigung dieses Zusammenhangs ergab sich auch in Teilen bei den Schwartz-Dimensionen. Wenn auch insgesamt keine wirklich nennenswerten Zusammenhänge seiner Dimensionen mit dem Zufriedenheitsniveau im Eurobarometer aufgetreten sind, zeigen jedoch die Ergebnisse hinsichtlich der WVS-Daten, daß das Arbeitszufriedenheitsniveau insgesamt höher war bei vorherrschender Intellektueller Autonomie ($r = .42$, $p < .05$) und im Gegensatz dazu geringer war, je wichtiger Konservatismus/Eingebettetsein angesehen wird ($r = -.43$, $p < .05$). Das Arbeitszufriedenheitsniveau war somit auch hier tendenziell höher in Kulturen, in denen das Individuum eine größere Rolle als die Gruppe spielt.

Betrachtet man die Zusammenhänge, die sich mit den Dimensionen der GLOBE-Studie ergaben, zeigten sich ebenfalls überwiegend Ergebnisse, die in die gleiche Richtung gingen wie die Ergebnisse mit den Dimensionen von Hofstede, Schwartz und Triandis. Jedoch ist das Bild insgesamt hier etwas heterogener und unterscheidet sich etwas, je nachdem ob die Praktiken- oder die Werte-Dimensionen betrachtet werden. Einerseits zeigte sich aber auch hier erst einmal, daß das Arbeitszufriedenheitsniveau geringer war in Ländern, in denen ein hoher InGroup-Kollektivismus vorhanden ist ($r = -.59$, $p < .01$ und $r = -.77$, $p < .01$). Somit bestätige auch dies die Ergebnisse, die sich schon mit den inhaltlich ähnlichen Dimensionen der anderen Kulturmodelle zeigten. Insgesamt läßt sich somit festhalten, daß scheinbar gegenüber einem Interviewer im Mittel eine höhere Arbeitszufriedenheit in Kulturen berichtet wird, in denen die Menschen sich mehr über ihre eigene Person definieren. Im Gegensatz dazu äußern sich somit die Befragten im Mittel tendenziell geringer zufrieden in Kulturen, in denen die Gruppe und ihr Zusammenhalt im Vordergrund steht (vgl. zur graphischen Verdeutlichung beispielhaft das Streudiagramm hinsichtlich der Individualismus-Dimension von Hofstede und des Arbeitszufriedenheitsniveaus im WVS in Abbildung 15). Die hier gefundenen Ergebnisse stützen somit deutlich den in Hypothese 1 angenommenen Zusammenhang zwischen Individualismus/Kollektivismus und dem Arbeitszufriedenheitsniveau.

Schaut man sich die Zusammenhänge mit den weiteren Kulturdimensionen an, so bestätigen auch die Korrelationen hinsichtlich der Machtdistanz der GLOBE-Studie die Ergebnisse, die mit der Hofstede-Dimension gemessen wurden. Das Arbeitszufriedenheitsniveau war demnach auch hier tendenziell geringer in Kulturen, in denen einen hohe Distanz zwischen Menschen mit großer Macht und jenen mit geringer Macht herrscht ($r = -.63$, $p < .01$ und $r = -.84$, $p < .01$). Bezüglich dieser Machtthematik zeige sich auch eine weitere Bestätigung, wenn man die Dimensionen des Schwartz-

Modells betrachtet. Auch dort bestand eine negative Korrelation zwischen der Hierarchie-Dimension und dem im WVS gemessenen Arbeitszufriedenheitsniveau ($r = -.36, p < .10$). Insgesamt bestätigen die Ergebnisse somit die Annahmen hinsichtlich Machtdistanz, die in Hypothese 2 zum Ausdruck kamen.

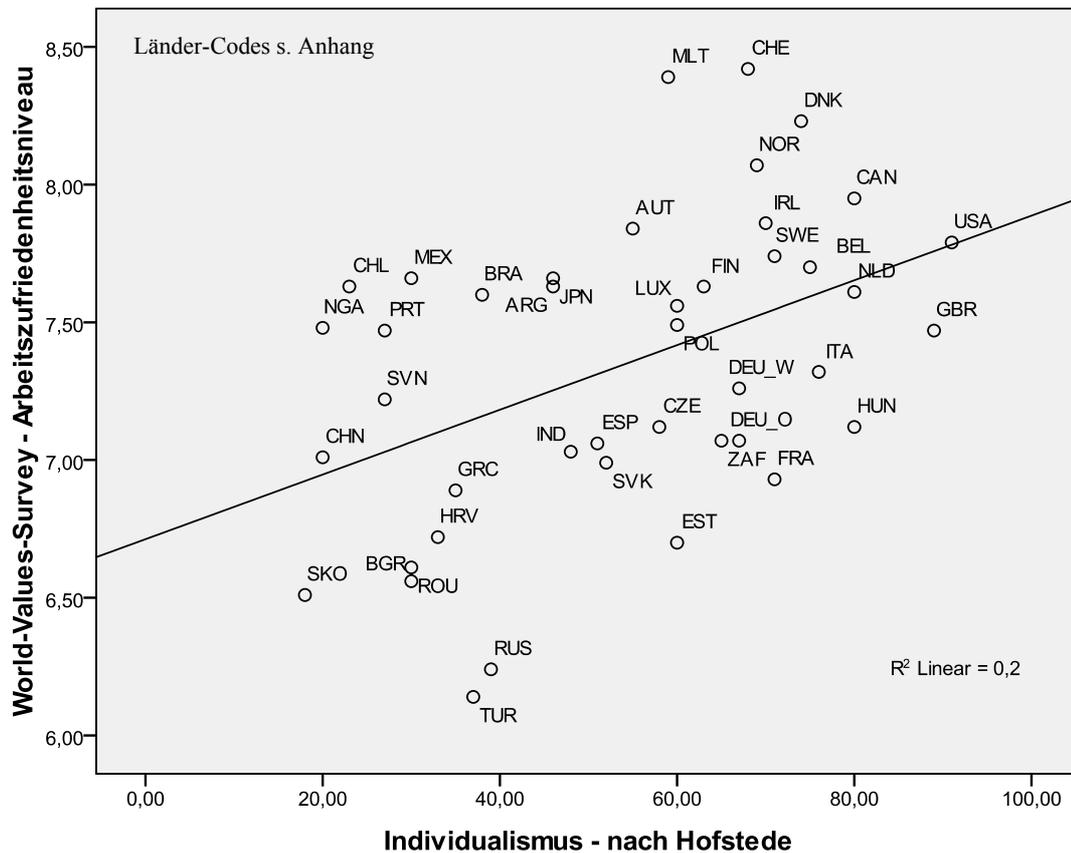


Abbildung 15: Zusammenhang des Arbeitszufriedenheitsniveaus im WVS und der Individualismus-Dimension (nach Hofstede)

Bezüglich der Unsicherheitsvermeidungs-Dimension des GLOBE-Modells schienen sich – zumindest auf den ersten Blick – unerwartete Zusammenhänge zu ergeben. Diesbezüglich existierte erstaunlicherweise eine positive Korrelation mit dem Arbeitszufriedenheitsniveau, wenn man die Praktiken in den Ländern betrachtete ($r = .55, p < .01$ und $r = .70, p < 01$). Demnach schien es so, als würde hier gelten, daß das Arbeitszufriedenheitsniveau höher ist in Ländern, in denen eine hohe Unsicherheitsvermeidung vorherrschend ist. Dies würde dann in einem Widerspruch zu den Ergebnissen mit den Hofstede-Dimensionen stehen. Führt man sich jedoch erneut vor Augen, daß bereits House et al. (2004) zeigen konnten, daß ihre Unsicherheitsvermeidungs-Dimensionen augenscheinlich etwas anderes messen als die

Unsicherheitsvermeidungs-Dimension von Hofstede (vgl. Kapitel 3.1), passen auch die hier gefundenen Ergebnisse wieder ins Bild. Die Unsicherheitsvermeidung, so wie sie von Hofstede gemessen wurde, korreliert nämlich insgesamt negativ mit der tatsächlich vorhandenen Unsicherheitsvermeidung (Praktiken) in einem Land – so wie es im GLOBE-Modell gemessen wurde – und wiederum positiv mit dem Wunsch nach mehr Unsicherheitsvermeidung (Werte). Somit war durchaus zu erwarten, daß sich eine positive Korrelation zwischen der Unsicherheitsvermeidung (Praktiken) des GLOBE-Modells und dem Arbeitszufriedenheitsniveau ergab. Die Erfassung der Unsicherheitsvermeidung als Werte-Dimension im GLOBE-Modell, die in der Tat das Gleiche zu messen scheint wie Hofstedes Dimension, korrelierte dann wie erwartet auch wieder negativ mit dem Arbeitszufriedenheitsniveau ($r = -.56$, $p < .01$ und $r = -.82$, $p < .01$) Die Ergebnisse mit den Dimensionen des GLOBE-Modells deuten somit insgesamt in die gleiche Richtung wie die Ergebnisse anhand der Hofstede Dimensionen. Es wurde ein tendenziell geringeres Arbeitszufriedenheitsniveau in den Kulturen berichtet, in denen Risiken und Unsicherheiten von den Menschen nicht erwünscht sind. Auch diese Ergebnisse bestätigen somit die Überlegungen, die hinsichtlich dieser Kulturdimension in Hypothese 2 zum Ausdruck kamen.

Auch wenn eine Untersuchung der Zusammenhänge mit anderen Kulturdimensionen nicht explizit im Fokus der hier vorliegenden Studie steht, sollen der Vollständigkeit halber auch die Zusammenhänge zwischen dem Arbeitszufriedenheitsniveau und den weiteren Kulturdimensionen der GLOBE-Studie kurz erwähnt werden. So bestanden signifikante Zusammenhänge mit den anhand der Praktiken gemessenen Dimensionen der Zukunftsorientierung ($r = .57$, $p < .01$ und $r = .87$, $p < .01$) sowie der Leistungsorientierung ($r = .42$, $p < .05$ und $r = .57$, $p < .05$). Bezüglich der ersten Dimensionen zeigte sich dann ein negativer Zusammenhang, wenn man die Wunschvorstellung der Befragten betrachtete ($r = -.45$, $p < .05$ und $r = -.68$, $p < .01$). Dieser entgegengesetzte Zusammenhang von Praktiken und Werte der Zukunftsorientierung ist jedoch auch hier nicht verwunderlich, da diese beiden Kennwerte auch selbst – wie die meisten anderen Dimensionen des GLOBE-Modells – wiederum negativ korreliert sind (vgl. Kapitel 3.1). Neben Zukunfts- und Leistungsorientierung wurde jedoch schließlich auch ein höheres Arbeitszufriedenheitsniveau in Ländern berichtet, in denen der Wunsch nach größerer geschlechterbezogener Gleichheit stark ausgeprägt ist ($r = .53$, $p < .01$ und $r = .55$, $p < .01$).

Insgesamt zeigt sich somit, daß die Ergebnisse dieser bivariaten Korrelationsanalysen vor allem die Hypothese 1 deutlich bestätigen können und somit die auf Basis der bisherigen Forschungsliteratur aufgestellte Annahme, daß in individualistischen Ländern ein höheres Arbeitszufriedenheitsniveau berichtet wird, während es in kollektivistischen Ländern eher gemäßigt ist, untermauert werden kann. Darüber hinaus können jedoch ebenso die eher losen Überlegungen der Hypothese 2 bestätigt werden, die lediglich auf den korrelativen Beziehungen der kulturellen Wertedimensionen untereinander beruhen.

Um auch die Überlegungen des destruktiven Testens mit einzubeziehen, wurden die Zusammenhänge des Arbeitszufriedenheitsniveaus ebenfalls mit ausgewählten sozioökonomischen Indikatoren geprüft (vgl. Tabelle 6). Es zeigte sich dabei insgesamt ein eindeutiger Zusammenhang mit dem Wohlstand eines Landes. Dieser kann auch als Indikator für die Arbeitsplatzqualität angesehen werden kann, da diese in der Regel mit dem Reichtum eines Landes steigt.

Tabelle 6: Zusammenhang des Arbeitszufriedenheitsniveaus mit sozioökonomischen Indikatoren

	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem Eurobarometer - EB)
Sozioökonomische Indikatoren		
Bruttoinlandsprodukt 2004	.60*** (48)	.68*** (17)
HDI 2000	.43*** (43)	.61** (14)
Arbeitslosenquote	-.20 (45)	-.65** (15)
Quality-of-Live-Index	.61*** (27)	.72*** (13)

*** p < .01; ** p < .05; * p < .10; in Klammern jeweils die Anzahl der Fälle/Länder

Wie die Ergebnisse im Einzelnen zeigen, war das Arbeitszufriedenheitsniveau höher in Ländern mit einem höheren Bruttoinlandsprodukt ($r = .60$, $p < .01$ und $r = .68$, $p < .01$), mit einem höheren Human-Development-Index ($r = .43$, $p < .01$ und $r = .61$, $p < .05$) sowie mit einem höheren Lebensstandard ($r = .61$, $p < .01$ und $r = .72$, $p < .01$). Interessanterweise ergab sich jedoch kein eindeutig signifikanter Zusammenhang mit den Daten zur Arbeitslosenquote, so wie sie in dieser Studie operationalisiert wurden. Insgesamt kann jedoch festgehalten werden, daß das Arbeitszufriedenheitsniveau somit auch durch die ökonomische Situation der Länder beeinflusst wird und insgesamt höher in Ländern war, die eine höhere Arbeitsplatzqualität haben; ein Ergebnis, was insgesamt nicht verwunderlich ist und auch durchaus so erwartet wurde.

9.3.3 Der Zusammenhang von Kultur und Antworttendenz-Norm

Bevor der eigentlich in dieser Arbeit zu untersuchende Zusammenhang von Antworttendenz-Normen und dem Arbeitszufriedenheitsniveau betrachtet werden kann, sollten vorher gemäß den Hypothesen 3 bis 5 die Zusammenhänge der Antworttendenz-Normen mit der Kultur eines Landes untersucht und möglichst die bisher schon in der Literatur bekannten Zusammenhänge repliziert werden. Für diese Überprüfung wurden vor allem die verschiedenen, neu erstellten Antworttendenz-Indizes, die auf Länderebene anhand der Daten des WVS berechnet wurden, herangezogen. Tabelle 7 zeigt für diese neu gebildeten Indizes die deskriptiven Kennwerte und ihre Interkorrelation. Zusätzlich sind ebenfalls die Interkorrelationen mit den bereits bekannten Indizes von Hofstede (2001) und Harzing (2006) aufgeführt. Es fällt auf, daß die drei Zustimmung-Indizes des WVS trotz verschiedener Skalenlabel bzw. -stufen mit mittlerer Stärke miteinander positiv korrelierten.

Tabelle 7: Zusammenhang der Antworttendenz-Normen untereinander

	MW	StA	1	2	3	4	5	6	7	8	9
Zustimmungs-Index (NARS 1) 5er-Skala, Zustimmunglabel, WVS	0,382	0,15									
Zustimmungs-Index (NARS 2) 4er-Skala, Zustimmunglabel, WVS	0,261	0,09	.32*** (69)								
Zustimmungs-Index (NARS 3) 5er-Skala, Wichtigkeitslabel, WVS	0,580	0,21	.32** (50)	.25* (52)							
Zustimmungs-Index (ARS) Hofstede (2001)	190,75	18,07	.49*** (30)	.43*** (36)	.09 (25)						
Zustimmungs-Index (NARS) Harzing (2006)	0,288	0,10	.50** (22)	.37* (24)	.31 (22)	.55** (20)					
Extremtendenz-Index (ERS) WVS	0,359	0,11	.50*** (70)	.28** (70)	.33** (51)	.40** (31)	.13 (23)				
Extremtendenz-Index (ERS) Harzing (2006)	0,106	0,03	.26 (22)	.26 (24)	.28 (22)	.46** (20)	.35* (26)	.43** (23)			
Mittentendenz-Index (MRS) WVS	0,148	0,05	-.46*** (70)	-.17 (70)	-.06 (51)	-.30 (31)	-.28 (23)	-.49*** (71)	-.15 (23)		
Mittentendenz-Index (MRS) Harzing (2006)	0,217	0,04	-.15 (22)	-.46** (24)	-.57*** (22)	-.26 (20)	-.30 (26)	-.22 (23)	-.45** (26)	.08 (23)	
Index Soz. Erwünschtheit Bernardi (2006)	6,067	0,71	.19 (7)	.55 (9)	-.08 (6)	.53 (9)	-.56 (5)	.25 (7)	-.35 (5)	.57 (7)	.57 (5)

*** p < .01; ** p < .05; * p < .10; in Klammern jeweils die Anzahl der Fälle/Länder

Wichtig ist ebenfalls zu erwähnen, daß auch der Extremtendenz-Index des WVS relativ hoch positiv mit den Zustimmungstendenz-Indizes des WVS korreliert war. Dieser Effekt ist so nicht unbedingt zu erwarten gewesen. Er ist jedoch vor allem darauf

zurückzuführen, daß ein Großteil der Items im WVS – wie es unglücklicherweise sehr häufig in internationalen Studien der Fall ist – lediglich positiv und somit in die gleiche Richtung formuliert war. Die neu gebildeten Indizes korrelierten zusätzlich hoch mit den entsprechenden Indizes von Hofstede und von Harzing, was somit die konvergente Validität der neuen Indizes deutlich macht. Dies gilt sowohl für die jeweiligen Korrelationen der Zustimmung-Indizes als auch des Extremtendenz-Index'. Lediglich der Mittentendenz-Index des WVS stand in keinem Zusammenhang mit dem entsprechenden Index von Harzing.

Der Soziale-Erwünschtheits-Index von Bernardi (2006) hingegen korrelierte zwar auch durchaus mit den anderen Indizes, jedoch ließen sich aufgrund der extrem kleinen Stichprobe diesbezüglich keine signifikanten Zusammenhänge finden.

Die gemäß den Hypothesen zu betrachtenden Ergebnisse zum Zusammenhang dieser neu errechneten Antworttendenz-Normen und der kulturellen Werte sind schließlich in Tabelle 8 dargestellt. Bedeutsame Zusammenhänge ließen sich hier vor allem mit den Zustimmungstendenz-Normen finden. In gleicher Form wie in der bisherigen Forschung (vgl. Harzing 2006; Smith, 2004) zeigte sich, daß zustimmendes Antwortverhalten vor allem in eher kollektivistischen Kulturen auftrat. Betrachtet man den Index, der anhand aller 5er-Likert-Skalen mit Zustimmunglabel errechnet wurde, zeigt sich, daß sowohl im Zusammenhang mit der Dimension von Hofstede ($r = -.53$, $p < .01$) mit der Dimension von Triandis ($r = -.75$, $p < .01$), mit den Dimensionen von Schwartz (Konservatismus: $r = .51$, $p < .05$; Aff. Autonomie: $r = -.35$, $p < .10$; Int. Autonomie: $r = -.36$, $p < .10$) als auch mit der Dimension der Kulturpraktiken der GLOBE-Studie ($r = .66$, $p < .01$) der erwartete Zusammenhang auftrat. Je stärker somit der Gruppenzusammenhalt in einem Land ausgeprägt ist, desto höher ist im Mittel die Tendenz, in Befragungen seinem Gegenüber zuzustimmen und seine eigene Meinung eher zurückzuhalten.

Bezüglich des Zustimmungs-Index der 4er-Likert-Skalen ergaben sich im Prinzip die gleichen Zusammenhänge, die jedoch insgesamt etwas schwächer ausgeprägt waren. Auch hier zeigte sich, daß eine höhere Zustimmungstendenz-Norm eher in kollektivistischen Kulturen aufzufinden war. Noch schwächer waren schließlich die Zusammenhänge bezüglich der 5er-Likert-Skala mit Wichtigkeits-Label. Die Zusammenhänge gingen auch hier zwar weitestgehend in die erwartete Richtung, jedoch waren die Korrelationen insgesamt noch schwächer ausgeprägt und erreichten daher vielfach keine Signifikanz.

Hypothese 3 kann somit anhand der im WVS neu errechneten Zustimmungstendenz-Normen deutlich bestätigt werden. Insgesamt bestätigte sich der erwartete Zusammenhang von Individualismus/Kollektivismus und Zustimmungstendenzen, so daß auch anhand der Daten des WVS erneut gezeigt werden kann, daß zustimmendes Antwortverhalten vor allem in kollektivistischen Kulturen auftritt. Dieser Effekt läßt sich dabei vor allem bei den klassischen 5er-Likert-Skalen mit einem Zustimmungslabel finden; also genau den Skalen, die extrem häufig in der Einstellungs- und Meinungsforschung eingesetzt werden.

Tabelle 8: Zusammenhang von Antworttendenz-Normen und Kultur

	NARS 5er-Skala Zustimmung WVS	NARS 4er-Skala Zustimmung WVS	NARS 5er-Skala Wichtigkeit WVS	ERS 5er-Skala Zustimmung WVS	MRS 5er-Skala Zustimmung WVS	SD Bernardi
Kulturdimensionen nach Hofstede	N = 52	N = 59	N = 43	N = 52	N = 52	N = 11
Individualismus	-.53***	-.26**	-.21	-.08	.07	-.73**
Machtdistanz	.60***	.50***	.37**	.17	-.21	.57*
Unsicherheitsvermeidung	.07	.33**	.36**	.21	.20	.68**
Kulturdimensionen nach Triandis	N = 40	N = 44	N = 32	N = 40	N = 40	N = 9
Individualismus	-.75***	-.48***	-.12	-.34**	.34**	-.63*
Kulturdimensionen nach Schwartz	N = 24	N = 28	N = 22	N = 24	N = 24	N = 6
Konservatismus	.51**	.25	.13	.10	-.24	-.38
Affektive Autonomie	-.35*	-.34*	-.08	-.06	.09	.10
Intellektuelle Autonomie	-.36*	.01	-.08	.03	.25	.69
Hierarchie	.40*	-.01	-.37*	-.12	-.36*	.10
Gleichheitsdenken	-.47**	-.13	.16	-.02	.28	-.01
Kulturdimensionen der GLOBE-Studie: Praktiken	N = 39	N = 45	N = 30	N = 39	N = 39	N = 11
Machtdistanz	.25	.38**	.44**	.23	-.06	.70**
Unsicherheitsvermeidung	-.44***	-.49***	-.30	-.34**	-.00	-.15
Institut. Kollektivismus	-.24	-.28*	-.42**	-.32*	.27*	-.23
InGroup Kollektivismus	.66***	.51***	.36*	.20	-.28*	.51
Kulturdimensionen der GLOBE-Studie: Werte	N = 39	N = 45	N = 30	N = 39	N = 39	N = 11
Machtdistanz	.23	-.12	-.03	.10	-.18	-.70**
Unsicherheitsvermeidung	.76***	.47***	.32*	.38**	-.30*	.56*
Institut. Kollektivismus	.19	.20	.30	.32**	-.32**	.60**
InGroup Kollektivismus	-.15	-.06	-.05	-.01	.06	.21

*** p < .01; ** p < .05; * p < .10

Auch hinsichtlich der Zusammenhänge der Zustimmungstendenz-Indizes mit den Kulturdimensionen Machtdistanz und Unsicherheitsvermeidung zeigen sich weitestgehend die erwarteten Ergebnisse. Alle drei Zustimmungstendenz-Indizes

korrelierten positiv mit Hofstedes Machtdistanz ($r = .60, p < .01$; $r = .50, p < .01$ und $r = .37, p < .05$). Auch bezüglich der GLOBE-Dimension zeigten sich zumindest bei zwei der Indizes signifikante Zusammenhänge ($r = .38, p < .05$ und $r = .44, p < .05$). Die Dimensionen Hierarchie und Gleichheitsdenken von Schwartz korrelierten ebenfalls erwartungsgemäß mit dem Index der 5er-Likert-Skalen mit Zustimmungslabel ($r = .40, p < .10$ und $r = -.47, p < .05$). Jedoch existierte unerwarteterweise ebenfalls eine negative Korrelation des Index der 5er-Skalen mit Wichtigkeitslabel mit der Hierarchie-Dimension nach Schwartz ($r = -.37, p < .10$), die so nicht direkt zu erklären ist. Betrachtet man jedoch alle Ergebnisse zusammen, zeigt sich durchaus der erwartete Zusammenhang. Hypothese 4 kann somit hinsichtlich Machtdistanz als bestätigt angesehen werden. Insgesamt gilt jedoch auch hier, daß vor allem zwischen dem Index der 5er-Likert-Skalen mit Zustimmungslabel und der Kulturdimension Machtdistanz die höchsten Zusammenhänge auftraten.

Bezüglich der Dimensionen Unsicherheitsvermeidung korrelierten immerhin zwei der drei Zustimmungskindizes mit Hofstedes Wert ($r = .33, p < .05$ und $r = .36, p < .05$) und zusätzlich alle drei Indizes bei der GLOBE-Studie positiv mit – erwartungsgemäß – der Werte-Dimension der Studie ($r = .76, p < .01$; $r = .47, p < .01$ und $r = .32, p < .10$). Die Zusammenhänge der Zustimmungstendenz-Indizes mit den Praktiken ($r = -.44, p < .01$ und $r = -.49, p < .01$) waren auch hier aufgrund des Negativ-Zusammenhangs der Praktiken- und Werte-Dimensionen der GLOBE-Studie untereinander entgegengesetzt. Die Ergebnisse sind somit zwar nicht völlig eindeutig, deuten jedoch insgesamt auch in die erwartete Richtung. Überraschend ist jedoch in der Tat, daß gerade der Index der 5er-Likert-Skalen mit Zustimmungslabel, der bisher immer die stärksten Effekte zeigte, überhaupt nicht mit Hofstedes Dimension korrelierte ($r = .07, p > .10$). Insgesamt läßt sich aber dennoch festhalten, daß die Mehrzahl der Zusammenhänge die erwartete Richtung zeigte und somit Hypothese 4 auch hinsichtlich der Dimension Unsicherheitsvermeidung weitestgehend bestätigt werden kann.

Die Zusammenhänge mit Extremtendenz- und Mittentendenz-Normen sind insgesamt weniger eindeutig. Hinsichtlich der Extremtendenz zeigten sich über die verschiedenen Kulturmodelle hinweg überhaupt keine gleichgearteten Zusammenhänge. Vermutet wurde gemäß Hypothese 5, daß extremes Antwortverhalten eher in individualistischen Kulturen auftreten würde. Diese Vermutung kann hier jedoch überhaupt nicht bestätigt werden. Das Rating von Triandis korrelierte erstaunlicherweise sogar negativ mit der Extremtendenz-Norm ($r = -.34, p < .05$). Möglicherweise sind diese unerwarteten

Ergebnisse darauf zurückzuführen, daß der ERS-Indikator nicht optimal operationalisiert werden konnte.

Auch wenn man die Mittentendenzen in den einzelnen Ländern betrachtet, ergeben sich keine klaren Zusammenhänge mit der jeweiligen Kultur der Länder. Im Wesentlichen zeigten sich hier nur die entgegengesetzten Korrelationen verglichen mit der Extremtendenz-Norm. Dies ist aufgrund der Tatsache, daß eine Mittentendenz durchaus als das Gegenteil einer Extremtendenz angesehen werden kann und auch die beiden Indizes negativ miteinander korreliert waren (vgl. Tabelle 8), nicht überraschend. Klare interpretierbare Zusammenhänge zeigten sich für die Mittentendenzen aber insgesamt auch nicht. Hypothese 5 kann daher nicht bestätigt werden und reiht sich somit in die bisherigen Forschungsergebnisse ein, die hinsichtlich Extrem- und Mittentendenzen auch kein eindeutiges Ergebnis fanden.

Zusammenhänge zwischen Kultur und Sozialer-Erwünschtheits-Norm hatte auch bereits Bernardi (2006) in seiner Studie berichtet, jedoch bezog er sich dabei lediglich auf die Werte von Hofstede. Die Analysen in Tabelle 8 bezüglich Hofstedes Dimensionen stellen somit zwar erst einmal nur eine Wiederholung dar und erzielten folglich die gleichen Ergebnisse: Antwortverhalten im Sinne Sozialer Erwünschtheit ist stärker bei hohem Kollektivismus ($r = -.73$, $p < .05$), hoher Machtdistanz ($r = .57$, $p < .10$) und hoher Unsicherheitsvermeidung ($r = .68$, $p < .05$). Zusätzlich wurden diese Tendenzen jedoch in Teilen durch die Dimensionen der GLOBE-Studie und der Dimension von Triandis in den hier neu durchgeführten Berechnungen bestätigt. Auch bezüglich dieser Dimensionen zeigte sich die gleiche Tendenz, daß sozial erwünschtes Antwortverhalten eher in kollektivistischen Kulturen, in Kulturen mit hoher Machtdistanz und in Kulturen mit hoher Unsicherheitsvermeidung auftritt. Insgesamt sollte diesen Ergebnissen jedoch nicht zu viel Bedeutung beigemessen werden, da sie leider nur auf sehr kleinen Länderstichproben beruhten. Hypothese 6 kann daher insgesamt in Teilen als bestätigt angesehen werden.

Anhand der Analysen zum Zusammenhang von Kultur und Antworttendenz-Norm, die in Teilen die bereits bekannten Ergebnisse der Forschung replizieren sollten, zeigt sich somit alles in allem, daß sich auch in den Berechnungen dieser Studie hier vor allem hinsichtlich der Norm zur Zustimmungstendenz klare Zusammenhänge ergaben. Weder Extrem- noch Mittentendenzen noch Soziale Erwünschtheit zeigten hingegen ein klares Zusammenhangsprofil. Dieses Ergebnis unterstützt somit auch die Entscheidung, den

Fokus der hier vorliegenden Studie vor allem auf die Untersuchung einer Art des Antwortverhaltens, der Zustimmungstendenz, gelegt zu haben.

9.3.4 Der Zusammenhang von Antworttendenz-Normen und Arbeitszufriedenheitsniveau

Während die bisherigen, dargestellten Resultate vor allem Ergebnisse der Forschung replizierten, beziehen sich die folgenden Ergebnisse nun auf den zentralen Bestandteil dieser Studie: den Zusammenhang von allgemeinen Antworttendenz-Normen und der Antwort auf eine Arbeitszufriedenheitsfrage.

Gemäß den Überlegungen zu Hypothese 7 wurde angenommen, daß auch allgemeine Antworttendenzen einen Einfluß auf das Arbeitszufriedenheitsniveau haben müßten. Betrachtet man die Ergebnisse der Analysen bezüglich der Zustimmungstendenz-Normen in Tabelle 9, zeigt sich, daß dies in der Tat wie erwartet der Fall war. Alle fünf Indizes der Zustimmungstendenz-Normen standen in einem deutlichen Zusammenhang mit dem Arbeitszufriedenheitsniveau sowohl aus dem WVS als auch aus dem EB. Lediglich der Zustimmungstendenz-Index von Hofstede zeigte keine signifikanten Ergebnisse, da die Anzahl der Länder, die betrachtet werden konnten, zu gering war und auch die Effekte insgesamt hier etwas geringer waren. Jedoch zeigten auch die Ergebnisse von Hofstedes Index in die gleiche Richtung. Hypothese 7 kann somit bezüglich des Einflusses von Zustimmungstendenz-Normen bestätigt werden. Diese stehen in einem deutlichen Zusammenhang zum Arbeitszufriedenheitsniveau eines Landes.

Tabelle 9: Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau

	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem Eurobarometer - EB)
Zustimmungs-Index NARS 1 5er-Skala, Zustimmungslabel, WVS	-.47*** (46)	-.68*** (16)
Zustimmungs-Index NARS 2 4er-Skala, Zustimmungslabel, WVS	-.38*** (47)	-.59** (16)
Zustimmungs-Index NARS 3 5er-Skala, Wichtigkeitslabel, WVS	-.47*** (42)	-.71*** (16)
Zustimmungs-Index von Harzing (2006)	-.47** (23)	-.68** (11)
Zustimmungs-Index von Hofstede (2001)	-.29 (25)	-.13 (14)

*** p < .01; ** p < .05; * p < .10; in Klammern jeweils die Anzahl der Fälle/Länder

Hinsichtlich des Einflusses von Extremtendenz- oder Mittentendenz-Normen sowie Sozialen-Erwünschtheits-Normen ließen sich jedoch so gut wie keine Zusammenhänge finden (vgl. Tabelle 10). Lediglich die Indizes von Harzing (2006) korrelierten signifikant mit dem Arbeitszufriedenheitsniveau, jedoch auch nur mit dem Niveau, das im EB gemessen wurde. Dieser Befund ließ sich trotz einer sogar größeren Stichprobe nicht beim Zufriedenheitsniveau des WVS finden, so daß der beim EB gemessene Effekt nicht bestätigt werden kann. Umgekehrt zeigten auch die neu am WVS errechneten Indizes keinerlei bedeutsame Zusammenhänge, weder mit dem Arbeitszufriedenheitsniveau des WVS noch dem Niveau des EB. Hinsichtlich des Index zur Sozialen Erwünschtheit ließ sich der Einfluß aufgrund der geringen Überlappung der beiden Datensätze nur anhand von 7 Ländern mit dem WVS prüfen. Daher konnte der dort durchaus vorhandene Effekt ($r = -.36$, $p < .10$) jedoch leider statistisch nicht abgesichert werden. Hypothese 7 kann somit hinsichtlich der weiteren Antworttendenz-Normen, d.h. Extrem-, Mitten- und Soziale-Erwünschtheits-Tendenz nicht bestätigt werden.

Betrachtet man alle Antworttendenz-Normen abschließend gemeinsam, ergibt sich somit lediglich hinsichtlich Zustimmungstendenz-Normen ein eindeutiger Zusammenhang zum Arbeitszufriedenheitsniveau eines Landes.

Tabelle 10: Zusammenhang der weiteren Antworttendenz-Normen mit dem Arbeitszufriedenheitsniveau

	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem Eurobarometer - EB)
Indizes für Extremtendenz-Normen		
Extremtendenz-Index ERS 5er-Skala, Zustimmungsetikette, WVS	-.07 (47)	.10 (17)
Extremtendenz-Index von Harzing (2006)	-.14 (23)	-.60* (11)
Indizes für Mittentendenz-Normen		
Mittentendenz-Index MRS 5er-Skala, Zustimmungsetikette, WVS	.01 (47)	-.41 (17)
Mittentendenz-Index von Harzing (2006)	.33 (23)	.69** (11)
Indizes für eine Soziale-Erwünschtheitsnorm		
Index Sozialer Erwünschtheit SD von Bernardi (2006)	-.36 (7)	Stichprobe zu klein

*** $p < .01$; ** $p < .05$; * $p < .10$; in Klammern jeweils die Anzahl der Fälle/Länder

Während Hypothese 7 lediglich annahm, daß es überhaupt einen Zusammenhang zwischen Antworttendenz-Normen und dem Arbeitszufriedenheitsniveau geben müßte,

kamen in Hypothese 9 die Überlegungen zur Richtung des Zusammenhangs zum Ausdruck. Nachdem die bisherigen signifikanten Ergebnisse nun zeigen konnten, daß in der Tat ein Zusammenhang von Zustimmungstendenz-Normen und dem Arbeitszufriedenheitsniveau existiert, stellte sich gemäß den Überlegungen dieser Hypothese die Frage, ob dieser dann positiver oder negativer Natur sei. Betrachtet man daher nun erneut die Ergebnisse in Tabelle 9, läßt sich diese Frage nach der Art des Zusammenhangs relativ einfach beantworten. Die Ergebnisse zeigen deutlich, daß alle Zusammenhänge negativer Art sind. Sowohl die neu am WVS gebildeten drei Indikatoren als auch die bereits in der Literatur veröffentlichten Kennwerte standen in einem klaren negativen Zusammenhang zum Arbeitszufriedenheitsniveau. Wenngleich man dabei zwar dem Zusammenhang der WVS-Zustimmungstendenz-Indikatoren mit dem WVS-Arbeitszufriedenheitsniveau vorwerfen könnte, daß beide Kennwerte aus ein und derselben Datenquelle stammten, werden diese Zusammenhänge jedoch auch deutlich von den Antworttendenz-Indikatoren, die aus völlig anderen, eigenständigen Untersuchungen stammten, bestätigt. Die Kritik einer reinen Single-Source-Messung kann somit entkräftet werden. In gleicher Weise sollte darauf hingewiesen werden, daß auch die Zusammenhänge, die sich mit den Daten des EB zeigen, auf Basis von jeweils völlig unabhängigen Untersuchungen entstanden sind. Diese gegenseitige Bestätigung der Zustimmungstendenz-Normen aus drei unabhängigen Datenquellen und der Arbeitszufriedenheitsniveaus aus ebenfalls zwei unabhängigen Quellen unterstreicht daher insgesamt die Bedeutsamkeit dieser Ergebnisse. Zustimmungstendenz-Normen und das Arbeitszufriedenheitsniveau hängen somit, wie in Anlehnung an die Überlegungen von Markus und Kitayama (1991) angenommen, in negativer Weise zusammen. Ein hohes Arbeitszufriedenheitsniveau zeigt sich somit nur in Kulturen, in denen Zustimmungsnormen kaum oder gar nicht existent sind. Umgekehrt könnte man es auch so beschreiben, daß in Kulturen, in denen die Menschen in der Kommunikation eher ein hohes Maß an Zurückhaltung und Bescheidenheit zeigen, im Mittel sich eher gemäßigt zur Arbeitszufriedenheitsfrage geäußert wird. Man bedenke hier erneut, daß auch in den Ländern mit vergleichsweise geringerer Zufriedenheit das Zufriedenheitsniveau im Mittel immer noch im positiven Bereich liegt. Bei hoher kultureller Zustimmungsnorm äußern sich alle Befragten im Mittel somit nicht unzufrieden über ihre Arbeit, sondern bezeichnen sich gegenüber dem Interviewer lediglich in etwas gemäßigterer Form als zufrieden.

Die dieser Arbeit zugrunde liegenden kritischen Überlegungen, ob Arbeitszufriedenheitsniveaus überhaupt direkt zu vergleichen sind, können folglich anhand der hier vorliegenden Ergebnisse hinsichtlich ihrer Daseinsberechtigung eindrucksvoll bestätigt werden. Arbeitszufriedenheitsniveaus in einzelnen Ländern unterscheiden sich offensichtlich nicht nur durch die jeweiligen Arbeitsbedingungen, sondern sind auch in einem hohen Maße von der jeweiligen Kommunikationsnorm, die in der Kultur befolgt werden soll, beeinflusst. Zustimmungsnormen führen daher dazu, daß das gemessene Niveau grundsätzlich niedriger angesiedelt ist. Eine direkte Vergleichbarkeit der Niveaus hinsichtlich der Frage nach der wahren Zufriedenheit der Menschen in einem Land mit ihrem Arbeitsplatz ist somit eigentlich nicht mehr gegeben.

Betrachtet man die Ergebnisse in Tabelle 9 genauer, fällt zusätzlich noch ein weiterer Aspekt auf. Die korrelativen Zusammenhänge, die sich mit den Daten des EB ergaben haben, waren insgesamt höher als die Zusammenhänge mit den Daten des WVS. Dieser Effekt war so nicht zu erwarten gewesen. Bei der Suche nach einer Erklärung für diesen Unterschied fällt die jeweils unterschiedliche Stichprobe der beiden Studien ins Auge. Während das Eurobarometer lediglich Länder der Europäischen Union erfaßte, sind im WVS Länder der gesamten Welt berücksichtigt worden. Welcher Unterschied kommt jedoch durch diese verschiedene Stichprobenwahl zum Ausdruck? In Bezug zu der hier vorliegenden Fragestellung ist bedeutsam, daß die Länder der Europäischen Union insgesamt vor allem eine bessere wirtschaftliche Entwicklung durchlaufen haben als viele der Länder des WVS. Es kann daher angenommen werden, daß dieses höhere wirtschaftliche Niveau auch einen Einfluß auf die Arbeitsplatzqualität hat und daher in den Ländern der Europäischen Union insgesamt bessere Arbeitsplatzbedingungen vorliegen sollten. Da unbestritten ist, daß das gemessene Arbeitszufriedenheitsniveau neben der Beeinflussung durch Kommunikationsnormen auch einen wahren Kern besitzt und daher auch einen Ausdruck der jeweils tatsächlichen Arbeitsbedingungen in einem Land darstellt, könnte dieser Unterschied in der Ausstattung der Arbeitsplätze dazu geführt haben, daß der Einfluß von Kommunikationsnormen nur dort stärker zum Vorschein kommt, wo eine ausreichende Arbeitsplatzqualität schon erreicht ist. In den Ländern mit schlechteren Bedingungen wäre somit das Arbeitszufriedenheitsniveau stärker von der Arbeitsplatzqualität beeinflusst, während in den Ländern mit besseren Bedingungen dann stärker die Kommunikationsnormen durchschlagen. Bildlich gesprochen würde dies bedeuten, daß eine Kommunikationsnorm erst dann auf Befragte

einen deutlichen Einfluß ausübt, wenn diese – im Prinzip im Sinne der Maslowschen Bedürfnispyramide – nicht mehr um die Befriedigung ihrer Grundbedürfnisse fürchten müssen. Erst wenn man nicht mehr um die Sicherheit am Arbeitsplatz fürchten muß und nicht mehr unter extremen Bedingungen arbeiten muß, würden Kommunikationsnormen einen stärkeren Einfluß auf das Verhalten der Befragten zeigen.

Um diese Überlegungen genauer zu überprüfen, wurde daher eine Re-Analyse der WVS-Daten vorgenommen. Dazu mußte in diesem Datensatz, in dem bisher noch alle Länder gemeinsam analysiert wurden, eine Trennung in Länder mit besseren und Länder mit schlechteren Bedingungen erfolgen. Operationalisiert anhand des Bruttoinlandsprodukts (BIP) wurden daher die Länder des Datensatzes anhand eines Mediansplits in eine Gruppe sogenannter reicher Länder (d.h. mit hohem BIP) mit guten Arbeitsbedingungen und in eine weitere Gruppe mit sogenannten armen Ländern (d.h. mit niedrigem BP) mit schlechten Bedingungen eingeteilt. Während in der ersten Gruppe vornehmlich west-europäische Länder plus Kanada, Japan und der USA verblieben, fanden sich in der zweiten Gruppe viele osteuropäische Länder (Rumänien, Bulgarien, Ungarn, Weißrußland, Rußland, Polen, Slowakei, etc.) aber auch die asiatischen Länder (Indien, China, Südkorea) sowie süd- und mittelamerikanische (Brasilien, Argentinien, Chile, Mexiko) und afrikanische Länder (Nigeria, Süd-Afrika) wieder.

Wie in Tabelle 11 zu sehen ist, zeigen die Ergebnisse dieser erneuten Analyse der WVS-Daten, daß sich unterschiedliche Zusammenhänge in den armen und den reichen Ländern ergaben.

Tabelle 11: Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau bei armen und reichen Ländern des WVS

	reiche Länder	arme Länder
	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem World-Values-Survey - WVS)
Zustimmungs-Index NARS 1 5er-Skala, Zustimmungslabel, WVS	-.51** (22)	-.02 (24)
Zustimmungs-Index NARS 2 4er-Skala, Zustimmungslabel, WVS	-.37* (23)	-.04 (24)
Zustimmungs-Index NARS 3 5er-Skala, Wichtigkeitslabel, WVS	-.51** (22)	-.34 (20)
Zustimmungs-Index von Harzing (2006)	-.61** (13)	.08 (10)
Zustimmungs-Index von Hofstede (2001)	-.34 (19)	Fallanzahl zu klein

*** p < .01; ** p < .05; * p < .10; in Klammern jeweils die Anzahl der Fälle/Länder

Der negative Zusammenhang von Zustimmungstendenz-Normen und dem Arbeitszufriedenheitsniveau blieb in den reichen Ländern bestehen, verschwand jedoch in den armen Ländern fast vollständig. Dort zeigte sich überhaupt kein Zusammenhang mehr mit 3 der 4 Zustimmungs-Indizes. Zur grafischen Veranschaulichung ist in Abbildung 16 beispielhaft der Zusammenhang des ersten Zustimmungsindex NARS 1 mit dem Arbeitszufriedenheitsniveau sowohl in den armen als auch in den reichen Ländern dargestellt.

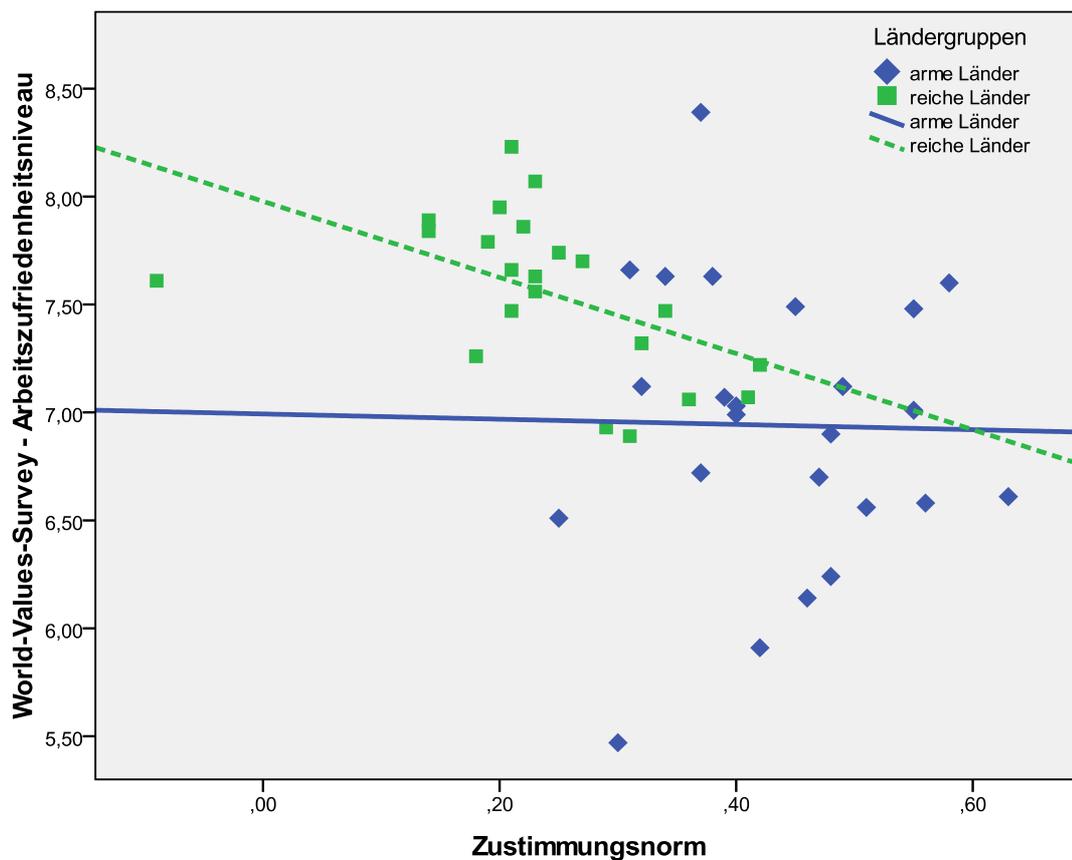


Abbildung 16: Zusammenhang der Zustimmungsnorm und des Arbeitszufriedenheitsniveaus in armen und reichen Ländern

Lediglich der Index, der anhand der Wichtigkeits-Skalen errechnet wurde, scheint noch mit dem Arbeitszufriedenheitsniveau in Zusammenhang zu stehen, jedoch in nicht signifikanter Weise. Es zeigte sich somit insgesamt eine Bestätigung der oben genannten Überlegungen. Es gilt demnach zwar weiterhin, daß eine hohe Zustimmungsnorm zu einem insgesamt gemäßigeren Arbeitszufriedenheitsniveau führt. Einschränkend muß jedoch nun hinzugefügt werden, daß dies nur in reichen Ländern

gilt. Dies könnte demnach so interpretiert werden, daß nur bei einem entsprechenden Wohlstand des Landes, der gleichbedeutend ist mit einer generellen Verbesserung der Arbeitsbedingungen, die Kommunikationsnormen die Messung des Arbeitszufriedenheitsniveaus verzerren. Wenn die Arbeitsbedingungen eher schlecht sind, scheinen diese Normen keinen Einfluß auszuüben. Es könnte angenommen werden, daß die Menschen dort scheinbar andere Sorgen – z.B. die Sicherheit am Arbeitsplatz – haben, als sich um das Befolgen von Kommunikationsnormen zu kümmern. Dieses neue Ergebnis des Vergleichs von armen und reichen Ländern deutete jedoch jetzt darauf hin, daß auch die weiteren Analysen vorrangig bei der Gruppe der reichen Länder durchgeführt werden sollten, da sich interessanterweise nur dort die angenommenen Zusammenhänge zeigen ließen.

Zwar konnten die bisherigen Ergebnisse die Annahmen der Hypothesen deutlich bestätigen. Im Sinne des destruktiven Testens sollte jedoch auch überprüft werden, inwieweit die hier in den reichen Ländern gefundenen Zusammenhänge zwischen der Zustimmungsnorm und dem Arbeitszufriedenheitsniveau nicht auch durch andere Aspekte erklärt werden könnten. Aus diesem Grund wurden die bivariaten Korrelationsanalyse in den reichen Ländern wiederholt und nun auch die sozioökonomischen Aspekte mit einbezogen, die ebenfalls bedeutsam mit dem Arbeitszufriedenheitsniveau korrelierten: das Bruttoinlandsprodukt, der Human-Development-Index und der Quality-of-Life-Index.

Die Ergebnisse zeigen, daß der Zusammenhang von Zustimmungsnormen und dem Arbeitszufriedenheitsniveau auch dann grundsätzlich bestehen bleibt, wenn man für die wirtschaftlichen Rahmenbedingungen und die Unterschiede im jeweiligen Lebensstandard der einzelnen Länder mit Hilfe von Partiellen Korrelationsanalysen kontrolliert (vgl. Tabelle 12). Bei fast allen Indikatoren blieb ein deutlicher negativer Zusammenhang bestehen. Lediglich der Zustimmungsnorm-Index der 4er-Skala und der Indikator von Hofstede büßten in ihrer Erklärungskraft ein. Alle anderen Zusammenhänge, vor allem mit den Indikatoren der 5er-Skalen des WVS, zeigten jedoch weiterhin eine stark negative Tendenz. Dies bedeutet folglich, daß, selbst wenn man für den Einfluß der sozioökonomischen Indikatoren in der statistischen Analyse kontrolliert, bleiben die bisher gefundenen Beziehungen bestehen. Auch diese Form des destruktiven Testens zeigt somit, daß das Arbeitszufriedenheitsniveau geringer bzw. gemäßiger ist, wenn in einem Land hohe Zustimmungsnormen vorherrschend sind.

Tabelle 12: Zusammenhang von Zustimmungstendenz-Normen und Arbeitszufriedenheitsniveau unter Kontrolle der sozioökonomischen Bedingungen

	Arbeitszufriedenheitsniveau (Ländermittelwerte – nur reiche Länder – aus dem World-Values-Survey - WVS)	Arbeitszufriedenheitsniveau (Ländermittelwerte aus dem Eurobarometer - EB)
Kontrollvariable Bruttoinlandsprodukt 2004		
Zustimmungs-Index NARS 1 5er-Skala, Zustimmungsetikett, WVS	-.37* (19)	-.61** (13)
Zustimmungs-Index NARS 2 4er-Skala, Zustimmungsetikett, WVS	-.15 (20)	-.43 (13)
Zustimmungs-Index NARS 3 5er-Skala, Wichtigkeitsetikett, WVS	-.55*** (19)	-.69*** (13)
Zustimmungs-Index von Harzing (2006)	-.34 (9)	-.50 (7)
Zustimmungs-Index von Hofstede (2001)	-.02 (15)	-.49 (10)
Kontrollvariable Human-Development-Index 2000		
Zustimmungs-Index NARS 1 5er-Skala, Zustimmungsetikett, WVS	-.28 (16)	-.50* (10)
Zustimmungs-Index NARS 2 4er-Skala, Zustimmungsetikett, WVS	-.00 (17)	-.30 (10)
Zustimmungs-Index NARS 3 5er-Skala, Wichtigkeitsetikett, WVS	-.44* (17)	-.54* (10)
Zustimmungs-Index von Harzing (2006)	-.41 (9)	-.58 (7)
Zustimmungs-Index von Hofstede (2001)	-.04 (15)	-.40 (10)
Kontrollvariable Quality-of-Life-Index		
Zustimmungs-Index NARS 1 5er-Skala, Zustimmungsetikett, WVS	-.23 (14)	-.49 (10)
Zustimmungs-Index NARS 2 4er-Skala, Zustimmungsetikett, WVS	.04 (15)	-.29 (10)
Zustimmungs-Index NARS 3 5er-Skala, Wichtigkeitsetikett, WVS	-.28 (14)	-.39 (10)
Zustimmungs-Index von Harzing (2006)	-.29 (9)	-.50 (7)
Zustimmungs-Index von Hofstede (2001)	.00 (15)	-.46 (10)

*** p < .01; ** p < .05; * p < .10; in Klammern jeweils die Anzahl der Fälle/Länder

9.3.5 Zusammenfassung der Ergebnisse der bivariaten Analysen

Die bisherigen sekundäranalytischen Berechnungen und deren zentrale Ergebnisse zeigen insgesamt folgendes Bild. Auf der einen Seite läßt sich der Zusammenhang von Kultur und Antworttendenzen bestätigen, der auch schon in früheren Studien herausgefunden wurde (Harzing, 2006; Smith, 2004): Zustimmungstendenzen treten vor allem in kollektivistischen Kulturen auf. Des weiteren konnte gezeigt werden, daß auch

diese neueren Daten der hier vorliegenden Arbeit den Zusammenhang von Individualismus/Kollektivismus und dem Arbeitszufriedenheitsniveau bestätigen. Dieses Niveau ist höher in Staaten, in denen die Kultur das Individuum höher bewertet als die Gruppe. Als neues Ergebnis dieser Studie kommt nun hinzu, daß eine hohe Norm zur Zustimmung dazu führt, daß insgesamt ein geringeres, d.h. gemäßigteres Arbeitszufriedenheitsniveau berichtet wird. Der gemäß den Hypothesen angenommene Zusammenhang der drei zentralen Konstrukte dieser Studie läßt sich somit für eine spezielle Konstellation bestätigen: für eine Form der Antworttendenzen, der Norm zur Zustimmung, und für die am meisten untersuchte Kulturdimension, nämlich Individualismus/Kollektivismus (vgl. Abbildung 17).

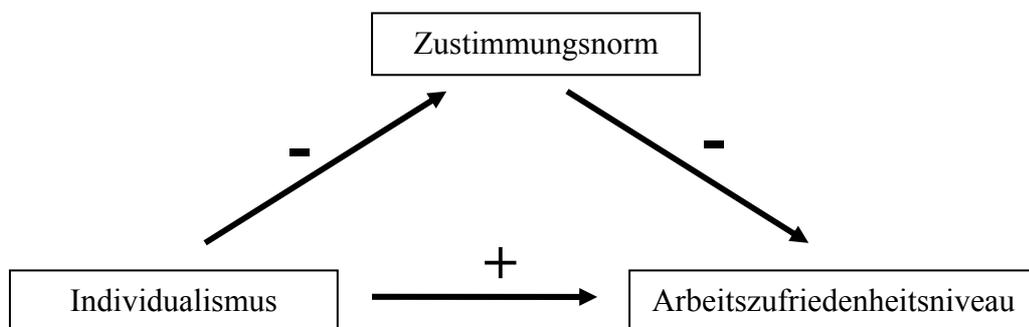


Abbildung 17: Bestätigter Gesamtzusammenhang von Individualismus, Zustimmungsnorm und Arbeitszufriedenheitsniveau

Im Folgenden soll nun gezeigt werden, inwieweit sich diese drei Konstrukte in einen gemeinsamen Zusammenhang setzen lassen, indem die Antworttendenz als wichtiger mediierender Aspekt angesehen werden kann, der den vermittelnden Faktor zwischen Kultur und dem Arbeitszufriedenheitsniveau darstellt.

9.4 Die Mediationsanalysen

In Bezug zu den Überlegungen, daß die Kultur eines Landes die Menschen in einer bestimmten Art und Weise sozialisiert, wie sie miteinander kommunizieren sollten, und dieses Kommunikationsverhalten dann auch mitentscheidend dafür ist, welches Arbeitszufriedenheitsniveau im Durchschnitt berichtet wird, wurde angenommen, daß Zustimmungstendenzen somit einen mediierenden Faktor darstellen würden. Um diese theoretische Annahme auch statistisch absichern zu können, wurde eine

Mediationsanalyse gemäß den Vorgaben von Baron und Kenny (1986) mit den drei zentralen Konstrukten dieser Studie durchgeführt.

9.4.1 Vorgehen der Mediationsanalysen

Bei dieser Mediationsanalyse wird generell geprüft, ob der anfänglich gefundene Zusammenhang zwischen einer unabhängigen Variable (im Beispiel hier die Kultur) und einer abhängigen Variable (im Beispiel hier das Arbeitszufriedenheitsniveau) sich durch den Einfluß der unabhängigen Variable auf einen Mediator (im Beispiel hier die Antworttendenz-Norm), der wiederum die abhängige Variable beeinflusst, erklären läßt. Ist dies der Fall, wäre der sogenannte totale Effekt der unabhängigen Variable auf die abhängige Variable dann nur ein Scheinzusammenhang, und der Mediator wäre die eigentliche Einflußgröße auf die abhängige Variable. Damit dies letztendlich gilt, müssen gemäß Baron und Kenny (1986) vier zentrale Voraussetzungen erfüllt sein, die, auf den hier vorliegenden Fall bezogen, Folgendes bedeuten würden:

- a) Kultur als unabhängige Variable muß einen deutlichen und möglichst signifikanten Effekt (der ‚totale Effekt‘) auf das Arbeitszufriedenheitsniveau, die abhängige Variable, haben.
- b) Kultur als unabhängige Variable muß ebenso einen deutlichen und möglichst signifikanten Effekt auf die Antworttendenz-Norm, den Mediator, haben.
- c) Die Antworttendenz-Norm als Mediator muß einen signifikanten Effekt auf das Arbeitszufriedenheitsniveau, die abhängige Variable, haben. Dieser Effekt muß im Mediationsmodell auch unter Kontrolle des statistischen Einflusses der unabhängigen Variable bestehen bleiben.
- d) Der zuvor als totaler Effekt gemessene signifikante Zusammenhang zwischen Kultur als unabhängige Variable und Arbeitszufriedenheitsniveau als abhängige Variable verringert sich im Mediationsmodell oder ist optimalerweise dort sogar gar nicht mehr existent, wenn gleichzeitig für den Effekt des in c) überprüften Zusammenhangs statistisch kontrolliert wird.

Zur Prüfung dieser Voraussetzungen wurden die gemäß Baron und Kenny (1986) jeweils notwendigen drei Regressionsanalysen mit der Kulturdimension, der Antworttendenz-Norm sowie dem Arbeitszufriedenheitsniveau berechnet. Da sich der

Einfluß von Antworttendenz-Normen auf das Arbeitszufriedenheitsniveau nur in reichen Ländern zeigen ließ (vgl. Tabelle 11), wurden diese im Folgenden berichteten Mediationsanalysen auch vorrangig nur für die reichen Länder durchgeführt. Von einer Mediation in den armen Ländern konnte nicht mehr ausgegangen werden, da sich dort ja bereits kein Effekt des Mediators, d.h. der Antworttendenz-Norm, auf die abhängige Variable zeigen ließ. Aufgrund der vorangegangenen Ergebnisse zum Zusammenhang der verschiedenen Antworttendenz-Normen zu Kultur und Arbeitszufriedenheitsniveau (vgl. Tabelle 8 und Tabelle 10), ist es ebenso nur sinnvoll, die Mediationseffekte ausschließlich für die Zustimmungsnorm zu berechnen. Als Indikator für die Zustimmungsnorm wurde aus den neu gebildeten Indikatoren vor allem der Index der 5er-Likert-Skalen mit Zustimmungslabel gewählt, da sich für diesen in den vorangegangenen Analysen die stärksten Effekte in Bezug auf die Kulturdimensionen einerseits und dem Arbeitszufriedenheitsniveau andererseits gezeigt hatten. Zur weiteren Absicherung wurden jedoch auch die länderbezogenen Zustimmungswerte aus der Studie von Harzing (2006) in die Berechnungen mit einbezogen. Insgesamt wurden jeweils alle Kombinationen aus den verschiedenen Kulturdimensionen, den beiden Zustimmungsnormen sowie den zwei Arbeitszufriedenheitsniveaus des World-Values-Surveys und des Eurobarometers getrennt voneinander berechnet.

9.4.2 Ergebnisse

Die erste Mediationsanalyse anhand des Individualismus-Kennwerts von Hofstede zeigt, daß der angenommene mediierende Zusammenhang der drei Konstrukte durchaus bestätigt werden kann (vgl. Abbildung 18). Der totale Effekt von Individualismus auf das Arbeitszufriedenheitsniveau ($\beta = .39$, $p = .08$) sank durch den Einbezug des Mediators auf ein deutlich niedrigeres, nicht signifikantes Niveau ($\beta = .20$, $p = .39$). Zusätzlich zeigten sich deutliche Effekte auf dem mediierenden Pfad. Individualismus hatte einen signifikanten Effekt auf die Zustimmungsnorm ($\beta = -.49$, $p = .03$) und diese Norm zeigte wiederum einen deutlichen Effekt auf das Arbeitszufriedenheitsniveau ($\beta = -.39$, $p = .11$). Zwar muß korrekterweise hinzugefügt werden, daß der Effekt der Zustimmungsnorm auf das Zufriedenheitsniveau nicht ganz auf einem Level von zumindest $p \leq .10$ signifikant war. Jedoch sollte bedacht werden, daß es sich aufgrund der Eingrenzung auf die reichen Länder um eine verkleinerte Stichprobe von nur 21 Ländern handelte. Da eine große umfangreiche Stichprobe in solchen Analysen auf dem

Länderniveau grundsätzlich nur schwer zu realisieren ist, sollte dieses Ergebnis daher nicht unbeachtet beiseite gelegt werden, sondern hingegen schon einmal als erster Indikator für die Bestätigung der Mediationshypothese angesehen werden. Hinsichtlich der einzelnen Einflußstärken fällt vor allem auf, daß der Effekt des Mediators auf die abhängige Variable mit $\beta = -.39$ deutlich größer war als der resultierende Effekt mit $\beta = .20$ der unabhängigen Variable auf die abhängige Variable. Dieses Ergebnis kann daher durchaus als Mediationseffekt angesehen werden und als ein erstes Resultat in Richtung des in der Hypothese 11 angenommenen Zusammenhangs gedeutet werden. Die Zustimmungsnorm kann gemäß diesem ersten Ergebnis somit fast vollständig den Effekt von Individualismus auf das Arbeitszufriedenheitsniveau medieren.

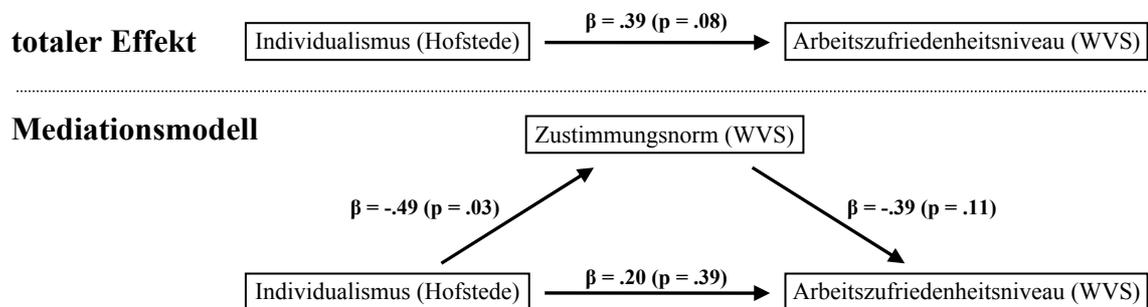


Abbildung 18: Mediationsanalyse mit Individualismus (Hofstede) in reichen Ländern; N=21

Um diesen gefundenen Zusammenhang jedoch methodisch abzusichern, bot es sich an, die gleichen Analysen mit weiteren, inhaltlich ähnlichen, jedoch unabhängigen Dimensionen zu wiederholen. Sollten beispielsweise die Individualismus-Indikatoren der GLOBE-Studie oder des Triandis-Ratings ähnliche Zusammenhänge zu Tage bringen, müßte dies als Bestätigung des hier in Teilen leider nicht komplett signifikanten Zusammenhangs verstanden werden.

Die Ergebnisse anhand der Dimension InGroup-Kollektivismus der GLOBE-Studie zeigen in der Tat ähnliche Ergebnisse wie die Dimension von Hofstede (vgl. Abbildung 19). Auch hier fiel der anfänglich signifikante totale Effekt der Kultur-Dimension auf das Arbeitszufriedenheitsniveau ($\beta = -.51, p = .03$) auf einen deutlich niedrigeren, nicht signifikanten Effekt im Mediationsmodell ($\beta = -.32, p = .25$). Der Einfluß der Kulturdimension auf die Zustimmungsnorm blieb auch hier deutlich bestehen ($\beta = .59, p < .01$), genauso wie es einen sichtbaren Zusammenhang zwischen der Zustimmungsnorm und dem Arbeitszufriedenheitsniveau gab ($\beta = -.33, p = .23$). Jedoch

muß hier ebenfalls angemerkt werden, daß der Einfluß der Zustimmungsnorm aufgrund der kleinen Stichprobe die Signifikanzgrenze nicht ganz erreichte. Dennoch zeigt sich ebenfalls anhand der GLOBE-Dimension, daß der bivariate Zusammenhang von InGroup-Kollektivismus und dem Zufriedenheitsniveau deutlich abnimmt und seine Signifikanz verliert, wenn die Zustimmungsnorm als mediierender Faktor mit in das Modell aufgenommen wird. Der Pfad über den Mediator zeigt daher insgesamt erneut deutliche Effekte.

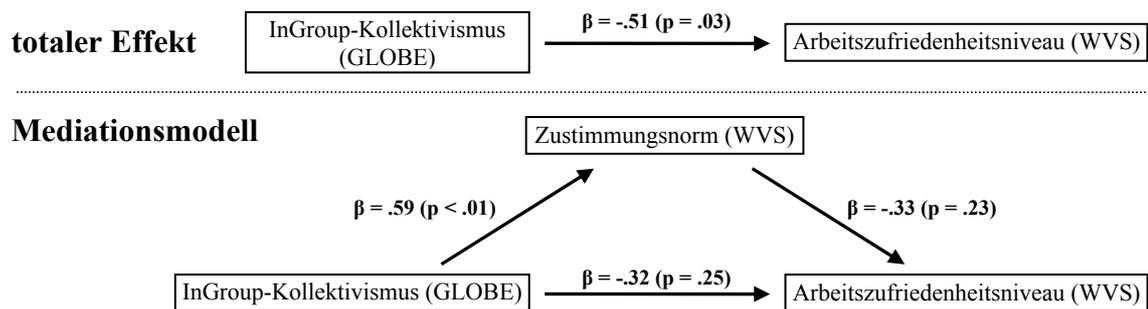


Abbildung 19: Mediationsanalyse mit InGroup-Kollektivismus (GLOBE-Studie) in reichen Ländern; N=18

Betrachtet man die Ergebnisse, die sich in der Mediationsanalyse mit dem Individualismus-Wert von Triandis ergeben, wird deutlich, daß diese Ergebnisse ebenfalls die gleiche Aussage zulassen. Hier zeigte der Pfad über die mediierende Variable nun auch die bisher vermißten, deutlichen Signifikanzen (vgl. Abbildung 20).

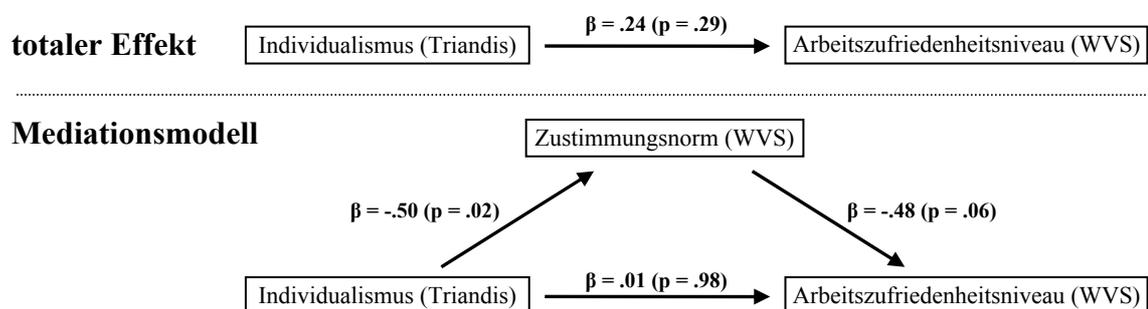


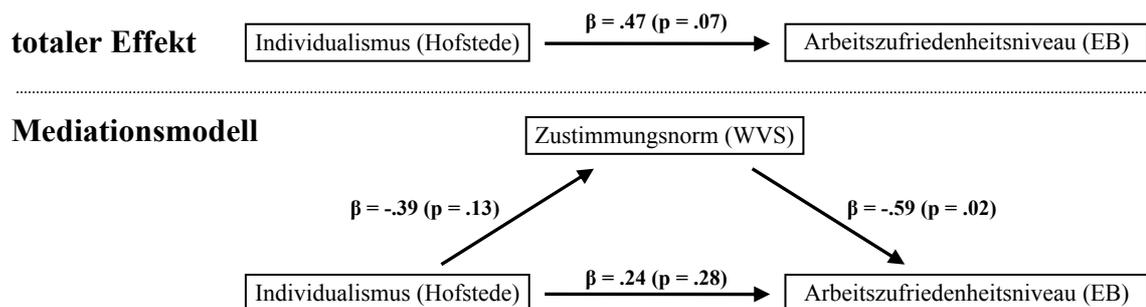
Abbildung 20: Mediationsanalyse mit Individualismus (Triandis) in reichen Ländern; N=21

Individualismus beeinflusste signifikant die Zustimmungsnorm ($\beta = -.50$, $p = .02$), die wiederum mit dem Arbeitszufriedenheitsniveau signifikant in Zusammenhang stand

($\beta = -.48$, $p = .06$). Der Einfluß der Kultur-Dimension auf das Arbeitszufriedenheitsniveau verschwand im Mediationsmodell mit dem Individualismus-Index von Triandis dann auch vollständig ($\beta = .01$, $p = .98$). Einschränkend muß jedoch hier erwähnt werden, daß der anfängliche totale Effekt hier auch nicht besonders groß war ($\beta = .24$, $p = .29$). Dennoch läßt sich festhalten, daß auch dieser dritte Indikator für die Individualismus-Unterschiede einen deutlichen Hinweis auf das Vorhandensein eines Mediationseffektes darstellt.

Ein weiteres bestätigendes Ergebnis erhält man jedoch auch, wenn nicht die Kultur-Dimension variiert wurde, sondern das Maß für das Arbeitszufriedenheitsniveau. Die Mediationsanalysen anhand des Zufriedenheitsniveaus aus dem Eurobarometer zeigen genau das gleiche Muster (vgl. Abbildung 21), das sich schon mit den Daten des World-Values-Surveys ergab. In allen drei Analysen, sowohl mit dem Hofstede- und dem Triandis-Maß als auch mit der GLOBE-Dimension verringerte der Einfluß der Zustimmungsnorm auf das Arbeitszufriedenheitsniveau deutlich den Effekt der Kultur-Dimension und konnte somit auch hier den Zusammenhang von Kultur und Zufriedenheitsniveau mediiieren. Darüber hinaus war zusätzlich bei dieser Analyse das Single-Source-Problem komplett ausgeschaltet, da alle drei Kennwerte aus voneinander unabhängigen Studien stammten.

In allen drei Analysen anhand des Zufriedenheitsniveaus des EB hatte die Zustimmungsnorm nicht nur einen deutlichen Effekt auf das Zufriedenheitsniveau, sondern war im Gegensatz zu den vorhergehenden drei Analysen anhand der WVS-Daten auch eindeutig signifikant. Der folglich nun geringere Einfluß der Kultur war dabei in zwei der drei Analysen (mit Ausnahme bei der GLOBE-Dimension) auch nicht mehr signifikant, wenn die Zustimmungsnorm zusätzlich berücksichtigt wurde. Anhand der Eurobarometer-Daten läßt sich daher insgesamt ebenfalls eindrucksvoll bestätigen, daß eine Zustimmungsnorm in der Tat als ein vermittelnder Faktor zwischen Kultur und dem Arbeitszufriedenheitsniveau angesehen werden kann.



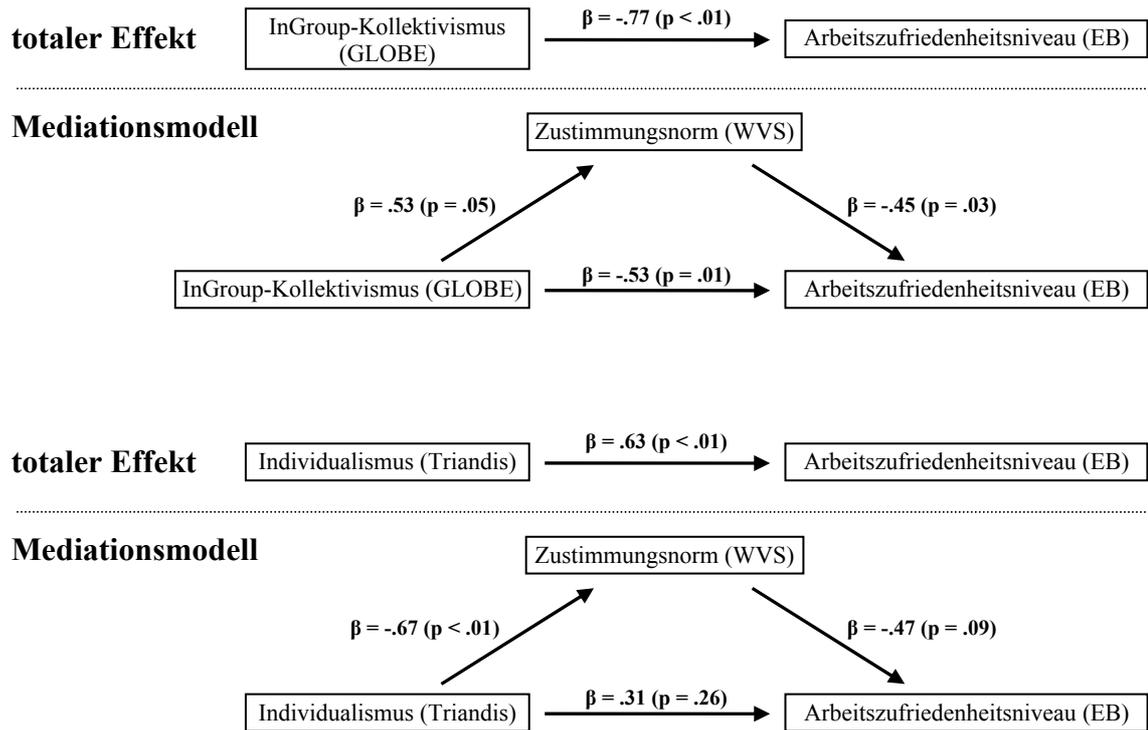


Abbildung 21: Mediationsanalysen mit Individualismus-Dimensionen und Zustimmungsnorm anhand der Eurobarometer-Daten, $N_1 = 16 / N_2 = 14 / N_3 = 16$

Als dritte Form der Variation konnte schließlich auch der Index der Zustimmungsnorm variiert werden. Dabei zeigen die Mediationsanalysen anhand der Zustimmungsnorm von Harzing (2006) ähnliche Muster von Effekten wie die vorherigen Analysen. Bei Berücksichtigung der Hofstede-Dimension im Mediationsmodell (vgl. Abbildung 22) verringerte auch hier der Einfluß der Zustimmungsnorm auf das Arbeitszufriedenheitsniveau den Effekt der Kultur-Dimension von $\beta = .37$ ($p = .24$) auf $\beta = .23$ ($p = .42$). Zwar war hier der Einfluß von Kultur auf die Zustimmungsnorm relativ niedrig ($\beta = -.27$, $p = .42$), jedoch kann insgesamt von einer Teil-Mediation ausgegangen werden.

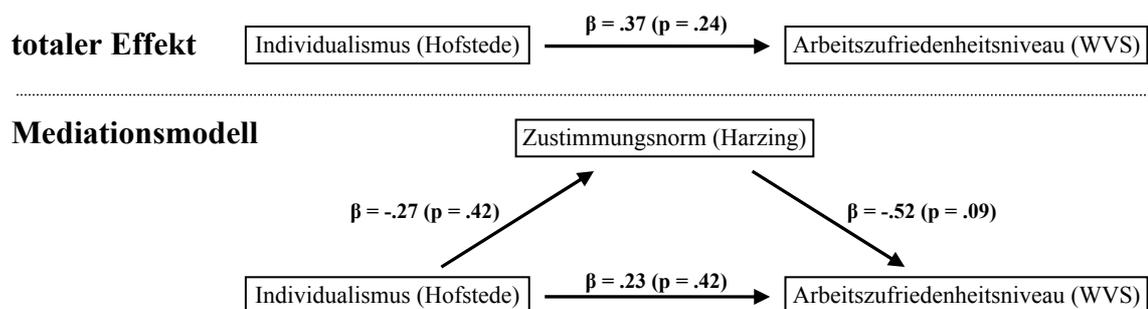


Abbildung 22: Mediationsanalyse mit Individualismus (Hofstede) und Zustimmungsnorm (Harzing) in reichen Ländern, $N=12$

Bei der erneuten Überprüfung des Einflusses des Harzing-Indikators auch anhand der Eurobarometer-Daten zeigte sich ebenfalls der erwartete Effekt, hier sogar in noch deutlicherer Form (vgl. Abbildung 23). Der Pfad über den Mediator war insgesamt stärker als der direkte Einfluß von Kultur auf das Zufriedenheitsniveau. Auch hier übte somit im Mediationsmodell – wie deutlich zu sehen ist – die Zustimmungsnorm den deutlich stärkeren Einfluß auf das Arbeitszufriedenheitsniveau im Vergleich zur Kulturdimension aus.

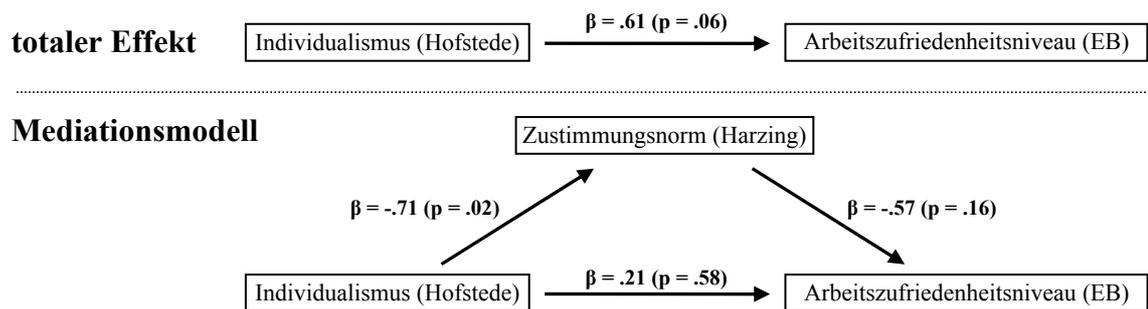


Abbildung 23: Mediationsanalyse mit Individualismus (Hofstede) und Zustimmungsnorm (Harzing) anhand der Eurobarometer-Daten, N=10

Faßt man die hier dargestellten Ergebnisse zusammen, dann zeigt sich anhand vieler unterschiedlicher Berechnungen der erwartete mediierende Zusammenhang: Die Berücksichtigung der Zustimmungsnorm in den Analysen führte insgesamt zu einer zum Teil deutlichen Verringerung des Einflusses der Kulturdimension. Zwar konnten nicht alle gefundenen Effekte komplett signifikant abgesichert werden, jedoch sollte bei solchen Analysen auf Länderebene stärker die Effektgröße als das reine Signifikanzniveau beachtet werden, da umfangreiche Länderstichproben nur schwer zu generieren sind und sie aufgrund der Tatsache, daß in Bezug zu den vorangegangenen Ergebnissen die Analysen nur in reichen Ländern durchgeführt wurden, automatisch in ihrer Größe beschränkt waren.

In der Regel jedoch besaß im Mediationsmodell trotz zum Teil fehlender Signifikanzen einerseits die Kulturdimension einen deutlichen Effekt auf die Zustimmungsnorm, dann besaß die Zustimmungsnorm den größeren Effekt auf das Arbeitszufriedenheitsniveau, und der Effekt der Kulturdimension auf das Arbeitszufriedenheitsniveau verringerte sich ebenfalls deutlich beim Vergleich des totalen Effekts zum Mediationsmodell hin. Die Vorgaben von Baron und Kenny (1986) wurden somit erfüllt.

Der Vollständigkeit halber sollte jedoch erwähnt werden, daß bei einer Variation von drei Kulturdimensionen, zwei Indizes der Zustimmungsnorm und zwei Maßen des Arbeitszufriedenheitsniveaus sich insgesamt 12 verschiedene Mediationsanalysen berechnen lassen. Bei den weiteren vier, hier nicht explizit dargestellten Analysen zeigte sich bei zwei Mediationsmodellen nicht komplett, aber zumindest in Teilen der erwartete Effekt (vgl. Anhang B). Jedoch konnten auch zwei Modelle keinen Mediationseffekt nachweisen. Insgesamt aber kann zusammenfassend festgehalten werden, daß 8 der insgesamt 12 möglichen Analysen zum Teil deutliche Effekte hinsichtlich einer Zustimmungsnorm als Mediator zeigten. Zustimmungsnormen sind somit durchaus in der Lage, als erklärender Faktor an die Stelle der bisherigen Blackbox zwischen Individualismus/Kollektivismus und dem Arbeitszufriedenheitsniveau zu treten. Die Hypothese 11 kann daher weitestgehend bestätigt werden.

Neben Individualismus/Kollektivismus korrelierte in den vorherigen bivariaten Analysen auch die Kulturdimension der Machtdistanz sowohl mit den Zustimmungsnormen als auch mit dem Arbeitszufriedenheitsniveau bedeutsam. Zwar konnte auf Basis der bisherigen Forschung keine eindeutige Hypothese für einen Mediationseffekt bzgl. Machtdistanz aufgestellt werden. Dennoch ist auch ein Ziel dieser Studie, die Ergebnisse hinsichtlich dieses Zusammenhangs explorativ zu überprüfen. Die Mediationsanalysen, die auch hier für jede mögliche Kombination aus zwei Kulturdimensionen (Hofstede, GLOBE), zwei Zustimmungsnorm-Indizes (WVS, Harzing) und zwei Zufriedenheitsniveaus (WVS, EB) getrennt berechnet wurden, zeigen jedoch, daß insgesamt von keinem mediiierenden Effekt der Zustimmungsnorm hinsichtlich der Machtdistanz, die in einem Land vorherrscht, ausgegangen werden kann (vgl. für eine beispielhafte Analyse Abbildung 24, für die weiteren sieben möglichen Analysen siehe Anhang B).

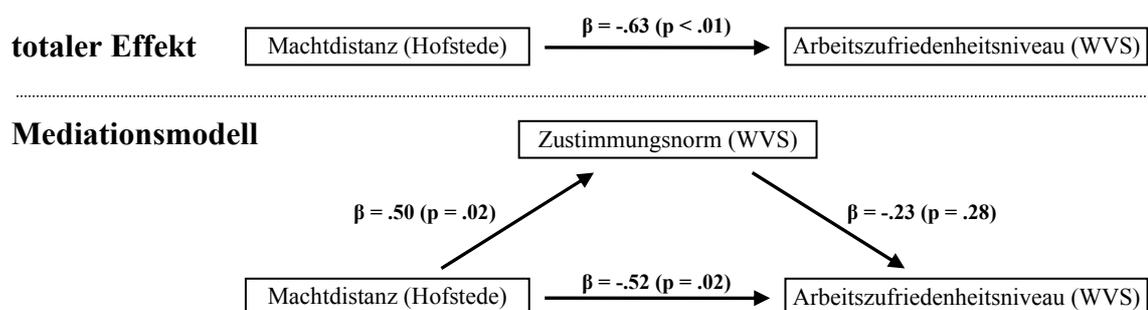


Abbildung 24: Mediationsanalyse mit Machtdistanz (Hofstede) in reichen Ländern, N=21

In allen Analysen, die ebenfalls wieder nur mit den reichen Ländern durchgeführt wurden, zeigte sich das gleiche Bild, welches beispielhaft hier nur am Zusammenhang der Machtdistanz nach Hofstede, dem Zustimmungsnorm-Index des WVS und dem Arbeitszufriedenheitsniveau aus dem WVS dargestellt ist. Der starke Zusammenhang zwischen der Kultur-Dimension und dem Zufriedenheitsniveau, der im totalen Effekt zum Ausdruck kam, blieb auch im Mediationsmodell bestehen, in dem zusätzlich noch der Einfluß der Zustimmungsnorm berücksichtigt wurde. Der Einfluß der Kultur blieb dabei nicht nur in diesen Modellen zumeist signifikant, sondern war in der Regel auch höher als der Einfluß der Zustimmungsnorm auf das Zufriedenheitsniveau. Letzterer war zwar in Teilen auch vorhanden, jedoch insgesamt eher schwach ausgeprägt. Dieser zumindest schwache Zusammenhang zwischen Zustimmungsnorm und Arbeitszufriedenheitsniveau sorgte dann immerhin dafür, daß der Zusammenhang von Kultur und Arbeitszufriedenheitsniveau im Mediationsmodell sich etwas gegenüber dem totalen Effekt verringerte. Andererseits konnte jedoch aufgrund des immer noch starken Zusammenhangs der Kulturdimension mit dem Zufriedenheitsniveau nicht von einem Mediationseffekt ausgegangen werden. In Kulturen, die von großen Machtunterschieden geprägt sind, scheint es demnach zusätzliche Aspekte neben dem sozialisierten Kommunikationsverhalten zu geben, die einen eigenständigen und sogar insgesamt stärkeren Einfluß auf das von allen Befragten im Mittel berichtete Zufriedenheitsniveau haben. Zu denken wäre hier möglicherweise daran, daß sich im Mittelwert der Zufriedenheit in diesen Kulturen in der Tat stärker ein Teil der ‚wahren‘ Zufriedenheit zeigt, der in Kulturen mit hoher Machtdistanz möglicherweise deshalb tendenziell niedriger sein könnte, da Mitarbeiter sich vielleicht stärker wünschen, auch ihrem Chef einmal die Meinung sagen zu dürfen. Genauso gut könnte es aber auch sein, daß ein gewisses Maß an sozialem Miteinander auch im Arbeitsleben für die Zufriedenheit wichtig ist, welches durch eine hohe Form von Machtdistanz jedoch eher verhindert würde. Diese beiden Gedanken können jedoch nur erste Überlegungen hinsichtlich der nicht bestätigten Mediation sein. Die Gründe für den auch im Mediationsmodell immer noch sehr starken Zusammenhang von Machtdistanz und Zufriedenheit trotz der Kontrolle von Antworttendenzen müßten in zukünftigen Studien dann noch genauer beleuchtet und überprüft werden.

Abschließend wurden ebenfalls auch Mediationsanalysen mit der Kulturdimension Unsicherheitsvermeidung durchgeführt, da auch diese Dimension deutliche

Zusammenhänge in den vorherigen bivariaten Analysen gezeigt hatte. Auch bei dieser Kulturdimension zeigte sich jedoch ein ähnliches Bild wie bei den Analysen anhand der Machtdistanz-Dimension (siehe Anhang B). Es ergaben sich zwar teilweise gewisse Effektstärken bzgl. des Einflusses der Zustimmungsnorm auf das Arbeitszufriedenheitsniveau, jedoch blieb der stark negative Zusammenhang von Unsicherheitsvermeidung und Arbeitszufriedenheit auch in der Mediationsanalyse bestehen und war in der Regel größer als der Einfluß der Zustimmungsnorm auf das Zufriedenheitsniveau. Auch hier kann somit nicht von einem Mediationseffekt ausgegangen werden und es gelten die gleichen Überlegungen wie beim Einfluß von Machtdistanz. Scheinbar muß es weitere Faktoren geben, die vielleicht erklären können, warum das Zufriedenheitsniveau in Ländern mit dem Wunsch nach hoher Unsicherheitsvermeidung im Mittel niedriger ist. Zukünftige Studien könnten auch hier gegebenenfalls weiteren Aufschluß geben.

Gemäß den Überlegungen zu Hypothese 11 war ebenfalls angedacht, auch den Zusammenhang von Kultur, d.h. speziell Individualismus, mit der Extremtendenz-Norm und dem Arbeitszufriedenheitsniveau zu überprüfen. Eine Mediationsanalyse dieser drei Aspekte erübrigte sich jedoch, da weder ein klarer Bezug von Extremtendenzen und Kulturdimensionen beobachtet werden konnte (vgl. Tabelle 8), noch standen Extremtendenz-Normen in der hier vorliegenden Studie in einem bedeutsamen Bezug zum Arbeitszufriedenheitsniveau (vgl. Tabelle 10). Die Voraussetzungen für eine mögliche Mediation waren somit schon nicht mehr gegeben, weshalb die entsprechenden Mediationsanalysen dann auch nicht mehr durchgeführt wurden.

Auch wenn eine Analyse der Mediationseffekte in den sogenannten armen Ländern aufgrund des vorher nicht nachzuweisenden Zusammenhangs von Zustimmungsnorm und Zufriedenheitsniveau eigentlich nicht mehr sinnvoll war, wurden die Mediationsanalysen exemplarisch dennoch für diese Ländergruppe zumindest hinsichtlich der Kombination aus Individualismus/Kollektivismus und Zustimmungsnorm berechnet (vgl. Anhang B). Die Ergebnisse bestätigen jedoch die vorangegangenen Vermutungen. Ein Mediationseffekt kann bei diesen Ländern nicht nachgewiesen werden. Ein weiteres Nebenprodukt der Analysen und somit ein kleiner Seitenaspekt der Ergebnisse soll dennoch hier eine kurze Erwähnung finden. Bei den durchgeführten Berechnungen der Mediationsanalysen in den armen Ländern fiel

nämlich auf, daß interessanterweise nicht nur wie erwartet der Zusammenhang von Antworttendenz und Zufriedenheitsniveau in diesen Ländern nicht existierte, sondern unerwarteterweise auch kein eindeutiger Zusammenhang mehr zwischen den jeweils anderen beiden Faktoren sich zeigte. Die in der bisherigen Forschung gefundenen Zusammenhänge zwischen Individualismus und Zustimmungsnorm und Individualismus und Zufriedenheitsniveau, die auch hier in dieser Studie anhand der am Gesamtländerdatensatz durchgeführten Berechnungen durchaus belegt werden konnten, fielen überraschenderweise in vielen Fällen weg, wenn man nur die ökonomisch schwachen Länder betrachtete. Obwohl somit eigentlich von gut gesicherten Effekten bzgl. dieser Aspekte ausgegangen werden konnte, scheint es Unterschiede zwischen diesen beiden Gruppen von Ländern zu geben. Inwieweit dies möglicherweise eine grundsätzliche Bedeutung für die Forschung speziell in ökonomisch schwächeren Ländern haben könnte, wird im abschließenden Diskussionskapitel beleuchtet.

9.4.3 Zusammenfassung der Ergebnisse der Mediationsanalysen

Versucht man nun die Resultate aller Mediationsanalysen zusammenzufassen, wird deutlich, daß vor allem ein klares Ergebnis aus den zahlreichen Analysen resultiert: Zustimmungstendenzen können durchaus als Mediator verstanden werden und bilden somit einen verbindenden Faktor, der die Korrelation von Kultur und Arbeitszufriedenheitsniveau besser erklären kann. Dies gilt interessanterweise jedoch nur für die Kulturdimension Individualismus/Kollektivismus. Diese Dimension ist aber auch die am besten in der Literatur untersuchte und bestätigte Dimension, so daß sich auch vorab nur hinsichtlich dieser Dimension eine klare Mediationshypothese aufstellen ließ. Diese Hypothese kann in der Tat mehrheitlich bestätigt werden. Zwar hätte man sich auch eine Mediation hinsichtlich der Kulturaspekte von Machtdistanz und Unsicherheitsvermeidung unter Umständen vorstellen können, jedoch war diesbezüglich die bisherige Befundlage in der Literatur schon nicht eindeutig, so daß auch keine klaren Hypothesen vorab formuliert wurden. Hinsichtlich der Individualismus/Kollektivismus-Dimension jedoch sieht das Bild anders aus. Die Ergebnisse diesbezüglich können demnach die Überlegungen bezüglich der angenommenen Unterschiede zwischen individualistischen und kollektivistischen Kulturen (vgl. vor allem Markus & Kitayama, 1991) anhand der hier vorgebrachten

empirischen Daten untermauern. Kollektivistische Kulturen sind folglich durch eine hohe kommunikative Zurückhaltung zu charakterisieren. Diese äußert sich einerseits in einer starken Zustimmungstendenz, die dadurch entsteht, daß Menschen in dieser Art von Kultur in ihrer Sozialisation lernen, ihre eigene Meinung nicht zu stark in den Vordergrund zu spielen. Stattdessen wird verlangt, sich eher bescheiden und zurückhaltend zu äußern. Andererseits zeigt sich diese Form der kommunikativen Zurückhaltung, wie an den hier vorgelegten Ergebnissen zu sehen, auch bei der Messung von Arbeitszufriedenheit. Hier jedoch mit genau entgegengesetztem Effekt. Dieser Effekt, der auf den ersten Blick vielleicht überraschend war, läßt sich jedoch dann problemlos in Übereinkunft mit den Überlegungen von Markus und Kitayama (1991) bringen, wenn man dem Gedanken folgt, daß eigene Leistungen und eigener Stolz auf persönlich Erreichtes in kollektivistischen Kulturen nicht zur Schau gestellt werden bzw. werden sollen. Da eine Aussage zur eigenen Zufriedenheit mit der Arbeit durchaus als Urteil über das eigene Fortkommen im Job verstanden werden kann, läßt sich auch plausibel nachvollziehen, warum dieses Urteil im Mittel in kollektivistischen Kulturen, die von einem hohen Maß an Zustimmungsnorm geprägt sind, niedriger ist. Der Stolz auf den eigenen Job würde somit nur in individualistischen Kulturen herausgestellt, jedoch in kollektivistischen Kulturen eher zurückhaltend gezeigt.

Die hier in den Mediationsanalysen gefundenen Ergebnisse fungieren darüber hinaus ebenfalls als eine Erklärung für den in der Forschung bisher nicht klar erläuterten Zusammenhang, daß das Arbeitszufriedenheitsniveau in individualistischen Staaten höher und in kollektivistischen Staaten geringer ist. Als Grund kann folglich an Stelle der bisherigen Blackbox und ihrer weitestgehend unbestätigten Annahmen nun das Kommunikationsverhalten treten, welches jeweils in den einzelnen Kulturen erwünscht ist. Nicht die Kultur an sich führt in irgendeiner Form dazu, daß das Arbeitszufriedenheitsniveau höher oder niedriger in einem Land ist, sondern vielmehr ist dies in Teilen ein Effekt des in der Kultur sozialisierten Kommunikationsverhaltens. Sieht man jedoch dieses Verhalten nun als einen Grund für Unterschiede im Zufriedenheitsniveau an, ist dies dann auch gleichbedeutend mit der Erkenntnis, daß man eigentlich keine genaue Aussage mehr über den wahren Zufriedenheitsgrad der Befragten machen kann. Diese Einschränkung gilt vor allem in kollektivistischen Kulturen. Gerade in diesen Kulturen kreuzen die Befragten zwar auch eine Box der Likert-Skala an bzw. geben gegenüber dem Interview ein entsprechendes Urteil auf der Likert-Skala ab. Ihre wahre Zufriedenheit, die sie innerlich verspüren, mag jedoch

komplett anders ausgeprägt sein. Nach außen hin scheint es für Mitglieder dieser Kulturen jedoch aufgrund von Kommunikationsnormen nur angemessen zu sein, ein eher gemäßigteres Urteil abzugeben.

Daß die Arbeitsplatzqualität dennoch einen Einfluß auf das Arbeitszufriedenheitsniveau zu haben scheint und im Zufriedenheitsurteil selbstverständlich auch immer noch ein wahrer Kern der Zufriedenheit trotz der Überformung durch Antworttendenz-Normen enthalten ist, wird dadurch deutlich, daß sich sowohl die bivariaten als auch die mediierenden Zusammenhängen von Zustimmungsnorm und Arbeitszufriedenheitsniveau nur in Ländern mit höherem Wohlstand und besseren ökonomischen Bedingungen wiederfinden lassen. Wenn in Ländern mit schlechteren Bedingungen sich wie in den hier durchgeführten Analysen kein Einfluß der Zustimmungsnorm nachweisen läßt, könnte dies hier möglicherweise so interpretiert werden, daß scheinbar erst ein gewisses Niveau der Arbeitsplatzqualität erreicht werden muß, ab dem dann Kommunikationsnormen erst auf das Zufriedenheitsurteil durchschlagen und eine Verzerrung des in den Ländern berichteten Zufriedenheitsniveaus bewirken. Demnach könnte man hier erst einmal annehmen, daß Kommunikationsnormen erst dann wirken, wenn sozusagen die Grundbedürfnisse der Mitarbeiter erfüllt sind.

9.5 Die Mehrebenenanalyse

Die bisherigen Analysen der vorherigen Kapitel befaßten sich mit dem Zusammenhang von Antworttendenzen und der Arbeitszufriedenheit auf dem Länderniveau. Geprüft wurde somit, ob das Zufriedenheitsniveau (d.h. wie zufrieden beschreiben sich die Befragten im Durchschnitt) mit kulturellen Aspekten oder weiteren länderbezogenen Kennwerten zu erklären ist. Dazu wurden unter anderem für die Antworttendenzen und die Zufriedenheit Indikatoren eingesetzt, die aus gemittelten Individualdaten errechnet wurden. Die mittlere Antworttendenz aller Befragten eines Landes wurde dabei als ein Ausdruck des generell in einem Land akzeptierten Antwortverhaltens verstanden und somit als eine Antworttendenz-Norm interpretiert. Hinsichtlich einer dieser Normen, der Zustimmungsnorm, zeigte sich dabei, daß diese als ein Grund für die niedrigere Zufriedenheit angesehen werden kann, die von den Befragten im Durchschnitt berichtete wurde. Inwieweit jedoch das individuell unterschiedliche Antwortverhalten einzelner Personen ebenfalls in einem Bezug zur dann individuellen

Arbeitszufriedenheit der jeweiligen Person steht, kann aus diesen Analysen nicht geschlossen werden. Von einer Analyse auf dem Länderniveau kann nicht auf einen Zusammenhang auf dem Individualniveau geschlossen werden (Nezlek, Schröder-Abé & Schütz, 2006). Aus methodischer und mathematischer Sicht wäre es ohne Probleme möglich, daß auf dem individuellen Niveau ebenfalls ein positiver oder aber ein negativer oder auch überhaupt kein Zusammenhang zwischen beiden Konstrukten besteht.

Gemäß den Überlegungen zu Hypothese 10 wurde jedoch angenommen, daß die individuelle Tendenz eines Befragten zu zustimmendem Antwortverhalten sich in einem höheren Zufriedenheitsurteil desjenigen niederschlagen sollte und somit entgegengesetzt zum nationalen Einfluß wirken würde. Um diesen Zusammenhang zu prüfen, wurde eine Mehrebenenanalyse (Bryk & Raudenbush, 1992; Hox, 1995; Engel, 1998; Nezlek, Schröder-Abé & Schütz, 2006) mit Hilfe des Programms HLM 6 durchgeführt. Eine klassische Regressionsanalyse aller Individualdaten wäre in diesem Fall nicht angemessen gewesen, da es sich bei den hier zur Verfügung stehenden Daten um eine hierarchische Datenstruktur handelt, bei der die jeweiligen Individualdaten innerhalb der Länder genestet sind. Zwar hätte man möglicherweise eine Regressionsanalyse jeweils in jedem Land gesondert durchführen können, jedoch kann dann die Unterschiedlichkeit zwischen den Ländern nur schwer betrachtet und in der Interpretation berücksichtigt werden. Die Mehrebenenanalyse ist hingegen in der Lage, sowohl die Einflüsse länderbezogener Kennwerte als auch die Einflüsse individualpsychologischer Variablen gleichzeitig zu berücksichtigen. Sie vermeidet darüber hinaus die weiteren Nachteile klassischer Regressionsanalysen, die auf unterschiedlichen Datenleveln mit Hilfe von Aggregation und Disaggregation der Daten oftmals durchgeführt werden. Ein bedeutsamer Nachteil besteht darin, daß beispielsweise durch eine Aggregation von Individualdaten hin zu Länderkennwerten immer ein Großteil der in den Daten vorhandenen Varianz vernichtet wird. Stattdessen ist die Mehrebenenanalyse in der Lage, die Varianz der Individualdaten in einen Teil, der durch Länderunterschiede erklärt werden kann, und einen Teil, der durch individuelle Unterschiede hervorgerufen wird, aufzuteilen. Mit Hilfe der Mehrebenenanalyse besteht daher insgesamt die Möglichkeit, den Zusammenhang, der in den vorangegangenen bivariaten Analysen der Länderdaten gefunden wurde, anhand eines weiteren statistischen Verfahrens zu replizieren und gleichzeitig eine Aussage zu individualpsychologischen Zusammenhängen treffen zu können.

9.5.1 Vorgehen der Mehrebenenanalyse

In einer Mehrebenenanalyse, die sich oftmals – wie auch im hier vorliegenden Fall – auf zwei Datenebenen beschränkt, werden für jede Ebene einzelne Regressionsgleichungen in ein und derselben Analyse berechnet. Eine ausführliche Erläuterung der Vorgehensweise findet sich unter anderem bei Bryk und Raudenbush (1992), Hox (1995) oder aber Engel (1998). Hier sollen für ein besseres Verständnis der in dieser Studie durchgeführten Analyse nur die Grundzüge der Methode, übertragen auf die inhaltliche Fragestellung dieser Arbeit, dargestellt werden. Im einfachsten Modell einer Mehrebenenanalyse werden Unterschiede in den Gruppenmittelwerten anhand folgender Gleichungen dargestellt:

$$\text{Ebene 1 (Individuallevel): } y_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Ebene 2 (Länderniveau): } \beta_{0j} = \gamma_{00} + u_{0j}$$

Die individuelle Ausprägung einer Variable y (in der vorliegenden Studie entspricht dies der Arbeitszufriedenheit) wird dabei für i Individuen in j Ländern ermittelt. Die Ausprägung der abhängigen Variable y errechnet sich aus der Regressionskonstante β_{0j} , die dem Mittelwert von y in jedem Land j (also dem jeweiligen Arbeitszufriedenheitsniveau) entspricht, und kann anhand eines individuenbezogenen Fehlerterms r variieren. Der Ländermittelwert β_{0j} wird dann in der zweiten Gleichung durch den Gesamtmittelwert γ_{00} und eines länderbezogenen Fehlerterms u_{0j} modelliert und kann somit für jedes Land unterschiedlich sein. Die Varianz von r_{ij} entspricht dabei der Varianz der abhängigen Variable auf dem ersten Level (Individualniveau), während u_{0j} die Varianz der abhängigen Variable auf dem 2. Level (Länderniveau) darstellt. Dieses sogenannte Basismodell wird gewöhnlich als Ausgangspunkt der Analysen herangezogen und beschreibt lediglich eine Unterschiedlichkeit der Mittelwerte in den einzelnen Ländern. Es wird zusätzlich zur Berechnung des Intra-Class-Coefficient (ICC) benötigt, der anhand folgender Formel angibt, wieviel der Gesamtvarianz der abhängigen Individualvariable auf die unterschiedlichen Ebenen zurückgeführt werden kann:

$$\text{ICC} = \sigma_0^2 / (\sigma_0^2 + \sigma_r^2) \quad \text{mit} \quad \sigma_0^2 = \text{Varianz des Fehlerterms } u_{0j}$$
$$\sigma_r^2 = \text{Varianz des Fehlerterms } r_{ij}$$

Dieses Basismodell erklärt jedoch die vorhandene Varianz noch nicht, sondern teilt sie lediglich in die individuen- und die länderbezogene Varianz auf. In der Regel will man aber die Unterschiedlichkeit anhand spezifischer Einflußgrößen erklären, z.B. wie in

dieser Studie die Variabilität des Zufriedenheitsniveaus in Abhängigkeit von Zustimmungsnormen. Dieses würde in folgendem Modell zum Ausdruck kommen:

$$\text{Ebene 1 (Individuallevel): } y_{ij} = \beta_{0j} + r_{ij}$$

$$\text{Ebene 2 (Länderniveau): } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Zustimmungsnorm}) + u_{0j}$$

In diesem Modell wird auf Ebene 2 nun zusätzlich der Ländermittelwert der Zufriedenheit durch die jeweils unterschiedliche Zustimmungsnorm des Landes erklärt. Wenn man nun die unterschiedliche Zufriedenheit y_{ij} der Individuen ebenfalls anhand der individuellen Zustimmungstendenz erklären möchte, würde ein weiterer Prädiktor hinzugefügt, diesmal aber auf Ebene 1:

$$\text{Ebene 1 (Individuallevel): } y_{ij} = \beta_{0j} + \beta_{1j} (\text{Zustimmungstendenz}) + r_{ij}$$

$$\text{Ebene 2 (Länderniveau): } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Zustimmungsnorm}) + u_{0j}$$

Die individuelle Zufriedenheit y_{ij} wird somit nun anhand des Ländermittelwerts, der individuellen Zustimmungstendenzen und des Fehlerterms des Individuallevels modelliert, wobei der Ländermittelwert wiederum aufgrund der länderbezogenen Zustimmungsnorm und des Fehlerterms des Länderlevels variieren kann.

Nun kann es aber durchaus sein, daß die Stärke des Einflusses von individuellen Prädiktoren sich je nach Umgebungsbedingungen unterscheidet. Eine Zustimmungstendenz muß nicht zwingend in jedem Land die gleiche Einflußstärke besitzen. Die Stärke von β_{1j} würde somit über die Gruppen hinweg variieren, was durch eine weitere Gleichung auf Ebene 2 zum Ausdruck gebracht werden kann:

$$\text{Ebene 1 (Individuallevel): } y_{ij} = \beta_{0j} + \beta_{1j} (\text{Zustimmungstendenz}) + r_{ij}$$

$$\text{Ebene 2 (Länderniveau): } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Zustimmungsnorm}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + u_{1j}$$

Bei der Darstellung der zweiten Gleichung auf Ebene 2 zur Berechnung von β_{1j} wird anhand des zusätzlichen Fehlerterms u_{1j} jedoch erst einmal nur deutlich, daß β_{1j} über die Gruppen hinweg zufällig variieren kann. Wäre die Stärke des Einflusses – wie in Standard-Regressionsgleichungen grundsätzlich angenommen – in allen Gruppen gleich, wäre $\beta_{1j} = \gamma_{10}$, wobei dann $u_{1j} = 0$ wäre und β_{1j} somit gleichbedeutend mit der Gesamteinflußstärke (γ_{10}) in der gesamten Stichprobe wäre. Will man nun die Unterschiedlichkeit der Einflußstärke der individuellen Zustimmungstendenz aufgrund der Umgebungsbedingungen aber anhand eines konkreten Prädiktors erklären, würde ein entsprechender Prädiktor in die Gleichung der Regressionssteigung β_{1j} auf Ebene 2 eingefügt. Möglicherweise fühlen sich ja Personen, die zur Zustimmung tendieren, gerade in Umwelten, wo eine starke Norm existiert, stärker dazu gedrängt, ihrer

persönlichen Tendenz zu folgen, während dies in einer Umwelt, wo nur eine diffuse Norm existiert, eher geringer der Fall ist. Man würde somit die Unterschiedlichkeit der Einflußstärke der Zustimmungstendenz beispielsweise aufgrund der länderspezifischen Zustimmungsnorm erklären. Dadurch wird automatisch ein Interaktionseffekt zwischen den Prädiktoren der Individual- und der Länderebene (eine sogenannte cross-level-interaction) erzeugt:

$$\text{Ebene 1 (Individuallevel): } y_{ij} = \beta_{0j} + \beta_{1j} (\text{Zustimmungstendenz}) + r_{ij}$$

$$\text{Ebene 2 (Länderniveau): } \beta_{0j} = \gamma_{00} + \gamma_{01} (\text{Zustimmungsnorm}) + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} (\text{Zustimmungsnorm}) + u_{1j}$$

In diesem Gesamtmodell wird nun die individuell unterschiedliche Arbeitszufriedenheit somit durch a) die individuelle Zustimmungstendenz, b) die Zustimmungsnorm eines Landes, c) die Wechselwirkung aus individueller Zustimmungstendenz und Zustimmungsnorm sowie d) einer zufälligen Variabilität erklärt. Dieses Erklärungsmuster wird noch deutlicher dargestellt, wenn man eine andere Form der Darstellung (die single-equation-form) wählt, bei der die drei Gleichungen ineinander eingesetzt werden:

$$y_{ij} = \gamma_{00} + \gamma_{10} (\text{Zustimmungstendenz}) + \gamma_{01} (\text{Zustimmungsnorm}) + \gamma_{11}(\text{Zustimmungsnorm})(\text{Zustimmungstendenz}) + u_{1j}(\text{Zustimmungstendenz}) + u_{0j} + r_{ij}$$

In dieser Gleichung wird der erste Teil (bestehend aus: $\gamma_{00} + \gamma_{10} (\text{Zustimmungstendenz}) + \gamma_{01} (\text{Zustimmungsnorm}) + \gamma_{11}(\text{Zustimmungsnorm})(\text{Zustimmungstendenz})$), als fester Teil des Modells bezeichnet, während der zweite Teil (bestehend aus: $u_{1j}(\text{Zustimmungstendenz}) + u_{0j} + r_{ij}$), als der variable Modell-Teil bezeichnet wird. Die eben beschriebene Interaktion zwischen den beiden Ebenen drückt sich nun deutlicher im Term $\gamma_{11}(\text{Zustimmungsnorm})(\text{Zustimmungstendenz})$ aus, wodurch somit die Zustimmungsnorm als ein Moderator auf den Einfluß der individuellen Zustimmungstendenz in den jeweiligen Ländern wirkt.

Es lassen sich aus Sicht der Mehrebenenanalyse grundsätzlich verschiedene Modelle aufstellen; einerseits nur mit festen, andererseits nur mit variablen oder aber mit beiden Arten von Effekten. Die entsprechenden Terme sollten jedoch immer aufgrund von theoretischen Überlegungen festgesetzt werden, um nicht nur in einer rein empiristischen Vorgehensweise einfach das Modell zu wählen, was die meiste Varianz in den Daten erklärt, sondern zusätzlich die entsprechenden Effekte auch theoretisch

erklären zu können. Die Herleitung des hier eingesetzten Mehrebenenmodells folgte daher den theoretischen Überlegungen dieser Arbeit.

Graphisch darstellen ließe sich der eben beschriebene Zusammenhang, der auch in Hypothese 10 zum Ausdruck kommt, anhand des folgenden Schaubilds (vgl. Abbildung 25), in dem die direkten Effekte anhand der Pfeile deutlich werden und der Interaktionseffekt gestrichelt dargestellt ist.

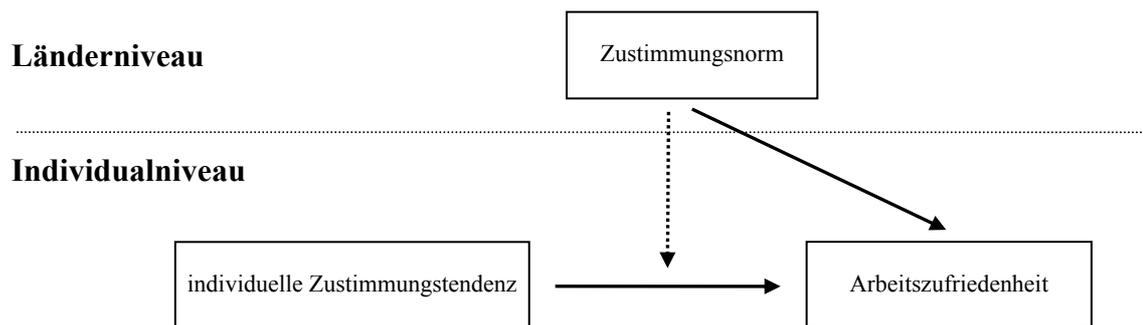


Abbildung 25: Mehrebenenanalytisches Modell des Einflusses von Zustimmungsnorm und Zustimmungstendenz

9.5.2 Stichprobe und Meßgrößen

Zum Einsatz kamen bei der Mehrebenenanalyse erneut die Daten des World-Values-Surveys (WVS). Diesmal gingen in die Berechnung jedoch nicht nur die gemittelten Länderkennwerte ein, sondern die Analyse erfolgte ebenfalls anhand aller Individualdaten. Die Stichprobe umfaßte letztendlich insgesamt die arbeitende Bevölkerung in jedem Land, da nur diesem Teil der Stichprobe die Frage zur Arbeitszufriedenheit gestellt wurde.

Die individuelle Zustimmungstendenz, die Zustimmungsnorm und die individuelle Arbeitszufriedenheit wurden dabei in gleicher Weise wie in den vorangegangenen Analysen operationalisiert und durch folgende Indikatoren dargestellt. Die individuelle Arbeitszufriedenheit wurde gemessen anhand der einzigen im WVS vorhandenen Frage „Overall, how satisfied or dissatisfied are you with your job?“. Die individuelle Zustimmungstendenz wurde auf die gleiche Weise gebildet, wie in Kapitel 9.2 beschrieben, so daß auch hier die Anzahl der zustimmenden Äußerungen im Vergleich zu den ablehnenden Äußerungen, relativiert an der Gesamtzahl der Items, errechnet

wurde. Diese Netto-Zustimmungstendenz (NARS) wurde hier für die 5er-Likert-Skalen mit Zustimmungslabel gebildet, da diese Skalenform gemäß den vorangegangenen Analysen die stärksten Effekte zeigte. Individuell unterschiedliche Kennwerte wurden dann erneut für jedes Land gemittelt, um die jeweilige Zustimmungsnorm zu erhalten. Die Werte dieser Norm entsprachen somit exakt den gleichen Werten wie in den vorherigen bivariaten Berechnungen sowie den Mediationsanalysen. Zusätzlich zu den drei hypothesenrelevanten Indikatoren wurden jedoch in dieser Analyse auf Individuallevel auch das Alter und das Geschlecht als Kontrollvariablen mitberücksichtigt, da sich in der Literatur zeigte, daß diese Aspekte in Zusammenhang mit zustimmenden Antwortverhalten stehen (vgl. Kapitel 6.3).

In den zusätzlichen Analysen zum destruktiven Testen wurden darüber hinaus als länderbezogene Indikatoren der Individualismus-Index von Hofstede und das Bruttoinlandsprodukt gemäß den Daten der Vereinten Nationen aus dem Jahre 2004 mit einbezogen (d.h. ebenfalls die gleichen Indikatoren, die auch in den vorherigen bivariaten Analysen genutzt wurden). Auf dem individuellen Level wurden schließlich noch Daten über das individuelle Einkommen und die personenbezogene Bildung als erklärende Faktoren auf ihre konkurrierende Einflußstärke hin getestet. Beide Indikatoren wurden dabei anhand von Daten des World-Values-Surveys operationalisiert, welcher für jeden Befragten eine Einteilung hinsichtlich dieser Aspekte in niedrig, mittel oder hoch aufweist.

9.5.3 Ergebnisse

Ergebnisse in der Gruppe der reichen Länder

Aufgrund der vorangegangenen Ergebnisse ist es sinnvoll, vor allem die Analysen der Mehrebenenmodelle innerhalb der Gruppe der reichen Länder genauer zu betrachten, da sich nur für diese ein Einfluß der Zustimmungsnorm in den bivariaten Analysen bestätigen ließ. Die ersten Berechnungen mit diesen Ländern anhand des einfachen Basismodells $y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$, indem noch keine erklärenden Variablen aufgenommen waren, ergaben einen ICC von 0,02643 und zeigen somit, daß insgesamt 2,64% der Varianz der Arbeitszufriedenheit auf dem Länderniveau zu finden war, wenn man nur die reichen Länder betrachtet. Dieser Wert erscheint auf den ersten Blick zwar recht niedrig. Er ist jedoch auch nicht sonderlich überraschend, da sich bereits in den deskriptiven Vergleichen des Arbeitszufriedenheitsniveaus, die mit allen Ländern des

World-Values-Surveys durchgeführt wurden (vgl. Kapitel 9.3.1), eine klare Verschiebung der Ländermittelwerte zum positiven Skalenende hin zeigte, die nur noch wenig Varianz auf dem Länderniveau zuließ. Darüber hinaus war aus psychologischer Sicht auch zu erwarten, daß der Großteil der Varianz des Zufriedenheitsurteils in individuellen Unterschieden zu finden ist.

In der eigentlichen Mehrebenenanalyse zur Testung der Hypothesen wurden in einem ersten Schritt die Kontrollvariablen Alter und Geschlecht in die Berechnungen aufgenommen. Während Alter am Gesamtmittelwert (*grand mean*) zentriert wurde, ging die Dummy-Variable des Geschlechts unzentriert in die Berechnungen ein. Beide Variablen wurden als fixe Effekte in das Modell aufgenommen, da keine Annahmen über den unterschiedlichen Einfluß der Kontrollvariablen in den einzelnen Ländern getätigt wurden. Die Varianzkomponenten dieser Variablen wurden demnach als fest bestimmt. Alle im Folgenden berichteten Koeffizienten stellen immer unstandardisierte Beta-Koeffizienten dar und beziehen sich auch immer auf die fixen Effekte mit robusten Standardfehlern der Mehrebenenanalyse anhand von HLM 6.

Die Ergebnisse dieses ersten Schritts zeigen (vgl. Tabelle 13), daß es keinen Zusammenhang mit dem Geschlecht gab ($\beta = -.017$, $p > .50$), jedoch war das Arbeitszufriedenheitsurteil generell höher, je älter die Befragten waren ($\beta = .018$, $p < .001$). Im nächsten Schritt wurde zur eigentlichen Überprüfung der inhaltlichen Hypothesen berechnet, ob auch in der Mehrebenenanalyse die Unterschiedlichkeit des Zufriedenheitsniveaus durch die jeweilige Zustimmungsnorm des Landes erklärt werden kann. Dazu wurde die Zustimmungsnorm, zentriert am Gesamtmittelwert, als Prädiktor für den Ländermittelwert β_{0j} aufgenommen. Die Ergebnisse können in der Tat zeigen, daß auch in dieser Analyse eine höhere Zustimmungsnorm mit einem niedrigeren Zufriedenheitsniveau einherging ($\beta = -1.423$; $p < .05$). Zusätzlich erhöhte sich der Modell-Fit in signifikanter Weise ($\chi^2(1) = 5.31$; $p < .05$). Die Mehrebenenanalyse kann somit das Ergebnis bestätigen, welches bereits in den bivariaten länderbezogenen Analysen zum Vorschein kam. Durch den Einbezug dieses länderbezogenen Prädiktors verringerte sich die Varianz auf dem Länderniveau von insgesamt $\sigma_0^2 = 0,09626$ auf nun $\sigma_0^2 = 0,07397$. Die Zustimmungsnorm konnte daher insgesamt 23,1% der Varianz auf dem Länderniveau erklären. Es kann folglich festgehalten werden, daß auch gemäß diesen Ergebnissen der Mehrebenenanalyse – wie in den vorangegangenen Analysen – gilt, daß das Arbeitszufriedenheitsniveau geringer ist in Ländern mit einer hohen Zustimmungsnorm.

Tabelle 13: Ergebnisse der Mehrebenenanalyse – reiche Länder

	unstand. Beta (Standardfehler)	finales unstand. Beta (Standardfehler)	Zuwachs im Modell- Fit Chi-Quadrat (df)
Schritt 1: Kontrollvariablen ($N_{\text{Länder}}=21$; $N_{\text{Individuen}}=26333$)			
Alter	.018*** (0.001)	.017*** (0.001)	$\chi^2 (2) = 331.75***$
Geschlecht	-.017 (0.055)	-.015 (0.055)	
Schritt 2: länderbezogener Prädiktor			
Zustimmungsnorm	-1.423* (0.716)	-1.426* (0.718)	$\chi^2 (1) = 5.31**$
Schritt 3: Individual-Prädiktor			
Zustimmungstendenz	.064 (0.080)	.067 (0.078)	$\chi^2 (1) = 4.31**$
Schritt 4: variabler Individual-Prädiktor			
Zustimmungstendenz	.067 (0.078)		$\chi^2 (2) = 70.60***$
Schritt 5: Cross-Level-Interaktion			
Norm x Tendenz		-.037 (0.889)	$\chi^2 (1) = 0.002$

*** p < .01; ** p < .05; * p < .10;

in Klammern: bei den Beta-Koeffizienten der Standardfehler, beim Chi-Quadrat-Wert die Anzahl der Freiheitsgrade

Des Weiteren wurden in Hypothese 10 Annahmen darüber aufgestellt, welcher Zusammenhang denn zwischen der individuellen Zustimmungstendenz, die ein Befragter besitzt, und seinem individuellen Arbeitszufriedenheitsurteil besteht. Dazu wurde als weiterer fixer Prädiktor auf dem individuellen Level die Zustimmungstendenz, zentriert am Gruppenmittelwert, in das Modell aufgenommen. Zwar erhöhte sich der Modell-Fit ($\chi^2 (1) = 4.31$; $p < .05$) durch den Einbezug des Prädiktors etwas, jedoch war dieser Einfluß der Zustimmungstendenz nicht signifikant ($\beta = .064$; $p > .10$). Er wies zwar immerhin – wie erwartet – eine positive Richtung auf. Die erklärte Varianz auf dem Individualniveau verringerte sich jedoch nur von $\sigma_r^2 = 3,70690$ auf nun $\sigma_r^2 = 3,70629$, was lediglich 0,02% entspricht. Der Zuwachs des Modell-Fits deutet aber darauf hin, daß es möglicherweise doch einen Effekt der Zustimmungstendenz auf dem Individuallevel geben kann. Dieser könnte sich jedoch von Land zu Land unterscheiden, so daß letztendlich kein Haupteffekt der Zustimmungstendenz zu messen ist, was sich dann hier in einem nicht-signifikanten Koeffizienten äußerte. Daher bot sich als nächster sinnvoller Schritt an, einen variablen Anteil dieses Prädiktors in das Modell mit aufzunehmen. Diese weitere Veränderung des Modells dahingehend, daß der

Einfluß der Zustimmungstendenz auch in den einzelnen Ländern variieren kann – d.h. dieser Steigungskoeffizient wurde als variabel gesetzt –, verbesserte das Modell weiter ($\chi^2(2) = 70.60; p < .01$). Somit ist ersichtlich, daß der Zusammenhang zwischen der individuellen Zustimmungstendenz und der Arbeitszufriedenheit in den jeweiligen Ländern unterschiedlich ist (siehe Abbildung 26 für die Darstellung der Zusammenhänge in allen 21 Ländern). In den meisten Ländern schien es zwar einen positiven Zusammenhang dieser beiden Aspekte zu geben. In manchen Ländern jedoch war dieser Zusammenhang auch deutlich negativ, was schließlich dazu geführt hat, daß ein Gesamteffekt nicht nachzuweisen war. Es wurde dann im nächsten Schritt angenommen, daß diese Unterschiedlichkeit des Einflusses der Zustimmungstendenz sich möglicherweise durch die jeweiligen Umgebungsbedingungen erklären ließe.



Abbildung 26: Zusammenhang der Zustimmungstendenz mit dem Arbeitszufriedenheitsurteil in den einzelnen reichen Ländern

Daher wurde abschließend schließlich noch geprüft, ob möglicherweise eine Interaktion in der Form vorlag, daß die individuelle Zustimmungstendenz nur in Ländern mit einer bestimmten Zustimmungsnorm einen Einfluß auf das Arbeitszufriedenheitsurteil hatte. Die Ergebnisse gemäß Tabelle 13 zeigen jedoch eindeutig, daß von keiner Interaktion auszugehen ist. Weder war der Prädiktor und somit die Interaktion signifikant

($\beta = -.037$; $p > .10$), noch verbesserte sich der Modell-Fit in deutlicher Weise gegenüber dem vorherigen Modell ($\chi^2(1) = 0,002$; $p > .10$). Die Zustimmungsnorm kann somit nicht als Moderator für die länderspezifischen Zusammenhänge erhalten. Zwar unterschieden sich die Zusammenhänge in den einzelnen Ländern, jedoch ist der Grund offenbar nicht darin zu sehen, daß eine Zustimmungsnorm je nach Befragtem einen unterschiedlichen Effekt erzeugt. Die Überlegungen zur Interaktion von persönlicher Zustimmungstendenz und der gesellschaftlich vorgeschriebenen Zustimmungsnorm können somit nicht bestätigt werden. Das Arbeitszufriedenheitsurteil eines Befragten, der beispielsweise eine hohe Zustimmungstendenz zeigt, wird somit nicht unterschiedlich verzerrt, je nachdem ob er in einem Umfeld mit einer hohen oder niedrigen Zustimmungsnorm lebt. Letztendlich können möglicherweise weitere Länderkriterien die unterschiedlichen Zusammenhänge erklären. Es sollte daher zukünftig nach diesen weiteren Variablen, die als Moderator für den unterschiedlichen Einfluß der Zustimmungstendenz erhalten können, gesucht werden. Dies jedoch wäre dann eine potentielle Aufgabe für weitere Studien, da es nicht im Fokus der hier vorliegenden Arbeit stand.

Die Ergebnisse in der Gruppe der reichen Länder zeigen somit insgesamt, daß auf der einen Seite der Einfluß des länderbezogenen Prädiktors, d.h. der Zustimmungsnorm, bestätigt werden konnte. Es ließen sich jedoch weder klare Zusammenhänge auf individuellem Niveau hinsichtlich der Zustimmungstendenz finden, noch lag eine Wechselwirkung zwischen der individuellen Zustimmungstendenz und der Zustimmungsnorm vor. Dennoch kann somit die Hypothese 10 immerhin in Teilen bestätigt werden kann. Die Mehrebenenanalyse kann somit das vorherige Ergebnis auf dem Länderniveau replizieren und zeigen, daß eine hohe Zustimmungsnorm mit einem gemäßigeren Arbeitszufriedenheitsniveau einhergeht. Der erwartete Zusammenhang auf dem individuellen Level kann jedoch nicht bestätigt werden, denn die individuelle Ausprägung einer Zustimmungstendenz stand in keinem Zusammenhang mit der individuellen Arbeitszufriedenheit. Eine genauere Diskussion dieses durchaus überraschenden Ergebnisses erfolgt im abschließenden Diskussionskapitel der Arbeit. Eine Interaktion und somit ein Einfluß der individuellen Zustimmungstendenz nur unter bestimmten Umgebungsbedingungen konnte ebenfalls nicht nachgewiesen werden. Diese unterschiedlichen Ergebnisse auf dem individuellen und dem Länderniveau machen dabei übrigens noch einmal deutlich, daß von einem Zusammenhang auf dem

höheren Analyselevel keineswegs auf einen Zusammenhang auf dem individuellen Level oder aber auch umgekehrt geschlossen werden kann, und bestätigen somit die Sinnhaftigkeit des Einsatzes der Mehrebenenanalyse.

Destruktives Testen

Um die gefundenen Zusammenhänge auf dem Länderlevel auch in dieser statistischen Mehrebenenanalyse gegenüber Alternativerklärungen abzusichern, wurden weitere Berechnungen unter Einbezug der alternativen Erklärungsfaktoren Individualismus und Bruttoinlandsprodukt (auf dem Länderlevel) sowie Bildung und Einkommen (auf dem Individuallevel) durchgeführt. Für diese Berechnungen wurden die bisher schon analysierten Faktoren, die in der oben beschriebenen Analyse keinen Einfluß zeigten, aus Gründen der Modellspartheit jedoch wieder aus der Analyse entfernt. Somit blieben im Ausgangsmodell dieser Berechnungen des destruktiven Testens lediglich die Zustimmungsnorm als länderbezogener Prädiktor und das Alter als fixer Individualprädiktor enthalten. Die einzelnen Alternativprädiktoren wurden dann – immer zentriert am Gesamtmittelwert – jeweils einzeln und nacheinander in das Modell aufgenommen.

Es zeigt sich, daß drei der vier Faktoren durchaus ebenfalls einen zusätzlichen signifikanten Erklärungswert über das Ausgangsmodell hinaus besaßen (vgl. Tabelle 14). So war das Arbeitszufriedenheitsniveau höher bei größerem Bruttoinlandsprodukt ($\beta = .000012$; $p < .10$). Das Bruttoinlandsprodukt verringerte die Varianz auf dem Länderniveau dabei von $\sigma_0^2 = 0,9238$ auf nun $\sigma_0^2 = 0,07611$, was insgesamt weiteren 17,6% entspricht. Der Individualismus-Indikator von Hofstede konnte jedoch keinen bedeutsamen Anteil von Varianz auf dem Länderlevel aufklären ($\beta = .003$; $p > .10$). Er verringerte lediglich die Varianz von $\sigma_0^2 = 0,09238$ auf nun $\sigma_0^2 = 0,08953$, d.h. nur um 3,08%. Dieses Ergebnis bedeutet somit, daß die Kulturdimension keinen Einfluß mehr zeigte, wenn gleichzeitig die Zustimmungsnorm in der Analyse mit berücksichtigt wurde. Die Zustimmungsnorm hingegen behielt mit $\gamma_{01} = -1.404$ weiterhin ihren Einfluß (wenngleich sie auch mit $p = .11$ aufgrund der kleinen Stichprobengröße nicht ganz das Signifikanz-Niveau erreicht). Dieses Ergebnis der Mehrebenenanalyse bestätigt somit die vorhergehenden Ergebnisse der Mediationsanalysen, in der ebenfalls bei gleichzeitiger Berücksichtigung von einer Zustimmungsnorm und der Kulturdimension Individualismus eher die Zustimmungsnorm den stärkeren Einfluß auf das Arbeitszufriedenheitsniveau besaß.

Tabelle 14: Ergebnisse des mehrbenenanalytischen destruktiven Testens – reiche Länder

	unstand. Beta neuer Faktor (Standardfehler)	finale unstand. Beta der Ausgangsfaktoren	Zuwachs im Modell-Fit, Chi-Quadrat (df)
Ausgangsmodell (N _{Länder} =21; N _{Individuen} =36702)			
Alter	.019*** (0.001)		
Zustimmungsnorm	-1.688** (0.764)		
Modell 1: Individualismus			
Individualismus	.003 (0.003)	Alter: $\beta_1 = .019***$ Zust.-Norm: $\gamma_{01} = -1.404$ (p = .11)	$\chi^2 (1) = 0.66$
Modell 2: Bruttoinlandsprodukt			
BIP	.000012* (0.000007)	Alter: $\beta_1 = .019***$ Zust.-Norm: $\gamma_{01} = -1.179^*$	$\chi^2 (1) = 3.89^{**}$
Ausgangsmodell (N _{Länder} =22; N _{Individuen} =32075)			
Alter	.020*** (0.001)		
Zustimmungsnorm	-1.809** (0.869)		
Modell 3: Einkommen			
Einkommen	.209*** (0.037)	Alter: $\beta_1 = .020***$ Zust.-Norm: $\gamma_{01} = -1.760^*$	$\chi^2 (1) = 205.38^{***}$
Ausgangsmodell (N _{Länder} =18; N _{Individuen} =11555)			
Alter	.014*** (0.002)		
Zustimmungsnorm	-0.885 (0.532)		
Modell 4: Bildung			
Bildung	.164*** (0.033)	Alter: $\beta_1 = .015***$ Zust.-Norm: $\gamma_{01} = -.891$ (p = .15)	$\chi^2 (1) = 47.34^{***}$

*** p < .01; ** p < .05; * p < .10;

in Klammern: bei den Beta-Koeffizienten der 1. Spalte der Standardfehler, beim Chi-Quadrat-Wert die Anzahl der Freiheitsgrade

Auf dem individuellen Niveau wurden jeweils neue Ausgangsmodelle berechnet, da sich die Datenbasis für jede Analyse unterschied. Dabei zeigte sich, daß sowohl das Einkommen ($\beta = .209$; $p < .01$) als auch der Bildungsstand der Befragten ($\beta = .164$; $p < .01$) in einem positiven Zusammenhang zur Arbeitszufriedenheit standen. Jedoch verringerte der Einbezug der Einkommensvariable die Varianz auf dem Individuallevel nur um 0,63%, von $\sigma_r^2 = 3,73686$ auf nun $\sigma_r^2 = 3,71307$. Ähnlich wenig verringerte der Einbezug der Variable zum Bildungsstand die Varianz auf dem Individuallevel nur von $\sigma_r^2 = 3,43071$ auf nun $\sigma_r^2 = 3,41598$, was lediglich 0,42% entspricht. Beide

Individualfaktoren stehen daher aufgrund der großen Stichprobengröße zwar in einem signifikanten Zusammenhang mit dem Arbeitszufriedenheitsurteil, erklären aber nur wenig die Unterschiedlichkeit des Urteils.

Hinsichtlich aller dieser vier Modelle mit den alternativen Erklärungsfaktoren ist jedoch trotz des signifikanten Einflusses der zusätzlichen Faktoren vor allem zu beachten, daß die Zustimmungsnorm immer ihren negativen Einfluß auf das Arbeitszufriedenheitsniveau behielt. Sowohl bei Berücksichtigung der Einkommensvariable und des Bruttoinlandsprodukts blieb der Einfluß dabei deutlich signifikant. Bei Berücksichtigung von Individualismus und der Bildungsvariable verlor die Zustimmungsnorm etwas von ihrer Einflußstärke und erreichte daher zwar nicht mehr ganz das Signifikanzniveau. Dies lag jedoch vor allem an der kleinen Länderstichprobe, die im Falle der Bildungsvariable bei Modell 2 zusätzlich aufgrund fehlender Ausgangsdaten von HLM 6 verringert wurde. Insgesamt ist jedoch festzuhalten, daß auch bei diesen Analysen im Sinne des destruktiven Testens somit der negative Einfluß der Zustimmungsnorm in den reichen Ländern bestehen blieb.

Ergebnisse in der Gruppe der ärmeren Länder

Um auch die Zusammenhänge in den wirtschaftlich ärmeren Ländern zu überprüfen und möglicherweise auch bezüglich dieser Gruppe die Ergebnisse der vorherigen bivariaten Korrelationsanalysen zu replizieren, wurde die gleiche schrittweise Mehrebenenanalyse auch anhand dieser zweiten Gruppe von Ländern durchgeführt. Die Varianzaufteilung des Basismodells ergab, daß für die ärmeren Ländern ein ICC von 0,05089 sich errechnen ließ und somit 5,09% der Varianz der Arbeitszufriedenheitsurteile sich auf das Länderniveau bezogen. Bei der eigentlichen schrittweisen Mehrebenenanalyse wurden die gleiche Vorgehensweise wie bei den reichen Ländern gewählt und erneut nacheinander die Kontrollvariablen Alter und Geschlecht, die Zustimmungsnorm, die Zustimmungstendenz und die Interaktion der letzten beiden Aspekte in das Modell mit aufgenommen. Die Ergebnisse zeigen jedoch (vgl. Tabelle 15), daß lediglich Alter als Kontrollvariable in einem deutlichen Zusammenhang zum Arbeitszufriedenheitsurteil stand ($\beta = .018$, $p < .001$). Alle anderen Variablen hatten keinen Erklärungsgehalt in den ärmeren Ländern, weder für das individuelle Zufriedenheitsurteil noch für das länderspezifische Zufriedenheitsniveau.

Tabelle 15: Ergebnisse der Mehrebenenanalyse – arme Länder

	unstand. Beta neuer Faktor (Standardfehler)	finales unstand. Beta (Standardfehler)	Zuwachs im Modell- Fit Chi-Quadrat (df)
Schritt 1: Kontrollvariablen ($N_{\text{Länder}}=19$; $N_{\text{Individuen}}=23258$)			
Alter	.018*** (0.002)	.018*** (0.002)	$\chi^2 (2) = 189.95^{***}$
Geschlecht	.073 (0.046)	.070 (0.046)	
Schritt 2: länderbezogener Prädiktor			
Zustimmungsnorm	-1.299 (1.389)	-1.301 (1.389)	$\chi^2 (1) = 0.89$
Schritt 3: Individual-Prädiktor			
Zustimmungstendenz	.044 (0.072)	.038 (0.084)	$\chi^2 (1) = 1.14$
Schritt 4: variabler Individual-Prädiktor			
Zustimmungstendenz	.036 (0.084)		$\chi^2 (2) = 17.11^{***}$
Schritt 5: Cross-Level-Interaktion			
Norm x Tendenz		.457 (0.583)	$\chi^2 (1) = 0.26$

*** p < .01; ** p < .05; * p < .10;

in Klammern: bei den Beta-Koeffizienten der Standardfehler, beim Chi-Quadrat-Wert die Anzahl der Freiheitsgrade

Die Mehrebenenanalyse kann einerseits somit auch hier die Ergebnisse der bivariaten Korrelationsanalysen erst einmal dahingehend bestätigen, daß in den ärmeren Ländern das Arbeitszufriedenheitsniveau nicht von einer Zustimmungsnorm beeinflusst wird. Andererseits jedoch zeigte sich auch in dieser Ländergruppe kein Zusammenhang auf dem individuellen Niveau. Lediglich der signifikante Zuwachs im Modell-Fit beim Einbezug des variablen Einflusses des Individual-Prädiktors zeigt eine Verbesserung des Modells und deutet darauf hin, daß es auch in dieser Ländergruppe einen deutlichen Unterschied zwischen den Ländern hinsichtlich des Zusammenhangs der individuellen Zustimmungstendenz und dem Arbeitszufriedenheitsurteil gab. Dieser Unterschied ließ sich jedoch wiederum nicht durch eine Wechselwirkung aus Zustimmungsnorm und individueller Kommunikationstendenz erklären, da auch der Koeffizient der Interaktion weder signifikant wurde, noch ein besserer Modell-Fit zustande kam.

Destruktives Testen

Die Analysen zum destruktiven Testen ergaben auch bei den ärmeren Ländern weitestgehend ähnliche Ergebnisse wie bei den reichen Ländern. Im Prinzip diente sie hier aber fast eher der Suche nach Faktoren, die überhaupt die Unterschiedlichkeit der Arbeitszufriedenheit in dieser Ländergruppe erklären können, da die bisherigen

Faktoren dazu ja nicht in der Lage waren. Auch hier wurde das Ausgangsmodell aufgrund der vorherigen Ergebnisse verkleinert und nur Alter als erklärende Variable im Ausgangsmodell weiter berücksichtigt. Es zeigt sich auch hier, daß Individualismus als Länderkennwert keinen Erklärungsgehalt für das Arbeitszufriedenheitsniveau besaß (vgl. Tabelle 16). Dies gleicht somit einerseits den Ergebnissen in den reichen Ländern. Es ist jedoch andererseits zumindest insofern interessant, da dieser Zusammenhang hier bei den ärmeren Ländern somit selbst dann nicht existierte, wenn die Zustimmungsnorm überhaupt nicht im Modell enthalten war.

Tabelle 16: Ergebnisse des mehrerebenenanalytischen destruktiven Testens – arme Länder

	unstand. Beta neuer Faktor (Standardfehler)	finales unstand. Beta der Ausgangsfaktoren	Zuwachs im Modell-Fit, Chi-Quadrat (df)
Ausgangsmodell (N _{Länder} =20; N _{Individuen} =24632)			
Alter	.018*** (0.003)		
Modell 1: Individualismus			
Individualismus	.005 (0.006)	Alter: $\beta_1 = .018***$	$\chi^2(1) = 0.52$
Modell 2: Bruttoinlandsprodukt			
BIP	.000023 (0.000043)	Alter: $\beta_1 = .019***$	$\chi^2(1) = 0.51$
Ausgangsmodell (N _{Länder} =20; N _{Individuen} =23200)			
Alter	.018*** (0.002)		
Modell 3: Einkommen			
Einkommen	.423*** (0.069)	Alter: $\beta_1 = .018***$	$\chi^2(1) = 419.67***$
Ausgangsmodell (N _{Länder} =16; N _{Individuen} =12270)			
Alter	.009** (0.003)		
Modell 2: Bildung			
Bildung	.428*** (0.100)	Alter: $\beta_1 = .010***$	$\chi^2(1) = 178.53***$

*** p < .01; ** p < .05; * p < .10;

in Klammern: bei den Beta-Koeffizienten der 1. Spalte der Standardfehler, beim Chi-Quadrat-Wert die Anzahl der Freiheitsgrade

Überraschender ist sicherlich auf den ersten Blick, daß das Bruttoinlandsprodukt bei dieser Gruppe von Ländern nicht zu einem höheren Zufriedenheitsniveau führte. Auf der einen Seite könnte man zwar nun denken, daß die bisherige Erklärung für den fehlenden Einfluß der Zustimmungsnormen in den ärmeren Ländern nicht mehr

zutreffend wäre. Diese besagte, daß erst ein entsprechendes Niveau der Arbeitsplatzqualität erreicht sein muß, bevor die Befragten sich an den geltenden Kommunikationsnormen orientieren würden. Auf der anderen Seite widerspricht jedoch dieses Ergebnis nicht zwingend diesen Überlegungen, denn dieses Ergebnis zeigt erst einmal nur, daß ein Zuwachs im Bruttoinlandsprodukt und somit eine bessere Arbeitsplatzqualität auf diesem geringen Wohlstandsniveau zu keinem Effekt bei der Zufriedenheit führt. Man könnte folglich fast mit der Unterscheidung von Herzberg (1967) argumentieren, wonach dieser Zuwachs an Hygienefaktoren nicht zu einer erhöhten Zufriedenheit führt. Somit könnten die Befragten durchaus auch weiterhin mehr mit ihrer Arbeitsumgebung beschäftigt sein als mit dem Befolgen von Kommunikationsnormen, was folglich die bisherige Erklärung weiter bestehen lassen würde. Die Unterschiede im Zufriedenheitsniveau in den ärmeren Ländern bleiben aber weiterhin weitestgehend ungeklärt und sollten in weiteren Studien speziell untersucht werden. Zustimmungsnormen können nach den bisherigen Ergebnissen dabei als erklärender Faktor jedoch ausgeschlossen werden.

Auf dem individuellen Niveau hingen auch bei dieser Gruppe von Ländern das persönliche Einkommen ($\beta = .423$; $p < .01$) und der Bildungsstand ($\beta = .428$; $p < .01$) neben dem Alter positiv mit dem Arbeitszufriedenheitsurteil zusammen. Jedoch verringerte die Einkommensvariable erneut kaum die Varianz, lediglich von $\sigma_r^2 = 5,50764$ auf nun $\sigma_r^2 = 5,40840$, was lediglich 1,80% entspricht. Gleiches galt für die Variable des Bildungsstands, bei der die Varianz sich nur von $\sigma_r^2 = 5,74113$ auf nun $\sigma_r^2 = 5,65761$ verringerte, was lediglich 1,45% entspricht.

9.5.4 Zusammenfassung der Ergebnisse der Mehrebenenanalysen

Die Ergebnisse der Mehrebenenanalysen bestätigen weitestgehend die bisherigen Erkenntnisse der vorangegangenen Analysen. Betrachtet man das Länderniveau und somit den Zusammenhang von Zustimmungsnormen und dem Arbeitszufriedenheitsniveau, dann führt das Vorhandensein solch einer Kommunikationsnorm auch gemäß diesen statistischen Analysen zu einem gemäßigeren Zufriedenheitsniveau. Dieser Effekt läßt sich speziell in Ländern mit hohem Wohlstand und somit guten Arbeitsbedingungen finden. Bei den ärmeren Ländern hingegen läßt sich auch bei den Mehrebenenanalysen, die besser die

hierarchische Struktur der Daten berücksichtigen, letztendlich kein Zusammenhang finden.

Betrachtet man die Unterschiedlichkeit des Arbeitszufriedenheitsurteils auf Länder- und auf Individualniveau, wird deutlich, daß die Varianz zwischen den Ländern zwar insgesamt viel kleiner als die Varianz ist, die auf die individuell unterschiedlichen Urteile zurückgeht. Jedoch läßt sich bei genau dieser länderbezogenen Varianz ein Großteil – nämlich bis zu 23% – von einer Kommunikationsnorm, d.h. im Speziellen der Zustimmungsnorm, erklären. Eine Zustimmungsnorm bleibt somit auch aufgrund dieser weiteren Form der Analyse ein wichtiger Einflußfaktor auf die unterschiedlich hohen Zufriedenheitsniveaus in den Ländern. Zusätzlich ist bedeutsam, daß trotz des Einbezugs von alternativen Erklärungsmöglichkeiten bei den Modellen des destruktiven Testens die Zustimmungsnorm weiterhin als bedeutender Prädiktor bestehen bleibt und somit über die anderen Faktoren hinaus einen bedeutsamen Teil der Varianz erklärt. Sie sollte damit auch gemäß den Ergebnissen dieser Analysen als ein wichtiger Faktor neben den anderen Einflußgrößen betrachtet werden.

Wenngleich die Hypothese 10 somit auf dem Länderniveau bestätigt werden konnte, zeigten sich keine klaren Zusammenhänge auf dem Individualniveau. Zwar scheint es zum Teil durchaus einen generellen Zusammenhang zwischen der individuellen Zustimmungstendenz und dem Zufriedenheitsurteil zu geben, jedoch ist dieser Einfluß nicht einheitlich, sondern unterscheidet sich je nach Land. Ein Haupteffekt ließ sich daher insgesamt nicht nachweisen. Welche Umgebungsaspekte jedoch schließlich ausschlaggebend für den unterschiedlichen Zusammenhang in den einzelnen Ländern sind, konnte in den bisherigen Analysen nicht endgültig geklärt werden. Die Zustimmungsnorm scheint aber dafür nicht als Moderator in Betracht zu kommen. Hier ist weiterer Forschungsbedarf nötig.

10 Experimentelle Überprüfung

Für die bisherigen Analysen, die die bivariaten Zusammenhänge, den mediierenden Effekt und die Einflüsse auf dem Länder- und dem Individualniveau prüften, wurden Sekundärdaten großer internationaler Studien genutzt. Diese haben den großen Vorteil, daß eine Vielzahl von Ländern einbezogen werden konnte und somit sich die externe Validität der Ergebnisse erhöht. Diesem Vorteil steht jedoch der Nachteil gegenüber, daß es in solch sekundäranalytischen Untersuchungen schwierig, wenn nicht sogar zum Teil unmöglich ist, alle potentiellen Störvariablen zu kontrollieren. Zwar dienten die Berechnungen des destruktiven Testens bereits dazu, diesen Nachteil in Teilen zu kompensieren und somit die interne Validität der Ergebnisse zu erhöhen. Es könnte jedoch durchaus weitere potentielle Einflußgrößen geben, deren schiere Anzahl in solch einer Analyse aber nicht umfassend geprüft werden kann. Aus diesem Grund sollte ein separat durchgeführtes Experiment die bisher gefundenen Ergebnisse bestätigen.

10.1 Vorgehen des Experiments

Die Durchführung des Experiments stützte sich insgesamt auf die bisherigen Erkenntnisse der vorangegangenen Berechnungen. Gemäß den Ergebnissen der Mediationsanalysen ließ sich schlußfolgern, daß die kulturelle Umgebung eines jeweiligen Landes dazu führt, daß eine spezielle Kommunikationsnorm sich innerhalb des Landes entwickelt, welche die Menschen dann dazu bewegt, in einer bestimmten Art und Weise auf Fragebögen zu antworten. Im Fall der hier vorliegenden Arbeit bedeutete dies konkret, daß kollektivistische Wertvorstellungen zu einer höheren Zustimmungsnorm führen, welche dann ihrerseits dazu führt, daß im Mittel eine eher gemäßigte Zufriedenheit berichtet wird. Um diesen Einfluß einer Zustimmungsnorm auch in einem experimentellen Setting überprüfen zu können, mußten daher vor allem folgende drei Aspekte geklärt werden:

- die Erfassung des unterschiedlichen kulturellen Denkens hinsichtlich Individualismus/Kollektivismus
- die Erfassung der Zustimmungsnorm
- die Erfassung des Zufriedenheitsurteils

Zur Erfassung der unterschiedlichen kulturellen Wertvorstellungen in einem Experiment muß im Prinzip die Kultur ins Labor geholt werden. Auf der einen Seite hätte man dazu Versuchspersonen aus verschiedenen Kulturen verschiedenen Gruppen zuordnen können und diese als Experimentalgruppen unterscheiden können. Dies würde jedoch nur einem quasi-experimentellen Design folgen, da Kultur nicht zufällig den Versuchspersonen zugewiesen werden kann. Auf der anderen Seite kann man daher – wie bereits in zahlreichen interkulturellen Experimenten geschehen – Kultur in der Form variieren, daß man bei Versuchspersonen das jeweils bedeutsame kulturelle Denken anhand von Primings situativ verstärkt und somit hervorhebt (vgl. Kapitel 5.4). Aufbauend auf den Überlegungen, daß in individualistischen Kulturen vor allem ein independentes Selbst eine große Bedeutung hat, während in kollektivistischen Kulturen das interdependente Selbst viel stärker ausgeprägt ist (vgl. u.a. Markus & Kitayama, 1991), wurde daher auch in dieser Studie anhand von Primings das jeweilige Selbstkonzept experimentell variiert.

Gemäß den theoretischen Überlegungen sollte das dann unterschiedlich aktivierte Selbstkonzept, welches der Erfassung der kulturellen Unterschiede hinsichtlich Individualismus und Kollektivismus dienen würde, dazu führen, daß auch ein unterschiedliches Kommunikationsverhalten von den Versuchspersonen präferiert werden sollte. Bei geprimtem individualistischen Denken (d.h. hohem independenten Selbstkonzept) sollte eine Zustimmungsnorm weniger zum Vorschein kommen als bei geprimtem kollektivistischen Denken (d.h. hohem interdependentem Selbstkonzept). Diese unterschiedliche Form der Zustimmungsnorm sollte dann schließlich dazu führen, daß die Personen auch ein im Mittel unterschiedliches Urteil hinsichtlich einer Arbeitszufriedenheitsfrage abgeben. Dabei sollte das Zufriedenheitsurteil dann in der kollektivistischen Gruppe signifikant geringer sein.

Der experimentelle Hypothesenkomplex läßt sich daher wie folgt zusammenfassen:

- a) Ein individualistisches Priming führt zu einer stärkeren Aktivierung des independenten Selbst, während ein kollektivistisches Priming eine stärkere Aktivierung des interdependenten Selbst bewirkt.
- b) Die kollektivistisch geprimte Versuchsgruppe, die ein stärkeres interdependentes Selbst aufweist, zeigt im Mittel ein stärkeres Zustimmungsverhalten als die individualistisch geprimte Vergleichsgruppe.

- c) Die kollektivistisch geprägte Versuchsgruppe mit einem stärkeren interdependenten Selbst berichtet im Mittel ein geringes Zufriedenheitsniveau als die individualistische Vergleichsgruppe.

Um diese Hypothesen zu prüfen, erhielten Versuchspersonen nacheinander verschiedene Aufgabenteile vorgesetzt (vgl. Anhang C). Diese wurden ihnen in gleicher Weise, wie dies auch in anderen Priming-Studien der Fall war (vgl. beispielsweise Kühnen & Hannover, 2000), als eigenständige Studien präsentiert, um die Probanden über die getesteten Zusammenhänge im Unklaren zu lassen und somit die Wirksamkeit des Primings subtiler zu gestalten. Der erste Teil bestand aus dem Primingverfahren, welchem ein erster Manipulationscheck zur Überprüfung seiner Wirksamkeit folgte. Als angeblich zweite Studie sollten die Versuchspersonen ein Urteil zu einem Zufriedenheitsszenario abgeben. In einer angeblich dritten Studie erhielten sie schließlich Itemsets, die der Messung von Antworttendenzen dienten. Dieser Teil enthielt zusätzlich einen zweiten Manipulationscheck, gefolgt von abschließenden soziodemographischen Fragen.

10.2 Stichprobe und Meßgrößen

Stichprobe

Bei den Versuchspersonen handelte es sich um Studenten der Universität zu Köln, die in zufälliger Weise auf dem Campus rekrutiert wurden und in einen ruhigen, abgetrennten Bereich gebeten wurden, um an der Studie teilzunehmen. Dadurch wurde, ähnlich wie in früheren Priming-Studien (Brewer & Gardner, 1996; Gardner et al., 1999; Kühnen & Hannover, 2000; Kühnen, Hannover & Schubert, 2001; Roeder & Hannover, 2002; Suh, Diener & Updegraff, 2008), dafür gesorgt, daß die Stichprobe relativ homogen bezüglich Faktoren wie Alter, Bildungsstand oder auch Einkommen war, so daß die gemessenen Effekte ursächlich auf das Priming zurückzuführen wären. Dabei wurden insgesamt vor allem Studenten wirtschaftsnaher Fächer rekrutiert, da bei diesen davon auszugehen war, daß diese sich besser in die in der Studie verwendete Szenariumssituation hineinversetzen konnten. Aus dem gleichen Grund wurde ebenfalls darauf geachtet, daß die Studenten sich mehrheitlich im zweiten Teil ihres Studiums befanden und wenn möglich ebenfalls schon Arbeitserfahrung in Form einer vorherigen Ausbildung, Praktika oder auch Nebenjobs gesammelt hatten.

Das Priming

Aufgrund der Ergebnisse der Meta-Analyse von Oyserman und Lee (2008), die die Effekte der geläufigsten interkulturellen Priming-Verfahren verglich, wurde das SDFP-Priming (Similarities and differences with family and friends task) eingesetzt, welches von Trafimow et al. (1991) entwickelt wurde. Dieses Priming hatte gemäß der Meta-Analyse insgesamt die stärksten Effekte in den bisherigen Studien gezeigt. Bei dieser Form des Primings sollen sich Probanden zwei Minuten Gedanken dazu machen, was sie von ihrer Familie und ihren Freunden unterscheidet bzw. was sie mit ihrer Familie und ihren Freunden gemeinsam haben. Erstere Variante dient dazu, die Probanden auf ihr eigenes Ich zu fokussieren und die independenten Wissensinhalte zu aktivieren, während letztere Variante den Wir-Bezug in den Vordergrund rückt und interdependente Wissensinhalte primt. Die erste Variante dient somit dazu, eine individualistische Experimentalgruppe zu kreieren, während die zweite Variante das kollektivistische kulturelle Denken erzeugt. Zusätzlich sollten die Versuchspersonen die Gedanken, die sie sich während der zwei Minuten gemacht hatten, niederschreiben, was die Wirksamkeit des Primings zusätzlich verstärken sollte.

Die Manipulationschecks

Um die Wirksamkeit des Primings überprüfen zu können, wurden zwei verschiedene Manipulationschecks in das Experiment eingebunden. Die erste Überprüfung fand anhand des von Miller et al. (1990) entwickelten sogenannten social-judgement-tasks (SJT) statt, bei dem Probanden die Angemessenheit des Verhaltens einer fiktiven Person Lisa einschätzen sollten, die es unterlassen hatte, ihrer Freundin Anne bei einem persönlichen Problem zu helfen. Ursprünglich als Maß zur Erforschung der Wahrnehmung sozialer Verantwortung von den Autoren entwickelt, setzten jedoch auch Gardner et al. (1999) diese Aufgabe bereits als Manipulationscheck in Priming-Studien ein. Eine kollektivistisch geprimte Versuchsperson sollte dabei im Vergleich zu einer individualistische geprimten Person das in der Geschichte gezeigte Verhalten eher als unangemessen beurteilen und die fiktive Person Lisa bestrafen wollen.

Am Ende des Fragebogens wurde in Anlehnung an die Vorgehensweise von Gardner et al. (1999) zur erneuten Überprüfung als weiterer Manipulationscheck eine Kurzform des Twenty Statement Test (TST) (Kuhn & McPartland, 1954) eingesetzt, bei dem Versuchspersonen mehrere Aussagen zu der Frage „Wer bin ich?“ in vorgefertigten Antwortsätzen geben müssen, die alle mit „Ich bin ...“ beginnen. Angewendet wurde

im Experiment die auf zehn Aussagen verkürzte Form dieses Tests (TenST), der ebenfalls bereits erfolgreich in verschiedenen Studien eingesetzt wurde (Khanlou, 2005; Kinket & Verkuyten, 1997; Kiss, 2004). Die von den Probanden gegebenen Aussagen wurden zur Überprüfung der Primingwirksamkeit anhand des von Rhee et al. (1995) entwickelten Kategorisierungsschemas durch vier unabhängige Rater entweder als unabhängige oder interdependente Aussage eingeordnet. Kollektivistisch geprägte Versuchspersonen sollten beim TenST vergleichsweise mehr interdependente Aussagen tätigen.

Die Maße der Antworttendenz

Im Experiment sollte vor allem die Zustimmungstendenz erfaßt werden. Darüber hinaus sollte aus explorativen Gründen jedoch ebenfalls die Extremtendenz und die Soziale Erwünschtheit gemessen werden. Für die Erfassung der von den Probanden gezeigten Zustimmungstendenz konnte nicht auf ein speziell für diesen Zweck entwickeltes Verfahren zurückgegriffen werden, da ein Instrument zur Messung von Zustimmungsnormen in der Literatur bisher nicht zu existieren scheint. Da die bisher in der Forschung angewendete Vorgehensweise darin bestand, daß die Zustimmungsnorm durch Auszählen der entsprechenden Antwortkategorien anhand von inhaltlich möglichst heterogenen Items ermittelt wurde, wurde eine kurze aber dennoch heterogene Itemzusammenstellung zur Bildung der Zustimmungssindizes benötigt. Die Wahl fiel daher schließlich auf eine Kurzform des Portrait-Value-Questionnaires (PVQ), die von Schwartz (2007) für den European Social Survey (ESS) entwickelt wurde. Diese Kurzform erfaßt anhand von 21 Items alle zehn verschiedenen Dimensionen seines Kulturmodells. Sie wurde vor allem deshalb ausgewählt, da aufgrund der entgegengesetzten Anordnung der Wertedimensionen in Schwartz' Wertemodell es unwahrscheinlich ist, daß eine Versuchsperson auf allen Dimensionen hohe Werte erreicht. Da sich die Werte in seinem Modell diametral gegenüberliegen, dürfte im Normalfall eine durchgehende Zustimmung zu allen Items nicht erfolgen. Geschieht dies dennoch, kann davon ausgegangen werden, daß die Versuchsperson generell zu zustimmendem Verhalten tendiert. Ein Zustimmungsbias würde somit dann dadurch offensichtlich. Da geschlechterspezifische Versionen dieses Fragebogens existieren, wurde auch im Experiment den weiblichen bzw. männlichen Probanden jeweils die entsprechende Version des Fragebogens vorgelegt. Die Zustimmungsnorm wurde anhand dieses Itemsets schließlich gemäß der üblichen Vorgehensweise gebildet, indem

pro Versuchsperson die Anzahl der Äußerungen auf der positiven Seite der Skala (in diesem Fall „ist mir sehr ähnlich“ und „ist mir ähnlich“) den ablehnenden Äußerungen gegenübergestellt wurde, dann anhand der Gesamtanzahl der Items relativiert wurde und schließlich für die jeweiligen Untersuchungsgruppen gemittelt wurde.

Auch wenn sich in den vorangegangenen Untersuchungen keine klaren Ergebnisse hinsichtlich der Extrem-/Mittentendenzen zeigten, wurden diese der Vollständigkeit halber aber ebenfalls im Experiment erhoben. Dazu wurde speziell das von Greenleaf (1992b) entwickelte Instrument eingesetzt, welches explizit zur Erfassung von Extremtendenzen erstellt wurde. Bei diesem Instrument sollen die Versuchspersonen Aussagen zu 16 inhaltlich voneinander völlig unabhängigen Items machen. Die Extremtendenz wurde mit Hilfe des Instruments in der bereits bekannten Form erhoben, indem die Anzahl der Äußerungen auf den äußersten Skalenstufen anhand der Gesamtanzahl relativiert wurde und erneut für die jeweiligen Untersuchungsgruppen gemittelt wurde. Neben der expliziten Messung von extremem Antwortverhalten hat dieses Instrument jedoch einen weiteren Vorteil, selbst wenn dieser von Greenleaf (1992b) so nicht intendiert wurde. Da es nämlich die notwendigen Voraussetzungen besitzt (d.h. heterogenes Itemset und Zustimmungsskala), konnte das Instrument auch zusätzlich zur Messung von zustimmendem Antwortverhalten eingesetzt werden. Somit wurde anhand dieses Itemsets ebenfalls ein Zustimmungsmaß gemäß der oben genannten Vorgehensweise errechnet. In umgekehrter Weise wurde schließlich auch anhand des PVQ ein Maß für die Extrem- bzw. Mittentendenz errechnet, um noch einen weiteren Indikator zur Überprüfung der Hypothesen einsetzen zu können.

Trotz uneindeutiger vorheriger Ergebnisse hinsichtlich Sozialer Erwünschtheit wurde dennoch auch diese Form der Antworttendenz anhand von speziellen Skalen gemessen, um auch ihren Einfluß experimentell überprüfen zu können. Zur Ermittlung von sozial erwünschten Antworten wurden insgesamt zwei Skalen verwendet. Zum Einen wurde die von Winkler, Kroh und Spiess (2006) entwickelte deutsche Kurzform des „Balanced Inventory of Desirable Responding“ (BIDR) von Paulhus (1984) eingesetzt. Zum Anderen ging die vom Zentrum für Umfragen, Methoden und Analysen (ZUMA) entwickelte deutsche Kurzform der ursprünglich 33 Items umfassenden Marlowe-Crowne Social Desirability Scale (MC-SDS) in den Fragebogen mit ein (ZUMA-Informationssystem, 2007). Letztere mißt soziale Erwünschtheit als eindimensionales Konstrukt, während das BIDR zwischen zwei Dimensionen der sozialen Erwünschtheit, der Fremd- und Selbsttäuschung, unterscheidet. King und Bruner (2000) stellten

bezüglich der beiden Skalen im Übrigen fest, daß diese beiden nur minimal miteinander korrelieren. Beide Skalen dienten dazu, sie hinsichtlich ihrer Effekte auf das Arbeitszufriedenheitsniveau zu prüfen.

Die Messung der Zufriedenheit

Für die Messung der Zufriedenheit fiel die Wahl statt einer Frage nach der tatsächlichen Zufriedenheit der Versuchspersonen in ihrem jetzigen Job auf die Bewertung eines Szenarios, um die zu beurteilende Situation für alle Versuchspersonen identisch zu halten. Da es sich bei der Stichprobe darüber hinaus nicht um Arbeitnehmer im klassischen Sinne handelte, wurden Überlegungen angestellt, wie ein Maß für die Arbeitszufriedenheit in angemessener Weise operationalisiert werden konnte. Als Inhalt des Szenarios wurde daher nicht die Beschreibung eines Arbeitsplatzes, sondern vielmehr eine Beschreibung einer Praktikumsstelle dargestellt, da davon auszugehen war, daß sich die studentischen Versuchspersonen am besten in solch eine Situation hineinversetzen konnten. Bei der Beschreibung des Szenarios wurde darauf geachtet, daß sowohl Aspekte aufgenommen wurden, die eher wichtig für den Einzelnen sind (Art der Tätigkeit, Bezahlung, Zeitumfang, etc.), als auch eher gruppenbezogene Aspekte (Teamklima, Gruppenbüro, etc.). Dieses Praktikumsangebot sollten die Studenten dann hinsichtlich der Tatsache bewerten, wie zufrieden sie bei dieser Praktikumsstelle wären. Das Szenario wurde dabei in einem Pre-Test an N = 41 Studenten auf seine Verständlichkeit und Einsetzbarkeit hin geprüft. Die Fragen des Pre-Tests dahingehend, ob die Studenten sich in die Praktikumsbeschreibung hineinversetzen und sich mit ihr identifizieren konnten und ob diese verständlich formuliert war, erreichten vollkommen zufriedenstellende Werte.

Die Bewertung des Szenarios fand im Experiment letztendlich anhand von drei Aspekten statt: anhand eines Gesamturteils für die Zufriedenheit, anhand einer Bewerbungsbereitschaft für dieses Praktikum und anhand verschiedener Zufriedenheitsfacetten, die an die SAZ-Skala von Fischer und Lück (1972) angelehnt waren. Darüber hinaus wurde am Ende des Fragebogens zusätzlich nach der allgemeinen Zufriedenheit mit dem Studium gefragt, um ein weiteres Zufriedenheitsmaß, welches noch näher an der Erlebniswelt der Studenten ist, mitzerheben.

10.3 Ergebnisse

Die Stichprobe umfaßte insgesamt 195 Studenten, mit ausschließlich deutscher Herkunft, die aus den Semester 3 bis 17 ($M = 6.77$) stammten, von denen 57% männlich waren und die im Alter von 19 bis 31 Jahren ($M = 23.91$) waren. 77% der Stichprobe studierten das Fach BWL, 15% VWL und weitere 8% Wirtschaftspädagogik. Insgesamt 102 (52%) Studenten wurden independent und 93 (48%) interdependent geprint. Fast alle Studenten wiesen bereits einen starken Kontakt zum Arbeitsleben auf. Zwar hatten nur 28% bereits eine Ausbildung absolviert, jedoch arbeiteten 82% der Befragten während der Semesterferien und insgesamt 73% sogar während des Semesters. Anhand der Tatsache, daß zusätzlich 90% der Studierenden angaben, daß Praktika in ihrem Studium wichtig seien, war somit insgesamt davon auszugehen, daß die Beschreibung einer Praktikumsstelle für die Studenten grundsätzlich ein realistisches Szenario darstellen sollte. Die Eignung der Praktikumsbeschreibung als Szenario wurde dennoch im Experiment anhand von drei Fragen überprüft. Dabei zeigte sich auch an der positiven Bewertung der Versuchspersonen hinsichtlich der Skalen (Skala: jeweils von 1 = ‚stimme überhaupt nicht zu‘ bis 7 = ‚stimme voll und ganz zu‘) zur Verständlichkeit der Beschreibung ($M = 5.88$), der Fähigkeit, sich in diese hineinversetzen zu können ($M = 5.62$) und der Passung zum Studium ($M = 5.69$), daß das Praktikumszenario als geeignet angesehen werden konnte.

Hinsichtlich der Hypothesenprüfung ist bedeutsam, daß bezüglich keiner der oben genannten soziodemographischen Aspekte (Geschlecht, Alter, Studienfach, Arbeitserfahrung) ein signifikanter Unterschied zwischen der independent/individualistisch geprinten Gruppe (IND-Gruppe) und der interdependent/kollektivistisch geprinten Gruppe (KOLL-Gruppe) bestand; mit Ausnahme der Semesterzahl: Dort unterschieden sich die Gruppen geringfügig ($M_{IND} = 6.38$; $M_{KOLL} = 7.19$; $t(180.102) = 1.72$; $p < .10$; $d = .249$).

Manipulationscheck SJT

Die Ergebnisse des ersten Manipulationschecks anhand des social-judgement-tasks deuteten insgesamt auf eine Wirksamkeit des Primings hin. Alle Effekte besaßen die erwartete Richtung (vgl. Tabelle 17), wobei eine der Variablen deutlich signifikant bei einer mittleren Effektstärke war und zwei der Variablen zumindest kleine Effekte zeigten. Die kollektivistisch geprinte Gruppe gab deutlich stärker an, daß man

verpflichtet gewesen wäre, in der beschriebenen Situation zu helfen ($M_{IND} = 0.48$, $M_{KOLL} = 0.71$, $t(192.99) = 3.34$, $p < .01$, $d = .48$). Wenngleich die beiden anderen Effekte aufgrund der kleinen Stichprobe zwar nicht signifikant wurden, zeigten sie dennoch in die erwartete Richtung, daß die kollektivistisch geprägten Versuchspersonen das Verhalten weniger nachvollziehbar fanden ($M_{IND} = 1.98$, $M_{KOLL} = 1.83$, $t(193) = 0.92$, $p = .362$, $d = .13$) und eher der Meinung waren, daß die Person für ihr Verhalten gerügt werden sollte ($M_{IND} = 0.61$, $M_{KOLL} = 0.68$, $t(192.52) = 1.01$, $p = .313$, $d = .15$). Hinsichtlich des dritten Items muß jedoch durchaus kritisch angemerkt werden, daß gegenüber dem Originalwortlaut des Items im Englischen („Should Lisa be punished?“) die deutsche Übersetzung, die in dieser Studie verwendet wurde, aufgrund des Wortes „gerügt“ als etwas unglücklich bezeichnet werden kann. Mehrere Versuchspersonen hatten Anmerkungen zu dieser Formulierung gemacht, was möglicherweise dazu geführt haben könnte, daß der Effekt auf diesem Item letztendlich nicht so stark war.

Aufgrund dieser Ergebnisse des ersten Manipulationschecks kann somit insgesamt aber zumindest in vorsichtiger Weise davon ausgegangen werden, daß das Priming, zu dem die Versuchspersonen zufällig zugeteilt wurden, wirksam gewesen ist und bei der IND-Gruppe in der Tat stärker unabhängige Wissensinhalte und bei der KOLL-Gruppe stärker interdependente Wissensinhalte aktiviert worden sind.

Tabelle 17: Ergebnisse des SJT Manipulationschecks

	N	M	SD	T	df	p	d
Wie nachvollziehbar findest du Lisas Verhalten? (Skala: 1 = ‚überhaupt nicht nachvollziehbar‘ bis 7 = ‚außerordentlich nachvollziehbar‘)							
IND-Gruppe	102	1.98	1.12	.92	193	.362	.13
KOLL-Gruppe	93	1.83	1.20				
Wäre Lisa deiner Meinung nach verpflichtet gewesen, Anne zu helfen? (Skala: 0 = nein, 1 = ja)							
IND-Gruppe	102	.48	.50	3.34	192.99	.001	.48
KOLL-Gruppe	93	.71	.46				
Sollte Lisa für ihr Verhalten gerügt werden? (Skala: 0 = nein, 1 = ja)							
IND-Gruppe	102	.61	.49	1.01	192.52	.313	.15
KOLL-Gruppe	93	.68	.47				

Die Überprüfung anhand eines zweiten Manipulationschecks am Ende des Fragebogens, der die Nachhaltigkeit des Primings messen sollte, zeigte jedoch überraschenderweise andere Ergebnisse. Für die Berechnung der Ergebnisse wurden alle Aussagen zum TenST durch vier unabhängige Rater in entweder individuenbezogene (d.h. individualistische) oder aber gruppenbezogenen (d.h. kollektivistische) Aussagen eingeordnet. Der entsprechende Kappa-Koeffizient von $\kappa = .926$ zeigte dabei eine hohe Übereinstimmung der Rater an. Betrachtet man die Ergebnisse des TenST als zweitem Manipulationscheck im Einzelnen, zeigten sich interessanterweise keine signifikanten Unterschiede zwischen den Experimentalgruppen (vgl. Tabelle 18). Zwar gaben die Versuchspersonen in beiden Gruppen jeweils insgesamt mehr independente Aussagen als interdependente Aussagen an (independent: $M_{IND} = 0.65$, $M_{KOLL} = 0.64$; interdependent: $M_{IND} = 0.35$, $M_{KOLL} = 0.36$), jedoch zeigte sich nicht der erwartete Unterschied zwischen den beiden Priminggruppen. Eine Wirksamkeit des Primings, die bis zum Ende der Befragung anhielt, konnte somit anhand des TenST nicht nachgewiesen werden.

Tabelle 18: Ergebnisse des TenST Manipulationschecks

	N	M	SD	T	df	p	d
Anteil individualistischer Aussagen (Skala: 0 = ‚keine individualistische Aussage‘ bis 1 = ‚ausschließlich individualistische Aussagen‘)							
IND-Gruppe	102	0.65	0.19	.36	193	.718	.05
KOLL-Gruppe	93	0.64	0.21				
Anteil kollektivistische Aussagen (Skala: 0 = ‚keine kollektivistische Aussage‘ bis 1 = ‚ausschließlich kollektivistische Aussagen‘)							
IND-Gruppe	102	0.35	.19	.36	193	.718	.05
KOLL-Gruppe	93	0.36	.21				
Verhältnis von individualistischer vs. kollektivistische Aussagen (Skala: -1 = ‚ausschließlich kollektivistische Aussagen‘ bis 1 = ‚ausschließlich individualistische Aussagen‘)							
IND-Gruppe	102	.30	.38	.36	193	.718	.05
KOLL-Gruppe	93	.28	.42				

Betrachtet man die Ergebnisse der beiden Manipulationschecks zusammen, scheint auf der einen Seite das Ergebnis des SJT den Ergebnissen des TenST zu widersprechen. Gemäß dem ersten Manipulationscheck konnte durchaus von einer Wirksamkeit des Primings ausgegangen werden, gemäß dem zweiten Manipulationscheck wäre sie aber nicht vorhanden gewesen. Bei diesen unterschiedlichen Ergebnissen ist jedoch zu

beachten, daß diese Manipulationschecks an zwei unterschiedlichen Stellen der Befragung eingesetzt wurden. Die Ergebnisse können daher durchaus auch so interpretiert werden, daß direkt nach der Durchführung des Primings ein entsprechender Effekt anhand des unmittelbar anschließenden SJT gemessen werden konnte. Dieser Effekt hat jedoch möglicherweise innerhalb des Experiments dann soweit nachgelassen, daß er am Ende des Fragebogens anhand des TenST nicht mehr nachgewiesen werden konnte. Demnach könnte dennoch von einer Wirksamkeit des Primings ausgegangen werden, bei der jedoch nicht ganz klar ist, wie lange genau diese Wirkung anhielt. Diese potentielle Möglichkeit des Nachlassens der Primingwirkung könnte dabei durchaus bedeutsam sein für die Ergebnisse der anderen Testskalen.

Die Antworttendenz-Maße

Bezüglich der Maße zur Erfassung der Antworttendenz wurden sowohl anhand der verkürzten PVQ-Skala von Schwarz als auch anhand der Skala von Greenleaf entsprechende Werte für das zustimmende und das extreme Antwortverhalten berechnet. Die Werte für die Soziale Erwünschtheit ergaben sich aus den Kurzskalen der Marlowe-Crowne- und der BIDR-Skala. Die Ergebnisse bzgl. dieser verschiedenen Antworttendenzmaße zeigen jedoch insgesamt, daß anhand keiner der Indizes ein signifikanter Unterschied zwischen den beiden Priming-Gruppen erfassen werden konnte (vgl. Tabelle 19).

Lediglich beim NARS-Index der Greenleaf-Skala ($M_{\text{IND}} = 0.02$, $M_{\text{KOLL}} = 0.05$, $t(193) = -0.94$, $p = .347$, $d = .14$), beim Extremtendenz-Index der PVQ-Skala ($M_{\text{IND}} = 0.18$, $M_{\text{KOLL}} = 0.20$, $t(193) = -0.92$, $p = .362$, $d = .13$) sowie bei der Teilskala des BIDR hinsichtlich Selbsttäuschung ($M_{\text{IND}} = 11.11$, $M_{\text{KOLL}} = 11.76$, $t(190) = 1.54$, $p = .126$, $d = .22$) wurden zumindest kleine Effekte deutlich. Während die Ergebnisse der Greenleaf-Skala erwartungsgemäß zeigten, daß kollektivistisch geprägte Versuchspersonen immerhin eine leicht höhere Zustimmungstendenz im Mittel besaßen, ist das Ergebnis des Extremtendenz-Index der PVQ-Skala und auch der Selbsttäuschungsskala eher überraschend. Hier wäre eher zu erwarten gewesen, daß im Gegensatz zu den tatsächlichen Ergebnissen die individualistisch geprägte Gruppe stärkere Extremtendenzen und auch eine höhere Selbsttäuschung gezeigt hätte. Insgesamt waren jedoch alle diese drei geringen Effekte bei der geringen Größe der Stichprobe nicht signifikant.

Tabelle 19: Ergebnisse der Antworttendenz-Maße

	N	M	SD	T	df	p	d
Index NARS der PVQ-Skala (Wertebereich: -1 = ‚hohes Ablehnungsverhalten‘ bis 1 = ‚hohes Zustimmungsverhalten‘)							
IND-Gruppe	102	0.32	0.20	.471	193	.638	.07
KOLL-Gruppe	93	0.31	0.19				
Index NARS der Greenleaf-Skala (Wertebereich: -1 = ‚hohes Ablehnungsverhalten‘ bis 1 = ‚hohes Zustimmungsverhalten‘)							
IND-Gruppe	102	.02	.22	-0.94	193	.347	.14
KOLL-Gruppe	93	.05	.20				
Index Extremtendenz-Norm der PVQ-Skala (Wertebereich: 0 = ‚keine Extremtendenz‘ bis 1 = ‚hohe Extremtendenz‘)							
IND-Gruppe	102	.18	.14	-0.92	193	.362	.13
KOLL-Gruppe	93	.20	.16				
Index Extremtendenz-Norm der Greenleaf-Skala (Wertebereich: 0 = ‚keine Extremtendenz‘ bis 1 = ‚hohe Extremtendenz‘)							
IND-Gruppe	102	.29	.18	-0.55	193	.585	.08
KOLL-Gruppe	93	.30	.20				
Index Mittentendenz-Norm der PVQ-Skala (Wertebereich: 0 = ‚keine Mittentendenz‘ bis 1 = ‚hohe Mittentendenz‘)							
IND-Gruppe	102	.40	.15	-0.29	190.278	.774	.04
KOLL-Gruppe	93	.41	.12				
Index Mittentendenz-Norm der Greenleaf-Skala (Wertebereich: 0 = ‚keine Mittentendenz‘ bis 1 = ‚hohe Mittentendenz‘)							
IND-Gruppe	102	.34	.14	-0.38	193	.707	.05
KOLL-Gruppe	93	.35	.16				
Index Soziale Erwünschtheit Marlowe-Crowne (Wertebereich: 4 = ‚geringe Soziale Erwünschtheit‘ bis 28 = ‚hohe Soziale Erwünschtheit‘)							
IND-Gruppe	102	12.40	3.14	-0.64	193	.524	.09
KOLL-Gruppe	93	12.10	3.54				
Index Soziale Erwünschtheit BIDR - Fremdtäuschung (Wertebereich: 3 = ‚geringe Soziale Erwünschtheit‘ bis 21 = ‚hohe Soziale Erwünschtheit‘)							
IND-Gruppe	102	9.68	3.88	-0.39	192	.695	.06
KOLL-Gruppe	93	9.46	3.97				
Index Soziale Erwünschtheit BIDR - Selbsttäuschung (Wertebereich: 3 = ‚geringe Soziale Erwünschtheit‘ bis 21 = ‚hohe Soziale Erwünschtheit‘)							
IND-Gruppe	102	11.11	2.79	1.54	190	.126	.22
KOLL-Gruppe	93	11.76	3.07				

Der Effekt, der aufgrund des Primings erwartet worden wäre, ließ sich somit an keinem der ermittelten Indizes statistisch nachweisen. Trotz der scheinbar stärkeren Aktivierung von einerseits independenten und andererseits interdependenten Kognitionsinhalten

wirkte sich dies offensichtlich nicht auf das generelle Kommunikationsverhalten der Versuchspersonen aus; zumindest nicht bezüglich der Meßwerte, die in dieser Studie erfaßt wurden.

Die Zufriedenheitsurteile

Die Zufriedenheitsurteile zum Praktikumsszenario, die sowohl anhand eines Gesamturteils als auch anhand von verschiedenen Facetten erhoben wurden, sowie die Bewerbungsintention zeigten trotz der vorangegangenen Ergebnisse der Antworttendenz-Maße interessanterweise jedoch zum Teil deutliche Unterschiede zwischen den Priming-Gruppen (vgl. Tabelle 20).

Tabelle 20: Ergebnisse der Zufriedenheitsurteile

	N	M	SD	T	df	p	d
Allgemeine Zufriedenheit (Wertebereich: 1 = ‚überhaupt nicht zufrieden‘ bis 7 = ‚außerordentlich zufrieden‘)							
IND-Gruppe	101	5.76	0.75	-2.14	162.90	.034	.32
KOLL-Gruppe	92	5.48	1.05				
Gesamt-Zufriedenheit anhand aller Facetten (Wertebereich: 1 = ‚überhaupt nicht zufrieden‘ bis 7 = ‚außerordentlich zufrieden‘)							
IND-Gruppe	102	5.68	.70	-1.88	193	.062	.32
KOLL-Gruppe	93	5.48	..80				
Bewerbungswahrscheinlichkeit (Wertebereich: 1 = ‚überhaupt nicht wahrscheinlich‘ bis 7 = ‚sehr wahrscheinlich‘)							
IND-Gruppe	100	5.60	1.13	-2.14	167.70	.034	.27
KOLL-Gruppe	92	5.18	1.51				
Zufriedenheit mit dem Studium (Wertebereich: 1 = ‚überhaupt nicht zufrieden‘ bis 7 = ‚außerordentlich zufrieden‘)							
IND-Gruppe	102	5.36	1.09	-2.56	176.09	.01	.37
KOLL-Gruppe	92	4.91	1.33				

Die Zufriedenheit mit dem Praktikumsplatz war bei den individualistisch geprägten Versuchspersonen nicht nur signifikant höher hinsichtlich des globalen Allgemeinurteils ($M_{IND} = 5.76$, $M_{KOLL} = 5.48$, $t(162.90) = -2.14$, $p < .05$, $d = .32$) sondern auch hinsichtlich des Gesamtmaßes, welches anhand der einzelnen Facetten errechnet wurde ($M_{IND} = 5.68$, $M_{KOLL} = 5.48$, $t(193) = -1.88$, $p < .10$, $d = .32$). Die aus den einzelnen Facetten gebildete Gesamtzufriedenheitsskala erreichte dabei im Übrigen ein zufriedenstellendes Cronbachs Alpha von $\alpha = .80$. Die Versuchspersonengruppe, deren

independenten Selbstkonzeptaspekte stärker aktiviert wurden, gab demnach anhand beider Kennwerte ein deutlich höheres Zufriedenheitsurteil für diese spezifische Praktikumssituation ab, die ja für alle Befragten konstant gehalten wurde.

Auch hinsichtlich der zusätzlich abgefragten Bewerbungswahrscheinlichkeit zeigte sich der gleiche Effekt ($M_{\text{IND}} = 5.60$, $M_{\text{KOLL}} = 5.18$, $t(167.70) = -2.14$, $p < .05$, $d = .27$). Auch wenn dieses Item zwar keine direkte Zufriedenheitsmessung darstellte, wurde mit ihm dennoch auch eine Art Beurteilung des Praktikumsszenarios erhoben. Hierbei bewerteten die individualistisch geprägten Personen dieses Szenario ebenfalls signifikant besser. Selbst bei der Frage nach der allgemeinen Zufriedenheit mit dem Studium gab diese Priming-Gruppe ein signifikant höheres Urteil ab ($M_{\text{IND}} = 5.36$, $M_{\text{KOLL}} = 4.19$, $t(176.09) = -2.56$, $p < .05$, $d = .37$).

Betrachtet man diese Ergebnisse hinsichtlich der Zufriedenheitsmaße im Gesamten, erscheinen sie auf den ersten Blick erst einmal überraschend, da sich ja bezüglich der Antworttendenz-Maße keine Unterschiede zwischen den Priming-Gruppen feststellen ließen. Demnach war gemäß den theoretischen Vorüberlegungen eigentlich nicht mehr davon auszugehen, daß die Zufriedenheitsmaße unterschiedlich ausfallen sollten, denn eigentlich hätte das Priming zu einer bestimmten Form des Antwortverhaltens führen sollen, die dann wiederum verantwortlich ist für die Antworten zu den Zufriedenheitsfragen. Dennoch liegt der Schluß nahe, daß dieser Unterschied im angekreuzten Zufriedenheitsurteil durch das Priming zustande gekommen sein muß, da die Versuchspersonen zufällig den beiden Experimentalgruppen zugeteilt wurden und sich weitestgehend keine Unterschiede zwischen den Gruppen anhand von soziodemographischen Aspekten feststellen ließen. Der einzige Unterschied bestand somit eigentlich nur im Priming. Lediglich die unterschiedliche Semesteranzahl der beiden Experimentalgruppen könnte möglicherweise für die Unterschiede in den Zufriedenheitsurteilen noch verantwortlich gewesen sein. Dies wäre vor allem für die Aussage zur Zufriedenheit mit dem Studium eine plausible Annahme. Eine erste Bestätigung dieser Vermutung zeigte sich durch die Berechnung der Korrelation zwischen diesen Zufriedenheitsurteilen und der Semesteranzahl, welche mit $r = -.40$ ($p < .001$) einen deutlichen Zusammenhang für die Zufriedenheit mit dem Studium ergab. Studenten in höheren Semestern gaben folglich deutlich schlechtere Zufriedenheitsurteile ab. Wie jedoch weitere Korrelationsanalysen zeigen konnten, standen die anderen Zufriedenheitsmaße in überhaupt keinem bedeutsamen Zusammenhang mit der Semesteranzahl ($r_{\text{allg. Gesamtzufriedenheit}} = .12$, $p = .09$;

$r_{\text{Gesamtzufriedenheit Facetten}} = .09, p = .19$; $r_{\text{Bewerbungswahrscheinlichkeit}} = .01, p = .91$). Um dennoch den Einfluß des Primings unabhängig von der Semesteranzahl noch einmal erneut zu prüfen, wurden in einer weiteren Analyse jeweils die Zufriedenheitsurteile der beiden Priming-Gruppen einer Kovarianzanalyse unterzogen (vgl. Tabelle 21).

Tabelle 21: Kovarianzanalyse der Zufriedenheitsurteile

	F	df	p	η^2
Allgemeine Zufriedenheit				
Effekt des Primings	3.978	1	.048	.021
Effekt der Semesteranzahl	1.779	1	.142	.011
Gesamt-Zufriedenheit anhand aller Facetten				
Effekt des Primings	2.997	1	.085	.015
Effekt der Semesteranzahl	1.187	1	.277	.006
Bewerbungswahrscheinlichkeit				
Effekt des Primings	4.683	1	.032	.024
Effekt der Semesteranzahl	.020	1	.888	.000
Zufriedenheit mit dem Studium				
Effekt des Primings	4.237	1	.041	.022
Effekt der Semesteranzahl	33.305	1	.000	.148

Die Ergebnisse zeigen, daß die Semesteranzahl nur hinsichtlich der Zufriedenheit mit dem Studium einen deutlichen Einfluß hatte ($\eta^2 = .15, p < .001$). Dieser Einfluß war darüber hinaus auch deutlich stärker als der Einfluß des Primings ($\eta^2 = .022, p < .05$). Der Unterschied bei diesem Zufriedenheitsmaß ist daher in der Tat eher auf die nicht intendierte Unterschiedlichkeit der Experimentalgruppen hinsichtlich der Semesterzahl zurückzuführen. Bei allen anderen Zufriedenheitsmaßen jedoch unterschied sich das Urteil nur bezüglich der Priming-Bedingungen. Die Zufriedenheitsurteile, die für das Praktikums-Szenario abgegeben wurden, konnten somit nur durch die Unterschiedlichkeit des Primings hervorgerufen worden sein.

10.4 Die Diskussion der experimentellen Ergebnisse

Ziel des Experiments war die Überprüfung der argumentativen Kausalkette, daß eine bestimmte Form des kulturellen Denkens zu einer bestimmten Form des Kommunikationsverhaltens führt, welche schließlich die Höhe eines abgegebenen Zufriedenheitsurteils beeinflußt. Diese Überlegungen finden in den Ergebnissen des

Experiments in Teilen ihre Bestätigung. Jedoch läßt sich die Kausalkette nicht vollständig durch die experimentelle Überprüfung bestätigen.

Das Priming, welches independente und interdependenten Wissensinhalte aktivierte, führte in der Tat in der erwarteten Weise dazu, daß die sozusagen individualistische Versuchspersonengruppe ein positiveres Zufriedenheitsurteil abgab als die kollektivistische Gruppe. Dabei ist erneut zu betonen, daß die Beschreibung des Praktikumsszenarios für beide Gruppen komplett identisch war. Ein unterschiedliches Zufriedenheitsurteil wäre somit nicht zu erwarten gewesen und kann daher nur auf das unterschiedliche Priming zurückgeführt werden, da dies der einzige Aspekt war, anhand dessen sich die Experimentalgruppen unterschieden. Hinsichtlich eines zentralen Aspekts der theoretischen Überlegungen kann folglich das Experiment eine Bestätigung liefern: In individualistischen Kulturen werden unabhängig von der zu bewertenden Situation höhere Zufriedenheitsurteile abgegeben als in kollektivistischen Kulturen. Die für gewöhnlich durchgeführte Interpretation von gemessenen Zufriedenheitswerten als vollständig ‚wahre‘ Zufriedenheitswerte und deren direkter Vergleich zwischen Ländern und Kulturen scheint somit auch gemäß den Ergebnissen dieses Experiments nicht als sinnvoll zu sein.

Die Überlegung, daß diese Unterschiedlichkeit der Zufriedenheitsurteile vor allem durch eine unterschiedliche Kommunikationsnorm beeinflusst ist, kann jedoch anhand der experimentellen Ergebnisse so direkt nicht bestätigt werden. Dies kann verschiedene Gründe gehabt haben, die im Folgenden diskutiert werden. Auf den ersten Blick erscheint es so, daß die argumentative Kausalkette in ihrer bisherigen Form so nicht mehr zu halten wäre. Zwar führte die unterschiedliche Ausgangssituation im Priming am Ende zum erwarteten unterschiedlichen Effekt bei der Zufriedenheit, jedoch läßt das vermittelnde Bindeglied ja die erwarteten Unterschiede vermissen. Man müßte demnach eigentlich annehmen, daß zwar individualistische Kulturen höhere Zufriedenheitsurteile als kollektivistische Kulturen abgeben, jedoch dieser Unterschied noch durch einen weiteren Aspekt neben der Zustimmungsnorm verursacht sein müßte. Eine komplette Zurückweisung der Kausalkette und der inhaltlichen Überlegungen könnte unter Umständen aber verfrüht sein, da auch verschiedene methodische Aspekte dazu geführt haben könnten, daß nicht alle Effekte wie erwartet gemessen werden konnten. Ein Hauptgrund für die fehlenden Unterschiede bei den Antworttendenz-Maßen ist mit großer Wahrscheinlichkeit im möglichen Nachlassen des Primings zu suchen. Die Ergebnisse des zweiten Manipulationschecks zeigten ja, daß am Ende des Fragebogens,

d.h. also zumindest ab der Position des TenST, von keiner Primingwirksamkeit mehr ausgegangen werden konnte. Wann genau das Priming in welche Stärke davor schon nachgelassen haben könnte, läßt sich verständlicherweise nicht mit Sicherheit sagen. Wenn jedoch davon ausgegangen würde, daß die Wirkung des Priming nicht zufällig abrupt beim TenST auf einmal nachgelassen hätte, sondern stattdessen kontinuierlich im Laufe des Experiments weniger wurde, wäre durchaus verständlich, warum zwar bei den Zufriedenheitsmaßen noch ein Unterschied zu finden war, jedoch bei den Skalen der Antworttendenz-Indizes nicht mehr. Während die Zufriedenheitsmaße nämlich direkt nach dem Priming und dem ersten Manipulationscheck von den Versuchspersonen beurteilt wurden, gaben die Versuchsteilnehmer das Urteil zu den Antworttendenz-Indizes erst nach einer kleinen zeitlichen Pause und direkt vor dem TenST ab. Möglicherweise war daher schon bei den Skalen des PVQ von Schwartz (2007) sowie des Instruments von Greenleaf (1992b) die Wirkung des Primings schon soweit zurückgegangen, daß keine Unterschiede zwischen den Gruppen mehr gemessen werden konnten.

Wollte man die einzelnen Instrumente auf ihre Anwendbarkeit hin kritisch beleuchten, könnte man möglicherweise auch hinsichtlich des TenST anmerken, daß es sich bei diesem Instrument um einen nicht ganz so einfach anwendbaren Manipulationscheck handelt. Grundsätzlich bedarf die Auswertung des Instruments nämlich immer der Interpretation von textlichen Aussagen durch Rater. Diese läßt aufgrund der sprachlichen Ungenauigkeiten und des unterschiedlichen sprachlichen Verständnisses der jeweiligen Rater immer Spielräume zu. Zwar wird versucht, diese unterschiedliche Interpretation durch Schulungen und einheitliche Kategorisierungsschemata zu vereinheitlichen. Jedoch besteht immer auch die Möglichkeit, daß nicht gefundene Unterschiede zwischen Experimentalgruppen in kleinen Teilen auch durch die unterschiedliche Kategorisierung im Auswertungsprozeß bedingt sind. Man könnte beim TenST hinsichtlich eines gänzlich anderen Aspektes aber auch kritisch die grundsätzliche Frage stellen, ob ein Manipulationscheck, der Unterschiede in ich- und wir-bezogenem Denken messen soll, selbst aber ausschließlich Aussagen erfordert, die mit ‚Ich‘ beginnen, nicht auch in einer gewissen Form als eine Art Priming wirkt. Er würde dann gegebenenfalls das experimentelle Priming im Prinzip konterkarieren. Dem entgegen steht jedoch eindeutig, daß der TST bzw. der TenST als eines der gängigsten Verfahren angesehen werden kann, die zur Überprüfung dieses Primings von interkulturellem Denken eingesetzt wurden (Oyserman & Lee, 2008). In sehr vielen der

bisherigen Studien stellt dieses Verfahren den üblichen Manipulationscheck dar und zeigte in den veröffentlichten Studien immer wieder deutliche Effekte. Trotz der potentiell möglichen Kritik an diesem Verfahren ist es daher nicht sehr wahrscheinlich, daß in Wahrheit in dem hier vorliegenden Experiment auch am Ende des Fragebogens noch Unterschiede zwischen den Experimentalgruppen bestanden hätten, die lediglich nur durch den TenST nicht zum Vorschein kamen. Plausibler bleibt somit die Annahme, daß die Wirkung des Primings über die Zeit nachgelassen hat.

Kritisch beleuchten ließen sich auch die Meßskalen, die zur Erfassung der Antworttendenzen eingesetzt wurden. Hierbei ist vor allem das Fehlen von Effekten hinsichtlich der Zustimmungstendenz von Bedeutung, da diesbezüglich klare Hypothesen für die Wirksamkeit des individualistischen/kollektivistischen Primings bestanden hatten. Das Fehlen der Effekte bei Extrem-/Mittentendenzen sowie Sozialer Erwünschtheit könnte hingegen auch durch rein inhaltliche Überlegungen erklärt werden, denn auch die frühere Forschung hat hinsichtlich dieser Indizes keine klaren Ergebnisse in Abhängigkeit von Individualismus/Kollektivismus finden können. Was jedoch das Fehlen der Effekte bei zustimmendem Antwortverhalten angeht, könnte möglicherweise die PVQ-Skala nicht geeignet für die Erfassung des Index gewesen sein. Auf der einen Seite erfaßt sie nur Werte-bezogene Aspekte. Diese waren möglicherweise nicht heterogen genug. Auf der anderen Seite enthält die Antwortskala auch nicht die üblichen ‚stimme-zu vs. stimme-nicht-zu‘-Label, sondern eine Ähnlichkeitseinschätzung. Beides kann daher dazu geführt haben, daß Zustimmungsverhalten nicht durch diese Skala angemessen erfaßt werden konnte. Ein besseres Instrument ließ sich jedoch in der Recherche für dieses Experiment nicht finden. Neben der PVQ-Skala wurde aber ja auch die Skala von Greenleaf (1992b) eingesetzt. Auch bei dieser zeigten sich jedoch nicht die erwarteten Effekte. Diese Skala könnte unter Umständen dahingehend kritisiert werden, daß sie eigentlich nur zur Messung von extremem Antwortverhalten in der bisherigen Forschung erfolgreich eingesetzt wurde, jedoch nicht zur Messung von zustimmendem Antwortverhalten. Für die Messung von Zustimmungstendenzen sollte die Skala aber auch eigentlich die notwendigen Voraussetzungen (d.h. heterogenes Itemset, Stimme-zu-Label) besessen haben. Die Überlegung, daß in Wahrheit Unterschiede im Kommunikationsverhalten im Experiment erzeugt wurden, und nur nicht durch die ungeeigneten Instrumente gemessen werden konnten, ist daher zwar vielleicht für den PVQ noch plausibel, läßt sich jedoch hinsichtlich des Greenleaf-Instruments nur bedingt halten. Letzteres wurde

auch weiter hinten im zeitlichen Ablauf des Experiments eingesetzt und wurde von den Versuchspersonen erst kurz vor dem TenST bearbeitet. Insgesamt sprechen diese Überlegungen auch eher dafür, von einem Nachlassen des Primings auszugehen.

Das Nachlassen des Primings wird auch dann als Erklärung plausibler, wenn man bedenkt, daß das Experiment aufgrund der Vielzahl der integrierten Skalen insgesamt bis zu 30 Minuten dauerte, z.T. nahmen sich einige Versuchspersonen sogar mehr Zeit. Verbunden mit der Tatsache, daß das Experiment vor Beginn einer Prüfungsphase mit den Studenten durchgeführt wurde und somit die Teilnehmer möglicherweise im Zeitstreß sowie weniger konzentriert beim Ausfüllen der Skalen waren, könnte die Dauer des Experiments in der Tat dafür gesorgt haben, daß die Wirkung des Primings nach und nach abfiel. Eine Verkürzung des im Experiment eingesetzten Fragebogens wurde zwar vorab ausführlich abgewogen, ließ sich jedoch aufgrund der verschiedenen zu erfassenden Konstrukte letztendlich nicht durchführen. Eine erneute experimentelle Durchführung sollte diesen Aspekt jedoch genauer beachten.

Trotz dieser möglichen methodischen Kritikpunkte sollte jedoch nicht außer Acht gelassen werden, daß das Vorhandensein der Unterschiede in den Zufriedenheitsurteilen weiterhin existent bleibt und nur schwer durch andere zufällige Effekte erklärt werden könnte, da diese ja gerade im Experiment kontrolliert wurden. Von daher bleibt als Erklärung lediglich die Verschiedenheit des Primings. Insgesamt sollten daher die Ergebnisse des Experiments durchaus als eine weitere Bestätigung der vorangegangenen Überlegungen dieser Arbeit verstanden werden. Lediglich bezüglich der Erklärungskraft von zustimmendem Antwortverhalten bleiben aufgrund der methodischen Probleme weitere Fragen offen, die in einer zusätzlichen Studie überprüft werden könnten.

11 Diskussion

Ziel der Arbeit ist es, einen Beitrag zur Problematik der interkulturellen Arbeitszufriedenheitsmessung zu liefern. Der zentrale Fokus bezog sich dabei auf die Fragestellung, inwieweit die Absolutwerte zu Fragen nach der Arbeitszufriedenheit über Ländergrenzen hinweg verglichen werden können. Zur Beantwortung dieser Fragestellung wurde im theoretischen Teil sowohl die Bedeutung für die unternehmerische Praxis als auch der Bezug zu den grundsätzlichen Fragen der interkulturellen Forschung dargestellt. Die hier behandelte Thematik betrifft im Prinzip alle weltweit tätigen Unternehmen, die internationale Befragungen ihrer Belegschaft durchführen. Jedoch ist die Problemstellung weder im Unternehmenskontext noch im Bereich der Forschung bisher ausgiebig genug gewürdigt worden. In Unternehmen, die die Arbeitszufriedenheit der Mitarbeiter als Steuerungsinstrument regelmäßig im Rahmen von Mitarbeiterbefragungen erheben, werden vielfach die gemessenen Werte der einzelnen länderspezifischen Standorte ohne großes Hinterfragen in direkter Weise in Beziehung gesetzt und durch den Vergleich der absoluten Zufriedenheitsniveaus unmittelbare Rückschlüsse auf die Unternehmenslage vor Ort gezogen. In ähnlicher Weise lassen sich auch – wie es am Beispiel der prominenten Studie des Gallup Instituts (2003) zu Beginn der Arbeit exemplarisch dargestellt wurde – selbst im forschungsnahen Umfeld oftmals Aussagen finden, bei denen international erhobene Daten zur Arbeitszufriedenheit ohne großes Hinterfragen miteinander verglichen werden. Hinsichtlich dieser Vorgehensweise wurde jedoch im Rahmen der theoretischen Ausführungen dargelegt, daß sie als unangemessen gelten muß. Begründet wurde dies vor allem mit Hilfe der bisherigen Erkenntnisse der interkulturellen Forschung, gemäß deren Ergebnissen verschiedene Länder in Befragungen ein jeweils spezifisches Antwortverhalten zeigen, welches durch die kulturellen Gegebenheiten vor Ort zu erklären ist. Diese methodische Problematik, die sich generell auf alle Arten von interkulturellen Befragungen bezieht, sollte daher auch einen Einfluß auf die Arbeitszufriedenheitsmessung im Rahmen von Mitarbeiterbefragungen haben. Da sich zu diesem Zusammenhang aber überraschenderweise keine bedeutsamen Studien in der Literatur finden ließen, stellte die Betrachtung des Einflusses von Antworttendenzen auf die Höhe des Arbeitszufriedenheitsniveaus einen zentralen Aspekt der hier vorliegenden Arbeit dar. Darüber hinaus sollten jedoch die Ergebnisse zu diesem Zusammenhang

auch in die übrigen Erkenntnisse der bisherigen Forschung eingeordnet werden. Der Zusammenhang von Antworttendenz und Arbeitszufriedenheit wurde daher ebenfalls auf seinen Bedeutungsgehalt in Bezug auf die beiden weiteren, jeweils bivariaten Zusammenhänge von einerseits Kultur und Antworttendenz und andererseits Kultur und Arbeitszufriedenheit geprüft. Gemäß den theoretischen Überlegungen ließ sich annehmen, daß der Gesamtzusammenhang sich derart gestalten sollte, daß die Kultur eines Landes als ein wichtiger, zentraler Sozialisierungsfaktor ein länderspezifisches Kommunikationsverhalten erzeuge und dieses Verhalten dann wiederum einen bedeutsamen Einfluß auf ein Arbeitszufriedenheitsurteil habe. Diese Perspektive, die sich mit den Zusammenhängen der drei zentralen Konstrukte Kultur, Antworttendenz und Arbeitszufriedenheit auf dem Länderniveau beschäftigte, wurde zusätzlich ergänzt durch die Betrachtung der Einflüsse auf dem Individuallevel. Auch dort war anzunehmen, daß persönliche Kommunikationstendenzen sich auf das persönlich abgegebene Arbeitszufriedenheitsurteil eines Befragten auswirken sollten.

11.1 Erkenntnisse aus den empirischen Analysen

Die empirischen Analysen der Arbeit ließen sich in mehrere Abschnitte untergliedern, die einerseits jeweils einzelne Aspekte der Fragestellungen prüften und andererseits auch zur Beantwortung der Gesamtfragestellung der gegenseitigen Überprüfung von Ergebnissen der jeweils anderen empirischen Berechnungen dienten. Ein großer Teil der Analysen wurde dabei anhand von Sekundärdaten durchgeführt. Neben der Darstellung der rein deskriptiven Länderunterschiede galt dies vor allem für die Überprüfung der bivariaten Zusammenhänge der drei zentralen Konstrukte dieser Arbeit. Auch die Mediationsanalysen, die den Gesamtzusammenhang der drei Konstrukte untereinander prüften, wurden anhand dieser Sekundärdaten durchgeführt. In gleicher Weise bedienten sich ebenfalls die durchgeführten Mehrebenenanalysen dieser bereits veröffentlichten Daten, jedoch bezogen sie als weiteren Schritt auch die Individualdaten der Befragten mit ein. Zusätzlich zu diesen umfangreichen Sekundäranalysen wurden jedoch ebenfalls die anhand der Sekundärdaten ermittelten Zusammenhänge mit Hilfe eines Experiments überprüft und somit weitere Primärdaten zur Überprüfung der theoretischen Überlegungen herangezogen.

Nationale Unterschiede im Arbeitszufriedenheitsniveau

Die deskriptiven Ergebnisse, die anhand von zwei umfangreichen internationalen Datensätzen (dem World-Values-Survey, WVS, und dem Eurobarometer, EB) errechnet wurden, konnten zeigen, daß sich das Arbeitszufriedenheitsniveau zwischen einzelnen Ländern zum Teil deutlich unterscheidet. Dabei zeigte sich bei allen Ländermittelwerten jedoch auch die typische Verschiebung der Zufriedenheitswerte hin zum positiven Ende der Skala. Dieser Effekt, der regelmäßig auf dem Individuallevel gemessen werden kann, zeigte sich somit hier auch auf dem Länderlevel. In den Daten des WVS wurden insgesamt 6,4% der Varianz der Zufriedenheitsurteile allein durch die Zugehörigkeit zu einem bestimmten Land erklärt, im EB waren es insgesamt 4,6% der Varianz. Diese Werte erscheinen auf den ersten Blick zwar recht niedrig. Es muß jedoch bedacht werden, daß bei den dazu durchgeführten varianzanalytischen Berechnungen die Unterschiedlichkeit der Urteile auf dem Individuallevel komplett in die Berechnung mit eingeht und die interindividuelle Varianz innerhalb der Länder als eine Art Fehlervarianz mitberücksichtigt wird. Selbst Hofstede (2001) geht in seinen einflußreichen Analysen aufgrund dieser statistischen Einschränkung davon aus, daß man die varianzanalytisch gemessenen Effekte lediglich als sehr konservativen Schätzer des tatsächlichen Effekts ansehen sollte. Er selbst berichtet in seinen Analysen ebenfalls nur von 4,2% Varianzaufklärung seiner Daten durch die Zugehörigkeit der Befragten zu einem Land, betont jedoch, daß „this, however, was 16 times as much as could be expected on the basis of pure chance“ (S. 50). Vor diesem Hintergrund sind die hier gemessenen Anteile der Varianzaufklärung ebenfalls als ein bedeutsamer Effekt anzusehen, der allein durch die Zugehörigkeit zu einem bestimmten Land zustande kommt. Darüber hinaus zeigte sich, daß diese Zufriedenheitsniveaus auch ein relativ stabiles Maß darstellen und über die Zeit hinweg, die zwischen den Erhebungszeitpunkten der beiden großen internationalen Befragungen lagen, sowie über die unterschiedliche Erhebungsform hinweg, die in den beiden Befragungen gewählt wurde, nur bedingt variieren. Die Ländermittelwerte der Arbeitszufriedenheit aus beiden Datensätzen korrelierten in dieser Studie hier in sehr hohem Maß mit insgesamt Spearman-Rho = .67 ($p < .001$) miteinander. Länder, die in der ersten Befragung einen vorderen Rang in der Zufriedenheit einnehmen, finden sich daher mit hoher Wahrscheinlichkeit auch in der zweiten Befragung auf einem der vorderen Plätze wieder. Diese relativ stabile Positionierung der Länder in der Zufriedenheitsrangliste konnte in dieser Studie in Teilen mit der jeweiligen Kultur der Länder in Verbindung

gebracht werden und schließlich mehrheitlich durch die kulturell bedingte Zustimmungsnorm erklärt werden.

Kulturdimensionen als potentieller Erklärungsfaktor

Ein erster Anhaltspunkt für die unterschiedlich hohen Zufriedenheitsniveaus ließ sich in der jeweiligen Kultur des Landes finden. Die Korrelationsanalysen hinsichtlich des Zusammenhangs der Kulturdimensionen und der Zufriedenheitsniveaus konnten zeigen, daß das Niveau vor allem bei ganz bestimmten kulturellen Bedingungen höher war. Der bisher in Teilen in der Literatur schon vermutete Zusammenhang von Individualismus/Kollektivismus und dem Arbeitszufriedenheitsniveau konnte anhand der Berechnungen bestätigt werden. Diesbezüglich zeigte sich, daß das Zufriedenheitsniveau höher in individualistischen Ländern ausfällt und eher gemäßigt in kollektivistischen Ländern. Darüber hinaus ergaben sich ebenso bedeutsame Zusammenhänge mit den Kulturdimensionen der Machtdistanz und der Unsicherheitsvermeidung. Hier zeigte sich, daß je stärker die Machtunterschiede zwischen den Mitgliedern einer Kultur ausgeprägt sind und je mehr Unsicherheiten im täglichen Leben von den Menschen abgelehnt und vermieden werden, desto geringer fällt tendenziell auch das Arbeitszufriedenheitsurteil aus. Diese Zusammenhänge zeigten sich dabei weitestgehend in übereinstimmender Weise, egal welches Kulturmodell herangezogen wurde. Unabhängig von der jeweiligen Operationalisierung von Kultur scheinen somit hohe Zufriedenheitsurteile nur in bestimmten kulturellen Umgebungen abgegeben zu werden. Die Ergebnisse der experimentellen Überprüfung konnten diesen Aspekt zumindest hinsichtlich der Individualismus/Kollektivismus-Unterscheidung bestätigen. Auch das reine Primat der kulturellen Verschiedenheit, was lediglich eine kurzfristige Verfügbarmachung eines typischen Denkmusters dieser Kulturen darstellte, führte dazu, daß die Zufriedenheitsurteile bezüglich ein und derselben Situation unterschiedlich ausfielen.

Kulturelle Wertvorstellungen an sich können jedoch noch nicht als eine wirkliche Erklärung für das Zustandekommen von unterschiedlich hohen Zufriedenheitsurteilen angesehen werden. Ein direkter Bezug zwischen der durchschnittlichen Beurteilung des Arbeitsplatzes und der Tatsache, daß in manchen Ländern beispielsweise das Individuum stärker als die Gruppe im Mittelpunkt des Denkens steht, drängt sich nicht auf. Im Prinzip steht zwischen diesen beiden Konstrukten somit eine Art Blackbox, bei der nicht klar ist, welche Faktoren in ihrem Inneren für den erfaßten Zusammenhang

von Kultur und Zufriedenheitsurteilen verantwortlich sind. Insoweit besitzen diese bivariaten Korrelationsergebnisse daher in Teilen nur eine Art deskriptiven Status. Eine konkrete Erklärung deutet sich erst dann an, wenn man mit einbezieht, daß spezielle kulturelle Umgebungen auch eine bestimmte Art des Kommunikationsverhaltens zum Vorschein bringen. In gleicher Weise, wie bereits die frühere Forschung zeigen konnte, stellte sich in den Analysen dieser Studie ein deutlicher Zusammenhang zwischen der Kultur eines Landes und der jeweiligen durchschnittlichen Antworttendenz heraus. Der Einfluß, den die Kultur auf das Antwortverhalten zu nehmen scheint, ist dabei vor allem für eine Art der Antworttendenz, der Zustimmungstendenz, von Bedeutung. Hinsichtlich Extrem- und Mittentendenzen sowie dem Antworten gemäß sozialer Erwünschtheit konnten hingegen ähnlich wie in früheren Untersuchungen keine klaren Zusammenhänge erkannt werden. Beim zustimmenden Antwortverhalten zeigte sich jedoch, daß dieses im Mittel höher ist in Kulturen, die durch hohen Kollektivismus, hohe Machtdistanz und hohe Unsicherheitsvermeidung gekennzeichnet sind. Am deutlichsten und am stärksten waren die Zusammenhänge zwischen Zustimmungsnorm und Kultur hinsichtlich der Individualismus/Kollektivismus-Dimension. Zu beachten ist, daß die Struktur der Zusammenhänge von Zustimmungsnorm und Kultur exakt das umgekehrte Muster im Vergleich zu den Zusammenhängen der Kulturdimensionen mit dem Arbeitszufriedenheitsniveau darstellte. Während das Arbeitszufriedenheitsniveau höher in individualistischen Kulturen und niedriger in Kulturen mit hoher Machtdistanz und Unsicherheitsvermeidung war, waren die Ergebnisse zur Zustimmungstendenz genau umgekehrt hinsichtlich dieser Kulturdimensionen, d.h. diese war niedrig in individualistischen Kulturen, aber hoch in Kulturen mit hoher Machtdistanz und hoher Unsicherheitsvermeidung. Diese entgegengesetzten Ergebnisse waren gemäß den theoretischen Überlegungen so erwartet worden. Sie machten es darüber hinaus aber auch umso plausibler, die Zustimmungstendenz als einen potentiellen Erklärungsfaktor für das Entstehen von eher hohen bzw. niedrigen Zufriedenheitsurteilen in bestimmten Kulturen zu betrachten.

Zustimmungsnormen als besserer Erklärungsfaktor

Um diese Hypothese hinsichtlich der Erklärungskraft von Zustimmungstendenzen genauer prüfen zu können, wurden in einem ersten Schritt die bivariaten Zusammenhänge der Antworttendenzen mit dem Zufriedenheitsniveau untersucht. Dabei sollte an dieser Stelle noch einmal betont werden, daß die Indizes der einzelnen

Länder jeweils den Mittelwert der individuellen Antworttendenzen darstellten. Dieser Mittelwert wurde als ein Indikator für eine Art Kommunikationsnorm interpretiert. Wenn in einem Land von den Menschen erwartet wird, sich klar und deutlich zu äußern, würde sich dies in einem höheren Mittelwert der individuellen Extremtendenzen niederschlagen. In gleicher Weise resultiert ein höherer Wert der gemittelten Zustimmungstendenzen, wenn als kulturell angemessenes Kommunikationsverhalten von den Menschen erwartet wird, die eigene Meinung eher zurückzustellen und sich stattdessen eher bescheiden auszudrücken. Insoweit konnten daher die auf Basis der Individualdaten gebildeten Antworttendenz-Indizes als Antworttendenz-Normen verstanden werden.

Deutliche Zusammenhänge zeigten sich bei den bivariaten Zusammenhängen von Antworttendenz-Normen und dem Zufriedenheitsniveau dann auch hier nur bezüglich der länderspezifischen Zustimmungsindizes, jedoch nicht für die Extrem-/Mittentendenzen oder die Soziale Erwünschtheit. Bei den Zustimmungsindizes ergaben sich insgesamt nicht nur bei einem, sondern bei allen fünf in dieser Studie betrachteten Indizes deutliche Zusammenhänge mit dem Arbeitszufriedenheitsniveau. Diese Zusammenhänge waren wie gemäß den Hypothesen angenommen in der Tat negativer Art. Dies bedeutet, daß sehr hohe Zufriedenheitsurteile im Mittel von den Befragten nur in Ländern abgegeben werden, in denen zustimmendes Antwortverhalten nicht weit verbreitet ist; bzw. in umgekehrter Weise gilt somit, daß das Vorhandensein von Zustimmungsnormen in einem Land dazu führte, daß dort tendenziell gemäßigte Zufriedenheitsurteile von allen Befragten abgegeben werden.

Genau dieser negative Zusammenhang wurde gemäß der Argumentation für Hypothese 9 erwartet, wenngleich er auf den ersten Blick möglicherweise etwas kontraintuitiv gewirkt haben könnte. In Anlehnung an die Überlegungen von Markus und Kitayama (1991) wurde in der Argumentation dort aber ausführlich dargelegt, daß die Bescheidenheit im kommunikativen Miteinander, die gemäß den Erkenntnissen dieser beiden Autoren vor allem für kollektivistische Länder typisch ist, sich einerseits in einem stärkeren zustimmenden Antwortverhalten zeigt und andererseits – sozusagen als zwei Seiten der gleichen Medaille – in einem gemäßigeren Urteil über die eigene Arbeitsstelle ausdrückt. Letzteres käme unter anderem dann dadurch zustande, daß diese Art von zurückhaltender Kommunikation, die die Menschen in solch einer Kultur von Kindheit an lernen, dazu führt, daß die Menschen sich insgesamt bescheidener über ihre Arbeit äußern. Urteile über den eigenen Job können dabei als eine Art eigene

Leistungszuschreibung über das Fortkommen bei der Arbeit verstanden werden und drücken daher auch eine Art Stolz auf das bisher persönlich Erreichte aus. Stolz zu äußern wird jedoch gerade in kollektivistischen Kulturen mit hohen Zustimmungsnormen nicht als angemessenes Kommunikationsverhalten angesehen. Menschen in solch einer Kultur stellen ihn daher nicht zur Schau. Aus diesem Grund ist nachvollziehbar, daß die in einer Kultur geltende Zustimmungsnorm den Menschen es im Prinzip verbietet, sich sehr positiv und zu euphorisch über die Zufriedenheit mit dem eigenen Job zu äußern.

Während die bivariaten Analysen sozusagen die Voraussetzungen dafür prüften, daß Zustimmungsnormen als der bessere Erklärungsfaktor für unterschiedlich hohe Zufriedenheitsniveaus im Vergleich zu kulturellen Dimensionen angesehen werden können, waren daran anschließend die Mediationsanalysen auch statistisch in der Lage zu zeigen, daß bei gleichzeitiger Betrachtung der Kulturdimensionen und der Zustimmungsnormen die letzteren in der Tat einen vermittelnden Faktor darstellen, der erklären kann, warum je nach Kultur unterschiedlich hohe Zufriedenheitsurteile von den Befragten abgegeben werden. Faßt man die Ergebnisse aller Mediationsanalysen zusammen, zeigen sich deutliche Belege für die Annahme, daß kulturelle Bedingungen zu einer bestimmten Zustimmungsnorm führen und diese Norm für ein dann geringeres Zufriedenheitsniveau verantwortlich ist. Einschränkend muß jedoch hinzugefügt werden, daß dieser Zusammenhang nur für die kulturellen Unterschiede hinsichtlich individualistischen/kollektivistischen Denkens gemessen werden konnte. Konkret bedeutet dies, daß die kollektivistische Sozialisierung von Menschen zu einem stärkeren Zurückstellen der eigenen Meinung und einer bescheideneren Kommunikation (ausgedrückt in einem hohen Zustimmungsindex) führt, so daß eine Arbeitszufriedenheitsfrage von diesen Menschen mit einem gemäßigeren Urteil beantwortet wird.

Eine vermittelnde Rolle und somit ein ähnlicher Sozialisierungseffekt konnte jedoch hinsichtlich der Dimensionen Machtdistanz und Unsicherheitsvermeidung nicht gefunden werden. Diese behielten ihren Erklärungsgehalt, selbst wenn in die Analysen die Zustimmungsnorm mit einbezogen wurde. Der genaue Grund für den Zusammenhang dieser Kulturdimensionen mit dem Arbeitszufriedenheitsniveau konnte daher nicht in dieser Studie gefunden werden. Jedoch ist er nicht bei Zustimmungsnormen zu suchen. Weiterer Forschungsbedarf ist somit diesbezüglich notwendig, um die sogenannte Blackbox bezüglich dieser beiden Kulturdimensionen

genauer beleuchten zu können. Hinsichtlich der auch in der bisherigen Forschung am besten untersuchten Unterscheidung in individualistische und kollektivistische Nationen kann nach den Ergebnissen dieser Studie jedoch insgesamt die sozialisierte Zustimmungsnorm als der bessere Erklärungsfaktor für die unterschiedlich hohen Zufriedenheitsniveaus angesehen werden. Zwar konnte in der experimentellen Überprüfung aufgrund methodischer Schwierigkeiten diese Ansicht nicht vollständig gestützt werden, jedoch lieferten vor allem die Sekundäranalysen in ihren Ergebnissen dafür überzeugende Argumente.

Weitere Einflußgrößen als konkurrierende Erklärungsfaktoren

Zwar bestand der Fokus der hier vorliegenden Arbeit vorrangig in der Betrachtung von Kommunikationsnormen. Dies bedeutet jedoch nicht im Umkehrschluß, daß der Einfluß anderer Faktoren auf das Arbeitszufriedenheitsniveau geleugnet werden sollte. Es wurde immer davon ausgegangen, daß auch andere – vor allem wirtschaftliche Faktoren – ebenfalls einen bedeutsamen Einfluß auf das Zufriedenheitsniveau besitzen. Die Analysen zeigten, daß vor allem der Wohlstand eines Landes, der auch als eine Art Ausdruck der allgemeinen Arbeitsplatzqualität verstanden werden kann, in einem deutlichen Verhältnis zum Arbeitszufriedenheitsniveau der Länder steht. Je stärker die Wirtschaftskraft des Landes ist, je besser es ökonomisch entwickelt ist und je höher die generelle Lebensqualität ist, desto positiver fällt somit auch das Arbeitszufriedenheitsurteil der Befragten im Mittel aus. Ökonomische Indikatoren stellen folglich weiterhin einen bedeutsamen Faktor für die Unterschiedlichkeit der Niveaus dar. Interessanterweise zeigte sich jedoch ein durchaus überraschendes Ergebnis, wenn man den Wohlstand eines Landes und die Zustimmungsnormen als konkurrierende Erklärungsfaktoren betrachtet. Der Einfluß von Kommunikationsnormen auf das Arbeitszufriedenheitsurteil bestand nämlich erstaunlicherweise nur innerhalb der wirtschaftlich stärkeren Länder. Bei den wirtschaftlich ärmeren Ländern konnte kein Zusammenhang zwischen Zustimmungsnormen und dem Arbeitszufriedenheitsniveau festgestellt werden. Dieses Ergebnis war so nicht erwartet worden, zeigt jedoch, daß der Wohlstand eines Landes als wichtiger Moderator für den Einfluß von Kommunikationsnormen angesehen werden kann. Wenn folglich Zustimmungsnormen nur in einer der Ländergruppen einen deutlichen Einfluß ausüben, wurde zuerst versucht, dies dahingehend zu interpretieren, daß Befragte sich erst dann um das Befolgen von Normen kümmern, wenn eine

ausreichende Lebens- oder auch Arbeitsplatzqualität erreicht ist. Man hätte damit in den Daten eine Art Maslowsche Bedürfnispyramide wiederfinden können. Das Befolgen von Normen wäre diesbezüglich ein Ausdruck des sozialen Bedürfnisses gewesen. Dieses Bedürfnis wäre folglich erst dann zum Vorschein gekommen, wenn, wie in den reichen Ländern der Fall, die Grundbedürfnisse sowie das Sicherheitsbedürfnis der Menschen befriedigt sind. In ärmeren Ländern hätte hingegen der Zuwachs an Lebens- und Arbeitsplatzqualität dann stärker das Arbeitszufriedenheitsniveau erklären sollen. Diese Vermutung konnte jedoch bezüglich der ärmeren Länder überraschenderweise nicht durch die Analysen bestätigt werden (vgl. die Mehrebenenanalysen in Kapitel 9.5.3). In diesen Ländern zeigt sich nämlich kein bedeutender Zusammenhang von wirtschaftlichem Wohlstand und dem Arbeitszufriedenheitsniveau, was wiederum Zweifel an der eben genannten, durchaus plausiblen Erklärung aufkommen läßt. Verbunden mit der Tatsache, daß interessanterweise die Mediationsanalysen in den ärmeren Ländern darüber hinaus noch nicht einmal die bisher scheinbar gesicherten Zusammenhänge von Individualismus und Zustimmungsnormen einerseits und Individualismus und Arbeitszufriedenheit andererseits in klarer und deutlicher Form replizieren konnten (vgl. Kapitel 9.4.2), deutet sich eine weitere mögliche Erklärung an. Wenn scheinbar bekannte Zusammenhänge von Konstrukten sich dennoch nicht in den Daten manifestieren, könnte dies nämlich unter anderem dadurch zustande gekommen sein, daß die Befragten mehr oder weniger zufällig die vorgegebenen Skalen ausgefüllt haben. Diese Überlegung stützt sich auf das Argument, daß unter Umständen in ärmeren Ländern die Befragten nicht so vertraut mit der klassischen Form der Datenerhebung in Form von Fragebögen und Likert-Skalen sind (van de Vijver & Poortinga, 1997). Die fehlende Erfahrung der Menschen, ihr Empfinden in ein zahlenbasiertes Urteil auf einer 5er-, 7er- oder auch 10er-Skala umzuwandeln und somit im Sinne des Modells des semantischen Urteilsprozesses (Fischer, 1989, 1991) den Weg von der inneren Modellbildung hin zur äußeren Modellbildung finden zu müssen, kann möglicherweise dazu führen, daß die erwarteten Effekte sich in den Daten nicht manifestieren können. Diese Überlegungen müssen jedoch erst einmal spekulativ bleiben, da entsprechende Informationen, die dies hätten überprüfen können, nicht vorlagen. Die weiteren Folgen, die jedoch dieses Ergebnis generell für die Forschung haben könnte, werden in der abschließenden kritischen Betrachtung genauer diskutiert. Zwar bleiben bezüglich der Ergebnisse in ärmeren Ländern somit weiterhin Unklarheiten bestehen, es läßt sich aber dennoch auf jeden Fall festhalten, daß sich

zumindest innerhalb der Gruppe der wirtschaftlich stärkeren Länder ein deutlicher Effekt des Wohlstandsniveaus auf das Arbeitszufriedenheitsniveau nachweisen ließ. Je höher der Wohlstand ausgeprägt ist, desto höher ist auch das Arbeitszufriedenheitsniveau. Die wirtschaftlichen Bedingungen in den einzelnen Ländern stellen daher – wie zu erwarten war – ebenfalls einen wichtigen Erklärungsfaktor für Unterschiede in der Arbeitszufriedenheit dar.

Noch bedeutsamer vor dem Hintergrund der in dieser Arbeit behandelten Problemstellung ist jedoch, daß sowohl in den Partiellen Korrelationsanalysen als auch in den Mehrebenenanalysen ein Einfluß der Zustimmungsnorm trotz des Einbezugs der Wohlstandsindikatoren bestehen blieb. Beide Faktoren erklären jeder für sich einen großen Teil der Varianz der Zufriedenheitsunterschiede. Zustimmungsnormen stellen jedoch folglich einen eigenständigen Faktor zur Erklärung von Unterschieden im Arbeitszufriedenheitsniveau dar. Es wird dadurch deutlich, daß neben den wirtschaftlichen Rahmenbedingungen eines Landes daher die Kommunikationsnorm als weitere Einflußgröße bei der Interpretation von Länderunterschieden zukünftig berücksichtigt werden sollte.

Uneindeutige Zusammenhänge auf dem individuellen Niveau

Auch auf dem individuellen Niveau sollten gemäß den theoretischen Überlegungen Zusammenhänge zwischen dem generellen Antwortverhalten eines Befragten und seinem Urteil zur Frage nach der Arbeitszufriedenheit bestehen. Ein eindeutiges Ergebnis konnten die empirischen Analysen hier jedoch nicht liefern. Einen klaren Zusammenhang, entweder wie angenommen in positiver Richtung oder aber auch in negativer Richtung, scheint es nicht zu geben. Stattdessen unterschied sich die Richtung des Zusammenhangs auf individuellem Level je nach Land. Während in der Mehrheit der Fälle ein Zusammenhang nicht vorhanden zu sein schien, existierte in wenigen Ländern ein deutlich positiver Zusammenhang und in wenigen Ländern aber auch ein deutlich negativer Zusammenhang zwischen individuellen Zustimmungstendenzen und dem persönlichen Arbeitszufriedenheitsurteil. Dieser unterschiedliche und zum Teil entgegengesetzte Wirkzusammenhang ist als der Grund dafür anzusehen, daß ein klarer Trend und somit ein signifikanter Prädiktor über alle Länder hinweg nicht nachzuweisen war. Die unterschiedlichen Einflüsse ließen sich jedoch auch nicht wie vermutet durch die jeweils unterschiedlichen kommunikativen Rahmenbedingungen der Länder erklären. Eine entsprechende Interaktion zwischen der Zustimmungsnorm und

der persönlichen Tendenz des jeweiligen Befragten konnte bei den Berechnungen der Mehrebenenanalysen nicht gefunden werden. Hinsichtlich dieses Aspekts bleibt daher weiterer Forschungsbedarf, um zu klären, unter welchen Bedingungen genau die persönliche Antworttendenz eines Befragten ein Arbeitszufriedenheitsurteil verzerren könnte.

Als deutliche Prädiktoren für das persönliche Zufriedenheitsniveau ergaben sich in dieser Studie auf dem individuellen Level vielmehr das Alter, der Bildungsstand und das Einkommensniveau ab. Diese Einflußgrößen wurden als Kontrollvariablen in den jeweiligen Analysen mit berücksichtigt. Bei den beiden letzteren sozioökonomischen Aspekten muß jedoch kritisch angemerkt werden, daß diese nur in einer recht groben Form in den internationalen Datensätzen operationalisiert wurden und nur eine recht unspezifische Information über die relative Position des Befragten in Bezug zur Vergleichsgruppe enthielten. Ein länderübergreifende valide Messung dieser Aspekte ist jedoch grundsätzlich mit vielen Problemen behaftet (vgl. unter anderem European Foundation for the Improvement of Living and Working Conditions, 2006), so daß diese Variablen in dieser bestmöglichen Operationalisierung dennoch in die Analysen mit aufgenommen wurden. Insgesamt waren die Ergebnisse bezüglich dieser drei sozioökonomischen Einflußgrößen auch wenig überraschend. Wenn diese in den statistischen Analysen mit berücksichtigt wurden, blieb jedoch dennoch der negative Zusammenhang zwischen der Zustimmungsnorm und den Arbeitszufriedenheitsurteilen bestehen, so daß sie ebenfalls nur als zusätzliche Erklärungsfaktoren auf dem individuellen Level angesehen werden sollten.

11.2 Schlußfolgerungen für die Unternehmenspraxis

Aus den Ergebnissen dieser Arbeit lassen sich einige wichtige Schlußfolgerungen für den Umgang mit Arbeitszufriedenheitsurteilen in der Unternehmenspraxis ziehen. Die Messung der Arbeitszufriedenheit wird in Unternehmen regelmäßig in der Durchführung von Mitarbeiterbefragungen als Managementtool eingesetzt und dient dem Unternehmen dabei als Informationsquelle über die Arbeitsplatzqualität und als Steuerungsgröße für zukünftige Verbesserungsmaßnahmen. In international tätigen Unternehmen soll sie darüber hinaus dem Vergleich der weltweit verstreuten Standorte dienen. Dabei wird von Seiten der Unternehmen angestrebt, die Ergebnisse solcher Mitarbeiterbefragungen über Ländergrenzen hinweg in Beziehung setzen zu können.

Als erste Voraussetzung ist dafür von Bedeutung, daß gewährleistet ist, daß das Meßinstrument, welches in der Regel aus einer umfangreichen Skala zur Erfassung der Arbeitszufriedenheit besteht, in verschiedenen Ländern und Kulturen gleichermaßen eingesetzt werden kann und somit die notwendige Konstrukt- und Itemäquivalenz besitzt. Frühere Studien (vgl. beispielsweise Borg, 1986; Müller, 2006) konnten diesbezüglich bereits zeigen, daß Instrumente zur Messung der Arbeitsplatzqualität durchaus den notwendigen Äquivalenzkriterien genügen können und daher von einer internationalen Einsetzbarkeit allgemein ausgegangen werden kann. Die Beurteilung der Arbeitswelt in Form von verschiedenen Zufriedenheitsfacetten und auch bezüglich unterschiedlicher Arbeitswerte vollzieht sich gemäß diesen Studien bei den Mitarbeitern unterschiedlichster Kulturen nach weitestgehend den gleichen Prinzipien. Die erste Voraussetzung kann somit durchaus als erfüllt angesehen werden. Es sei jedoch angemerkt, daß dennoch einige wichtige Aspekte bei der Erstellung dieser Meßinstrumente beachtet werden müssen, um einen meßäquivalenten Fragebogen zur Erfassung der Arbeitszufriedenheit überhaupt erstellen zu können (vgl. beispielsweise die Vorschläge von Müller, 2006). Auch wenn diese notwendige Voraussetzung schließlich erfüllt ist, ist jedoch immer noch keine Aussage darüber getroffen worden, ob denn überhaupt die anhand des sorgfältig erstellten Fragebogens gemessenen Werte in ihrer absoluten Höhe zwischen den Ländern verglichen werden können. Die interkulturelle Meßäquivalenz des Fragebogens garantiert nur die Angemessenheit der verwendeten Fragen in den einzelnen Ländern. Aussagen über die interkulturelle Vergleichbarkeit der Arbeitszufriedenheitsniveaus lassen sich daraus nicht ableiten. Schlußfolgerungen zu diesem zweiten Problem, welches völlig unabhängig von der Meßäquivalenz betrachtet werden muß, können aber aus dieser Arbeit hier gezogen werden.

Die Ergebnisse dieser Arbeit setzen dem von den Unternehmen angestrebten Vergleich jedoch klare Grenzen. Die Erkenntnis, daß unterschiedliche Kulturen ein unterschiedliches Zustimmungsverhalten in Befragungen zeigen und daß dieses Verhalten in einem klaren Bezug zur Beurteilung der Arbeitszufriedenheit steht, setzt hinter den Vergleich von absoluten Werten der Arbeitszufriedenheit aus verschiedenen Ländern ein großes Fragezeichen bzw. macht ihn in Teilen sogar unmöglich. Eine Mitarbeiterbefragung strebt in der Regel die Messung der mitarbeiterspezifischen Beurteilung von objektiven Merkmalen des Arbeitsplatzes an. Die Höhe der berichteten Arbeitszufriedenheit in den einzelnen Ländern wird aber gemäß den in dieser Arbeit

diskutierten Ergebnissen nicht nur durch rein objektive Gegebenheiten im Arbeitsleben und durch die Arbeitsplatzqualität beeinflusst, sondern in gleicher Weise durch die in einem Land gültige Kommunikationsnorm. Lediglich die objektiven Gegebenheiten stellen aber das eigentliche Befragungsinteresse des Unternehmens dar. Klare Aussagen über die Beurteilung des Arbeitsplatzes von Seiten der Mitarbeiter lassen sich jedoch aus den gemessenen Rohdaten einer Mitarbeiterbefragung aufgrund der Erkenntnisse der hier durchgeführten Analysen folglich nicht mehr so ohne weiteres ableiten, da Menschen sich zu dieser Frage nur vor dem Hintergrund ihres kulturell geforderten Antwortverhaltens äußern. Der in Befragungen gemessene Wert enthält somit einerseits zwar durchaus Informationen über die Bewertung des Arbeitsplatzes aber andererseits ebenso Informationen über das allgemeine Kommunikationsverhalten in der spezifischen Kultur. Da der Anteil dieser Informationsanteile aber zwischen den Kulturen variieren kann, drückt letztendlich der gleiche quantitative Wert von beispielsweise 7,3 in Deutschland mit großer Wahrscheinlichkeit etwas anderes aus als in einem asiatischen Land. Für die Unternehmenspraxis bedeutet dies, daß Zufriedenheitswerte aus verschiedenen Standorten nicht direkt miteinander in Beziehung gesetzt werden können, denn die Aussagekraft dieses Vergleichs ist sehr fraglich. Es muß deutlich gesagt werden, daß ein direkter Rückschluß auf die entweder bessere oder schlechtere Arbeitsplatzqualität von Standorten anhand solcher Vergleiche nicht möglich ist. Wird dies dennoch vorgenommen, werden im Prinzip Äpfel mit Birnen verglichen. Diese umgangssprachliche Beschreibung macht umso deutlicher, daß es folglich erst recht als unangemessen bezeichnet werden muß, wenn aus solchen Vergleichen auch noch Maßnahmen im Sinne der strategischen Personalentwicklung abgeleitet werden. Nichtsdestotrotz finden sich auch in den Veröffentlichungen renommierter Beratungsunternehmen immer wieder in dieser Form direkte Vergleiche von internationalen Mitarbeiterurteilen. Neben der bereits mehrfach zitierten Art, wie das Beratungsunternehmen Gallup seine Daten oftmals präsentiert, findet sich diese unangemessene Darstellung beispielsweise auch beim ebenso prominenten Unternehmen Towers Perrin (vgl. Towers Perrin, 2005). Auch der Bericht dieses Unternehmens aus dem Jahre 2005 enthält zwar einerseits einen Hinweis auf die Bedeutsamkeit von kulturell bedingten Antworttendenzen und die Empfehlung, die Werte nur innerhalb eines Landes zu interpretieren (vgl. S. 8). Andererseits enthält jedoch der gleiche Bericht sogar auf der gleichen Seite eine Grafik, in der die Werte aus den einzelnen Ländern deutlich im direkten Vergleich nebeneinander dargestellt sind.

Dies führt somit in Teilen den oben genannten, eigenen Hinweis des Beratungsunternehmens selbst wieder ad absurdum. Dieser, in dem Bericht sogar grafisch aufbereitete, Vergleich ist insgesamt dem Erkenntnisinteresse nicht förderlich, da er eine Vergleichbarkeit suggeriert, die so nicht gegeben ist.

Um dennoch einen Vergleich, der ja von den Unternehmen generell gewollt ist und für die Unternehmenssteuerung vielfach gebraucht wird, sinnvoll durchführen zu können, wäre eigentlich eine Trennung der beiden Informationsanteile in dem gemessenen Wert wünschenswert. Diese Trennung ließe sich jedoch methodisch und statistisch nur schwer bewerkstelligen. Aus diesem Grund wird oftmals empfohlen, so auch bei Müller (2006), die gemessenen Länderkennwerte nur in Bezug zu nationalen Vergleichswerten zu interpretieren und diese Benchmarks als Relativierungsgröße für die in der eigenen Befragung gemessenen Absolutwerte heranzuziehen. Es ließen sich beispielsweise Benchmarks im längsschnittartigen Zeitverlauf aus früheren Befragungen desselben Unternehmens oder aber querschnittsmäßig aus Befragungen anderer Unternehmen betrachten. Aus Sicht der in dieser Arbeit betrachteten Thematik von Antworttendenzen löst dieser Vergleich mit Benchmarks aber die grundlegende Problematik nur bedingt und nur unter bestimmten Voraussetzungen. Auf der einen Seite kann dadurch zwar durchaus in Teilen besser eingeschätzt werden, wie ein gemessener Wert auf der gesamten Zufriedenheitsskala vom Verständnis her einzuordnen ist, weil ungefähr eingeschätzt werden kann, ob der gemessene Wert grundsätzlich als hoch oder niedrig anzusehen ist. Auf der anderen Seite ist der zahlenmäßige Wert des Benchmarks jedoch auch kein rein absolutes Standardisierungsmaß, da es ebenfalls aus einem ‚wahren‘ Konstruktanteil und einem auf die Kommunikationsnorm bezogenen Anteil besteht. Ein Beispiel soll die Problematik verdeutlichen: Man könnte annehmen, daß sich durchaus im Zeitverlauf auch eine tatsächliche Besserung der Arbeitsplätze im Unternehmen vollzogen habe. Wenn aber nun der Vergleichswert aus früheren Befragungen als Benchmark für die heutige Befragung herangezogen wird, würde der Vergleich der Werte die tatsächliche qualitative Verbesserung mit der quantitativen Abschätzung der Verzerrtheit vermischen. So, wie bei der Durchführung von standardisierten Intelligenztests die Einordnung des aktuell gemessenen Werts im Vergleich zur Normstichprobe nur deshalb erfolgen kann, weil von einer eigenständigen Veränderung des Konstrukts über die Zeit und über verschiedene Gruppen hinweg nicht ausgegangen wird, kann auch die Einordnung der aktuellen Mitarbeiterbefragungswerte anhand des zeitlich früheren Benchmarks nur dann korrekt geschehen, wenn davon ausgegangen

wird, daß sich das Ausmaß der durch die Zustimmungsnorm bewirkten Verzerrung über die Zeit nicht verändert hat. Vorausgesetzt wird bei solch einem Vergleich somit, daß die Verzerrung eines Meßwerts durch Kommunikationsnormen über die Zeit konstant ist. In gleicher Weise würde ein Vergleich mit einem Benchmark, der aus Befragungen anderer Unternehmen gebildet wird, nur dann eine sinnvolle Aussage zulassen, wenn vorausgesetzt werden kann, daß das allgemeine Zustimmungsverhalten in allen Unternehmen vollkommen gleich ist. Diese Voraussetzung der Konstanz des Zustimmungsverhaltens, die für beide Arten von Benchmarks gilt, muß jedoch nicht zwingend gegeben sein (vgl. Hui & Triandis, 1985b). Das Ausmaß der Verzerrung durch Zustimmungstendenzen könnte sowohl zeitlich als auch unternehmensbezogen variieren. Letztere Variation ist dabei sogar die wahrscheinlichere. Verschiedene Unternehmen sind nie komplett gleich hinsichtlich ihrer Mitarbeiterschaft. Da aber zustimmendes Antwortverhalten auch mit der Bildung, dem Alter und auch dem Geschlecht der Befragten gemäß früheren Forschungsstudien zu variieren scheint (vgl. Kapitel 6.3), kann nicht automatisch von dem gleichen Ausmaß von Zustimmungstendenzen in den Unternehmen ausgegangen werden. Immerhin besser als Vergleichsmaßstab wäre daher der zeitlich früher gemessene Zufriedenheitswert desselben Unternehmens geeignet; vorausgesetzt, auch hier hat sich die Mitarbeiterschaft nicht zwischenzeitlich stark verändert. Ist dies in der Tat nicht der Fall, sollte ein Vergleich der Werte von heute und früher eine Aussage hinsichtlich der Verbesserung oder Verschlechterung der Arbeitsplatzqualität weitestgehend fehlerfrei ermöglichen. Der Grund für die Durchführbarkeit des Vergleichs läge dann vor allem darin, daß grundsätzlich nicht davon auszugehen ist, daß Kommunikationsnormen über die Zeit hinweg sehr labil wären und sich leicht verändern würden. Diese bilden sich vielmehr in lang andauernden Sozialisationsprozessen und verändern sich, ähnlich wie kulturelle Wertvorstellungen, nur sehr langsam über die Generationen.

Das Heranziehen von Benchmarks löst somit zumindest etwas die Interpretationsproblematik hinsichtlich der Meßwerte innerhalb eines Landes, aber es ist nicht in der Lage, die Vergleichsprobleme über die Ländergrenzen hinweg komplett auszuschalten. Man führe sich erneut das sehr plastische Beispiel von Fischer (2004) vor Augen (vgl. Kapitel 5.6.2), in dem die Messung der Schwanzlänge eines Elefanten und einer Maus anhand von Absolutwerten und ipsativen Werten verglichen wird. Die Kernaussage dieses Beispiels übertragen auf den hier betrachteten Kontext würde bedeuten, daß der Vergleich von Befragungswerten mit nationalen Benchmarks nur

dann sinnvoll funktionieren kann, wenn diese nationalen Benchmarks überall gleich sind. So wie die Körperfülle der Tiere, die im Beispiel von Fischer (2004) sozusagen als Benchmark herangezogen wurde, sich deutlich zwischen dem Elefanten und der Maus unterscheidet, ist davon auszugehen, daß sich auch nationale Befragungsbenchmarks durchaus stark unterscheiden können. Wenn dies jedoch der Fall ist, kommen Vergleiche von Absolutwerten und von standardisierten Werten möglicherweise zu völlig entgegengesetzten Aussagen. Aus diesem Grund muß auch hier als Fazit festgehalten werden, daß Vergleiche mit Benchmarks zwar für ‚Intra-Länder‘-Vergleiche durchaus genutzt werden können, jedoch für ‚Inter-Länder‘-Vergleiche nicht herangezogen werden sollten (vgl. Chan, 2003; Fischer, 2004; Hicks, 1970;).

Ein letztes Problem, welches sich durch den Einsatz von Benchmarks als Standardisierungsgrößen ebenso nicht lösen läßt, ist, daß trotz dieses Wertevergleichs das Ausmaß des ‚wahren‘ Zufriedenheitswerts dennoch nicht genau erfaßt werden kann. Setzt man die gemessenen Werte mit den Benchmarks direkt in Beziehung und subtrahiert die Werte voneinander, sind lediglich nur Aussagen in der Form ‚besser/schlechter als der Benchmark‘ möglich. Der Absolutheitsgehalt des gemessenen Werts geht dabei vollständig verloren. Auch dies wäre daher keine angemessene Lösung.

Insgesamt gilt daher, daß der Vergleich von aktuell gemessenen Zufriedenheitswerten anhand von Benchmarks eine in Teilen sinnvolle Lösung darstellt, die jedoch keineswegs das Allheilmittel für die Vergleichsproblematik darstellt. Da der Einbezug von Benchmarks dennoch das bisher noch am besten funktionierende Vorgehen darstellt, ist auf jeden Fall anzuraten, diesen Vergleich in allen interkulturellen Befragungen zumindest mit zu betrachten. Die optimalere Lösung wäre sicherlich eine Art Formel, die als Maßzahl den Wert ergäbe, den man jeweils von einem gemessenen Wert abziehen müßte, um an den ‚wahren‘ Meßwert des Konstruktes zu gelangen. Wie diese Art von Formel aber gefunden werden könnte, dazu gibt es in der bisherigen Forschungsliteratur noch keine Ansätze. Bekannt ist zwar, daß Zustimmungstendenzen eine wichtige Rolle spielen und beachtet werden müssen. Wie man aber trotz ihres Auftretens die ‚wahren‘ Meßwerte finden kann, ist immer noch weitestgehend unklar. Auch diese Arbeit kann bezüglich der Messung von Arbeitszufriedenheit dazu noch keine endgültige Lösung präsentieren. Weiterer Forschungsbedarf ist daher diesbezüglich notwendig.

11.3 Schlußfolgerungen für die Forschung

Neben den Implikationen, die die Ergebnisse dieser Arbeit vor allem für die Befragungsdurchführung in Unternehmen besitzen, lassen sich auch verschiedene Erkenntnisse für die Forschung ableiten. Eine erste Schlußfolgerung läßt sich aus dem Teil der Ergebnisse ziehen, der in erster Linie eine Replikation der bereits in der Forschung untersuchten Zusammenhänge darstellte. So bestätigen die hier ermittelten Zusammenhänge von Antworttendenzen und Kulturdimensionen die Ergebnisse, die ebenfalls schon in anderen Studien aufgetreten sind, und unterstreichen damit die Gültigkeit der bisherigen Überlegungen einiger interkultureller Forscher, Antworttendenzen als ein Ausdruck von Kultur und somit als ein Indikator für die Art des sozialen Miteinanders in den einzelnen Ländern anzusehen. Auch diese Arbeit hier kann zeigen, daß ein klarer Zusammenhang zwischen den kulturellen Werten einer Gesellschaft und der Art und Weise besteht, wie dort miteinander kommuniziert wird. Antworttendenzen sind somit nicht nur ein rein methodischer Bias, sondern besitzen in der kulturellen Sozialisation der Mitglieder einer Kultur eine theoretische Basis, die deren Ursprung erklären kann. Dennoch stellen sie im konkreten Meßkontext wiederum eine nicht gewollte Verzerrung von Werten dar, wenn nicht das Antwortverhalten selbst im Fokus des Interesses steht, sondern vielmehr die inhaltliche Bedeutung der gemessenen Konstrukte. In der in dieser Arbeit betrachteten Thematik betraf dies speziell die Frage nach der angemessenen Interpretation von Arbeitszufriedenheitswerten aus verschiedenen Ländern. Die Ergebnisse zeigten deutlich, daß je nach Stärke der Antworttendenz in einer Kultur auch im Mittel ein anderes Zufriedenheitsurteil das Resultat der Befragung war.

In dieser Hinsicht unterstreichen die Ergebnisse auch die Bedeutung der Überlegungen, die Fischer (1989, 1991) in seinem Modell der semantischen Urteilsbildung ausdrückt. Die Resultate dieser Arbeit liefern eine Bestätigung dafür, daß bei der Messung von Arbeitszufriedenheitsurteilen oder auch von anderen allgemeinen Einstellungen der Aspekt der Kommunikation vom Befragten in Richtung des Fragenden von großer Bedeutung ist. Dieser Teil des semantischen Urteilsprozesses, die äußere Modellbildung, ist in der bisherigen Messung von Arbeitszufriedenheitsurteilen im interkulturellen Kontext eher selten berücksichtigt worden. Oftmals wird wie selbstverständlich davon ausgegangen, daß das Kreuz auf einer Befragungsskala die ‚wahre‘ Einstellung bzw. Zufriedenheit des Befragten ausdrückt. Die Ergebnisse hier zeigen jedoch, daß dieser

Schluß zu voreilig ist. Vor allem im interkulturellen Kontext kommt dabei zum Tragen, daß Menschen in Befragungssituationen gegenüber einem Interviewer je nach kulturellem Hintergrund offensichtlich unterschiedlich ihre Meinung äußern. Dieser Aspekt mag in Untersuchungen, die nur in einer einzigen kulturellen Umgebung durchgeführt werden, von geringerer Bedeutung sein, da dort angenommen werden kann, daß diese Art der kommunikativen Verzerrung für alle Befragten gleich ist und daher die Werte untereinander wiederum verglichen werden können. Beim Vergleich über Länder- und Kulturgrenzen hinweg gilt dies jedoch dann nicht mehr. Aus diesem Grund kann nur empfohlen werden, ein stärkeres Augenmerk auch auf den Prozeß der äußeren Modellbildung zu legen. In dieser Phase des semantischen Urteilsprozesses, die entscheidend ist für die letztendlichen Daten, mit denen der Forscher arbeitet, können selbstverständlich viele verschiedene Faktoren einen Einfluß haben. Antworttendenzen sind nur einer davon. Für diesen Faktor jedoch können die Ergebnisse hier zeigen, daß er in der Tat eine gewichtige Rolle spielt. Die Vergleichbarkeit von Arbeitszufriedenheitsurteilen aus verschiedenen Ländern ist aufgrund dieses Faktors im Prinzip nicht mehr ohne Weiteres möglich. In dieser Hinsicht sind die Überlegungen dieser Arbeit auch bedeutsam für die grundsätzliche Problematik jeglicher interkultureller Forschung.

Genau genommen zeigt sich ein klarer Effekt aber nur für eine Form der Antworttendenz, nämlich die Zustimmungstendenz. Durch die unterschiedlichen Ergebnisse je nach Antworttendenz-Maß wird deutlich, daß eine präzise Trennung zwischen den verschiedenen Formen von Antworttendenzen notwendig und sinnvoll ist. Dies zeigt, daß es durchaus sehr entscheidend ist, wie das Konstrukt der Zustimmungstendenz genau operationalisiert wird. Bisherige Studien (vgl. u.a. Harzing, 2006) haben dabei zum Teil keine Unterscheidung zwischen verschiedenen Skalenlabeln vorgenommen. Die verschiedenen Indizes, die in der Untersuchung dieser Arbeit bewußt getrennt für die jeweiligen Skalenlabel gebildet wurden, zeigten jedoch unterschiedlich starke Effekte. Deutliche Effekte ergaben sich vor allem bei der Skalenbezeichnung „stimme zu vs. stimme nicht zu“, d.h. bei genau der Operationalisierung, die am Nächsten am inhaltlichen Verständnis des Konstrukts liegt. Weniger klare Zusammenhänge zeigten sich hingegen mit dem Zustimmungsindeks der Skalenbezeichnung „wichtig vs. unwichtig“. Es ist demnach keineswegs unbedeutend, anhand welcher Skalen ein Zustimmungsindeks überhaupt gebildet wird. Daß der Index der 4er-Likert-Skala darüber hinaus insgesamt geringere Effekte zeigte, betont auch die

Bedeutung des Umfangs der Antwortmöglichkeiten einer Skalen und bestätigt somit die Ergebnisse früherer Studien, die ebenfalls zeigen konnten, daß je nach Anzahl der Skalenstufen sich eine unterschiedlich starke Antworttendenz nachweisen ließ (vgl. Clarke III, 2001; Hui & Triandis, 1989).

Vor allem jedoch zeigt der nicht eindeutig vorhandene Zusammenhang auf dem Individualniveau zwischen dem Zustimmungsverhalten, gemessen anhand von stimme-zu-Skalen, und dem Urteil bei Zufriedenheitsskalen, daß es offensichtlich je nach Skalenbezeichnung unterschiedliche Antworteffekte geben kann. Man hätte vermuten können, daß eine allgemeine Zustimmungstendenz eines Befragten sich bei einer Frage nach der Zufriedenheit auch automatisch in einem höheren Urteil ausdrücken müßte. Diese vorschnelle Annahme stellte sich jedoch als falsch heraus, denn eine klare Korrelation – zumindest unabhängig von der Zugehörigkeit der Befragten zu einem Land – konnte in den hier durchgeführten Analysen nicht nachgewiesen werden. Im Prinzip scheinen daher die Urteile auf beiden Skalen aus individualpsychologischer Sicht in keinem klaren Bezug zueinander zu stehen. In den meisten Ländern zeigte sich fast kein Zusammenhang. Warum in wenigen Ländern dennoch ein positiver und in manchen auch ein negativer Zusammenhang zu existieren scheint, muß in weiterer Forschung noch genauer beleuchtet werden. Hier kann jedoch dennoch insgesamt festgehalten werden, daß es durchaus von entscheidender Bedeutung für das Auftreten allgemeiner Tendenzen im Beantworten von Fragen sein kann, welche Skalenbezeichnung gewählt wird und ob konkrete Itemaussagen dem Befragten vorgelegt werden, denen er zustimmen kann, oder ob es sich um eher offene Frageformulierungen in Form eines „wie beurteilen Sie folgendes“ handelt.

Des weiteren betreffen die Ergebnisse aber auch in grundlegender Art und Weise einen weiteren Aspekt, der in zukünftiger Forschung beachtet werden sollte. Da Menschen aus verschiedenen Kulturen offenbar nicht nur generell ein unterschiedliches Befragungsverhalten besitzen, sondern diese Ansicht, wie man sich auf Fragen zu äußern hätte, auch noch in Verbindung mit speziellen Fragenformen unterschiedliche Effekten zeigen kann, stellt sich die Frage, wie qualitativ hochwertige Daten in interkulturellen Befragungen zukünftig erhoben werden können. Dieses Problem scheint insbesondere bei der Messung anhand von explizit geäußerten Urteilen bei direkt formulierten Items und weniger bei indirekten Erhebungsformen zu bestehen, wie Müller (2006) bezüglich seiner eigenen Ergebnisse bereits in Anlehnung an die Unterscheidung von Argyris und Schön (1978) in ‚espoused values‘ und ‚values in use‘

diskutierte. ‚Espoused values‘ bezeichnen dabei die nach außen explizit geäußerten Wertvorstellung, welche stark von sozialen und kulturellen Normen geprägt sein können, während ‚values in use‘ die tatsächlich gelebten Wertvorstellungen darstellen. Beim direkten Meßansatz in Fragebögen werden jedoch eher Erstere erhoben und somit stärker die Auffassung des Befragten erfaßt, wie er denken und fühlen sollte. Letztere können vermutlich besser durch indirekte Verfahren gemessen werden. In seinen Ergebnisse konnte nämlich diesbezüglich Müller (2006) zeigen, daß seine indirekte Erfassung der Wichtigkeit von Arbeitszufriedenheitsfacetten für die Gesamtzufriedenheit durch eine Korrelationsanalyse der Facetten und des Gesamturteils zu völlig anderen Ergebnissen kam als eine direkte Messung, die explizit nach der Wichtigkeit von Aspekten der Arbeit fragt (wie sie beispielsweise von Hofstede, 2001, eingesetzt wurden).

Auch in dieser Arbeit hier zeigte sich hinsichtlich der absoluten Höhe der Arbeitszufriedenheit, daß diese ebenfalls im Sinne einer Art ‚espoused value‘ stark dadurch verzerrt zu sein scheint, was die Menschen einer Kultur für eine angemessene Antwort auf die gestellte Zufriedenheitsfrage ansehen. Offensichtlich steht solch eine Frage in einem Konflikt mit den grundlegenden Ansichten der Menschen darüber, was in ihrer Kultur in welcher Form überhaupt geäußert werden dürfte. Markus und Kitayama (1999) diskutieren diesbezüglich hinsichtlich eines Experiments von Yoshida, Kojo und Kaku (1982), daß in der dort betrachteten eher kollektivistischen Kultur interessanterweise gerade der Mitschüler als der nettere und auch übrigens als der kompetentere Sportler eingeschätzt wurde, der auf konkrete Nachfrage hin seine Leistungen eher in einem moderaten Licht schilderte, im Vergleich zu dem Mitschüler, der mit vollem Stolz von seinen tollen Leistungen berichtete. Diese Ergebnisse, die sich auf Grundschüler in Japan bezogen, konnten ebenfalls in China in ähnlicher Weise bei College-Studenten repliziert werden (Bond, Leung & Wan, 1982). Sollte diese Wahrnehmung in der Tat typisch sein für alle eher kollektivistischen Kulturen, wäre nahezu offensichtlich, daß ein Befragter auf eine Zufriedenheitsfrage in solch einer Kultur eigentlich überhaupt kein extrem hohes Urteil abgeben könnte, weil es ihm durch die geltenden Kommunikationsvorschriften im Prinzip nicht erlaubt würde. Er würde sich dadurch erst recht in ein schlechtes Licht stellen und antwortet daher eher in gemäßigerer Form.

Dieses fast paradoxe Antwortverhalten führt selbstverständlich dann zu einem umfassenderen Dilemma der Befragungsforschung. Bei einem gemäßigten

Zufriedenheitsurteil in einem kollektivistischen Land bestünde nämlich immer die Möglichkeit, daß es einerseits tatsächlich nur eine mäßige Zufriedenheit ausdrückt oder aber andererseits nur die nach außen geschilderte Form eines in Wahrheit sehr positiven Urteils darstellt. Man könnte daher sogar die grundlegende Frage aufwerfen, inwieweit Befragungen, die dem direkten Ansatz folgen, in Kulturen, in denen hohe Zustimmungsnormen existieren, überhaupt zu sinnvollen Ergebnissen kommen können. Demnach würde es gemäß den in dieser Arbeit ermittelten Ergebnissen nicht sinnvoll sein, vor allem in Kulturen mit hohem Kollektivismus solch eine direkte Frage nach der Zufriedenheit zu stellen. Sie würde mehr das gängige Kommunikationsverhalten in dieser Kultur widerspiegeln als das eigentlich intendierte inhaltliche Urteil zur Arbeitsplatzqualität. In stark individualistischen Ländern könnte man hingegen eher die Gewißheit haben, daß das abgegebene Urteil tatsächlich der Meinung der Befragten entspricht. Dort würde dann auch der Einsatz von Mitarbeiterbefragungen nach bisherigem Muster eher Sinn machen, während sie in kollektivistischen Ländern möglicherweise fragwürdige Ergebnisse ermitteln würden. Da jedoch weder die Einteilung in individualistische/kollektivistische Länder noch die Unterscheidung in hohe und niedrige Zustimmungstendenz einem klaren ja/nein-Urteil entspricht, sondern vielmehr ein Kontinuum darstellt, läßt sich auch keine klare und eindeutige Empfehlung ableiten, in welchen Ländern eine direkt gestellte Arbeitszufriedenheitsfrage sinnvoll ist und in welchen nicht. Es sollte jedoch zumindest bei der Interpretation der Ergebnisse diese Problematik immer mit berücksichtigt werden. Da insgesamt anzunehmen ist, daß vor allem der direkte Ansatz zu größeren Verzerrungen und somit zu vermehrten Interpretationsproblemen führen müßte, sollte in zukünftiger Forschung der Fokus möglicherweise eher auf die Frage gelegt werden, wie Arbeitszufriedenheit auch in indirekter Form in interkulturellen Befragungen erfaßt werden könnte.

Eine weitere Schlußfolgerung, die bei zukünftigen Studien beachtet werden sollte, läßt sich aus dem durchaus überraschenden Ergebnis ziehen, daß sich ein deutlicher Unterschied in den gemessenen Effekten ergab, wenn unterschiedliche Ländergruppen betrachtet wurden. So kam mehr oder weniger zufällig durch die Nutzung zweier verschiedener Sekundärdatensätze die Tatsache zum Vorschein, daß eine Zustimmungsnorm nur in reicheren, d.h. wirtschaftlich stärkeren, Ländern zum Vorschein, während in den ärmeren, d.h. wirtschaftlich schwächeren, Ländern kein Zusammenhang festgestellt werden konnte. Wenngleich – wie bereits erwähnt – durchaus weiterhin Unklarheiten bezüglich der Ergebnisse in ärmeren Ländern bestehen,

läßt sich dennoch anhand dieser Ergebnisse erkennen, daß es offensichtlich einen diskontinuierlichen Einfluß des Wohlstandsfaktors auf den Zusammenhang von Zustimmungsnormen und Arbeitszufriedenheitsniveau gibt. Die beiden Gruppen von Ländern sollten daher durchaus auch in zukünftiger Forschung als getrennte Gruppen betrachtet werden, anstatt den Einfluß von Wohlstand durch beispielsweise Partielle Korrelationen oder Kovarianzanalysen kontrollieren zu wollen. Durch diese Art von statistischen Verfahren wird nämlich im Prinzip nur der kontinuierliche, lineare Einfluß einer Variablen kontrolliert. Solch eine Art eines diskontinuierlichen Zusammenhangs können diese Verfahren jedoch nicht aufdecken. Weitere Forschung ist dennoch notwendig, um den genauen Grund für diese unterschiedlichen Effekte zu ermitteln. Die Einteilung der reichen und armen Länder anhand des Bruttoinlandsproduktes könnte letztendlich inhaltlich verschiedene Schlußfolgerungen zulassen. In dieser Arbeit wurde der Wohlstand eines Landes als Indikator für die durchschnittliche Arbeitsplatzqualität interpretiert. Darüber hinaus korreliert dieser Indikator aber ebenfalls mit weiteren arbeitsbezogenen Aspekten, die möglicherweise eine andere Erklärung der verschiedenen Effekte in den Ländergruppen liefern könnten. Ebenso könnte der Indikator jedoch beispielweise auch als Ausdruck des Bildungsniveaus verstanden werden. Ein geringeres Bildungsniveau könnte dann vermutlich eine Erklärung dafür darstellen, warum in manchen Ländern eine geringere Erfahrung mit der Beantwortung von Fragebögen besteht (van de Vijver & Poortinga, 1997). Die fehlende Erfahrung wiederum würde dann dazu führen, daß in ärmeren Ländern sich hinsichtlich der untersuchten psychologischen Konstrukte keine klaren Ergebnisse messen lassen. Dies sollte in weiteren Studien genauer untersucht werden.

Grundsätzlich unterstreichen diese je nach Ländergruppe spezifischen Ergebnisse aber auch noch einmal die Bedeutung einer durchdachten Auswahl von Länderstichproben in interkulturellen Studien. Wenn bestimmte Länder in internationalen Studien nur deshalb in die Untersuchung mit einbezogen wurden, weil zufällig Daten für sie vorhanden waren, sollte die Übertragung und Generalisierung der Erkenntnisse immer nur unter größter Vorsicht vorgenommen werden. Die Ergebnisse dieser Arbeit zeigen, daß, je nachdem welche Ländergruppen betrachtet werden, man zum Teil zu völlig unterschiedlichen Ergebnissen gelangt. Ein großes Problem der interkulturellen Forschung ist dabei sicherlich immer, daß eine Zufallsstichprobe von Ländern in der Regel nicht zu realisieren ist. Um das damit verbundene Problem der unklaren externen Validität der Resultate möglichst gering zu halten, sollte daher immer eine möglichst

große oder eine zumindest hinsichtlich der zu untersuchenden Konstrukte möglichst heterogene Länderstichprobe berücksichtigt werden. Dies wurde in den Analysen dieser Arbeit durch die Nutzung des World-Values-Surveys, des derzeit umfassendsten internationalen Datensatzes, angestrebt.

Kritisch hinterfragt werden könnten aufgrund der Erkenntnisse dieser Arbeit auch möglicherweise grundsätzlich die bisherigen Ergebnisse psychologischer Forschung, nicht nur im interkulturellen Bereich. Wenn, wie in dieser Studie der Fall, selbst scheinbar gut gesicherte Erkenntnisse in einer bestimmten Ländergruppe auf einmal nicht mehr repliziert werden können, drängt sich die Frage auf, inwieweit die bisherige Forschung und der derzeitige Erkenntnisstand vielleicht vielfach nur ein Abbild der eher westlichen Welt darstellt. Da der überwältigende Teil der Forschung – zumindest, was die Publikation in den einschlägigen Fachzeitschriften angeht – lediglich in westlichen Industrienationen durchgeführt wird, könnte es in der Tat sein, daß dadurch ein insgesamt verzerrtes Bild psychologischer Wirkungszusammenhänge existiert. So konnte in dieser Studie hier der eigentlich gut bestätigte Zusammenhang von Kulturdimensionen und dem Kommunikationsverhalten im Sinne von Zustimmungsnormen in der Gruppe der ärmeren Länder nicht klar und deutlich bestätigt werden. Selbst ein Zusammenhang zwischen Kulturdimensionen und Arbeitszufriedenheit und zwischen Wohlstand und Arbeitszufriedenheit konnte überraschender nicht eindeutig ermittelt werden. Als Gründe für diese fehlende Replikation wäre einerseits an inhaltliche Aspekte oder andererseits auch an methodische Aspekte zu denken. In inhaltlicher Hinsicht könnten sich möglicherweise die Zusammenhänge psychologischer Konstrukte untereinander tatsächlich in anderen Teilen der Welt von den Zusammenhängen in der westlichen Welt komplett unterscheiden. Möglicherweise existieren aber doch die gleichen Zusammenhänge und lediglich die Art der methodischen Vorgehensweise ist in anderen Kulturkreisen nicht angemessen, so daß dies die unterschiedlichen Ergebnisse aus rein methodischer Sicht erklären könnte. In dieser Hinsicht besteht sicherlich noch erheblicher Forschungsbedarf. Diese Arbeit hier kann nur einen ersten Schritt diesbezüglich darstellen und zumindest auf die große Bedeutung dieser Thematik hinweisen, auch wenn sie selbst noch keine direkten Antworten zu dieser allgemeinen Problemstellung liefern kann.

11.4 Einschränkungen der Arbeit

In Bezug zu den bereits im theoretischen Teil der Arbeit erläuterten Schwierigkeiten, die speziell bei Forschungsvorhaben im interkulturellen Bereich auftreten, müssen auch zu den hier durchgeführten Analysen und Berechnung einzelne kritische Anmerkungen gemacht werden.

Studien anhand interkultureller Datensätze können von vorneherein immer nur einem quasi-experimentellen Design folgen. Damit ist vor allem die Problematik von Alternativerklärungen verbunden. Da sich verschiedene Länder anhand von schier unzähligen Aspekten unterscheiden lassen, ist es fast unmöglich alle möglichen Einflußfaktoren in einem methodischen Design im interkulturellen Bereich zu berücksichtigen. In dieser Arbeit wurden explizit zwei große Faktorengruppen auf ihre Einflußstärke hinsichtlich des Arbeitszufriedenheitsniveaus geprüft, auf der einen Seite kulturelle Wertedimensionen und auf der anderen Seite verschiedene Arten von Antworttendenzen. Für diese beiden Arten von Einflußfaktoren ergaben sich in vielen Fällen signifikante Zusammenhänge mit der abhängigen Variable, so daß beide Faktoren schließlich als bedeutsame Einflußfaktoren angesehen werden können. Daß es nicht doch noch weitere Aspekte gibt, die ein unterschiedlich hohes Arbeitszufriedenheitsniveau erklären könnten, kann nie komplett ausgeschlossen werden. Die in dieser Arbeit durchgeführten Analysen zum destruktiven Testen konnten jedoch zumindest für einzelne mögliche Alternativerklärungen zeigen, daß trotz des Einbezugs der anderen Faktoren der Einfluß einer Zustimmungsnorm immer noch von Bedeutung war. Somit wurden die gemessenen Zusammenhänge immerhin schon gegenüber einigen wichtigen Alternativhypothesen getestet und ihre Bedeutsamkeit schon in Teilen abgesichert. Da aber darüber hinaus immer noch weitere Einflußgrößen existieren könnten, kann selbstverständlich nie von einer alleinigen Kausalität der in dieser Arbeit untersuchten Aspekte ausgegangen werden. Als weitere Tatsache kommt noch hinzu, daß die untersuchten unabhängigen Variablen untereinander ebenfalls zum Teil deutliche Interkorrelationen aufweisen, was die Aufdeckung und Separierung potentieller Effekte und daher auch klare Aussage über die jeweiligen Einflüsse und ihre kausale Richtung beträchtlich erschwert. Die Mediationsanalysen stellten diesbezüglich schon einen weiteren, bewußt gewählten Schritt dar, um die Beziehungen der jeweiligen unabhängigen Variablen untereinander in einen angemesseneren Gesamtrahmen zu setzen.

Die Interpretation in Form von kausalen Zusammenhängen ist auch allein daher immer nur mit Vorsicht durchzuführen, da eine Sekundäranalyse, die wie hier als Querschnittsdesign angelegt ist, zweifelsohne streng kausale Aussagen niemals treffen kann. Dennoch kann aufgrund der theoretisch inhaltlichen Konzeption der untersuchten Konstrukte zumindest in vorsichtiger Weise von einem kausalen Zusammenhang ausgegangen werden. Die Kultur eines Landes kann aufgrund ihrer tiefen Verwurzelung und der grundlegenden Beeinflussung aller Lebensbereiche als die basalste Komponente angesehen werden. Ein Teilaspekt der Kultur, nämlich die Sozialisation hinsichtlich des kommunikativen Verhaltens, läßt sich dann in den jeweiligen Antworttendenzen finden. Diese wiederum beziehen sich auf das allgemeine Kommunikationsverhalten von Befragten in allen Befragungen und zu allen Arten von Fragen. Aus diesem Grund ist die Beantwortung einer Zufriedenheitsfrage ebenso wiederum nur als Teilmenge dieses gesamten Kommunikationsverhaltens anzusehen. So gesehen kann daher aus dieser theoretisch inhaltlichen Perspektive heraus durchaus eine kausale Argumentationskette aufgestellt werden, in der die Kultur die allgemeine Kommunikationstendenz beeinflusst, welche wiederum der Grund für spezifische Befragungsurteile sein kann.

Hinsichtlich der Schwierigkeit einer Kausalaussage könnte möglicherweise auch die Auswahl der Indikatoren und deren Jahresbezug kritisiert werden. Nicht immer waren die in die Analysen einbezogenen Indikatoren hinsichtlich der überprüften Kausalrichtung in der zeitlich korrekten Abfolge angeordnet, d.h. die unabhängigen Indikatoren waren nicht immer früher als die abhängigen Indikatoren erhoben worden. Dennoch zeigte sich insgesamt, daß der Einbezug der Indikatoren als gerechtfertigt angesehen werden kann, da sie entweder auf der einen Seite relativ gering über die Zeit variierten oder aber auf der anderen Seite mit weiteren, inhaltlich ähnlichen Indikatoren im Sinne der konvergenten Validität stark zusammenhingen. Beispielsweise ist aus diesen Gründen auch die Nutzung der Kulturdimensionen der GLOBE-Studie angemessen, die zwar erst aus jüngerer Zeit stammen, aber dennoch als valide Maßzahlen genutzt werden konnten und auf ihren Zusammenhang mit dem zum Teil früher ermittelten Arbeitszufriedenheitsniveau geprüft werden konnten.

Insgesamt bedingten diese genannten Aspekte jeweils kausalitätsbezogene Einschränkungen, so daß aus diesem Grund daher gezielt eine experimentelle Analyse der zuvor gemessenen Zusammenhänge durchgeführt wurde, um die Ergebnisse auch in dieser Art und Weise abzusichern. Hinsichtlich der experimentellen Überprüfung kann aber nicht unerwähnt bleiben, daß diese zwar in Teilen einen kausalen Effekt zeigen

konnte, jedoch die gesamte Argumentationskette nicht komplett bestätigt werden konnte. Zwar war allein das Primen von Selbstkonzeptbestandteilen, die mit individualistischem und kollektivistischem Denken zusammenhingen, dazu in der Lage, unterschiedlich hohe Zufriedenheitsurteile bezüglich ein und derselben Beurteilungssituation zu erzeugen. Unterschiede in den Antworttendenzen, die eigentlich die Ursache für die verschiedenen Zufriedenheitsurteile darstellen sollten, traten jedoch nicht auf. Diese fehlenden Effekte ließen sich jedoch durchaus auf die methodische Abfolge und die damit verbundenen Schwierigkeiten der Priming-Experimente sowie vor allem auf die zu gering anhaltende Dauer des Primingeffekts zurückführen und wurden daher nicht als Widerlegung der ursprünglich aufgestellten Kausalkette angesehen. Ausführliche Überlegungen von Levine et al. (2003) können diesbezüglich übrigens zeigen, daß das Primen von Aspekten des Selbst offensichtlich doch nicht gänzlich ohne Probleme zu sein scheint. Mehrere ihrer durchgeführten Priming-Experimente zur Veränderung von independenten/interdependenten Aspekten zeigten sogar überhaupt keine Wirkung; zumindest nicht in Hinsicht auf die gängigen Skalen zur Messung von Selbstkonstrukten (wie z.B. die Skala von Singelis, 1994). Möglicherweise muß daher das Primen als eine Alternative dafür, Kultur ins Labor zu holen, zukünftig kritischer betrachtet werden. Die Ergebnisse des eigenen Experiments konnten aber zumindest hinsichtlich eines Teils der Argumentationskette signifikante Ergebnisse liefern. In weiteren Studien sollte unter Beachtung der kritischen Anmerkungen von Levine et al. (2003) jedoch gezeigt werden, daß auch diese noch fehlende Lücke in der kausalen Argumentation vollständig geschlossen werden kann.

In methodischer und statistischer Hinsicht könnte auch den Mediationsanalysen, sowie in Teilen den Mehrebenenanalysen, vorgeworfen werden, daß die darin gemessenen Effekte nicht immer das notwendige Signifikanzniveau erreichten. Dieser Aspekt war aber vor allem auf die notgedrungenenmaßen geringe Stichprobengröße zurückgeführt worden. Ein Grundproblem von interkultureller Forschung auf dem Länderniveau besteht genau darin, daß es ungemein schwer ist, auf eine umfangreiche Länderstichprobe Zugriff zu haben. Daher wurde argumentiert, daß nicht allein die Signifikanz als ausschlaggebende Größe angesehen werden sollte, sondern vielmehr die dazugehörigen Effektgrößen sowie das gesamte Muster der Effekte betrachtet werden sollte. Um somit trotz zum Teil fehlender Signifikanzen dennoch Vertrauen in die gefundenen Effekte haben zu können, wurden die Analysen anhand weiterer, inhaltlich gleicher, aber statistisch komplett unabhängiger Indikatoren wiederholt. Da sich dabei

in überzeugender Weise das gleiche Muster zeigte, wurde dies als Indiz dafür interpretiert, daß die Effekte sich nicht zufällig ergeben haben. Diese Variation der Indikatoren war somit dazu in der Lage, die zum Teil fehlenden Signifikanzen auszugleichen.

Bezüglich der Operationalisierung der abhängigen Variable kann schließlich noch kritisch angemerkt werden, daß diese in den Sekundäranalysen nur durch ein Item gemessen wurde. Dies war jedoch vor allem den internationalen Datensätzen geschuldet, auf die zurückgegriffen wurde. Da im World-Values-Survey nur ein Item enthalten war, welches eindeutig Arbeitszufriedenheit gemessen hat, konnte auf keine umfassendere Skala zurückgegriffen werden. Dennoch wurde trotz dieses Nachteils der World-Values-Survey aufgrund seiner umfassenden Länderanzahl eingesetzt. Wie im theoretischen Teil der Arbeit bereits erwähnt, ist jedoch auch die Messung anhand dieses einen Globalurteils durchaus aussagekräftig und muß nicht grundsätzlich einen Nachteil darstellen. In zukünftigen Studien wäre es aber sicherlich sinnvoll, auch Skalen mit mehreren Items und somit auch weitere Zufriedenheitsfacetten hinsichtlich ihrer Beeinflussung durch Kommunikationsnormen zu untersuchen.

11.5 Fazit

Interkulturelle Arbeitszufriedenheitsmessungen werden auch in Zukunft weiter von großer Bedeutung sein, vor allem im Unternehmenskontext in Form von Mitarbeiterbefragungen. Wie bei allen Befragungen, die über Landesgrenzen hinweg Daten erheben, sind solche Befragungen mit vielerlei Herausforderungen und Problemen verbunden. Ein spezifisches Problem wurde durch die Ergebnisse dieser hier vorliegenden Arbeit in deutlicher Weise in den Mittelpunkt gerückt. Die Art und Weise, wie Menschen in einem spezifischen Land auf Fragebögen reagieren, hat einen deutlichen Einfluß auf die Interpretierbarkeit von länderübergreifenden Resultaten. Der Vergleichbarkeit von Arbeitszufriedenheitsurteilen aus unterschiedlichen Kulturen sind gemäß den Ergebnissen dieser Arbeit somit Grenzen gesetzt worden. Diese Urteile stellen kein absolut sauberes und unverfälschtes Maß von Mitarbeitermeinungen dar, sondern sind in Teilen Ausdruck eines kulturell sozialisierten Kommunikationsverhaltens. Aus ihnen läßt sich somit zwar etwas über den sozialen und kommunikativen Umgang der Menschen in der jeweiligen Kultur lernen, ein Rückschluß auf die tatsächliche Meinung eines Mitarbeiters ist aber nur bedingt

möglich. Die Ergebnisse der Arbeit zeigen, daß vor allem bei der Betrachtung von Zufriedenheitsurteilen aus einer bestimmten Gruppe von Ländern, d.h. solchen mit relativ hohem Wohlstand und gleichzeitig starken Kommunikationsnormen, Schlußfolgerungen nur mit Vorsicht gezogen werden sollten. Zwar wird die Sinnhaftigkeit von Mitarbeiterbefragungen durch diese Arbeit nicht grundsätzlich in Frage gestellt, dennoch sollten Interpretationen von Befragungsergebnissen vor allem in dieser Art von Ländern nur unter Berücksichtigung der jeweiligen Kommunikationsnorm vorgenommen werden. In manchen Ländern sagen Mitarbeiter offensichtlich nicht unbedingt genau das, was sie eigentlich denken, sondern äußern nur eine kulturell gefärbte Meinung. Während in den klassisch westlichen Industrienationen, die mehrheitlich eine stark individualistische Kultur besitzen, Fragen an die Belegschaft nach der Zufriedenheit mit der Arbeit weiterhin brauchbare Resultate hervorbringen sollten, stellt sich vor allem die Frage bei den stark kollektivistischen Staaten, inwieweit das Urteil im Fragebogen überhaupt korrespondiert mit der ‚wahren‘ Einschätzung der Befragten. Zustimmungsnormen führen in diesen Ländern dazu, daß die persönliche, innere Einschätzung mit der nach außen abgegebenen Beurteilung in keinem direkten Bezug stehen muß. Im Übrigen könnte auch für wirtschaftlich schwächere Länder die Validität von Fragebogenurteilen aufgrund der Ergebnisse dieser Arbeit mit einem Fragezeichen versehen werden. Hier ist der Grund dafür jedoch weniger in der Kultur und dem damit verbundenen Kommunikationsverhalten zu suchen, sondern andere Aspekte, z.B. möglicherweise eine geringere Befragungserfahrung oder auch Bildung, könnten die Ursache dafür sein, daß klassische Befragungsverfahren dort gegebenenfalls nicht zu brauchbaren Daten führen.

Die Ergebnisse der Arbeit weisen somit insgesamt darauf hin, daß ein Vergleich von Arbeitszufriedenheitsniveaus aus unterschiedlichen Ländern nur mit äußerster Vorsicht durchgeführt werden sollte und immer den jeweiligen kulturellen und kommunikativen Kontext berücksichtigen sollte. Dieser Hinweis ist in seiner Bedeutung jedoch nicht nur auf den Arbeitskontext beschränkt, sondern läßt sich durchaus auch für alle andere Arten von interkulturellen Befragungen geben. Auch dort sollte noch stärker die Frage nach der Interpretierbarkeit der Daten gestellt werden.

Eine Lösung für die Problematik des interkulturellen Vergleichs von Befragungsdaten läßt sich nicht unbedingt leicht finden. Diesbezüglich zeigte sich, daß Benchmarks in Mitarbeiterbefragungen einen Teil der verloren gegangenen Interpretationssicherheit zurückbringen könnten, obwohl sie jedoch das Ursprungsproblem nicht vollständig

lösen. Interkulturelle Fragebogenstudien, nicht nur hinsichtlich der Messung von Arbeitszufriedenheit, erfassen eben nicht nur rein inhaltliche Unterschiede in den untersuchten Konstrukten, sondern bilden immer auch das jeweilige Kommunikationsmuster ab, welches sich in einer Kultur über die Zeit etabliert hat. In zukünftiger Forschung sollte daher möglicherweise auch über Alternativen in der Befragungsdurchführung nachgedacht werden. Dabei ließe sich an eine andere Art des Frageninhalts denken, an eine mehr indirekte Erfassung von Zufriedenheitsurteilen oder auch möglicherweise an eher verhaltensbezogene Maße, die weniger von Kommunikationsnormen beeinflusst sein sollten.

Literatur

- Aaker, J. L., & Lee, A. Y. (2001). "I" seek pleasures and "we" avoid pains: The role of self-regulatory goals in information processing and persuasion. *Journal of Consumer Research*, 28, 33–49.
- Adler, S., Skov, R.B. & Salvemini, N.J. (1985). Job characteristics and job satisfaction: When cause becomes consequence. *Organizational behaviour and Human Decision Processes*, 35, 266-278.
- Aiken, L. S. & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.
- Anderson, C.A. & Anderson, K.B. (1996). Violent crime rate studies in philosophical context: A destructive testing approach to heat and southern culture of violence effects. *Journal of Personality and Social Psychology*, 70, 740-756.
- Argyle, M., Henderson, M., Bond, M. H., Iizuka, Y. & Contarello, A. (1986). Cross-cultural variations in relationship rules. *International Journal of Psychology*, 21, 287-315.
- Argyris, C. & Schön, D.A. (1978). *Organizational learning: A theory of action perspective*. Reading, MA: Addison-Wesley.
- Aycan, Z. (2000). Cross-cultural industrial and organizational psychology: Contributions, past developments, and future directions. *Journal of Cross-Cultural Psychology*, 31, 110-128.
- Aycan, Z. & Kanungo, R.N. (2001). Cross-cultural industrial and organizational psychology: A critical appraisal of the field and future directions. In N. Anderson, D.S. Ones, H.K. Sinangil & C. Viswesvaran (Hrsg.), *Handbook of industrial, work and organizational psychology* (Band 1, S. 385-408). London: Sage Publications.
- Bachman, J.G. & O'Malley, P.M. (1984). Yeasaying, nay-saying, and going to extremes: Black–white differences in response styles. *Public Opinion Quarterly*, 48, 491–509.
- Bagozzi, R.P. & Yi, Y. (1991). Multitrait-multimethod matrices in consumer research. *Journal of Consumer Research*, 17, 426-439.
- Baron, R.M. & Kenny, D.A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51 (6), 1173-1182.
- Basabe, N., Paez, D., Valencia, J., Gonzales, J.L., Rime, B. & Diener, E. (2002). Cultural dimension, socioeconomic development, climate, and emotional hedonic level. *Cognition & Emotion*, 16 (1), 103-125.

- Baumgartner, C. & Udris, I. (2006). Das „Zürcher Modell“ der Arbeitszufriedenheit – 30 Jahre „still going strong“. In L. Fischer (Hrsg.), *Arbeitszufriedenheit – Konzepte und empirische Befunde* (S. 111-134). Göttingen: Hogrefe.
- Baumgartner, H. & Steenkamp, J.B.E.M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38 (2), 143-156.
- Bernardi, R.A. (2006). Associations between Hofstede's cultural constructs and social desirability response Bias. *Journal of Business Ethics*, 65, 43-53.
- Berry, J.W. (1969). On cross-cultural comparability. *International Journal of Psychology*, 4, 119-128.
- Berry, J.W. (1997). Introduction to methodology. In H.C. Triandis & R.W. Brislin (Hrsg.), *Handbook of cross-cultural psychology: Methodology* (Band 2, S. 1-27). Boston: Allyn and Bacon, Inc.
- Berry, J.W., Poortinga, Y.H., Segall, M.H. & Dasen, P.R. (1992). *Cross-cultural psychology: Research and applications*. New York: Cambridge University Press.
- Bochner, S. (1994). Cross-Cultural Differences in the Self-Concept. A Test of Hofstede's Individualism/Collectivism Distinction. *Journal of Cross-Cultural Psychology*, 25 (2), 273-283.
- Bond, M.H. & Cheung, T. (1983). College students' spontaneous self-concept: The effect of culture among respondents in Hong Kong, Japan, and the United States. *Journal of Cross-Cultural Psychology*, 14, 153-171.
- Bond, M.H., Leung, K. & Wan, K.-C. (1982). The social impact of self-effacing attributions: The Chinese case. *Journal of Social Psychology*, 118, 157-166.
- Bond, M.H. & Yang, K. S. (1982). Ethnic affirmation versus crosscultural accommodation: The variable impact of questionnaire language on Chinese bilinguals from Hong Kong. *Journal of Cross-Cultural Psychology*, 13, 169-185.
- Borg, I. (1986). A cross-cultural replication on Elizur's facets of work values. *Multivariate Behavioral Research*, 21, 401-410.
- Borg, I. (2000). *Führungsinstrument Mitarbeiterbefragung*. Göttingen: Verlag für Angewandte Psychologie.
- Borg, I. (2003). Affektiver Halo in Mitarbeiterbefragungen. *Zeitschrift für Arbeits- und Organisationspsychologie*. 24 (1), 1-11.
- Brett, J.M., Tinsley, C.H., Janssens, M., Barsness, Z.I. & Lytle, A.L. (1997). New approaches to the study of culture in industrial/organizational psychology. In P.C. Early & M. Erez (Hrsg.). *New perspectives on industrial/organizational psychology* (S. 75-129). San Francisco: The New Lexington Press.

- Brewer, M.B. & Gardner, W.L. (1996). Who is this “we”? Levels of collective identity and self-representations. *Journal of Personality and Social Psychology*, 71, 83–93.
- Brickman, P. & Campbell, D.T. (1971). Hedonic relativism and planning the good society. In M.H. Appley (Hrsg.), *Adaptation-level theory* (S. 28-305). New York: Academic Press.
- Brickman, P., Coates, D. & Janoff-Bulman, R. (1978). Lotterywinners and accident victims: Is happiness relative? *Journal of Personality and Social Psychology*, 37, 917-927.
- Brief, A.P. (1998). *Attitudes in and around organizations*. Thousand Oaks, CA: Sage.
- Briley, D. A., & Wyer, R. S. (2002). The effect of group membership salience on the avoidance of negative outcomes: Implications for social and consumer decisions. *Journal of Consumer Research*, 29, 400–415.
- Bruggemann, A., Groskurth, P. & Ulich, E. (1975). Arbeitszufriedenheit. *Schriften zur Arbeitspsychologie*, 17. Bern, Stuttgart: Huber.
- Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park: Sage.
- Bungard, W. (1997). Mitarbeiterbefragungen als Instrument modernen Innovations- und Qualitätsmanagements. In W. Bungard & I. Jöns (Hrsg.). *Mitarbeiterbefragungen* (S. 5-14). Weinheim: Psychologische Verlagsunion.
- Byrne, B.M. & Campbell, T.L. (1999). Cross-cultural comparisons and the presumption of equivalent measurement and theoretical structure – A look beneath the surface. *Journal of Cross-Cultural Psychology*, 30 (5), 555-574.
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, 56 (2), 81-105.
- Campbell, J.D., Trapnell, P.D., Heine, S.J., Katz, I.M., Lavalley, L.F. & Lehman, D.R. (1996). Self-concept clarity: Measurement, personality correlates, and cultural boundaries. *Journal of Personality and Social Psychology*, 70, 141–156.
- Carr, L.G. & Krause, N. (1978). Social status, psychiatric symptomatology, and response bias. *Journal of Health and Social Behaviour*, 19 (1), 86-91.
- Cartwright, S. & Cooper, C.L. (1992). *Mergers and acquisitions: The human factor*. Oxford: Butterworth-Heinemann.
- Chan, W. (2003). Analyzing ipsative data in psychological research. *Behaviormetrika*, 30, 99-121.
- Chen, C., Lee, S.Y. & Stevenson, H.W. (1995). Response style and cross-cultural comparisons of rating scales among east Asian and North American students. *Psychological Science*, 6 (3), 170–175.

- Cheung, G.W. & Rensvold, R.B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, 31 (2), 187-212.
- Chiu, R.K. & Kosinski, F.A. (1999). The role of affective dispositions in job satisfaction and work strain: Comparing collectivist and individualist societies. *International Journal of Psychology*, 34 (1), 19-28.
- Chun, K.T., Campbell, J.B. & Yoo, J.H. (1974). Extreme response style in crosscultural research. *Journal of Cross-Cultural Psychology*, 5 (4), 465-480.
- Church, A.T. (1987). Personality research in a non-Western culture: The Phillipines. *Psychological Bulletin*, 102 (2), 272-292.
- Clarke III, I. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior and Personality*, 15 (1), 137-152.
- Clarke III, I. (2001). Extreme response style in cross-cultural research. *International Marketing Review*, 18 (3), 301-324.
- Closs, S. J. (1996). On the factoring and interpretation of ipsative data. *Journal of Occupational and Organizational Psychology*, 69, 41-47.
- Collett, P. & O'Shea, G. (1976). Pointing the way to a fictional place: A study of direction giving in Iran and England. *European Journal of Social Psychology*, 6, 447-458.
- Cornwell, J.M. & Dunlop, W.P. (1994). On the questionable soundness of factoring ipsative data: A response to Saville and Willson (1991). *Journal of Occupational and Organizational Psychology*, 67, 89-100.
- Cranny, C.J., Smith, P.C. & Stone, E.F. (1992). *Job satisfaction: How people feel about their jobs and how it affects their performance*. New York: Lexington Books.
- Cross, S.E., Bacon, P.L. & Morris, M.L. (2000). The relational-interdependent self-construal and relationships. *Journal of Personality and Social Psychology*, 78, 791-808.
- Crowne, D.P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Cunningham, W., Cunningham, I.C.M. & Green, R.T. (1977). The ipsative process to reduce response set bias. *Public Opinion Quarterly*, 41, 379-394.
- de Jong, M.G., Steenkamp, J.B.E.M., Fox, J.P. & Baumgartner, H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of Marketing Research*, 45 (1), 104-115.
- Deresky, H. (1994). *International management: Managing across borders and cultures*. New York: HarperCollins College Publishers.

- Deutscher Bundestag (Hrsg.) (2002). *Schlussbericht der Enquete-Kommission "Globalisierung der Weltwirtschaft"*. Opladen: Leske + Budrich.
- Diener, E. (1995). A value based index for measuring national quality of life. *Social Indicators Research*, 36, 107-127.
- Diener, E., Diener, M. & Diener, C. (1995). Factors predicting the subjective well-being of nations. *Journal of Personality and Social Psychology*, 69 (5), 851-864.
- Diener, E., Scollon, C.K.N., Oishi, S., Dzokoto, V. & Suh, M.E. (2000). Positivity and the construction of life satisfaction judgements: Global happiness is not the sum of its parts. *Journal of Happiness Studies*, 1, 159-176.
- Drasgow, F. & Kanfer, R. (1985). Equivalence of psychological measurement in heterogenous populations. *Journal of Applied Psychology*, 70 (4), 662-680.
- Dunham, R.B. & Smith, F.J. (1979). *Organizational Surveys: An internal assessment of organizational health*. Glenview: Scott, Foresman and Company.
- Easterlin, R.A. (1974). Does economic growth improve the human lot? Some empirical evidence. In P.A. David & M.W. Reder (Hrsg.), *Nations and Households in Economic Growth* (S. 89-125). New York: Academic Press.
- Engel, U. (1998). *Einführung in die Mehrebenenanalyse*. Opladen: Westdeutscher Verlag.
- England, G.W. & Harpaz, I. (1983). Some methodological and analytic considerations in cross-national comparative research. *Journal of International Business Studies*, 14, 49-59.
- Erez, M. (1994). Toward a model of cross-cultural industrial and organizational psychology. In H.C. Triandis, M.D. Dunnette & L.M. Hough (Hrsg.), *Handbook of industrial and organizational psychology* (Band 4, S. 559-608). Palo Alto, CA: Consulting Psychologists Press.
- Eurobarometer (n.d.). [01.10.2008] verfügbar unter:
http://ec.europa.eu/public_opinion/index_en.htm
- European Foundation for the Improvement of Living and Working Conditions (2006). Construction of income bands for the 4th European Working Conditions Survey. [01.04.2008] verfügbar unter:
<http://www.eurofound.europa.eu/docs/ewco/4EWCS/4EWCSincomebands.pdf>
- Eysenck, H. & Eysenck, S. (1975). *The Eysenck Personality Questionnaire manual*. London: Hodder & Stoughton.
- Fischer, L. (1989). *Strukturen der Arbeitszufriedenheit. Zur Analyse individueller Bezugssysteme*. Göttingen: Hogrefe.
- Fischer, L. (1991). *Arbeitszufriedenheit. Beiträge zur Organisationspsychologie*, 5. Stuttgart: Verlag für Angewandte Psychologie.

- Fischer, L. & Belschak, F. (2006). Objektive Arbeitszufriedenheit? Oder: Was messen wir, wenn wir nach der Zufriedenheit mit der Arbeit fragen? In L. Fischer (Hrsg.), *Arbeitszufriedenheit – Konzepte und empirische Befunde* (S. 80-108). Göttingen: Hogrefe.
- Fischer, L. & Lück, H.E. (1972). Entwicklung einer Skala zur Messung von Arbeitszufriedenheit (SAZ). *Psychologie und Praxis*, 16, 64-76.
- Fischer, R. (2004). Standardization to account for cross-cultural response bias: A classification of score adjustment procedures and review of research in JCCP. *Journal of Cross-Cultural Psychology*, 35, 263-282.
- Fischer, R. (2006). Congruence and functions of personal and cultural values: Do my values reflect my culture's values? *Personality and Social Psychology Bulletin*, 32 (1), 1419-1431.
- Fischer, R., Fontaine, J., van de Vijver, F. & van Hemert, D. (2006). Response styles and the validity of cross-cultural comparisons: What is style and what is bias? Presentation at the 18th International Congress of the International Association for Cross-Cultural Psychology (IACCP), Spetses, Greece.
- Gardner, W.L., Gabriel, S. & Lee, A.Y. (1999). "I" value freedom, but "we" value relationships: Self-construal priming mirrors cultural differences in judgement. *Psychological Science*, 10 (4), 321-326.
- Georgas, J. & Berry, J.W. (1995). An ecocultural taxonomy for cross-cultural psychology. *Cross-Cultural Research: The Journal of Comparative Social Science*, 29 (2), 121-157.
- Georgas, J., van de Vijver, F.J.R. & Berry, J.W. (2004). The ecocultural framework, ecosocial indices, and psychological variables in cross-cultural research. *Journal of Cross-Cultural Psychology*, 35 (1), 74-96.
- Greenleaf, E.A. (1992a). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research*, 29, 176-188.
- Greenleaf, E.A. (1992b). Measuring extreme response style. *Public Opinion Quarterly*, 56, 328-351.
- Grice, H.P. (1975). Logic and conversation. In: P. Cole & J.L. Morgan (Hrsg.), *Syntax and semantics* (Band 3, S. 41-58). New York: Academic Press.
- Griffeth, R.W. & Hom, P.W. (1987). Some Multivariate Comparison of Multinational Managers. *Multivariate Behavioral Research*, 22, 173-191.
- Grimm, S.D. & Church, A.T. (1999). Across-cultural study of response biases in personality measures. *Journal of Research in Personality*, 33, 415-441.
- Gudykunst, W.B. & Matsumoto, Y. (1996). Cross-cultural variability of communication in personal relationships. In W.B. Gudykunst, S. Ting-Toomey & T. Nishida

- (Hrsg.). *Communication in personal relationships across cultures* (S. 19-56). Thousand Oaks, CA: Sage.
- Gurwitz, P.M. (1987). Ipsative rescaling: An answer to the response set problem in segmentation analysis. *Journal of Advertising Research*, 27, 37-42.
- Hackman, J.R. & Oldham, G.R. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, 16, 250-279.
- Haire, M., Ghiselli, E.E. & Porter, L.W. (1966). *Managerial thinking: An international study*. New York: Wiley.
- Harb, C. & Smith, P.B. (2008). Self-Construals Across Cultures. *Journal of Cross-Cultural Psychology*, 39 (2), 178-197.
- Harrison, G.L. (1995). Satisfaction, tension and interpersonal relations: A cross-cultural comparison of managers in Singapore and Australia. *Journal of Managerial Psychology*, 10 (8), 13-19.
- Harzing, A. (2005). Does the use of English-language questionnaires in cross-national research obscure national differences? *International Journal of Cross Cultural Management*, 5 (2), 213-224.
- Harzing, A. (2006). Response styles in cross-national survey research. *International Journal of Cross Cultural Management*, 6 (2), 243-266.
- Heine, S.J., Lehman, D.R., Peng, K.P. & Greenholz, J. (2002). What's wrong with cross-cultural comparisons of subjective Likert scales? The reference group effect. *Journal of Personality and Social Psychology*, 82, 903-918.
- Herzberg, F. (1967). *Work and the nature of man*. Cleveland, OH: World Book.
- Hicks, L. E. (1970). Some properties of ipsative, normative and forced-choice normative measures. *Psychological Bulletin*, 74, 167-184.
- Ho, D.Y.F. (1976). On the concept of face. *American Journal of Sociology*, 81, 867-890.
- Ho, D.Y.F. (1996). Filial piety and its psychological consequences. In M.H. Bond (Hrsg.). *Handbook of Chinese Psychology* (S. 155-165). Hongkong: Oxford University Press.
- Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills: Sage Publications.
- Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Thousand Oaks, CA: Sage Publications.
- Hofstede, G. (2002). The pitfalls of cross-national survey research: A reply to the article by Spector et al. on the psychometric properties of the Hofstede values survey module 1994. *Applied Psychology: An International Review*, 51 (1), 170-173.

- Hofstede, G. (2005). *Cultures and Organizations – Software of the mind*. New York: McGraw Hill.
- Hofstede, G. & Bond, M.H. (1988). The Confucius connection: From cultural roots to economic growth. *Organizational Dynamics*, 16 (4), 4-21.
- Hofstede, G. & McCrae, R.R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-Cultural Research*, 38 (1), 52-88.
- Holtgraves, T. (1997). Styles of language use: Individual and cultural variability in conversational indirectness. *Journal of Personality and Social Psychology*, 73, 624-637.
- House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W. & Gupta, V. (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Thousand Oaks, CA: Sage Publications.
- House, R.J., Wright, N.S. & Aditya, R.N. (1997). Cross-cultural research on organizational leadership: A critical analysis and a proposed theory. In P.C. Early & M. Erez (Hrsg.). *New perspectives on industrial/organizational psychology* (S. 535-625). San Francisco: The New Lexington Press.
- Hox, J. (1995). *Applied multilevel analysis*. Amsterdam: TT-Publikaties.
- Hu, H.C. (1944). The Chinese concepts of face. *American Anthropologist*, 46, 45–64.
- Huang, X. (2002). *Just a matter of taste? – One nation’s job satisfiers are another nation’s job dissatisfiers*. Veenendaal: Universal Press.
- Huang X. & van de Vliert, E. (2002). Intrinsic job rewards at Country-level and individual-level codetermine job satisfaction. *Journal of International Business Studies*, 33, 2, 385-394.
- Huang, X. & van de Vliert, E. (2004). Job level and national culture as joint roots of job satisfaction. *Applied Psychology: An international review*, 53 (3), 329-348.
- Hui, C.H. (1988). Measurement of individualism-collectivism. *Journal of Research in Personality*, 22, 17-36.
- Hui, C.H. & Triandis, H.C. (1985a). Measurement in cross-cultural psychology – A review and comparison of strategies. *Journal of Cross-Cultural Psychology*, 16 (2), 131-152.
- Hui, C.H. & Triandis, H.C. (1985b). The instability of response sets. *Public Opinion Quarterly* 49 (2), 253-260.
- Hui, C.H. & Triandis, H.C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20 (3), 296-309.
- Hui, C.H. & Yee, C. (1999). The impact of psychological collectivism and workgroup atmosphere on Chinese employees’ job satisfaction. *Applied Psychology: An International Review*, 48, 175-185.

- Hui, C.H., Yee, C. & Eastman, K.L. (1995). The relationship between individualism-collectivism and job satisfaction. *Applied Psychology: An International Review*, 44, 276-282.
- Inglehart, R. & C. Welzel (2004). What insights can multi-country surveys provide about people and societies? *Comparative Politics Newsletter*, APSA (Summer): 1-12.
- ISSP (n.d.). International Social Survey Programme. [01.10.2008] verfügbar unter: <http://www.issp.org/>
- Jahoda, G. (1980). Theoretical and systematic approaches in cross-cultural psychology. In H.C. Triandis & R.W. Brislin (Hrsg.) *Handbook of cross-cultural psychology: Perspectives* (Band 1, S. 69-141). Boston: Allyn & Bacon, Inc.
- Jöns, I. (2002). Fusionen und Akquisitionen aus der Sicht von Mitarbeitern und Führungskräften. *Wirtschaftspsychologie*, 8 (3), 22-26.
- Johnson, S.R. (1996). The multinational opinion survey. In A.I. Kraut (Hrsg.), *Organizational surveys: Tools for assessment and change* (S. 310-329). San Francisco, CA: Jossey-Bass Publishers.
- Johnson, T.P., Kulesa, P., Cho, Y.I. & Shavitt, S. (2005). The relation between culture and response styles. *Journal of Cross-Cultural Psychology*, 36, 264-277.
- Johnson, T., O'Rourke, D., Chavez, N., Sudman, S., Warnecke, R., Lacey, L. & Horm, J. (1997). Social cognition and responses to survey questions among culturally diverse populations. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz & D. Trewin (Hrsg.). *Survey Measurement and Process Quality* (S. 87-113). New York: John Wiley & Sons.
- Judge, T.A. & Hulin, C.L. (1993). Job satisfaction as a reflection of dispositions: A multiple-source causal analysis. *Organizational Behavior and Human Decision Processes*, 56, 388-421.
- Judge, T.A., Locke, E.A., Durham, C.C. & Kluger, A.N. (1998). Dispositional effects on job and life satisfaction: The role of core self evaluations. *Journal of Applied Psychology*, 83, 17-43.
- Judge, T.A., Parker, S., Colbert, A.E., Heller, D. & Ilies, R. (2001). Job satisfaction: A cross-cultural review. In N. Anderson, D.S. Ones, H.K. Sinangil & C. Viswesvaran (Hrsg.), *Handbook of Industrial, Work and Organizational Psychology (Band 2), Organizational Psychology* (S. 25-52). London: Sage Publications.
- Kagitcibasi, C. & Poortinga, Y.H. (2000). Cross-cultural psychology – Issues and overarching themes. *Journal of Cross-Cultural Psychology* 31 (1), 129-147.
- Kahneman, D. (1999). Objective happiness. In D. Kahnemann, E. Diener & N. Schwarz (Hrsg.). *Well-Being: The foundation of hedonic psychology* (S. 3-25). New York: Russell Sage Foundation.

- Kashima, E. & Kashima, Y. (1998). Culture and language: The case of cultural dimensions and personal pronoun use. *Journal of Cross-Cultural Psychology*, 29, 461-486.
- Kimmelmeier, M. & Cheng, B. (2004). Language and self-construal priming: A replication and extension in a Hong Kong sample. *Journal of Cross-Cultural Psychology*, 35, 705-712.
- Khanlou, N. (2005). Cultural identity as part of youth's self-concept in multicultural settings. *International Journal of Mental Health & Addiction*, 3 (2), 1-14.
- Kim, U., Park, Y.S. & Park, D. (2000). The challenge of cross-cultural psychology. *Journal of Cross-Cultural Psychology*, 31 (1), 63-75.
- King, M.F. & Bruner, G.C. (2000). Social desirability bias: A neglected aspect of validity testing. *Psychology & Marketing*, 17 (2), 79-103.
- Kinicki, A.J., McKee-Ryan, F.M., Schriesheim, C.A. & Carson, K.P. (2002). Assessing the construct validity of the Job Descriptive Index: A review and meta-analysis. *Journal of Applied Psychology*, 87 (1), 14-32.
- Kinket, B. & Verkuyten, M. (1997). Levels of ethnic self-identification and social context. *Social Psychology Quarterly*, 60 (4), 338-354.
- Kirkman, B.L. & Shapiro, D.L. (2001). The impact of cultural values on job satisfaction and organizational commitment in self-managing work teams: The mediation role of employee resistance. *Academy of Management Journal*, 44, 557-569.
- Kiss, P. (2004). Hungary and Europe: Identity and representation - Comprehensive country report prepared for the EURONAT project. ELTE Institute for Psychology, Eötvös Loránd University, Budapest/Ungarn. [15.01.2009] verfügbar unter <http://www.iue.it/RSCAS/Research/EURONAT/200412Rep.EURONAT-H.pdf>.
- Kitayama, S., Markus, H.R., Matsumoto, H. & Norasakkunit, V. (1997). Individual and collective processes in the construction of self: Self-enhancement in the United States and self-criticism in Japan. *Journal of Personality and Social Psychology*, 72, 1245-1267.
- Knowles, E.S. & Nathan, K.T. (1997). Acquiescent responding in self-reports: Cognitive style or social concern. *Journal of Research in Personality*, 31, 293-301.
- Koopmann, G. & Franzmeyer, F. (2003). Weltwirtschaft und internationale Arbeitsteilung. *Informationen zur politischen Bildung*, 280, 12-26.
- Kühnen, U. & Hannover, B. (2000). Assimilation and contrast in social comparisons as a consequence of self-construal activation. *European Journal of Social Psychology*, 30, 799-811.
- Kühnen, U. & Hannover, B. (2003). Kultur, Selbstkonzept und Kognition. *Zeitschrift für Psychologie*, 211 (4), 212-224.

- Kühnen, U., Hannover, B., Roeder, U., Schubert, B., Shah, A., Upmeyer, A. & Zakaria, S. (2001). Cross-cultural variations in identifying embedded figures: Comparisons from the US, Germany, Russia, and Malaysia. *Journal of Cross-Cultural Psychology, 32*, 365–371.
- Kühnen, U., Hannover, B. & Schubert, B. (2001). The semantic-procedural interface model of the self: The role of self-knowledge for context-dependent versus context-independent modes of thinking. *Journal of Personality and Social Psychology, 80*, (3), 397-409.
- Kuhn, M.H. & McPartland, T.S. (1954). An empirical investigation of self-attitudes. *American Sociological Review, 19* (1), 68-76.
- Lalwani, A.K., Shavitt, S. & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology, 90* (1), 165-178.
- Law, K.S. & Wong, C.S. (1999). Multidimensional constructs in structural equation analysis: An illustration using the job perception and job satisfaction constructs. *Journal of Management, 25* (2), 143-160.
- Law, K.S., Wong, C.S. & Mobley, W.H. (1998). Toward a taxonomy of multidimensional constructs. *Academy of Management Review, 23* (4), 741-755.
- Lawler, E.E. (1973). *Motivation in work organizations*. Monterey, CA: Brooks/Cole Publishing.
- Lee, C. & Green, R.T. (1991). Cross-cultural examination of the Fishbein behavioral intentions model. *Journal of International Business Studies, 22*, 289-305.
- Leung, K. & Bond, M.H. (1989). On the empirical identification of dimensions for cross-cultural comparison. *Journal of Cross-Cultural Psychology, 20*, 133-151.
- Leung, K., Bond, M.H., de Carrasquel, S., Munoz, C., Hernandez, M., Murakami, F., et al. (2002). Social axioms: The search for universal dimensions of general beliefs about how the world functions. *Journal of Cross-Cultural Psychology, 33*, 286-302.
- Levin, J. & Montag, I. (1987). The effect of testing instructions for handling social desirability on the Eysenck Personality Questionnaire. *Personality and Individual Differences, 8*, 163-167.
- Levine, T.R., Bresnahan, M.J., Park, H.S., Lapinski, M.K., Wittenbaum, G.M., Shearman, S.M., Lee, S.Y., Chung, D. & Ohashi, R. (2003). Self-Construal scales lack validity. *Human Communication Research, 29* (2), 210-252.
- Lincoln, J.R., Hanada, M. & Olson, J. (1981). Cultural orientation and individual reactions to organizations: A study of employees of Japanese-owned firms. *Administrative Science Quarterly, 26*, 93-115.

- Lincoln, J.R. & Kalleberg, A.L. (1985). Work organization and workforce commitment: A study of plants and employees in the U.S. and Japan. *American Sociological Review*, 50, 738-760.
- Little, T.D. (2000). On the comparability of constructs in cross-cultural research – A Critique of Cheung and Rensvold. *Journal of Cross-Cultural Psychology*, 31 (2), 213-219.
- Liu, C., Borg, I. & Spector, P.E. (2004). Measurement equivalence of the German Job Satisfaction Survey used in a multinational organization: Implications of Schwartz's culture model. *Journal of Applied Psychology*, 89, 1070-1082.
- Llorente, R. & Macias, E.F. (2005). Job satisfaction as an indicator of the quality of work. *The Journal of Socio-Economics*, 34, 656-673.
- Locke, E.A. (1976). The nature and causes of job satisfaction. In M.D. Dunnette (Hrsg.), *Handbook of industrial and organizational psychology* (S. 1297-1343). Chicago: Rand McNally.
- Lynn, R. & Martin, T. (1995). National differences for thirty-seven nations in extraversion, neuroticism, psychoticism and economic, demographic and other correlates. *Personality and Individual Differences*, 19 (3), 403–406.
- Marin, G., Gamba, R.J. & Marin, B.V. (1992). Extreme response style and acquiescence among Hispanics: The role of acculturation and education. *Journal of Cross-Cultural Psychology*, 23, 498-509.
- Markus, H. & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.
- Markus, H. & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross Cultural Psychology*, 29, 63–87.
- Marion-Landais, C.A. (1993). *A cross-cultural study of leader-member exchange quality and job satisfaction as correlates of intra-dyadic work-value congruence*. Unveröffentlichte Master Arbeit, University of South Florida, Tampa.
- Marsh, H.W. & Byrne, B.M. (1993). Confirmatory factor analysis of multi-group–multi-methodself-concept data: Between group and within-group variance constraints. *Multivariate Behavioral Research*, 28, 313-349.
- McCrae, R.R., & Costa, P.W. (1997). Personality trait structure as a human universal. *American Psychologist*, 52, 509-516.
- McFarlin, D.B. & Rice, R.W. (1991). Determinants of satisfaction with specific job facets: A test of Locke's model. *Journal of Business and Psychology*, 6 (1), 25-38.
- Miller, J.G., Bersoff, D.M. & Harwood, R.L. (1990). Perceptions of social responsibilities in India and the United States: Moral imperatives or personal decisions? *Journal of Personality and Social Psychology*, 58, 33-47.

- Miner, J.B. (1992). *Industrial-organizational psychology*. New York: McGraw-Hill.
- Morris, T. & Parvett, C.M. (1992). Management style and productivity in two cultures. *Journal of International Business Studies*, 23, 169-179.
- Morris, M. W., Williams, K. Y., Leung, K., Larrick, R., Mendoza, M. T., Bhatnagar, D., et al. (1998). Conflict management style: Accounting for cross-national differences. *Journal of International Business Studies*, 29, 729-747.
- MOW International Research Team (1987). *The meaning of working*. London: Academic Press.
- Müller, K. (2006). *Kulturelle Einflüsse auf die Arbeitszufriedenheit – Eine multinationale Untersuchung*. Taunusstein: Driesen.
- Nakane, C. (1970). *Japanese society*. Berkeley: University of California Press.
- Near, J.P. & Rechner, P.L. (1993). Cross-cultural variations in predictors of life satisfaction: An historical view of differences among West European countries. *Social Indicators Research*, 29, 109-121.
- Neuberger, O. (1974). *Messung der Arbeitszufriedenheit*. Stuttgart: Kohlhammer.
- Nezlek, J.B., Schröder-Abé, M. & Schütz, A. (2006). Mehrebenenanalysen in der psychologischen Forschung – Vorteile und Möglichkeiten der Mehrebenenmodellierung mit Zufallskoeffizienten, *Psychologische Rundschau*, 57 (4), 213-223.
- Ones, D.S., Viswesvaran, C. & Reiss, A.D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology*, 81, 660-679.
- Oyserman, D. & Lee, S.W.S. (2008). Does culture influence what and how we think? Effects of priming individualism and collectivism. *Psychological Bulletin*, 134 (2), 311-342.
- Paulhus, D. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598-609.
- Perlitz, M. & Seger, F. (1999). Strategische Unternehmensführung. In A. Kieser & W.A. Oechsler (Hrsg.) *Unternehmenspolitik* (S. 211-271). Stuttgart: Schäffer-Poeschel.
- Peters, A.M. & Boggs, S.T. (1987). Interactional routines as cultural influences upon language acquisition. In B.B. Schieffelin & E. Ochs (Hrsg.). *Language socialization across cultures* (S. 80-96). Cambridge: University Press.
- Phillips, D.L. & Clancy, K.J. (1970). Response biases in field studies of mental illness. *American Sociological Review*, 35 (3), 503-515.
- Poortinga, Y.H. (1989). Equivalence of cross-cultural data: An overview of basic issues. *International Journal of Psychology*, 23, 737-756.

- Podsakoff, P.M., MacKenzie, S.B., Lee, Y. und Podsakoff, N. P. (2003): Common Method Biases in Behavioral Research: A critical Review of the Literature and Recommended Remedies. *Journal of Applied Psychology*, 88 (5), 879-903.
- Ralston, D.A., Cunniff, M.K. & Gustafson, D.J. (1995). Cultural accommodation: The effect of language on the responses of bilingual Hong Kong Chinese managers. *Journal of Cross-Cultural Psychology*, 26, 714–727.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models* (2nd ed.). Thousand Oaks, CA: Sage.
- Ray, J.J. (1983). Reviving the problem of acquiescent response bias. *Journal of Social Psychology*, 121, 81-96.
- Rhee, E., Uleman, J.S., Lee, H.K. & Roman, R.J. (1995). Spontaneous self-descriptions and ethnic identities in individualistic and collectivistic cultures. *Journal of Personality and Social Psychology*, 69, 142–152.
- Riordan, C.M. & Vandenberg, R.J. (1994). A central question in cross-cultural research: Do employees of different cultures interpret work-related measures in an equivalent manner? *Journal of Management*, 20 (3), 643-671.
- Roeder, U.R. & Hannover, B. (2002). Kontextabhängigkeit als Dimension der Selbstkonstruktion. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 23, (3), 339-352.
- Ross, C.E. & Mirowsky, J. (1984). Socially desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior*, 25 (2), 189–197.
- Ryan, A.M., Chan, D., Ployhart, R.E. & Slade, L.A. (1999). Employee attitude surveys in a multinational organization: Considering language and culture in assessing measurement equivalence. *Personnel Psychology*, 52, 37-58.
- Ryan, A.M., Horvath, M. Ployhart, R.E., Schmitt, N. & Slade, L.A. (2000). Hypothesizing differential item functioning in global employee opinion surveys. *Personnel Psychology*, 53 (3), 531-562.
- Salancik, G.R. & Pfeffer, J.R. (1977). An examination of need-satisfaction models of job attitudes. *Administrative Science Quarterly*, 22, 427-456.
- Salancik, G.R. & Pfeffer, J.R. (1978). A social information processing approach to job attitudes. *Administrative Science Quarterly*, 23, 224-252.
- Scarpello, V. & Campbell, J.P. (1983). Job satisfaction: Are all the parts there? *Personnel Psychology*, 36, 577-600.
- Schaninger, C.M. & Buss, W.C. (1986). Removing response-style effects in attribute-determinance ratings to identify market segments. *Journal of Business Research*, 14, 237-252.

- Schieffelin, B.B. & Ochs, E. (1986). Language socialization. *Annual Review of Anthropology*, 15, 163-191.
- Schneider, B., Ashworth, S.D., Higgs, A.C. & Carr, L. (1996). Design, validity, and use of strategically focused employee attitude surveys. *Personnel Psychology*, 49, 695-705.
- Schuman, H. & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording, and context*. Thousand Oaks, CA: Sage.
- Schwarz, N. & Clore, G.L. (1983). Mood, misattribution, and judgement of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, 45, 513-523.
- Schwarz, N. & Strack, F. (1991). Evaluating one's life: a judgement model of subjective well-being. In F. Strack, M. Argyle & N. Schwarz (Hrsg.), *Subjective Well-Being: An interdisciplinary Perspective*, (S. 27-47). Oxford, UK: Pergamon Press.
- Schwartz, S.H. (1990). Individualism-collectivism – Critique and proposed refinements. *Journal of Cross-Cultural Psychology*, 21 (2), 139-157.
- Schwartz, S.H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (Hrsg.), *Advances in experimental social psychology* (S. 1-62). Orlando, FL: Academic Press.
- Schwartz, S.H. (1994). Beyond individualism/collectivism: New cultural dimensions of values. In U. Kim, H.C. Triandis, C. Kagitcibasi, S.C. Choi & G. Yoon (Hrsg.), *Individualism and collectivism: Theory, method and applications* (S. 85-119). Newbury Park, CA: Sage.
- Schwartz, S.H. (1999). A theory of cultural values and some implications for work. *Applied psychology: An international review*, 48 (1), 23-47.
- Schwartz, S.H. (2004). Mapping and interpreting cultural differences around the world. In H. Vinken, J. Soeters & P. Ester (Hrsg.), *Comparing cultures, Dimensions of culture in a comparative perspective* (S.43-73). Leiden, The Netherlands: Brill.
- Schwartz, S.H. (2007). A proposal for measuring value orientations across nations. In *European Social Survey Core Questionnaire Development*, Kapitel 7, 259-319. [29.03.2009] verfügbar unter:
http://www.europeansocialsurvey.org/index.php?option=com_docman&task=doc_view&gid=126&Itemid=80
- Schwartz, S.H. & Ros, M. (1995). Values in the West: A theoretical and empirical challenge to the individualism-collectivism cultural dimension. *World Psychology*, 1, 99-122.
- Segall, M.H. (1986). Assessment of social behavior. In W. J. Lonner & J.W. Berry (Hrsg.), *Field methods in crosscultural research* (S. 265-290). Beverly Hills, CA: Sage.

- Sego, D.J., Hui, C. & Law, K.S. (1997). Operationalizing cultural values as the mean of individual values: Problems and suggestions for research. In P.C. Earley & M. Erez (Hrsg.), *New perspectives on international industrial/organizational psychology* (S. 148-159). San Francisco, CA: New Lexington Press.
- Shiomi, K. & Loo, R. (1999). Cross-cultural response styles on the Kirton Adaptation–Innovation Inventory. *Social Behavior and Personality* 27 (4), 413–420.
- Singelis, T.M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, 20, 580–591.
- Six, B. & Felfe, J. (2006). Arbeitszufriedenheit im interkulturellen Vergleich. In: L. Fischer (Hrsg.), *Arbeitszufriedenheit – Konzepte und empirische Befunde* (S. 243-272), Göttingen: Hogrefe.
- Six, B. & Kleinbeck, U. (1989), Arbeitsmotivation und Arbeitszufriedenheit. In E. Roth, H. Schuler & A.B. Weinert (Hrsg.), *Enzyklopädie der Psychologie. Wirtschafts-, Organisations- und Arbeitspsychologie (Band 3). Organisationspsychologie* (S. 348-398). Göttingen: Hogrefe.
- Slocum, J.W., Jr. & Topichak, P.M. (1972). Do cultural differences affect job satisfaction? *Journal of Applied Psychology*, 56, 177-178.
- Smith, P.B. (2004). Acquiescent Response Bias as an Aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, 35 (1), 50-61.
- Smith, P.B., Dugan, S. & Trompenaars, F. (1996). National culture and the values of organizational employees: A dimensional analysis across 43 nations. *Journal of Cross-Cultural Psychology*, 27, 231-264.
- Smith, P.B. & Fischer, R. (2008). Acquiescence, Extreme Response Bias and Culture: A Multilevel Analysis. In F.J.R. van de Vijver, D.A. van Hemert, Y.H. Poortinga (Hrsg.) *Multilevelanalysis of Individuals and Culture* (S. 288-311). New York: Erlbaum.
- Smith, P.B. & Misumi, J. (1989). Japanese Management - A sun rising in the West? In C.L. Copper & I.T. Robertson (Hrsg.), *International review of industrial and organizational psychology* (Band 4). Chichester, England: Wiley.
- Smith, P.B., Peterson, M.F., Schwartz, S.H., Ahmad, A.H., Akande, D., Anderson, J.A., et al. (2002). Cultural values, sources of guidance and their relevance to managerial behaviors: A 47 nation study. *Journal of Cross-Cultural Psychology*, 33, 188-208.
- Smith, P.C., Kendall, L.M. & Hulin, C.L. (1969). *The measurement of satisfaction in work and retirement*. Chicago: Rand McNally.
- Söhnchen, F. (2007). Common Method Variance und Single Source Bias. In S. Albers, D. Klapper, U. Konradt, A. Walter und J. Wolf (Hrsg.): *Methodik der empirischen Forschung*, 2. Auflage (S. 135-150). Wiesbaden: Gabler.

- Sondergaard, M. (1994). Research note: Hofstede's consequences: A study of reviews, citations and replications. *Organization Studies*, 15 (3), 447-456.
- Souza-Poza, A. & Souza-Poza, A. (2000). Well-being at work: A cross-national analysis of the levels and determinants of job satisfaction. *Journal of Socio-Economics*, 29, 517-538.
- Spector, P.E. (1997). *Job satisfaction: Application, assessment, causes, and consequences*. Thousand Oaks, CA: Sage.
- Spector, P.E. & Cooper, C.L. (2002). The pitfalls of poor psychometric properties: A rejoinder to Hofstede's reply to us. *Applied Psychology: An International Review*, 51 (1), 174-178.
- Spector, P.E., Cooper, C.L. & Sparks, K. (2001). An international study of the psychometric properties of the Hofstede Values Survey Module 1994: A comparison of individual and country/province level results. *Applied Psychology: An International Review*, 50 (2), 269-281.
- Spector, P.E., Cooper, C.L., Poelmans, S., Allen, T.D., O'Driscoll, M., Sanchez, J.I., Siu, O.L., Dewe, P., Hart, P., Lu, L., de Moraes, L.F.R., Ostrognay, G.M., Sparks, K. Wong, P. & Yu, S. (2004). A cross-national comparative study of work-family stressors, working hours, and well-being: China and Latin America versus the Anglo world. *Personnel Psychology*, 57, 199-142.
- Strull, T. K. & Wyer, R. S. (1979). The role of category accessibility in the interpretation of information about persons: Some determinants and implications. *Journal of Personality and Social Psychology*, 37, 1660-1672.
- Staples, D.S. & Higgins, C.A. (1998). A study of the impact of factor importance weightings on job satisfaction measures. *Journal of Business and Psychology*, 13 (2), 211-232.
- Steenkamp, J.E.M. & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research*, 25, 78-90.
- Stening, B.W. & Everett, J.E. (1984). Response styles in a cross-cultural managerial study. *Journal of Social Psychology*, 122, 151-156.
- Stipek, D, Weiner, B. & Li, K. (1989). Testing some attribution-emotion relations in the People's Republic of China. *Journal of Personality and Social Psychology*, 56, 109-116.
- Suh, E.M., Diener, E. & Updegraff, J.A. (2008). From culture to priming conditions: Self-construal influences on life satisfaction judgements. *Journal of Cross-Cultural Psychology*, 39 (3), 3-13.
- Takahashi, K., Ohara, N., Antonucci, T.C. & Akiyama, H. (2002). Commonalities and differences in close relationships among the Americans and Japanese: A comparison by the individualism/collectivism concept. *International Journal of Behavioral Development*, 26 (5), 453-465.

- Ten Berge, J.M.F. (1999). A legitimate case of component analysis of ipsative measures, and partialling the mean as an alternative to ipsatization. *Multivariate Behavioral Research*, 34 (1), 89-102.
- The Gallup Organization (2003). *Das Engagement am Arbeitsplatz sinkt weiter*. Pressemeldung. The Gallup Organization.
- The Gallup Organization (2008). *Gallup-Engagement-Index 2008*. Pressemeldung. The Gallup Organization.
- Towers Perrin (2005). *Winning Strategies for a Global Workforce: Attracting, Retaining, and Engaging Employees for Competitive Advantage*. Towers Perrin Global Workforce Study. [01.09.2006] verfügbar unter: www.towersperrin.com
- Trafimow, D., Silverman, E.S., Fan, R.M. T. & Law, J.S.F. (1997). The effects of language and priming on the relative accessibility of the private self and the collective self. *Journal of Cross-Cultural Psychology*, 28, 107–123.
- Trafimow, D., Triandis, H.C. & Goto, S.G. (1991). Some tests of the distinction between the private self and the collective self. *Journal of Personality and Social Psychology*, 60, 649–655.
- Triandis, H.C. (1972). *The analysis of subjective culture*. New York: Wiley-Interscience.
- Triandis, H.C. (1994). Cross-cultural industrial and organizational psychology. In H.C. Triandis, M.D. Dunnette, L.M. Hough (Hrsg.), *Handbook of industrial and organizational psychology* (Band 4, S. 103-172). Palo Alto, CA: Consulting Psychologists Press.
- Triandis, H.C. (1995). *Individualism and collectivism*. Boulder, CO: Westview Press.
- Triandis, H.C., Carnevale, P., Gelfand, M., Robert, C., Wasti, A., Probst, T., et al. (2001). Culture and deception in business negotiations: A multilevel analysis. *International Journal of Cross-Cultural Management*, 1, 73–90.
- Triandis, H.C. & Gelfand, M.J., (1998). Converging Measurement of Horizontal and Vertical Individualism and Collectivism. *Journal of Personality and Social Psychology*, 74 (1), 118-128.
- Triandis, H.C., McCusker, C. & Hui, C.H. (1990). Multimethod Probes of Individualism and Collectivism. *Journal of Personality and Social Psychology*, 59 (5), 1006-1020.
- Triandis, H.C., & Suh, E. M. (2002). Cultural influences on personality. *Annual Review of Psychology*, 53, 133–160.
- Trilling, L. (1972). *Sincerity and authenticity*. London: Oxford University Press.
- United Nations (2006a). *Social Indicators – Indicators on income and economic activity*. [01.10.2006] verfügbar unter: <http://unstats.un.org/unsd/demographic/products/socind/inc-eco.htm>

- United Nations (2006b). *Social Indicators – Indicators on unemployment*. [01.10.2006] verfügbar unter:
<http://unstats.un.org/unsd/demographic/products/socind/unemployment.htm>
- United Nations (2008). *Human Development Report*. [15.01.2009] verfügbar unter:
<http://hdr.undp.org/en/statistics/data/>
- United Nations (2009). *Social Indicators – Indicators on income and economic activity*. [15.01.2009] verfügbar unter:
<http://unstats.un.org/unsd/demographic/products/socind/inc-eco.htm>
- Vandenberg, R.J. & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3 (1), 4-69.
- van de Vijver, F.J.R. & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks: Sage Publications.
- van de Vijver, F.J.R. & Leung, K. (2000). Methodological issues in psychological research on culture. In J.W. Berry, Y.H. Poortinga & J. Pandey (Hrsg.), *Handbook of cross-cultural psychology* (Band 1, S. 257-300). Boston: Allyn & Bacon, Inc..
- van de Vijver, F.J.R. & Poortinga, Y.H. (1997). Towards an integrated analysis of bias in cross-cultural assessment. *European Journal of Psychological Assessment*, 13 (1), 29-37.
- van de Vliert, E. (2006). Autocratic leadership around the globe – Do climate and wealth drive leadership culture? *Journal of Cross-Cultural Psychology*, 37 (1), 42-59.
- van de Vliert, E., Huang, X. & Levine, R.V. (2004). National wealth and thermal climate as predictors of motives for volunteer work. *Journal of Cross-Cultural Psychology*, 35 (1), 62-73.
- van de Vliert, E. & Janssen, O. (2002). "Better than" performance motives as roots of satisfaction across more and less developed countries. *Journal of Cross-Cultural Psychology*, 33 (4), 380-397.
- van de Vliert, E., Schwartz, S.H., Huisman, S.E., Hofstede, G. & Daan, S. (1999). Temperature, cultural masculinity, and domestic political violence – A cross-national study. *Journal of Cross-Cultural Psychology*, 30 (3), 291-314.
- van Herk, H., Poortinga, Y.H. & Verhallen, T.M.M. (2004). Response styles in rating scales – Evidence of Method Bias in Data from Six EU Countries. *Journal of Cross-Cultural Psychology*, 35 (3), 346-360.
- van Hemert, D.A., van de Vijver, F.J.R., Poortinga, Y.H. & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences*, 33, 1229-1249.

- Wanous, J.P. & Lawler, E.E. (1972). Measurement and meaning of job satisfaction. *Journal of Applied Psychology*, 56 (2), 95-105.
- Watkins, D., & Cheung, S. (1995). Culture, gender, and response bias: An analysis of responses to the selfdescription questionnaire. *Journal of Cross-cultural Psychology*, 26 (5), 490-504.
- Wegge, J. & van Dick, R. (2006). Arbeitszufriedenheit, Emotionen bei der Arbeit und organisationale Identifikation. In L. Fischer (Hrsg.), *Arbeitszufriedenheit – Konzepte und empirische Befunde* (S. 11-36). Göttingen: Hogrefe.
- Weiss, D.J., Dawis, R.V., England, G.W. & Lofquist, L.H. (1967). *Manual for the Minnesota Satisfaction Questionnaire*. Minneapolis: Industrial Relations Center, University of Minnesota.
- Weiss, H.M. (2002). Deconstructing job satisfaction: Separating evaluations, beliefs, and affective experiences. *Human Resource Management Review*, 12 (2), 167-292.
- Weiss, H.M. & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. *Research in Organizational Behavior*, 18, 1-74.
- Weiss, H.M. & Shaw, J.B. (1979). Social influences on judgements about tasks. *Organizational Behavior and Human Performance*, 24, 126-140.
- Werner, O. & Campbell, D.T. (1970). Translating, working through interpreters, and the problem of decentering. In R. Narroll & R. Cohen (Hrsg.), *A handbook of method in cultural anthropology* (S. 398-420). New York: American Museum of Natural History Press.
- Wilson, T.D. & Hodges, S.D. (1992). Attitudes as temporary constructions. In L. Martin & A. Tesser (Hrsg.) *The construction of social judgements* (Band 10, S. 37-65). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Winkler, J. D., Kanouse, D.E. & Ware, J.E., Jr. (1982). Controlling for acquiescence response set in scale development. *Journal of Applied Psychology*, 67, 555–561.
- Winkler, N., Kroh, M. & Spiess, M. (2006). Entwicklung einer deutschen Kurzskala zur zweidimensionalen Messung von sozialer Erwünschtheit. *German Institute for Economic Research. Discussion Papers*, 579.
- World-Values-Survey (n.d.). [01.10.2008] verfügbar unter:
<http://www.worldvaluessurvey.org/>
- World-Values-Survey-Association – WVSA (n.d.). World-Values-Survey Brochure. [01.10.2008] verfügbar unter:
http://margaux.grandvinum.se/SebTest/wvs/SebTest/wvs/articles/folder_publish_ed/article_base_110/files/WVSBrochure4.pdf
- Yoshida, T., Kojo, K. & Kaku, H. (1982). A study on the development of self-presentation in children. *Japanese Journal of Educational Psychology*, 30, 30-37.

Zax, M. & Takahashi, S. (1967). Cultural influences on response style: comparisons of Japanese and American College Students. *Journal of Social Psychology*, 71, 3–10.

Zerbe, W.J., & Paulhus, D.L. (1987). Socially desirable responding in organizational behavior: A reconception. *Academy of Management Review*, 12, 250-264.

ZUMA-Informationssystem (2007). Bonn: Gesis.

Anhang

Anhang A

Dargestellt sind hier die Abkürzungen der in den Abbildungen aufgeführten Länder, sowie deren Zustimmungs-Index, der anhand der 5er-Skalen mit Zustimmungslabel des WVS gebildet wurden.

Land	Abkürzung	Zustimmungs-Index NARS 1 des WVS
Argentinien	ARG	0,38
Belgien	BEL	0,27
Brasilien	BRA	0,58
Bulgarien	BGR	0,51
Chile	CHL	0,34
China	CHN	0,55
Dänemark	DNK	0,21
Deutschland-Ost	DEU_O	0,41
Deutschland-West	DEU_W	0,18
Estland	EST	0,47
Finnland	FIN	0,23
Frankreich	FRA	0,29
Griechenland	GRC	0,31
Großbritannien	GBR	0,21
Indien	IND	0,40
Irland	IRL	0,22
Italien	ITA	0,32
Japan	JPN	0,21
Kanada	CAN	0,20
Kroatien	HRV	0,37
Luxemburg	LUX	0,23
Malta	MLT	0,37
Mexiko	MEX	0,31
Niederlande	NLD	-0,09
Nigeria	NGA	0,55
Nordirland	NIRL	
Norwegen	NOR	0,23
Österreich	AUT	0,14
Polen	POL	0,45
Portugal	PRT	0,34
Rumänien	ROU	0,63
Russland	RUS	0,48
Schweden	SWE	0,25
Schweiz	CHE	

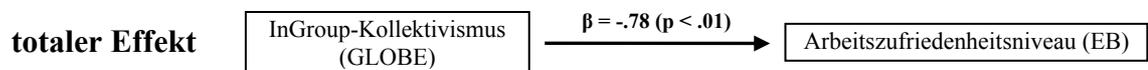
Slowakei	SVK	0,40
Slowenien	SVN	0,42
Spanien	ESP	0,36
Südafrika	ZAF	0,39
Südkorea	SKO	0,25
Tschechische Republik	CZE	0,32
Türkei	TUR	0,46
Ungarn	HUN	0,49
USA	USA	0,19

Anhang B

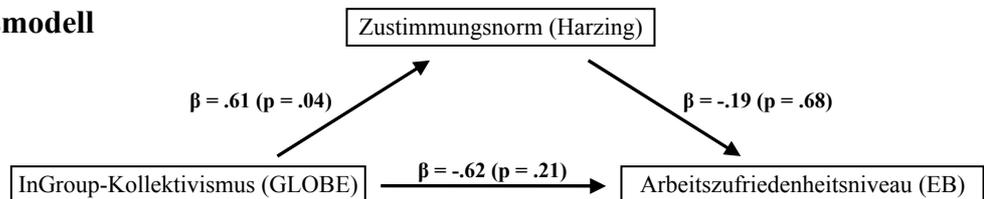
Dargestellt sind hier die weiteren Mediationsanalysen, die nicht im Fließtext dargestellt wurden.

1) weitere Mediationsanalysen anhand der Kulturdimension Individualismus/ Kollektivismus, der Zustimmungsnorm-Indizes und der Arbeitszufriedenheitsniveaus, berechnet nur für reiche Länder:

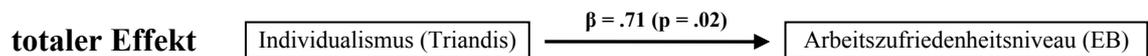
Länderanzahl: N=10



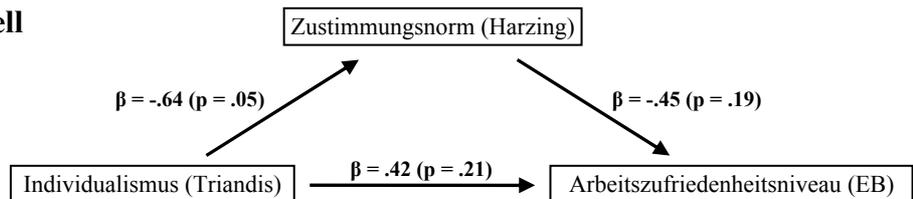
Mediationsmodell



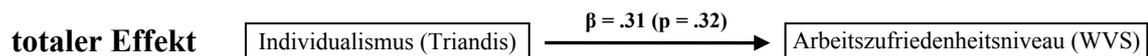
Länderanzahl: N=10



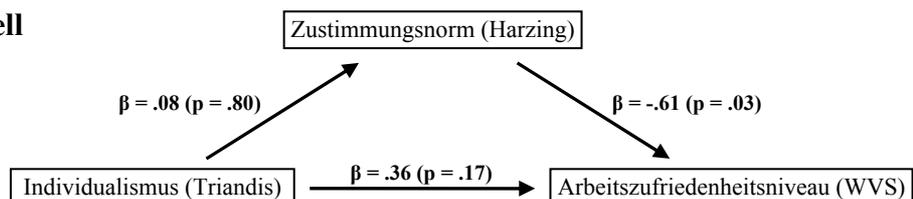
Mediationsmodell



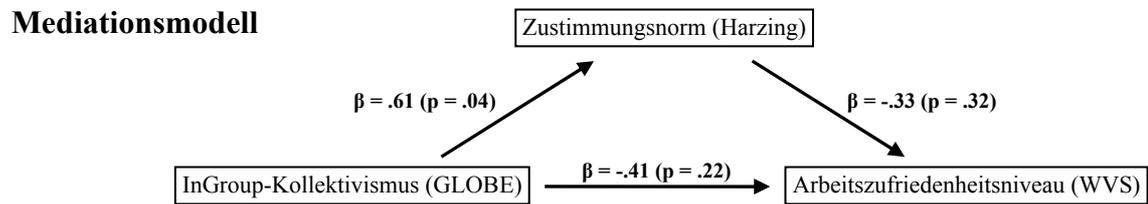
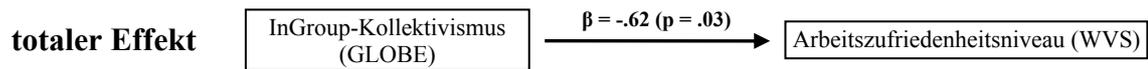
Länderanzahl: N=12



Mediationsmodell

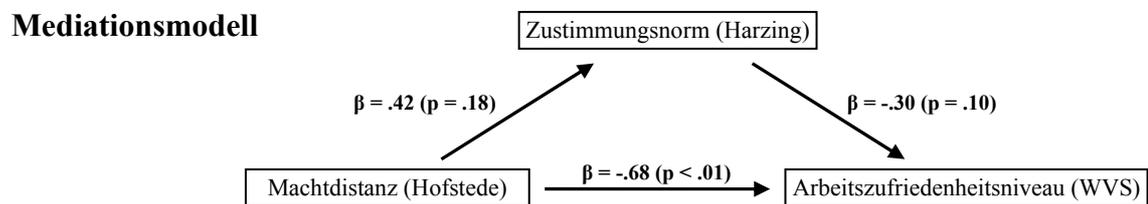
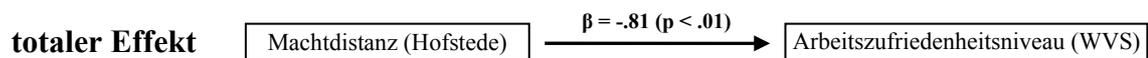


Länderanzahl: N=12

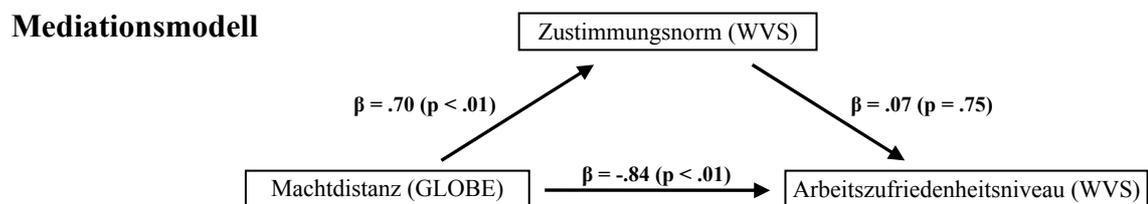
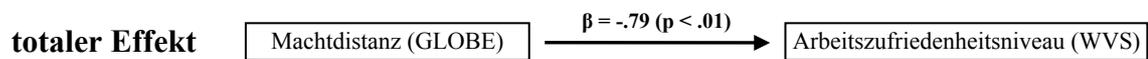


2) Mediationsanalysen anhand der Kulturdimension Machtdistanz, der Zustimmungsnorm-Indizes und der Arbeitszufriedenheitsniveaus, berechnet nur für reiche Länder:

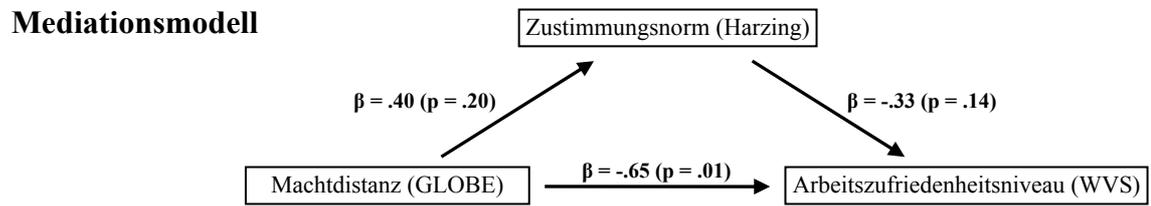
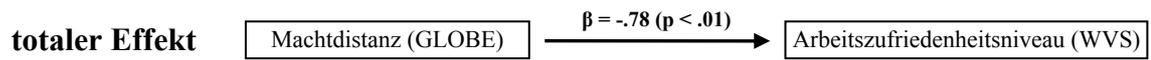
Länderanzahl: N=12



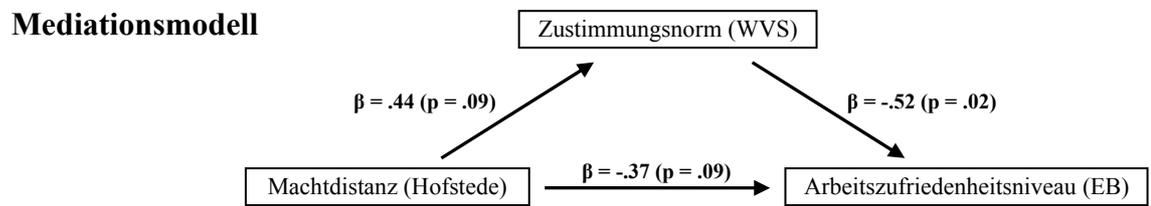
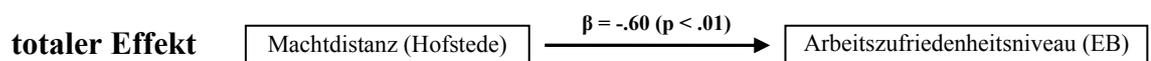
Länderanzahl: N=18



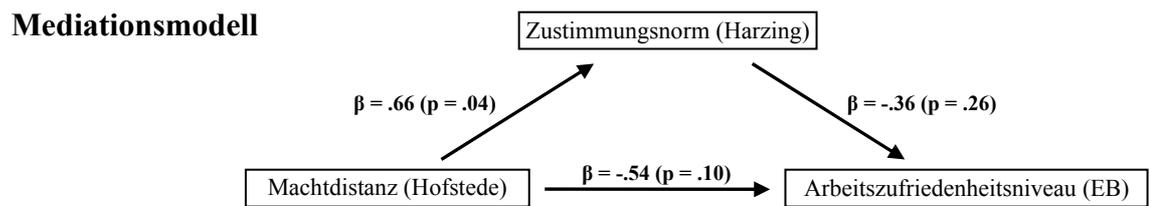
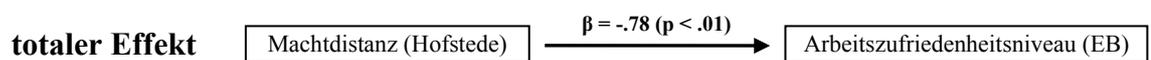
Länderanzahl: N=12



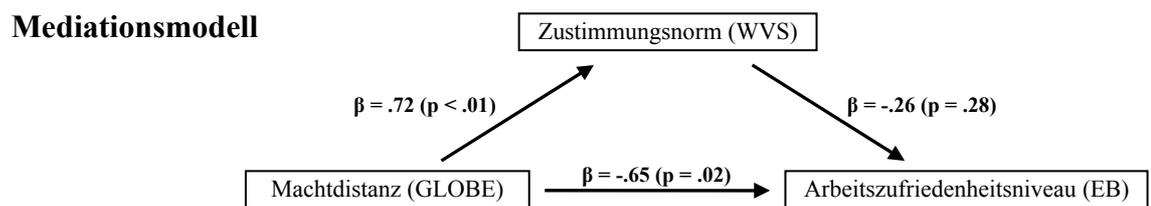
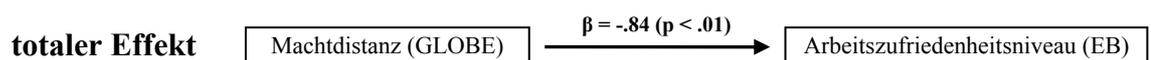
Länderanzahl: N=16



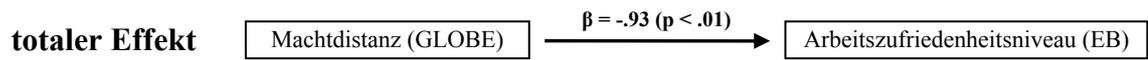
Länderanzahl: N=10



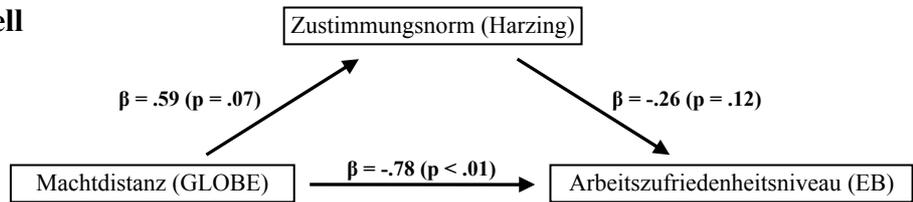
Länderanzahl: N=14



Länderanzahl: N=10

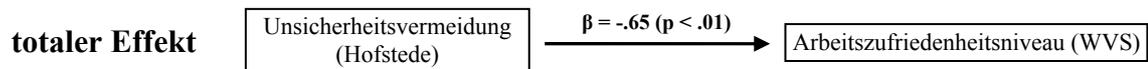


Mediationsmodell

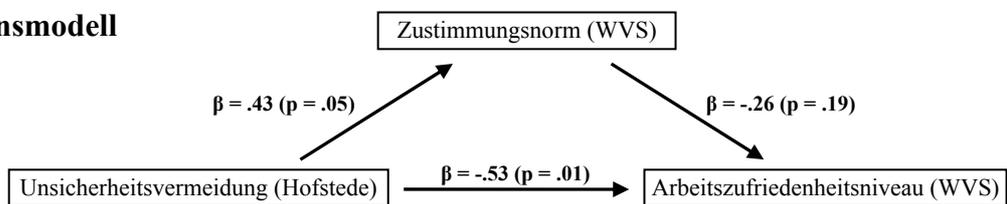


3) Mediationsanalysen anhand der Kulturdimension Unsicherheitsvermeidung, der Zustimmungsnorm-Indizes und der Arbeitszufriedenheitsniveaus, berechnet nur für reiche Länder:

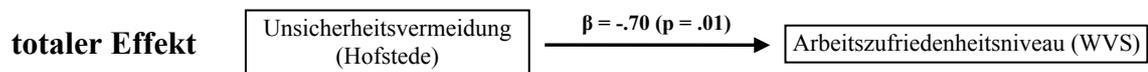
Länderanzahl: N=21



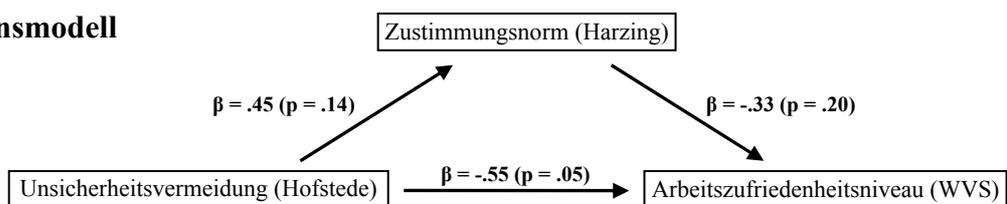
Mediationsmodell



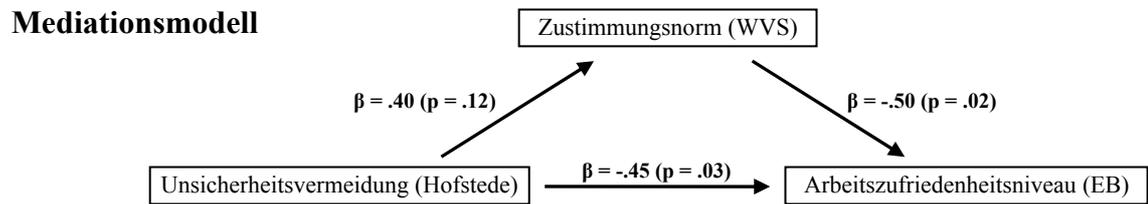
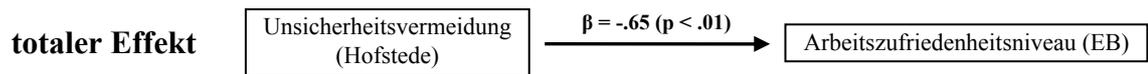
Länderanzahl: N=12



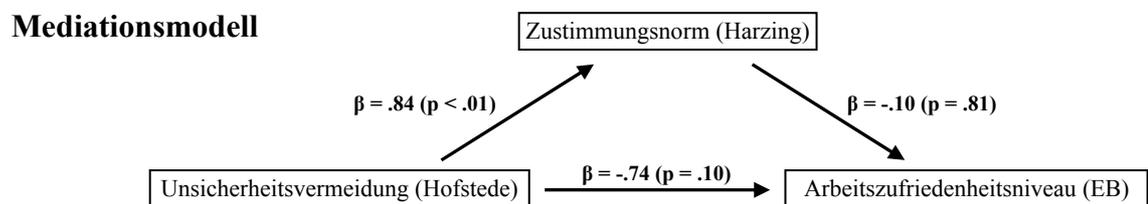
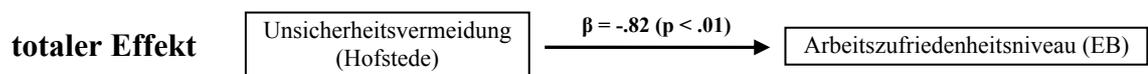
Mediationsmodell



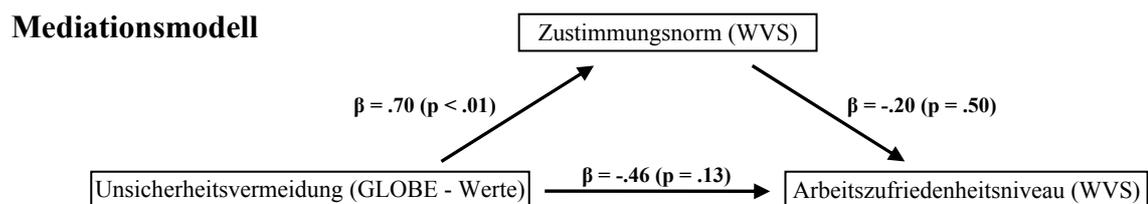
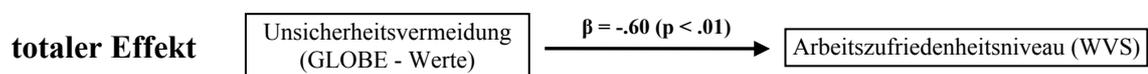
Länderanzahl: N=16



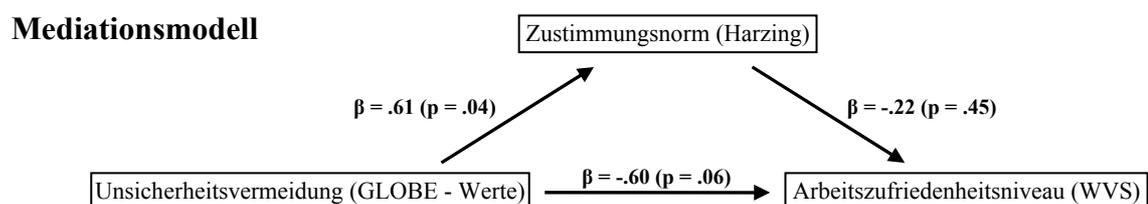
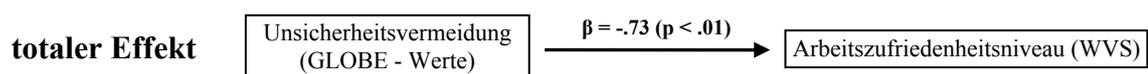
Länderanzahl: N=10



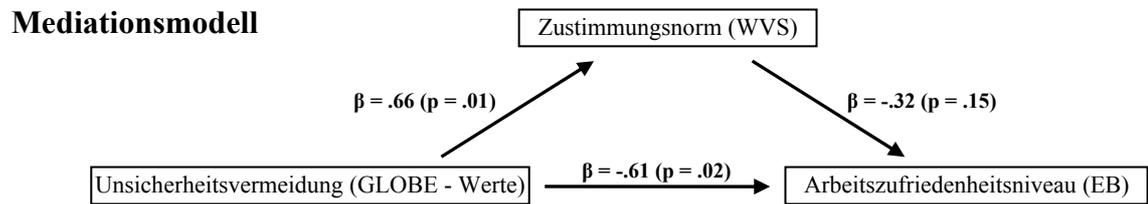
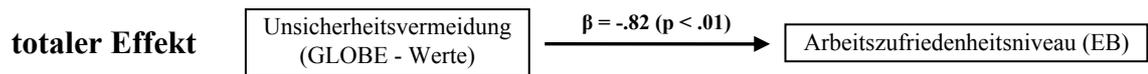
Länderanzahl: N=18



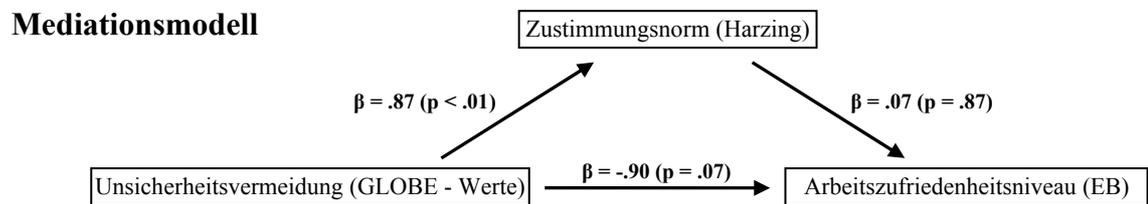
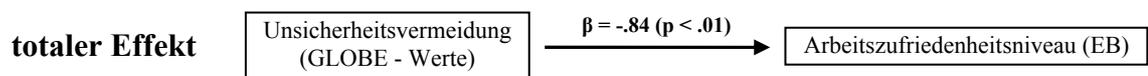
Länderanzahl: N=12



Länderanzahl: N=14

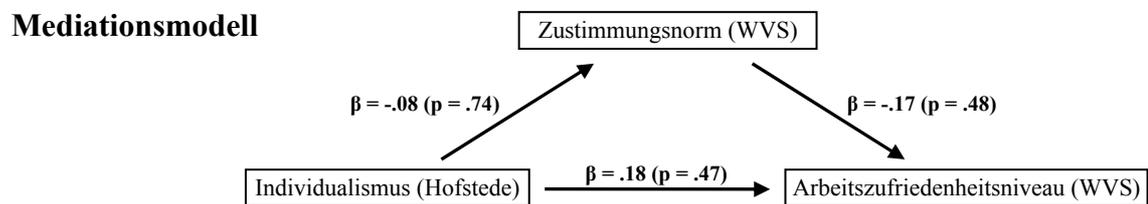
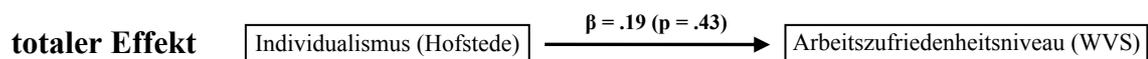


Länderanzahl: N=10

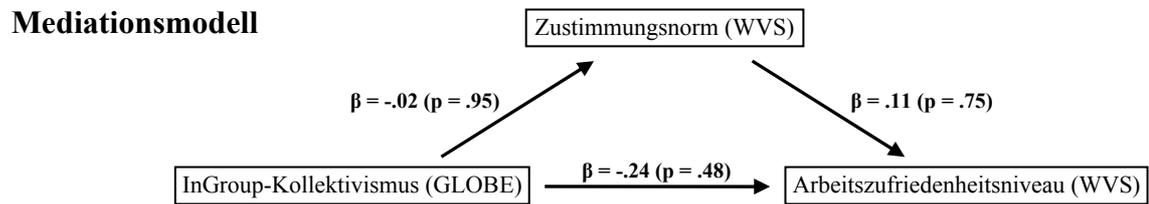
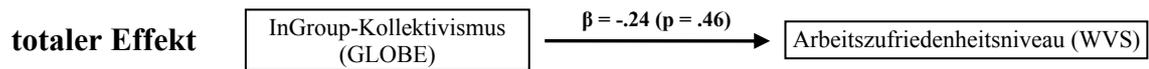


4) Mediationsanalysen, die nur bei armen Ländern berechnet wurden:

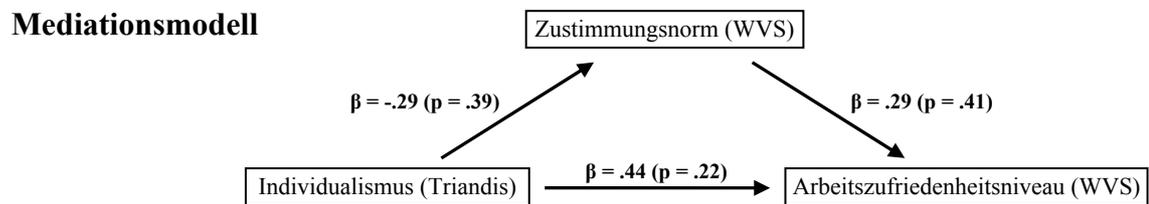
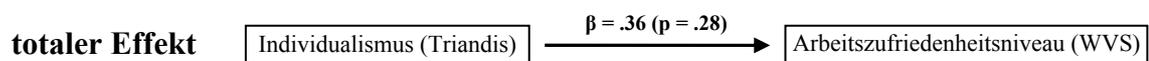
Länderanzahl: N=20



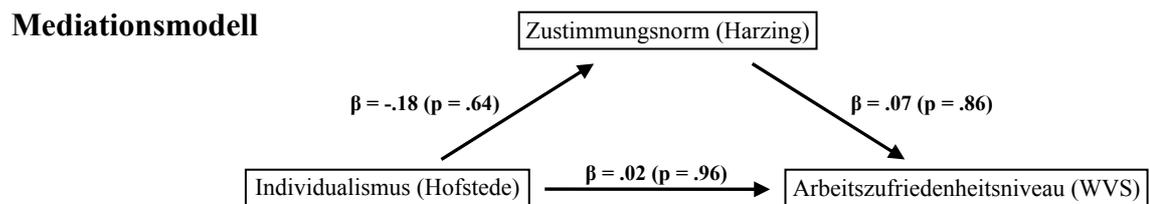
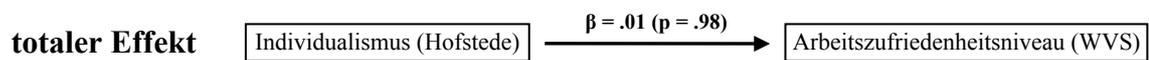
Länderanzahl: N=12



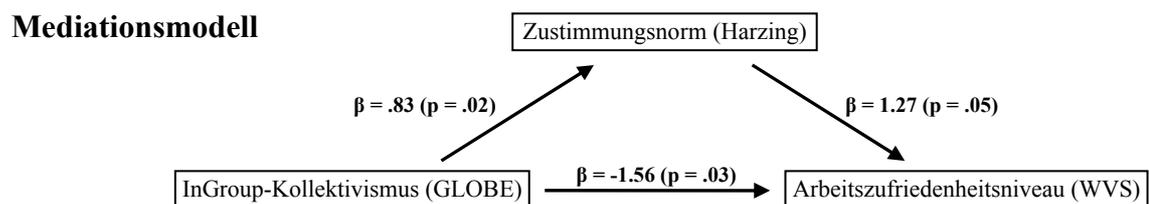
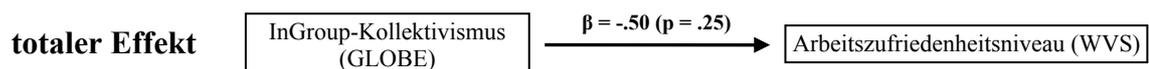
Länderanzahl: N=11



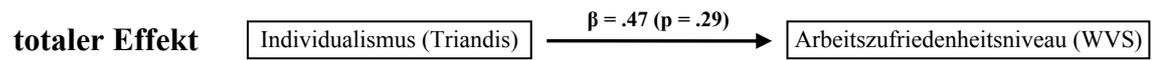
Länderanzahl: N=9



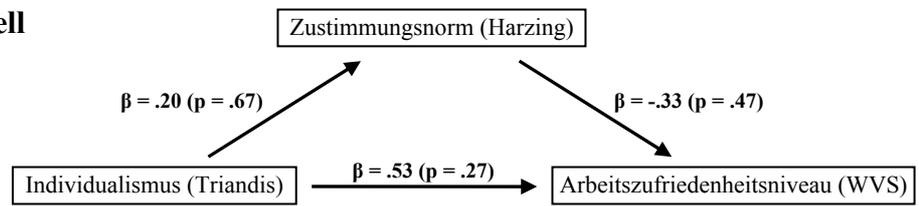
Länderanzahl: N=7



Länderanzahl: N=7



Mediationsmodell



Anhang C

Im Folgenden sind die einzelnen Seiten des Fragebogens aufgeführt, so wie er im Experiment eingesetzt wurde. Dargestellt ist jedoch lediglich die independente Variante des Primings sowie die weibliche Variante des PVQ.



Institut für Wirtschafts- und Sozialpsychologie

Lieber Teilnehmer, liebe Teilnehmerin,

vielen Dank, dass Du an unserer Untersuchungsreihe teilnimmst. Im Rahmen dieser Untersuchung werden Dir **drei voneinander unabhängige Studien** zu unterschiedlichen Themengebieten zur Beantwortung vorgelegt.

Bitte lies Dir jeweils die nun folgenden Instruktionen genau durch und versuche dann, die Fragen so zu beantworten, wie es am ehesten auf Dich zutrifft. Überlege bei den Fragen nicht erst, welche Antworten den „besten Eindruck“ machen, da es keine „richtigen“ oder „falschen“ Antworten gibt. Antworte stattdessen spontan.

Bitte bearbeite den Fragebogen in vorgegebener Reihenfolge. Bitte lasse nichts aus. Beachte bitte auch, dass Aufgaben auf den Rückseiten stehen, da der Fragebogen beidseitig gedruckt wurde.

Die Anonymisierung Deiner Daten wird selbstverständlich gewährleistet.

Für Deine Mitarbeit möchten wir Dir bereits im Vorfeld herzlich danken.



Institut für Wirtschafts- und Sozialpsychologie

Studie 1 (*Version „individualistisch“*)

Ziel der ersten Studie ist es, einen Einblick in das Sozialverhalten deutscher Studentinnen und Studenten, in ihrem persönlichen Umfeld, zu bekommen.

Bitte beachte, dass die erste Studie aus **zwei Teilaufgaben** besteht.

TEILAUFGABE 1

Mache Dir bitte zunächst zwei Minuten Gedanken darüber, was Dich von Deiner Familie und Deinen Freunden unterscheidet.

Was erwartest Du selbst von Dir?



Notiere bitte kurz in Stichpunkten, welche Gedanken Du Dir in den vorangegangenen zwei Minuten gemacht hast.



TEILAUFGABE 2

Bitte lese Dir nun den folgenden Text aufmerksam durch und beurteile die folgende Situation

Anne ist 30 Jahre alt und zeichnet gerne. Als sie die Zeitung aufschlägt liest sie, dass in einem Handel für Künstlerbedarf, der aufgelöst werden soll, ein großer Räumungsverkauf stattfindet. Anne erwartet einige Schnäppchen bei dem Räumungsverkauf machen zu können und will deshalb unbedingt zu dem Künstlerbedarfsladen gehen. Der Laden befindet sich in einer Straße am anderen Ende der Stadt. Da Anne nicht genau weiß, wo diese Straße ist, fragt sie ihre beste Freundin Lisa nach einer Wegbeschreibung. Sie erklärt Lisa, dass sie gerne früh zu dem Räumungsverkauf gehen würde, um noch eine große Auswahl vorzufinden. Deshalb braucht Anne eine schnelle Wegbeschreibung. Lisa allerdings ist gerade damit beschäftigt ein spannendes Buch zu lesen und will nicht dabei gestört werden. Aus diesem Grund hilft Lisa ihrer Freundin Anne nicht den Weg zu finden. Anne versucht daraufhin auf eigene Faust den Laden zu finden. Als sie aber nach langem Suchen dort ankommt, sind kaum noch Verkaufsartikel übrig.

Bitte gebe nun Deine Einschätzung der oben beschriebenen Situation ab.

1) Wie nachvollziehbar findest Du Lisas Verhalten?

absolut nachvollziehbar		teils/ teils			überhaupt nicht nachvollziehbar	
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

2) Wäre Lisa Deiner Meinung nach verpflichtet gewesen Anne zu helfen?

Ja Nein

3) Sollte Lisa für ihr Verhalten gerügt werden (bspw. durch verbale Zurechtweisung)?

Ja Nein



Studie 2

Eine Möglichkeit für Universitäten, die Studierenden stärker auf ihrem Weg ins Berufsleben zu unterstützen, ist, **offizielle Plattformen zur Praktikavermittlung** ins Leben zu rufen. Hierbei geht es darum, dass die auf diesen Plattformen zukünftig vorgestellten Ausschreibungen neben den Vorzügen des Praktikums die Arbeitssituation auch möglichst realistisch darstellen sollen, um sicherzustellen, dass beide Parteien im Nachhinein mit der Zusammenarbeit zufrieden sind.

Zu diesem Zweck wird derzeit in Zusammenarbeit mit Partner-Unternehmen erforscht, wie Arbeitsbedingungen im Rahmen eines Praktikums von Studenten wahrgenommen werden. Vor diesem Hintergrund bitten wir Sie, sich folgende **Praktikumsbeschreibung aufmerksam durchzulesen** und sich in diese Situation hineinzuversetzen um sie **anschließend bewerten** zu können.

Die Praktikumsbeschreibung ist bisher noch recht allgemein gehalten, wird aber später für die konkreten Unternehmensbereiche angepasst. **Gehen Sie folglich bitte davon aus, dass die Branche, sowie der Unternehmensbereich, in dem Sie das Praktikum absolvieren, Ihren persönlichen Vorstellungen entsprechen.**



Praktikumsbeschreibung

Wir sind ein großes, DAX notiertes Unternehmen, das engagierte, motivierte und leistungsstarke Praktikanten sucht. Eigeninitiative und Flexibilität werden in unserem Unternehmen groß geschrieben. Im Gegenzug für eine engagierte Mitarbeit Ihrerseits bieten wir Ihnen einen umfassenden Einblick und eine aktive Einbindung in die für Sie relevanten Unternehmensabläufe an. Sie werden eigenständig an Projekten von der Planung bis zur Durchführung mitarbeiten. Ziel ist es eine win-win Situation herzustellen, von der beide Seiten profitieren. Sie können je nach Wunsch und Studienausrichtung Ihr Praktikum bei uns in den folgenden Bereichen absolvieren: Marketing, Finanzen, Human Resources, Einkauf, Public Relations, Rechtswesen, Forschung & Entwicklung, etc.



Ihr Arbeitsplatz befindet sich in einem Büro mit anderen Kollegen, wodurch eine rege Kommunikation zwischen Ihnen und den anderen Mitarbeitern möglich ist. Auf Kommunikation und eine gute Betreuung seitens des Unternehmens wird viel Wert gelegt. Das Entwickeln von Problemlösungsvorschlägen hilft Ihrem Team bei der Entscheidungsfindung und gibt Ihnen die Möglichkeit, sich aktiv am Betriebsgeschehen zu beteiligen.

Aus bisherigen Projekten ist bekannt, dass bei einem hohen Arbeitspensum Überstunden geleistet werden müssen. Dass Sie dazu bereit sind, setzen wir voraus.

Ihre Leistung im Unternehmen ist uns eine adäquate Aufwandsentschädigung wert. Ihr Praktikumsgehalt entspricht dem, vergleichbarer Unternehmen, zusätzlich haben Sie die Chance, einen Bonus in Höhe eines Monatsgehaltes zu erhalten, wenn Ihre Praktikumsbewertung sehr gut ausfällt. Des Weiteren besteht für Sie die Möglichkeit, durch Ihren Praktikumsbetreuer für ein Praktikantenbindungsprogramm nominiert zu werden, welches nach dem Studium häufig zu einem Einstieg in dieses Unternehmen führt.

Es ist von ausschlaggebender Relevanz für die beschriebene Plattform, dass Sie sich mit einer solchen Beschreibung identifizieren können. Aus diesem Grund bitten wir Sie die folgenden Fragen möglichst sorgfältig und ehrlich zu beantworten.


**Was glauben Sie, wie zufrieden wären Sie mit folgenden Aspekten in dieser
Praktikumssituation?**

	außerordentlich zufrieden		teils/ teils			überhaupt nicht zufrieden		kann ich nicht beurteilen
Wie zufrieden wären Sie insgesamt mit diesem Praktikumsplatz?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	sehr wahrscheinlich		teils/ teils			überhaupt nicht wahrscheinlich		kann ich nicht beurteilen
Bitte bewerten Sie die Wahrscheinlichkeit, mit der Sie sich bei diesem Unternehmen um ein Praktikum bewerben würden?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Wie zufrieden wären Sie mit...	außerordentlich zufrieden		teils/ teils			überhaupt nicht zufrieden		kann ich nicht beurteilen
...der Möglichkeit Ihre Fähigkeiten einzusetzen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Möglichkeit Eigeninitiative zeigen zu können	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Möglichkeit praktische Erfahrungen zu sammeln	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Arbeitszeit	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Bezahlung	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...dem Einblick in das Unternehmen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Möglichkeit eines späteren Berufseinstiegs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
...der Attraktivität der Tätigkeiten	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	stimme voll und ganz zu		teils/ teils			stimme überhaupt nicht zu		kann ich nicht beurteilen
Ich finde die Praktikumsbeschreibung verständlich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich kann mich in die beschriebene Situation hinein versetzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Das angebotene Praktikum passt zu meinem Studium.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Studie 3 (Version „weiblich“)

Ziel dieser Studie ist es, die Beurteilung von Werturteilen und die Selbsteinschätzung von Personen zu untersuchen.

Im Folgenden werden Dir einige Personen beschrieben. Bitte entscheide, wie ähnlich oder unähnlich Dir die beschriebene Person ist.

	Ist mir sehr ähnlich	Ist mir ähnlich	Ist mir etwas ähnlich	Ist mir nur ein kleines bißchen ähnlich	Ist mir nicht ähnlich	Ist mir überhaupt nicht ähnlich
Es ist ihr wichtig, neue Ideen zu entwickeln und kreativ zu sein. Sie macht Sachen gerne auf ihre eigene originelle Art und Weise.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, reich zu sein. Sie möchte viel Geld haben und teure Sachen besitzen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie hält es für wichtig, dass alle Menschen auf der Welt gleich behandelt werden sollten. Sie glaubt, dass jeder Mensch im Leben gleiche Chancen haben sollte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, ihre Fähigkeiten zu zeigen. Sie möchte, dass die Leute bewundern, was sie tut.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, in einem sicheren Umfeld zu leben. Sie vermeidet alles, was ihre Sicherheit gefährden könnte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie mag Überraschungen und hält immer Ausschau nach neuen Aktivitäten. Sie denkt, dass im Leben Abwechslung wichtig ist.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie sucht das Abenteuer und geht gerne Risiken ein. Sie will ein aufregendes Leben haben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie glaubt, dass die Menschen tun sollten, was man ihnen sagt. Sie denkt, dass Menschen sich immer an Regeln halten sollten, selbst dann wenn es niemand sieht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, Menschen zuzuhören, die anders sind als sie. Auch wenn sie anderer Meinung ist als andere, will sie die anderen trotzdem verstehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, zurückhaltend und bescheiden zu sein. Sie versucht, die Aufmerksamkeit nicht auf sich zu lenken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



	Ist mir sehr ähnlich	Ist mir ähnlich	Ist mir etwas ähnlich	Ist mir nur ein kleines bisschen ähnlich	Ist mir nicht ähnlich	Ist mir überhaupt nicht ähnlich
Es ist ihr wichtig, Spaß zu haben. Sie gönnt sich selbst gerne etwas.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, selbst zu entscheiden, was sie tut. Sie ist gerne frei und unabhängig von anderen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr sehr wichtig, den Menschen um sie herum zu helfen. Sie will für deren Wohl sorgen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, sehr erfolgreich zu sein. Sie hofft, dass die Leute ihre Leistungen anerkennen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, dass der Staat ihre persönliche Sicherheit vor allen Bedrohungen gewährleistet. Sie will einen starken Staat, der seine Bürger verteidigt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, sich jederzeit korrekt zu verhalten. Sie vermeidet es, Dinge zu tun, die andere Leute für falsch halten könnten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, dass andere sie respektieren. Sie will, dass die Leute tun, was sie sagt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es ist ihr wichtig, ihren Freunden gegenüber loyal zu sein. Sie will sich für Menschen einsetzen, die ihr nahe stehen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie ist fest davon überzeugt, dass die Menschen sich um die Natur kümmern sollten. Umweltschutz ist ihr wichtig.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Tradition ist ihr wichtig. Sie versucht, sich an die Sitten und Gebräuche zu halten, die ihr von ihrer Religion oder ihrer Familie überliefert wurden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sie lässt keine Gelegenheit aus, Spaß zu haben. Es ist ihr wichtig, Dinge zu tun, die ihr Vergnügen bereiten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Inwieweit kannst du dich mit folgenden Aussagen identifizieren?

	stimme definitiv zu					stimme definitiv nicht zu
Wenn ich einen vollen Aschenbecher oder Mülleimer sehe, möchte ich ihn sofort geleert sehen.	<input type="radio"/>					
Ich bin ein häuslicher Mensch.	<input type="radio"/>					
Fernsehen ist meine bevorzugte Art der Unterhaltung.	<input type="radio"/>					
Egal wie schnell unser Einkommen steigt, wir scheinen uns nie weiter zu entwickeln.	<input type="radio"/>					
Ich versuche cholesterinhaltiges Essen zu vermeiden.	<input type="radio"/>					
Werbung beleidigt meine Intelligenz.	<input type="radio"/>					
In den Aktienmarkt zu investieren ist für die meisten Familien zu riskant.	<input type="radio"/>					
Um einen schlechten Atem zu vermeiden, sollte jeder Mensch Mundwasser benutzen.	<input type="radio"/>					
Fernsehwerbung bedient sich zu oft sexueller Reize.	<input type="radio"/>					
Um heutzutage erfolgreich zu sein ist ein Hochschulabschluss sehr wichtig.	<input type="radio"/>					
Meine Tage scheinen einem festen Ablauf zu folgen– z.B. esse ich jeden Tag zur gleichen Zeit.	<input type="radio"/>					
Ich mag es Orte zu bereisen, die ganz anders sind als mein Zuhause.	<input type="radio"/>					
Ich arbeite die meiste Zeit sehr hart.	<input type="radio"/>					
Ich mag es, wenn Menschen des anderen Geschlechts mich attraktiv finden.	<input type="radio"/>					
Ich werde wahrscheinlich nächstes Jahr mehr Geld ausgeben können als dieses Jahr.	<input type="radio"/>					
Ich esse mehr als ich sollte.	<input type="radio"/>					



Nimm bitte nun zu den folgenden menschlichen Verhaltensweisen Stellung. Was würdest Du sagen: *Inwieweit trifft der jeweilige Satz auf Dich persönlich zu?*

	trifft voll und ganz zu		teils/ teils			trifft überhaupt nicht zu	
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Mein erster Eindruck von Menschen stellt sich gewöhnlich als richtig heraus.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin mir oft unsicher in meinem Urteil.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich weiß immer genau, wieso ich etwas mag.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe schon mal zu viel Wechselgeld zurückbekommen und nichts gesagt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin immer ehrlich zu anderen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe gelegentlich mal jemanden ausgenutzt.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich sage immer, was ich denke.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin manchmal ärgerlich, wenn ich meinen Willen nicht bekomme.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich bin immer gewillt, einen Fehler, den ich mache zuzugeben.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich habe gelegentlich etwas mit Absicht gesagt, was die Gefühle anderer verletzen könnte.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



Wir bitten Dich, im Folgenden, die unten stehenden Satzanfänge zu ergänzen. Bitte überlege Dir möglichst zehn verschiedene Aussagen zu der Frage: "Wer bin ich?". Antworte, als würdest Du Dir selbst antworten und nicht jemand anderem. Schreibe die Antworten in der Reihenfolge auf, wie sie Dir einfallen. Gehe recht schnell vor.

Ich bin _____



SOZIODEMOGRAPHIE

Zum Schluss möchten wir Dich bitten, noch einige persönliche Angaben zu machen.

1. Wie alt bist Du? _____
 2. Du bist:
 - Männlich Weiblich
 3. Hast Du bereits eine Ausbildung absolviert?
 - Ja Nein
 4. Arbeitest Du in den Semesterferien?
 - Ja Nein
 5. Arbeitest Du während des Semesters?
 - Ja Nein
 6. Sind Praktika in Deinem Studium wichtig?
 - Ja Nein
 7. Wie viele Praktika hast Du bereits gemacht? _____
 8. Im wievielten Semester bist Du? _____
 9. Welchen Studiengang belegst Du? _____
 10. Wie zufrieden bist Du mit Deinem Studium?

außerordentlich zufrieden		teils/ teils		überhaupt nicht zufrieden
<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
 11. Welcher Nation fühlst Du Dich zugehörig? _____
 12. Was denkst Du, welchem Zweck diene diese Befragung?

-

VIELEN DANK