



**Universität Potsdam**

Humanwissenschaftliche Fakultät

Department Psychologie

Sozialpsychologie

Entwicklung und Validierung eines Inventars zur Erfassung positiver und negativer Attribute des Geschlechtsrollenselbstkonzepts

*Anja Berger*

*Dissertation zur Erlangung des akademischen Grades  
Doktor Philosophiae (Dr. phil.)*

Eingereicht  
bei der Humanwissenschaftlichen Fakultät  
der Universität Potsdam  
im Jahr 2010

Erstgutachterin: Prof. Dr. Barbara Krahe  
Zweitgutachterin: Prof. Dr. Birgit Elsner

Potsdam, September 2010

Dieses Werk ist unter einem Creative Commons Lizenzvertrag lizenziert:  
Namensnennung - Keine kommerzielle Nutzung - Weitergabe unter gleichen  
Bedingungen 3.0 Deutschland  
Um die Bedingungen der Lizenz einzusehen, folgen Sie bitte dem Hyperlink:  
<http://creativecommons.org/licenses/by-nc-sa/3.0/de/>

Online veröffentlicht auf dem  
Publikationsserver der Universität Potsdam:  
URL <http://opus.kobv.de/ubp/volltexte/2011/5121/>  
URN [urn:nbn:de:kobv:517-opus-51215](http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-51215)  
<http://nbn-resolving.org/urn:nbn:de:kobv:517-opus-51215>

# INHALTSVERZEICHNIS

<b>TABELLENVERZEICHNIS</b> .....	<b>IV</b>
<b>ABBILDUNGSVERZEICHNIS</b> .....	<b>V</b>
<b>ABKÜRZUNGEN</b> .....	<b>VI</b>
<b>ABSTRACT</b> .....	<b>VII</b>
<b>1 EINLEITUNG</b> .....	<b>1</b>
<b>2 THEORETISCHE GRUNDLAGEN</b> .....	<b>3</b>
<b>2.1 Das Selbst</b> .....	<b>3</b>
2.1.1 Konstruktion des Selbst (sozial, kulturell, kognitiv, Identität).....	4
2.1.2 Das Modell des dynamischen Selbst (multipel, flexibel, Korrektur).....	7
2.1.3 Die Bedeutung des Selbstwerts in der Selbstkonstruktion.....	7
<b>2.2 Das Geschlecht</b> .....	<b>8</b>
2.2.1 Definition von Geschlechtsrollen und Geschlechterstereotypen.....	9
2.2.2 Biologische Grundlagen von Geschlechtsrollen und Geschlechterstereotypen.....	13
2.2.3 Psychologische Theorien der Geschlechtsrollenstereotypisierung.....	14
2.2.3.1 Psychoanalytische Theorien (Freud, Erikson).....	15
2.2.3.2 Soziale Lerntheorien (Mischel, Bandura).....	16
2.2.3.3 Kognitive Entwicklungstheorien (Kohlberg, Kagan, Ullian).....	18
2.2.3.4 Geschlechterschematheorie / Geschlechtstypisierung (Bem).....	21
2.2.3.5 Soziale Identität.....	24
2.2.3.6 Rollentheorie des GSK.....	25
2.2.3.7 Psychologische Anpassung – Androgynität.....	28
2.2.4 Stabilität des Geschlechtsrollenselbstkonzepts.....	29
2.2.4.1 Individuelle Entwicklungsveränderungen über die Zeit.....	30
2.2.4.2 Flexibilität über verschiedene Situationen.....	33
2.2.5 Kulturelle Veränderungen.....	36
2.2.6 Korrelate des GSK auf der Ebene von Affekten, Kognitionen und Verhalten.....	41
<b>2.3 Empirische Erfassung des Geschlechtsrollenselbstkonzepts</b> .....	<b>45</b>
2.3.1 Übersicht der Entwicklung der Definition und Messinstrumente.....	47
2.3.1.1 Maskulinität und Femininität als bipolare Dimensionen.....	47
2.3.1.2 Maskulinität und Femininität als unabhängige Dimensionen.....	49
2.3.1.3 Kritik am BSRI.....	51
2.3.1.4 GSK als multidimensionales Konstrukt.....	57
2.3.1.5 Weitere Messinstrumente des GSK.....	59
2.3.2 Die Bedeutung negativer Aspekte der Geschlechtsrollenidentität.....	60
2.3.2.1 Selbstzuschreibungen: Theoretische Perspektiven.....	62
2.3.2.1.1 Die Perspektive der Stereotypenforschung: Stereotype Threat.....	63
2.3.2.1.2 Die Perspektive der Intergruppenforschung: Eigengruppenbevorzugung und Selbstkategorisierung.....	65
2.3.2.1.3 Valenz und Salienz sozialer Kategorien.....	67
2.3.2.1.4 Valenz als Charakteristik von Attributen.....	68
2.3.2.1.5 Valenz vs. Typizität.....	69
2.3.2.1.6 Das Konzept der sozialen Erwünschtheit.....	71
2.3.2.1.7 Soziale Erwünschtheit (negativer) Eigenschaften.....	74
2.3.2.2 Positive und negative Inhalte der Geschlechterstereotype.....	75
2.3.2.2.1 Valenz bei den Geschlechterstereotypen.....	76
2.3.2.2.2 Soziale Erwünschtheit (negativer) geschlechtsbezogener Eigenschaften.....	79
2.3.2.2.3 Zuschreibung negativer geschlechtsbezogener Eigenschaften.....	82
2.3.2.3 Berücksichtigung negativer Eigenschaften bei der Messung des GSK.....	84
2.3.2.3.1 Unmitigated Agency und Unmitigated Communion.....	85
2.3.2.3.2 Beziehungen positiver und negativer Teilskalen.....	87
2.3.2.3.3 Erlebens- und Verhaltenskorrelate von Unmitigated Agency und Communion.....	88
2.3.3 Korrelate positiver und negativer Maskulinität / Femininität.....	90
2.3.3.1 Maskulinität Positiv.....	92
2.3.3.2 Maskulinität Negativ.....	93
2.3.3.3 Femininität Positiv.....	95
2.3.3.4 Femininität Negativ.....	96
<b>2.4 GSK als Erwartungs-mal-Wert-Konstrukt</b> .....	<b>98</b>
2.4.1 Ursprung aus der Motivations- und Einstellungsforschung.....	98
2.4.2 Anwendung des Erwartungs-mal-Wert-Modells auf das GSK.....	99
2.4.3 Bedeutung des Produktes.....	100
<b>2.5 Zusammenfassung</b> .....	<b>101</b>

<b>2.6 Fragestellung und Hypothesen .....</b>	<b>102</b>
<b>2.7 Validierungskonstrukte und Hypothesen .....</b>	<b>104</b>
2.7.1 Hypothesen zu den Teilskalen des GSK .....	104
2.7.2 Hypothesen zu den Validierungskonstrukten .....	104
<b>2.8 Realisierung als Onlinestudien.....</b>	<b>105</b>
2.8.1 Datensammlung im Internet.....	105
2.8.2 Vergleich von Offline und Onlineforschung .....	105
2.8.3 Merkmale der webbasierten Forschung .....	107
2.8.4 Technische Realisation der Fragebögen .....	108
2.8.4.1 Design & Programmierung der Fragebögen .....	108
2.8.4.2 Bewerbung der Fragebögen .....	109
<b>3 EMPIRISCHE STUDIEN .....</b>	<b>111</b>
<b>3.1 Kritische Überprüfung des BSRI .....</b>	<b>112</b>
3.1.1 Studie 1: Einschätzung von Typizität und sozialer Erwünschtheit der BSRI-Items .....	112
3.1.1.1 Fragestellung .....	112
3.1.1.2 Methode .....	112
3.1.1.2.1 Stichprobe .....	112
3.1.1.2.2 Instrumente .....	114
3.1.1.2.3 Durchführung .....	114
3.1.1.3 Ergebnisse .....	114
3.1.1.3.1 Analyse der maskulinen Eigenschaften: soz. Erwünschtheit .....	115
3.1.1.3.2 Analyse der maskulinen Eigenschaften: Typizität.....	116
3.1.1.3.3 Analyse der femininen Eigenschaften: soz. Erwünschtheit .....	117
3.1.1.3.4 Analyse der femininen Eigenschaften: Typizität .....	118
3.1.1.3.5 Analyse der sozialen Erwünschtheit der Teilskalen .....	119
3.1.1.3.6 Analyse der Typizität der Teilskalen .....	121
3.1.1.4 Diskussion.....	122
3.1.2 Studie 2: Selbstbeschreibung mittels BSRI .....	124
3.1.2.1 Fragestellung .....	124
3.1.2.2 Methode .....	124
3.1.2.2.1 Stichprobe .....	124
3.1.2.2.2 Instrumente .....	125
3.1.2.2.3 Durchführung .....	125
3.1.2.3 Ergebnisse .....	125
3.1.2.3.1 Skalenskennwerte von Maskulinität und Femininität .....	125
3.1.2.3.2 Geschlechtsunterschiede bei Maskulinität und Femininität .....	127
3.1.2.3.3 Konfirmatorische Faktorenanalyse.....	128
3.1.2.4 Diskussion.....	130
<b>3.2 Neukonstruktion eines Instruments zur Erfassung positiver und negativer Aspekte des Geschlechtsrollenelbstkonzepts .....</b>	<b>131</b>
3.2.1 Studie 3: Konstruktionsstudie – Generierung neuer GSK-Items .....	131
3.2.1.1 Fragestellung .....	131
3.2.1.2 Methode .....	132
3.2.1.2.1 Stichprobe .....	132
3.2.1.2.2 Instrumente .....	133
3.2.1.3 Ergebnisse .....	134
3.2.1.4 Diskussion.....	136
3.2.2 Studie 4: Itemauswahl auf der Basis von Typizität, sozialer Erwünschtheit und differentieller Selbstzuschreibung von Männern und Frauen .....	138
3.2.2.1 Fragestellung .....	138
3.2.2.2 Methode .....	140
3.2.2.2.1 Stichprobe .....	140
3.2.2.2.2 Instrumente .....	142
3.2.2.2.3 Durchführung .....	142
3.2.2.3 Ergebnisse .....	143
3.2.2.3.1 Ergebnisse der Einschätzungen zur Typizität der neuen Items.....	143
3.2.2.3.2 Ergebnisse der Einschätzungen zur sozialen Erwünschtheit der neuen Items .....	149
3.2.2.3.3 Ergebnisse der Selbsteinschätzungen der neuen Items.....	152
3.2.2.3.4 Auswahl der Eigenschaften für die Endversion der Skala .....	157
3.2.2.4 Zusammenfassende Diskussion.....	158
3.2.3 Studie 5: Validierung der neuen Teilskalen .....	160
3.2.3.1 Fragestellung .....	160
3.2.3.2 Methode .....	160

3.2.3.2.1 Stichprobe .....	160
3.2.3.2.2 Instrumente .....	161
3.2.3.2.3 Durchführung .....	164
3.2.3.3 Ergebnisse .....	165
3.2.3.3.1 Explorative Faktorenanalyse der neuen Teilskalen .....	165
3.2.3.3.2 Konfirmatorische Faktorenanalyse der neuen Teilskalen .....	167
3.2.3.3.3 Statistische Kennwerte der Teilskalen .....	171
3.2.3.3.4 Re-Test-Reliabilität.....	172
3.2.3.3.5 Korrelative Zusammenhänge der Teilskalen.....	174
3.2.3.3.6 Statistische Kennwerte der Validierungskonstrukte .....	174
3.2.3.3.7 Korrelationen der Validierungskonstrukte mit den Teilskalen .....	175
3.2.3.3.8 Konstruktvaliditätsanalysen mittels hierarchischer Regressionen .....	176
3.2.3.3.9 Konvergente Validität .....	180
3.2.3.4 Diskussion.....	182
3.2.4 Studie 6: Replikation der neuen Teilskalen .....	183
3.2.4.1 Fragestellung .....	183
3.2.4.2 Methode .....	183
3.2.4.2.1 Stichprobe .....	183
3.2.4.2.2 Instrumente .....	184
3.2.4.2.3 Durchführung .....	185
3.2.4.3 Ergebnisse .....	185
3.2.4.3.1 Statistische Kennwerte der Teilskalen .....	185
3.2.4.3.2 Konfirmatorische Faktorenanalyse der neuen Teilskalen .....	186
3.2.4.3.3 Korrelative Zusammenhänge der Teilskalen.....	187
3.2.4.3.4 Statistische Kennwerte und Korrelationen der Validierungskonstrukte .....	187
3.2.4.3.5 Konstruktvaliditätsanalysen mittels hierarchischer Regressionen .....	188
3.2.4.4 Diskussion.....	190
3.2.5 Studie 7: Erweiterung der neuen Skalen um eine Erwartungs-mal-Wert-Komponente.....	191
3.2.5.1 Fragestellung .....	191
3.2.5.2 Methode .....	192
3.2.5.3 Ergebnisse .....	193
3.2.5.3.1 Korrelative Zusammenhänge der Teilskalen und der Wertekomponenten .....	193
3.2.5.3.2 Validitätsanalysen der individuellen Wertkomponente - Wichtigkeit .....	193
3.2.5.3.3 Gruppenvergleiche mittels Mediansplit zum Konzept Selbst * Wichtigkeit .....	196
3.2.5.4 Diskussion.....	199
<b>4 DISKUSSION .....</b>	<b>200</b>
<b>4.1 Notwendigkeit eines neuen Instruments: Die Grenzen des BSRI .....</b>	<b>202</b>
<b>4.2 Entwicklung und Validierung des PoNeGRI - Befunde zur Neukonstruktion .....</b>	<b>206</b>
4.2.1 Identifizierung positiver und negativer Attribute im GSK .....	206
4.2.2 Struktur des neuen Instruments.....	208
4.2.3 Konstruktvalidierung .....	213
<b>4.3 Generalisierbarkeit.....</b>	<b>216</b>
4.3.1 Onlineerhebung .....	216
4.3.2 Stichprobenwahl .....	218
<b>4.4 Einschränkungen, methodenkritische Aspekte und Ausblick .....</b>	<b>220</b>
<b>5 BIBLIOGRAFIE .....</b>	<b>VIII</b>

## TABELLENVERZEICHNIS

Tabelle 1. Übersicht bedeutsamer Theorien bezüglich der sozialen Geschlechtsentwicklung.....	14
Tabelle 2. Zusammenfassung der berichteten Geschlechtsunterschiede (in Auszügen).....	43
Tabelle 3. Bezeichnungen von Maskulinität und Femininität in verschiedenen Studien.....	52
Tabelle 4. Einsatz des BSRI und weiterer Instrumente in verschiedenen Nationen.....	56
Tabelle 5. Kategorien der Eigenschaften, die sich in ihrer sozialen Erwünschtheit für Männer und Frauen unterscheiden (entnommen aus: Prentice & Carranza, 2002).....	80
Tabelle 6. Zusammenhang von GSK-Aspekten mit Verhaltensweisen.....	90
Tabelle 7. Quellen und Zuordnung der Validierungskonstrukte.....	92
Tabelle 8. Validierungskonstrukte der vier Teilskalen der Geschlechtsrollen.....	105
Tabelle 9. Übersicht aller durchgeführten Studien.....	111
Tabelle 10. Demografie der Probanden aus Studie 1 zur Überprüfung des BSRI.....	113
Tabelle 11. Statistische Kennwerte der Einschätzung soz. Erwünschtheit mas. Items (BSRI).....	115
Tabelle 12. Statistische Kennwerte der Einschätzung der Typizität mas. Items (BSRI).....	116
Tabelle 13. Statistische Kennwerte der Einschätzung soz. Erwünschtheit fem. Items (BSRI).....	118
Tabelle 14. Statistische Kennwerte der Einschätzung der Typizität fem. Items (BSRI).....	118
Tabelle 15. Mittelwerte der Einschätzung der soz. Erwünschtheit der BSRI-Teilskalen FEM und MAS.....	119
Tabelle 16. Mittelwerte der Einschätzung der soz. Erwünschtheit der BSRI-Skalen FEM und MAS.....	120
Tabelle 17. Mittelwerte der Einschätzung der Typizität der BSRI-Teilskalen FEM und MAS.....	121
Tabelle 18. Mittelwerte der Einschätzung der Typizität der BSRI-Skalen FEM und MAS.....	122
Tabelle 19. Statistische Kennwerte der Selbsteinschätzung anhand mas. BSRI-Items.....	125
Tabelle 20. Statistische Kennwerte der Selbsteinschätzung anhand fem. BSRI-Items.....	126
Tabelle 21. Reliabilitäten und Mittelwerte von MAS und FEM (BSRI).....	127
Tabelle 22. Ausgewählte Eigenschaften nach Generierung pro Teilskala.....	135
Tabelle 23. Übersicht der Demografie zur Überprüfung der Typizität (3 Teilstichproben).....	141
Tabelle 24. Übersicht der Demografie zur Selbstbeschreibung (3 Teilstichproben).....	142
Tabelle 25. Itemstatistiken bzgl. der Typizität der neuen positiven maskulinen Teilskala.....	144
Tabelle 26. Itemstatistiken bzgl. der Typizität der neuen negativen maskulinen Teilskala.....	145
Tabelle 27. Itemstatistiken bzgl. der Typizität der neuen positiven femininen Teilskala.....	147
Tabelle 28. Itemstatistiken bzgl. der Typizität der neuen negativen femininen Teilskala.....	148
Tabelle 29. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen positiven maskulinen Teilskala.....	149
Tabelle 30. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen negativen maskulinen Teilskala.....	150
Tabelle 31. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen positiven femininen Teilskala.....	151
Tabelle 32. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen negativen femininen Teilskala.....	151
Tabelle 33. Itemstatistiken bzgl. der Selbstbeschreibung der neuen positiven maskulinen Teilskala.....	153
Tabelle 34. Itemstatistiken bzgl. der Selbstbeschreibung der neuen negativen maskulinen Teilskala.....	154
Tabelle 35. Itemstatistiken bzgl. der Selbstbeschreibung der neuen positiven femininen Teilskala.....	156
Tabelle 36. Itemstatistiken bzgl. der Selbstbeschreibung der neuen negativen femininen Teilskala.....	157
Tabelle 37. Auszuschließende Items.....	158
Tabelle 38. Auswahl der Items für die vier Teilskalen.....	158
Tabelle 39. Zuordnung und Vorhersagen der Validierungskonstrukte (Studie 5).....	160
Tabelle 40. Darstellung der Erfassung der Geschlechtsrollenitems.....	161
Tabelle 41. Übersicht der eingesetzten Validierungskonstrukte.....	165
Tabelle 42. Matrix der Faktorladungen der Eigenschaften (Explorative Faktorenanalyse, PC).....	166
Tabelle 43. Matrix der Faktorladungen der Eigenschaften (Explorative Faktorenanalyse, PAF).....	166
Tabelle 44. Modellprüfung mit den Items der neuen Skala.....	168
Tabelle 45. Skalenskennwerte aller vier neuen Teilskalen in der Selbstbeschreibung.....	171
Tabelle 46. Skalenskennwerte der vier Teilskalen in der Selbstbeschreibung zu T2.....	172
Tabelle 47. Korrelationen der Teilskalen zu T1 und T2.....	173
Tabelle 48. Korrelationen der Teilskalen.....	174
Tabelle 49. Statistische Kennwerte der Validierungskonstrukte für die vier Teilskalen.....	174
Tabelle 50. Korrelationen der Teilskalen mit den Validierungskonstrukten.....	175
Tabelle 51. Vorhersage von AVEM-BE (berufl. Ehrgeiz) durch Teilskalen (Mas+).....	176
Tabelle 52. Vorhersage von AMS - Hoffen auf Erfolg durch Teilskalen (Mas+).....	177
Tabelle 53. Vorhersage von Aggressivität (AQ) durch Teilskalen (Mas-).....	178
Tabelle 54. Vorhersage von Werteskala Macht durch Teilskalen (Mas-).....	178
Tabelle 55. Vorhersage von Empathie (Sorge) durch Teilskalen (Fem+).....	179

Tabelle 56. Vorhersage von sozialen Werten durch Teilskalen (Fem+) .....	179
Tabelle 57. Vorhersage von ESWS durch Teilskalen (Fem-) .....	180
Tabelle 58. Vorhersage von Ängstlicher Depressivität durch Teilskalen (Fem-) .....	180
Tabelle 59. Deskriptive Statistik und Korrelationen der BSRI-Skalen mit den GSK-Teilskalen .....	181
Tabelle 60. Darstellung der Erfassung der Geschlechtsrollenitems (Studie 6).....	184
Tabelle 61. Korrelationstabelle der Validierungskonstrukte für Studie 6 .....	185
Tabelle 62. Skalenskennwerte der 4 Teilskalen in der Selbstbeschreibung.....	185
Tabelle 63. Modellprüfung mit den Items der neuen Skala (Studie 6) .....	186
Tabelle 64. Korrelationen der Teilskalen .....	187
Tabelle 65. Statistische Kennwerte der Validierungskonstrukte für Studie 6.....	188
Tabelle 66. Korrelationen der Teilskalen mit den Validierungskonstrukten .....	188
Tabelle 67. Vorhersage von Leistungsorientierung durch Teilskalen (Mas+).....	189
Tabelle 68. Vorhersage von Aggressivität durch Teilskalen (Mas-).....	189
Tabelle 69. Vorhersage von Fürsorglichkeit durch Teilskalen (Fem+).....	190
Tabelle 70. Vorhersage von STAI durch Teilskalen (Fem-).....	190
Tabelle 71. Darstellung der Erfassung der Geschlechtsrollenitems .....	192
Tabelle 72. Korrelationen der Teilskalen und der individuellen Wichtigkeitskomponente .....	193
Tabelle 73. Korrelationen der gemittelten Produkte der Teilskalen mit den Validierungskonstrukten .....	193
Tabelle 74. Vorhersage von berufl. Ehrgeiz durch individuelle Wertkomponente (Mas+) .....	194
Tabelle 75. Vorhersage der Validierungskonstrukte durch Selbst- und Wichtigkeitszuschreibung .....	195
Tabelle 76. Mediane aus Selbst- und Wichtigkeitszuschreibung (N = 800).....	196

## ABBILDUNGSVERZEICHNIS

Abbildung 1. Kohlbergs (1966) Modell des Zusammenhangs von Geschlechtsrollenidentität, -attributen, und -stereotypen .....	19
Abbildung 2. Kagans (1964) Modell des Zusammenhangs von Geschlechtsrollenidentität, -attributen, und -stereotypen .....	20
Abbildung 3. Das Modell der sozialen Lerntheorie über den Zusammenhang von Geschlechtsrollenidentität, -attributen, und -stereotypen .....	20
Abbildung 4. Übersicht des Zusammenhangs von biologischem und psychologischem Geschlecht.....	21
Abbildung 5. Darstellung wichtiger Komponenten bei der Erfassung des GSK.....	46
Abbildung 6. Maskulinität und Femininität als bipolare Dimensionen .....	48
Abbildung 7. Maskulinität und Femininität als zwei unabhängige Dimensionen.....	49
Abbildung 8. Formel des Zusammenhangs von Einstellung, Bewertung und Erwartung nach Ajzen und Fishbein (Ajzen, 1985; Fishbein, 1967) .....	99
Abbildung 9. Mögliche Formel der Geschlechtsrolle als Funktion der eigenen Selbstzuschreibung in Abhängigkeit der persönlichen Bedeutsamkeit dieser Zuschreibung .....	99
Abbildung 10. Darstellung der Veränderung der sozialen Erwünschtheit über die Zeit.....	121
Abbildung 11. Selbstzuschreibung der BSRI-Eigenschaften .....	127
Abbildung 12. Konfirmatorische Faktorenanalyse des BSRI aus der Selbstbeschreibung.....	129
Abbildung 13. Online-Seite zur Generierung von Eigenschaften.....	133
Abbildung 14. Beziehungen ausgewählter Eigenschaften zueinander .....	135
Abbildung 15. Modelle zur Überprüfung der Beziehungen der Teilskalen zueinander .....	167
Abbildung 16. Faktorladungen aller Items der vier Teilskalen im Vier-Faktoren-Modell.....	170
Abbildung 17. Dreifachinteraktion von Valenz, Geschlechtstypizität und Geschlecht .....	172
Abbildung 18. Mittelwerte der Teilskalen zu T1 und T2 im Vergleich .....	173
Abbildung 19. Verteilung der Probanden in % in Abhängigkeit der Selbst- und Wichtigkeitseinschätzung .....	196
Abbildung 20. Vergleich der gemittelten 4-Item-Parameter aller vier Teilskalen .....	198

## ABKÜRZUNGEN

GSK	Geschlechtsrollenselbstkonzept
BSRI	Bem Sex Role Inventory
PAQ	Personal Attribute Questionnaire
EPAQ	Extended Personal Attribute Questionnaire
WWW	World Wide Web
Fem	Femininität
Mas	Maskulinität
M+ / Mas+	Maskulinität - positiv
M- / Mas-	Maskulinität - negativ
F+ / Fem+	Femininität - positiv
F- / Fem-	Femininität - negativ
AG	Agency
CM	Communion
UA	Unmitigated Agency
UC	Unmitigated Communion
FPI	Freiburger Persönlichkeitsinventar
PRF	Personality Research Form
SIT	Soziale Identitätstheorie
Vpn	Versuchsperson
GD	Gender Diagnosticity



## ABSTRACT

Das Geschlechtsrollen-Selbstkonzept, das sich im Laufe der Sozialisation in Auseinandersetzung mit den vorherrschenden Vorstellungen der umgebenden Kultur entwickelt, steht in Beziehung zu Affekten, Kognitionen und Verhaltensweisen in einer Vielzahl von Bereichen.

Bisherige GSK-Instrumente messen jedoch nahezu ausschließlich den positiven Aspekt von Maskulinität und Femininität. Die Definition des allgemeinen Selbstkonzepts gibt diese Limitierung auf positive Valenz nicht vor, und aus gesundheitspsychologischer Sicht sowie der Gruppenforschung ist die Bedeutung negativer Eigenschaften zur Selbstbeschreibung bekannt.

Vor diesem Hintergrund wurden sieben aufeinander aufbauende Studien durchgeführt mit dem Ziel ein neues Instrument zu entwickeln, deren Items zum einen kulturell aktuellen Eigenschaften zur Selbstbeschreibung entsprechen und zum anderen die Valenzunterschiede dieser Merkmalsbeschreibungen berücksichtigen.

Nach einer kritischen empirischen Überprüfung des deutschen BSRI, um Schwächen der Items ausschließlich positiver Valenz aufzudecken, wurde eine neue Skala entwickelt, die von Beginn an auch negative Selbstbeschreibungen berücksichtigte um der Komplexität des geschlechtlichen Selbst gerecht zu werden. Aufgrund der Einschätzungen zur Typizität und sozialen Erwünschtheit sowie mit ersten Resultaten aus der Selbstbeschreibung wurde die Auswahl der Items für die Teilskalen vorgenommen. In zwei weiteren Studien wurden schließlich die vier neu entwickelten Teilskalen des neuen GSK-Inventars einer Validierung unterzogen. Jeder der Teilskalen wurden theoriegeleitet spezifische Konstrukte zugeordnet und es konnte nachgewiesen werden, dass alle Teilskalen ihren eigenen Beitrag zur Vorhersage psychologischer Konstrukte leisten können. So standen beispielsweise die negativen maskulinen Eigenschaften in engerer Beziehung zu Aggressivität und machtbezogenen Werten als die positiven Aspekte der Maskulinität.

Als Ergebnis dieser Entwicklung stehen am Ende vier kurze, unabhängige, reliable Teilskalen, die positive als auch negative Aspekte von Maskulinität und Femininität abbilden und mittels sehr unterschiedlicher psychologischer Erlebens- und Verhaltenskonstrukte validiert wurden, die die Unabhängigkeit der Skalen belegen und diese für einen Einsatz in der Forschung empfehlen.

Die Einführung einer individuellen Wertkomponente im Zuge der Selbstbeschreibung, angelehnt an das bekannte Erwartungs-mal-Wert Modell der Motivations- und Einstellungsforschung, und die daraus mögliche multiplikative Verknüpfung von Selbsteinschätzung und persönlicher Wichtigkeit der Eigenschaften konnten den Aufklärungswert in Bezug auf unterschiedliche Validierungskonstrukte dagegen nicht verbessern und wurden daher nicht ins das Instrument integriert.



# 1 EINLEITUNG

In allen Gesellschaften erfolgt die Klassifikation von Frauen und Männern nicht ausschließlich aufgrund biologischer Unterschiede, sondern über kulturell und sozial geteilte Definitionen beider Geschlechtergruppen. Im Verlauf der individuellen Sozialisation werden Männer und Frauen mit diesen in der jeweiligen Gesellschaft geltenden Geschlechterstereotype und Geschlechtsrollenerwartungen beider Gruppen (Maskulinität und Femininität) konfrontiert und entwickeln ihre eigene Geschlechtsidentität (Alfermann, 1996), die kognitive, emotionale und verhaltensbasierte Komponenten enthält. Stereotype sind Überzeugungssysteme, die Verallgemeinerungen über die Charakteristiken und Rollenerwartungen einer Gruppe von Menschen, wie z. B. eine religiöse oder berufliche Gruppe (oder in diesem Bezug die Gruppe der Männer und Frauen) beinhalten (Archer & Lloyd, 2002). Sie sind ein notwendiges Instrument im Zuge der Organisation und Gestaltung des täglichen Lebens (Ashmore & del Boca, 1979).

Obwohl diese soziale Charakterisierung mit den biologischen Unterschieden von Männern und Frauen assoziiert ist, kann die eigene Zuschreibung dessen, was sozial als typisch männlich und weiblich gilt, unabhängig vom biologischen Geschlecht entwickelt werden. Diese individuelle Ausprägung von Maskulinität und Femininität wird zum Teil des eigenen geschlechtsbezogenen Selbstkonzepts, dahingehend, wie stark man sich mit den gesellschaftlich geteilten Charakteristiken beider Geschlechter identifiziert.

Das Wissen bezüglich der Geschlechterstereotype sowie die Anerkennung der eigenen Rolle als Mann oder Frau lassen an der Bedeutung des Geschlechts für die eigene Identität und das Selbstkonzept keinen Zweifel (Athenstaedt, Haas, & Schwab, 2004; Eagly, Wood, & Diekmann, 2000). Während die Vielfalt von Definitionen und Forschungsansätzen des psychologischen Geschlechts die Komplexität des Konstrukts offenbart, besteht weitgehende Übereinstimmung hinsichtlich der Repräsentation von Geschlechtsrollen, -normen und -kategorien im Selbst sowie im sozialen Leben: Individuen ziehen Geschlechterstereotype zur Selbstbeschreibung heran, die Erwartungshaltungen an Männer und Frauen seitens der Gesellschaft widerspiegeln (vgl. Ashmore, 1990; Athenstaedt, 2003; Bem, 1974; Cross & Madson, 1997; Deaux & LaFrance, 1998; Eagly & Steffen, 2000; Hannover, 2000; Spence & Buckner, 1995; 2000; Spence, Helmreich, & Stapp, 1974).

Zur empirischen Erfassung des Geschlechtsrollenelbstkonzepts (GSK) liegen verschiedene Instrumente vor, von denen die meisten im deutschsprachigen Raum nicht nur auf unterschiedlichen theoretischen Annahmen basieren, sondern auch auf verschiedenen Kriterien der Testkonstruktion (bezüglich Art der inhaltlichen Itemauswahl, Wahl der Gütekriterien, etc.). Eines ist jedoch allen gemein: konzeptionell messen sie nahezu ausschließlich den positiven Aspekt von Maskulinität und Femininität.

Die Definition des Selbstkonzepts gibt jedoch diese Limitierung auf positive Valenz nicht vor, wie beispielsweise von Woodhill & Samuel (2003) berechtigterweise angemerkt wird. Darüber hinaus ist aus der gesundheitspsychologischen Literatur bekannt, dass vor allem die negative Ausprägung von femininen Eigenschaften einen besonderen Beitrag zur Aufklärung

des Zusammenhangs der Geschlechtsrollen zu psychischen Krankheitsbildern (Helgeson, 1994a; Helgeson & Fritz, 1999) leisten kann. Die Gruppen- und Diskriminierungsforschung ihrerseits konnte vielfach belegen, dass negative Eigenschaften von der Fremd- als auch der Eigengruppe nicht nur kognitiv verfügbar sind, sondern auch zur Beschreibung aktiv herangezogen werden (Biernat, Vescio, & Green, 1996; Pickett, Bonner, & Coleman, 2002). Eine Person ist demnach durchaus in der Lage sich auch auf negativen Attributen zu positionieren. Nicht nur, dass es theoretisch nicht haltbar ist, die Selbstzuschreibung von Eigenschaften negativer Valenz per se auszuschließen, kann man zudem ausgehend von Ergebnissen früherer Studien die Existenz dieser negativen Selbstzuschreibungen nicht ignorieren. So wurde bei bisherigen Instrumenten wiederholt festgestellt, dass sich die Eigenschaften zur Bestimmung von Maskulinität und Femininität – obwohl theoretisch nicht so intendiert – in ihrer Valenz doch derart unterscheiden; dass einige dieser als positiv geltenden Eigenschaften als faktisch negative Attribute klassifiziert werden müssen.

Zusätzlich zu bestehenden inhaltlichen und theoretischen Unklarheiten ist der Aspekt der Messung selbst Fokus anhaltender Kritik. Die seit mitunter mehreren Jahrzehnten bestehenden Instrumente zur Messung des Geschlechtsrollenselbstkonzepts, welches als abhängig von der vorherrschenden Gesellschaft und Kultur gilt, müssen demnach den zeitlichen Veränderungen nach ihrer Entwicklung standhalten. Nicht bei allen Instrumenten ist jedoch geprüft, inwieweit dies auch heute noch gegeben ist.

Vor diesem Hintergrund verfolgt die vorliegende Arbeit folgende Zielsetzungen:

- (1) eine kritische Reflektion und Überprüfung der Aktualität bestehender Instrumente, insbesondere im Hinblick auf die Geschlechtstypizität der verwendeten Attribute.
- (2) In einem zweiten Schritt soll – aufbauend auf den Überlegungen des allgemeinen GSK – ein neues Instrument entwickelt und validiert werden, das zum einen den kulturell aktuellen Eigenschaften zur Selbstbeschreibung entspricht und zum anderen die Valenzunterschiede von Eigenschaften berücksichtigt.

Als Ergebnis der Konstruktion, die umfangreiche Item-, Skalen- und Modelanalysen mit einschließt, soll am Ende ein Inventar zur Erfassung des GSK mit vier Teilskalen stehen, die jeweils positive als auch negative Aspekte von Maskulinität und Femininität abbilden.

## 2 THEORETISCHE GRUNDLAGEN

Im folgenden Teil werden die Theorien, Instrumente und Forschungsergebnisse aus dem Bereich des Geschlechtsrollenselbstkonzepts vorgestellt. Da das GSK ein wichtiger Teil des Selbstkonzepts ist, werden in einem ersten Abschnitt zunächst die wesentlichen Aspekte der Selbstkonzeptforschung dargestellt. Im darauf folgenden Abschnitt wird dann der Bezug zum geschlechtsbezogenen Teil der Selbstkonstruktion hergestellt, der Gegenstand der vorliegenden Arbeit ist. In diesem Rahmen erfolgt die umfassende Betrachtung des Geschlechts als soziale Kategorie sowie die Darbietung der Inhalte und Theorien zur Entwicklung des Geschlechtsrollenselbstkonzepts. Zusätzlich wird der Bezug zur kulturellen und historischen Stabilität dieses Konzepts hergestellt.

In einem dritten Abschnitt werden schließlich Aspekte zur Messung des GSK dargestellt und diskutiert. Als wichtigster Bestandteil dieser Arbeit wird in diesem Überblick zunächst die Entwicklung der gängigsten Messinstrumente nachgezeichnet. Dabei wird insbesondere die Messung und die bis dato versäumte Erfassung auch negativer Aspekte des GSK, die aufgrund von Theorien der Gruppe und des Selbst nachweislich Teil des geschlechtsbezogenen Selbstkonzepts sind, diskutiert sowie die Notwendigkeit einer Überarbeitung und Neukonstruktion eines deutschen Instruments zur Messung des GSK herausgearbeitet. Dieses soll als ein ökonomisches und valides Instrument sowohl positive als auch negative Komponenten des GSK abbilden. In einem ausführlichen Validierungsprozess soll die Erweiterung eines solchen Instruments um die negativen Aspekte schließlich bestätigt werden.

### 2.1 Das Selbst

Ziel dieses Abschnitts ist die Darstellung der für die vorliegende Fragestellung relevanten Ausschnitte der Selbstforschung sowie der dazugehörigen Theorien und Ergebnisse. Dabei wird zunächst diskutiert, wie das Selbst definiert ist, auf welche Weise das Selbst als multiple und flexible kognitive Struktur beschrieben werden kann und inwieweit kulturelle Unterschiede in der Selbstkonstruktion existieren. In einem weiteren Schritt wird schließlich der spezifische Aspekt des geschlechtlichen Selbstkonzepts und dessen Bedeutung für das Individuum vor dem Hintergrund des globalen Selbst betrachtet.

Die Voraussetzung für das Selbst, seine Hülle, ist der physische Körper. Auch kulturübergreifend markiert die Erkenntnis des eigenen Körpers den Beginn der Wahrnehmung des Selbst (Baumeister, 1999). In vielen Sprachen spiegelt sich wider, dass das Selbst sowohl erkennendes Subjekt („I“) als auch betrachtetes Objekt sein kann („Me“). In der psychologischen Forschung wird das Selbst als eine Gedächtnisstruktur in Form eines assoziativen Netzwerkes definiert. Dabei umfasst diese kognitive Struktur das gesamte Wissen eines Individuums über sich selbst, kontinuierlich abgespeichert über die Spanne eines Lebens (Hannover, 1997; Linville & Carlston, 1994; Markus, 1977). Diese Wissens- und Erfahrungselemente des Selbst beeinflussen maßgeblich das Erleben und Verhalten einer Person in Abhängigkeit ihrer Konstruktion und Aktivierung.

### 2.1.1 Konstruktion des Selbst (sozial, kulturell, kognitiv, Identität)

Das psychologische Selbst und die eigene Identität werden im Laufe der Sozialisation definiert und geformt. Dieser Prozess ist dabei nicht durch einseitigen Einfluss der Sozialisationsagenten gekennzeichnet, da die fortschreitende individuelle Entwicklung ihrerseits wiederum soziale Interaktionen beeinflusst und lenken kann. Dieser Wirkungskreis, der dem Selbst eine mediatorische Rolle zuschreibt (Simon & Trötschel, 2007), führt im Ergebnis zu einer kognitiven Gedächtnisrepräsentation, welche die Gesamtheit der aus eigenen Erfahrungen mit sich und in Beziehung zu anderen Menschen abgeleiteten Verallgemeinerungen über das Selbst darstellt – dem Selbstkonzept oder auch Selbstschema (Hannover, 1997; Linville & Carlston, 1994; Markus, 1977). Dieses Selbstkonzept organisiert die Wahrnehmung und Verarbeitung zukünftiger Erfahrungen und steuert das Verhalten.

Unterschiedliche soziale Situationen, deren Vergleich und Interpretation sowie das daraus abgeleitete Verhalten führen zu individuell verschiedenen Selbstkonzepten – in Inhalt und Struktur. Hannover (1997) betrachtete in diesem Zusammenhang das Selbstkonzept auch nicht als einheitliche Struktur, sondern vielmehr als einen Komplex vielfältiger Teilaspekte des Selbst (sog. Selbstkonstrukte). Diese Selbstkonstrukte beinhalten Wissen über verschiedene Dimensionen einer Person, die insgesamt betrachtet bisweilen sogar widersprüchlich zueinander sein können.

Zu Beginn der Selbstforschung bestand Einigung darüber, das Selbst als eine Ganzheit zu betrachten, die individuell verschieden und unabhängig von sozialen Einflüssen ist. Im Verlauf der letzten Jahre rückten der Einfluss von anderen Personen oder Gruppen zunehmend in den Fokus der Selbstkonstruktion, da Individuen – wenn auch inter- sowie intraindividuell verschieden – erwiesenermaßen ihre Beziehung in ihrem sozialen Umfeld zur Selbstbeschreibung heranziehen (Brewer & Gardner, 1996; Markus & Kitayama, 1991; Singelis, 1994). Diese Beziehungen zu anderen repräsentieren den Teil ihres sozialen Selbst, was wiederum als nur ein Aspekt des globalen Selbst angesehen wird. Die Sozialisation nimmt in der Entwicklung des Selbstkonzepts eine entscheidende Rolle ein, da die in diesem Verlauf an ein Individuum gestellten Entwicklungsanforderungen von der Zugehörigkeit zu sozialen Gruppen und der kulturell geteilten Normen und Werte beeinflusst werden. In diesem Zusammenhang lassen sich Unterschiede in individualistischen und kollektivistischen Kulturen (Hofstede, 1980) dahingehend finden, dass beispielsweise kollektivistisch sozialisierte Individuen in viel stärkerem Maße von anderen Personen abhängen als vergleichsweise Mitglieder individualistischer Kulturen (Greenfield, Keller, Fuligni, & Maynard, 2003). Das kollektivistische Verhalten und die Verbundenheit zu anderen Individuen beschränken sich jedoch vor allem auf Interaktionen mit Vertretern der eigenen Gruppe (*ingroup*), während zur Fremdgruppe klar differenziert wird (Triandis, 1995). Triandis konnte in einer Vielzahl von Studien den Einfluss der Zugehörigkeit zu einer Gruppe auf Loyalität, Zielorientierung oder Verhaltensausrichtung zeigen (vgl. Triandis & Gelfand, 1998) und differenzierte das Selbst in drei Teilaspekte: das private, das öffentliche und das kollektive Selbst (Triandis, 1989). Während das öffentliche und das kollektive Selbst jeweils Informationen anderer über das Selbst oder über das Selbst in

Gemeinschaft enthalten und dabei auch im stärkeren Maße als kulturabhängig gelten, integriert das private Selbst Kognitionen über Verhalten und Eigenschaften der eigenen Person (Markus & Kitayama, 1991; Triandis, 1989).

Die Individualität einer jeden Person wird durch eine Vielzahl an Charakteristiken und Verhaltenseigenschaften definiert, die kontinuierlich adaptiert und in elaborierte Wissensstrukturen des Selbst neu organisiert werden. Mithin ist allgemein akzeptiert, dass die kognitiven Assoziationen des Selbstkonstrukts durch verschiedene Mechanismen aufgebaut werden, allen voran durch persönliche Erfahrungen, soziale Lernprozesse und mediale Einflüsse. Darüber hinaus besteht ebenfalls Einigung darüber, dass durch wiederholende Erfahrungen, Verstärkungsprozesse und unbewusste kognitive Verarbeitungsprozesse diese Selbstassoziationen verstärkt werden können, während ein Mangel an Lernerfahrungen zur Schwächung eben dieser Assoziationsverknüpfungen führen kann. Aufgrund der persönlichen Relevanz eigener personenbezogener Erlebnisse und Informationen sowie damit einhergehender wiederholter Aktivierung und Introspektion gelten Gedächtnisverknüpfungen das Selbst betreffend als allgemein stärker (für eine Zusammenfassung vgl. Perkins & Forehand, 2006).

Eine Vielzahl von Prozessen kann im Verlauf zur Individualisierung des Selbst beitragen: So werden zur Bewertung von Situationen sowohl internalisierte Standards als auch soziale Normen herangezogen, das Verhalten wird an Zielen ausgerichtet, überwacht sowie anschließend einer Bewertung unterzogen, die in der Entwicklung des Selbstwertgefühls resultiert. Dieses Selbstwertgefühl (= eine bewertete Einstellung gegenüber der eigenen Person in Bezug auf vielfältige Merkmale, Normen, Verhalten, Fähigkeiten, etc.) soll dabei durch verschiedene Strategien auf ein positives Niveau gebracht und dort gehalten werden. Dabei kommen wiederum sowohl individuelle Strategien (z. B. selbstwertdienliche Attributionen) als auch sozial (abwärts) ausgerichtete Vergleiche zum Einsatz (Simon & Trötschel, 2007). Die Herausforderung besteht darin, das individuelle Selbst und die soziale Welt klar voneinander zu trennen. Beide beeinflussen sich gegenseitig und gelten als miteinander verflochten. In der theoretischen Definition wurden vor allem Tajfel (1972) und Turner (1981) dieser zwei zentralen Ebenen des Selbst gerecht und integrierten diese in ihren Überlegungen zueinander.

Tajfel (1972) erkannte in diesem Zusammenhang jenen Teil des individuellen Selbstkonzepts an, der seine Grundlage aus dem Wissen und der Wertigkeit bezüglich der Mitgliedschaft in mindestens einer Gruppe zieht – die soziale Identität. Gerade die daraus resultierende Kategorisierung in Eigen- und Fremdgruppe wird als Ursache für die Entwicklung dieser sozialen Identität betrachtet. Als eine Konsequenz von Kategorisierungen werden Unterschiede zwischen Kategorien zwar überschätzt, was zu Problemen zwischen Wahrnehmung und Leben des Intergruppenverhaltens führen kann, gleichzeitig entlasten die Kategorien jedoch das kognitive System, indem sie die Welt übersichtlich ordnen (Ashmore & Del Boca, 1981). Im Vergleich zu Tajfel ging Turner (1981) einen Schritt weiter und davon aus, dass Individuen die Wahrnehmung ihrer selbst und anderer vor allem mittels abstrakter sozialer Kategorien strukturieren und diese Kategorien in das Selbstkonzept integrieren. Dabei gilt die

Summe aller sozialen Identifikationen, die eine Person heranzieht um sich selbst zu beschreiben, als soziale Identität.

Turner (1981) griff demnach die Bedeutung der sozialen Identität auf und fasste diese mit der persönlichen Identität in seinem Selbstkategorisierungsansatz zusammen. Turners Selbstkonzept lässt sich im Wesentlichen hinsichtlich dreier Formen von Identitäten zusammenfassen: der grundlegenden menschlichen (laut der man so ist „*wie alle anderen*“), der sozialen (die ein Individuum so sein lässt, „*wie einige andere*“) sowie der personalen (die die Individualität einer Person repräsentiert, so zu sein „*wie keine andere*“). Diese Theorie betrachtet eine Person sowohl als Mitglied einer Gruppe als auch als persönliche Identität, die durch interpersonelle und intragruppale Unterschiede definiert wird. Die soziale und die persönliche Identität stellen hypothetische kognitive Strukturen dar, die als zusammengeordnete Subsysteme den größten Teil des Selbstkonzepts ausmachen. Beide Identitäten können zur Erklärung der Veränderbarkeit des Selbstkonzepts herangezogen werden, da die – ursprünglich als beständig definierte – Selbstaspekte eben nicht ausschließlich für die Ausprägung des Selbstkonzepts verantwortlich gezeichnet werden können, sondern immer in den Zusammenhang mit der sozialen Identität einer Person gebracht werden müssen. Ein Individuum gilt dabei aufgrund seiner Merkmale, spezifischer situativer Gegebenheiten oder durch das Aufsuchen von Nähe und Gemeinschaft als Teil einer Gruppe. Es ist erwiesen, dass der Gruppe in der bestehenden Gesellschaft eine wichtige Bedeutung zuteil wird, da der Zusammenschluss und die Koordination von verschiedenen Personen die Effizienz und Möglichkeiten einzelner Individuen verbessern kann (Steiner, 1972).

Gerade in westlichen Kulturen steht jedoch dieser Gemeinschaftssinn in Konflikt mit dem persönlichen Bedürfnis nach Unabhängigkeit und Individualität, welches unter Umständen mit den Gruppenzielen unvereinbar ist und zugunsten der Anerkennung und Akzeptanz in der Gruppe zurückgestellt werden muss. Auf der anderen Seite kann die (erfolgreiche) Gruppenzugehörigkeit die soziale Identität stärken, die ihrerseits nicht nur durch eine höhere Bindung und Loyalität zur Gruppe gekennzeichnet ist, sondern zudem Belohnungen in sich birgt, die wiederum die persönliche Identität stärken und Kompromisse bezüglich der Gruppe rechtfertigen. Damit die soziale Identität möglichst positiv ausfällt, kommen im Vergleich zu anderen Gruppen vor allem Mechanismen der Verallgemeinerungen, Diskriminierung und Ablehnung der Fremdgruppe zum Einsatz. Damit wird schließlich auch der Bogen in der theoretischen Fundierung vom Selbst zur Gruppe und zu Vorurteilen und Stereotypen geschlagen.

Dabei stellt neben Kategorisierungen (aufgrund ethnischer Zugehörigkeit, sozioökonomischer Faktoren wie Verdienst, Status, etc.), die gewollt zur Gruppenbildung und positiven Identifizierung herangezogen werden, gerade das biologische Geschlecht eine der wichtigsten Grundlagen der Gruppenbildung dar. Zum einen ist das Geschlecht in seinen Ausprägungen limitiert und definiert damit die Anzahl von möglichen unterschiedlichen Gruppen. Zum anderen sind die geschlechtsbasierten Kategorien biologisch vorgegeben und beschränken daher die Freiwilligkeit der individuellen Gruppenwahl per se.



So bleibt zusammenfassend festzustellen, dass das Selbst zwar eine auf persönlichen Unterschieden im Vergleich zu anderen basierende individuelle Identität darstellt ("*One's self is precisely that, one's own, inherently differentiated and differentiating.*"; Turner & Onorato, 1999, S. 13), deren Einstellungen, Konzepte, Schemata oder Überzeugungen durch komplexe kognitive Prozesse geformt werden, es aber dennoch nicht losgelöst von den Erfahrungen als Mitglied einer sozialen Gruppe ist. Im Rahmen der Selbst-Repräsentations-Theorie betrachteten Deaux und Major (1987) das Geschlecht im Übrigen nicht notwendigerweise als biologisch gegeben, sondern als das, was durch das eigene Tun definiert wird. Mit einem verstärkten Fokus auf die auszuübenden Rollen in Abhängigkeit von sowohl Situation als auch Interaktionspartner ließen Deaux und Major eine gewisse Variabilität des Geschlechts zu, welches durch Erwartungen des Betrachters, situative Einflüsse wie auch vor allem durch das individuelle Selbstkonzept erst definiert und konstruiert wird.

### **2.1.2 Das Modell des dynamischen Selbst (multipel, flexibel, Korrektur)**

Wie bereits diskutiert, ist das Selbst strukturell gesehen nicht durch Einheitlichkeit charakterisiert, sondern stellt vielmehr eine komplexe kognitive Struktur dar mit spezifische Teilkonzepte, die situativ ausschnittsweise aktiviert werden können (Hannover, 1997; Kühnen & Hannover, 2003; Markus & Wurf, 1987). Dabei werden vor allem jene Aspekte des Selbstwissens zugänglich, die zum einen kürzlich aktiviert wurden, von Higgins (1990) beschrieben als *recent contextual priming*, oder solche, die im Zuge der Vergangenheit wiederholt aktiviert wurden. Diese situative Aktivierung von bestimmten Wissensbeständen ist Teil einer langen Debatte bezüglich des Aufbaus und der Funktionsweise des Selbst beziehungsweise darüber, inwieweit das Selbstkonzept durch Kontinuität und Konsistenz gekennzeichnet ist. Simon und Trötschel (2007) sahen das Selbstkonzept in engem Zusammenhang mit dem episodischen Gedächtnis, in Verknüpfung mit dem Wissen über persönliche Erfahrungen aus der Vergangenheit, was eine gewisse individuelle Kontinuität in der Struktur vermuten lässt. Auch der entwicklungspsychologische Ansatz unterstützt die Idee einer Kontinuität, da das Selbstkonzept nicht nur bereits im Kindesalter als ausgebildet gilt, sondern darüber hinaus regelrecht gegen diskonforme Informationen verteidigt wird (Brandstädter & Greve, 1994; Filipp, 1980). In einem ähnlichen Ansatz diskutierten Turner und Onorato (1999) das Selbstkonzept als verhältnismäßig festgelegte beständige Struktur im Informationsverarbeitungssystem. Dabei vertraten beide zwar die Vorstellung einer grundlegenden kognitiven Struktur, die durch zusammenhängende mentale Repräsentationen des Selbst gekennzeichnet ist, sprechen in ihren Ausführungen dem Gesamtsystem jedoch durchaus eine gewisse Veränderlichkeit zu.

### **2.1.3 Die Bedeutung des Selbstwerts in der Selbstkonstruktion**

Im Rahmen des sozialen Vergleichs gilt ein einfaches Prinzip: wird die Gruppe positiv bewertet, steigt auch der eigene Wert (Tajfel & Turner, 1986). Bezug nehmend zur Theorie von Festinger (1954) besteht die Motivation darin, die eigene Gruppe im Vergleich bewusst aufzuwerten, um den eigenen individuellen Selbstwert zu stärken oder um das Bedürfnis nach kognitiver

Kohärenz zu erfüllen (Abrams & Hogg, 2001; Tajfel, 1969). Die Untersuchungen zur Hypothese des Selbstwertes erfolgten nahezu ausschließlich mittels minimaler Gruppenparadigmen<sup>1</sup>, denen einfache Kategorisierungen in zwei Gruppen als ausreichende Ursache für Fremdgruppendifferenzierung zugrunde liegen.

Die Forschung konnte bisher in einer Vielzahl von Studien aussagekräftige Beweise vorlegen, dass Individuen eine möglichst positive Selbstdarstellung anstreben, sowohl in eigener Evaluation als auch gegenüber anderen (Devine, 1989; Greenwald & Banaji, 1995) und dass eine Beziehung zum Selbstwert besteht. Tajfel und Turner (1979) gingen davon aus, dass das Bedürfnis nach positivem Selbstwert und positiver sozialer Identität darin resultiert, die betreffenden Kategorien positiv zu bewerten. In diesem Zusammenhang wird deutlich, welche Rolle die Gruppe im Rahmen dieser Selbstbewertung spielt. Umfangreiche Untersuchungen bezüglich Eigen- und Fremdgruppen zeigten, dass soziale Vergleiche in jenen sozialen Kategorien vorgenommen werden, die relevant für die eigene Bewertung sind (Mummendey & Otten, 1998). Das Selbstwertgefühl der Gruppenmitglieder wird vor allem dadurch gestärkt, dass eine oder mehrere relevante Fremdgruppen im sozialen Vergleich herabgesetzt werden. Dabei gilt: je höher das Selbstwertgefühl der Gruppenmitglieder, desto größer ist die Wahrscheinlichkeit, die Eigengruppe zu bevorzugen und die Fremdgruppe zu diskriminieren, wobei tendenziell die Begünstigung der Eigengruppe eher geschieht als das Herabsetzen der Fremdgruppe (Mummendey & Otten, 1998).

## 2.2 Das Geschlecht

Gerade vor dem Hintergrund, dass Männer und Frauen zwei der wichtigsten Gruppen in der heutigen Gesellschaft darstellen, wird dem Gruppierungsmerkmal Geschlecht eine große Bedeutung für das eigene Selbstkonzept zuteil.

Das Geschlecht, biologisch und hormonell bestimmt und durch (Lern-)erfahrungen geformt, ist für die psychologische Entwicklung eines Menschen von großer Bedeutung. Es beeinflusst nicht nur die individuelle Wahrnehmung der Welt (Bem, 1981a; Bem, 1993; Deaux & LaFrance, 1998) und das Verhalten, sondern darüber hinaus als soziale Kategorie vor allem die Wahrnehmung eines Individuums als Mitglied einer Gruppe und die daraus resultierenden Erwartungen und Stereotypen (Alfermann, 1996). Das Geschlecht als Kategorie reflektiert dabei unter anderem soziale oder berufliche Rollen sowie beschreibende Eigenschaften und Einstellungen (Cejka & Eagly, 1999; Fiske & Stevens, 1993; Prentice & Carranza, 2002).

Das Geschlecht spielt eine große Rolle in der Identitätsbildung eines Menschen. Nicht nur, dass es als klassische Differenzierung in vielen Sprachen auf der ganzen Welt fungiert, charakterisieren Geschlechtsunterschiede darüber hinaus das Leben in vielfältiger Hinsicht. Zahlreiche Studien belegen wiederholt, auf welche Art und Weise Männer und Frauen in den vielen Aspekten des Lebens agieren, wahrnehmen oder fühlen.

Im folgenden Kapitel sollen unter anderem Definitionen, Geschlechterstereotype, faktisch existierende Geschlechtsunterschiede und Theorien zur Entwicklung des Geschlechts-

---

<sup>1</sup> Die bloße Kategorisierung von Personen in zwei Gruppen nach einem beliebigen Kriterium.

rollenelbstkonzepts vorgestellt werden. Gerade der Fokus auf Geschlechterstereotype nimmt in der psychologischen Forschung einen großen Teil ein, innerhalb dessen die Entstehung und Beeinflussung dieser Stereotype auf die Wahrnehmung von Frauen und Männer betrachtet wird (Cejka & Eagly, 1999; DeArmond et al., 2006; Deaux & Lewis, 1984; Hoyt, Simon, & Reid, 2009; Kite, Deaux, & Haines, 2008; Swim, 1994).

### 2.2.1 Definition von Geschlechtsrollen und Geschlechterstereotypen

Geschlechtsrollen und stereotype Einstellungen beeinflussen das Leben einer jeden Person. Sie steuern Erwartungen, Werte und das Verhalten. Die Forschung bezüglich Geschlechtsrollen, Stereotypen und ihrer Entwicklung verknüpft u. a. die Bereiche Soziologie, Anthropologie und Psychologie. Der Fokus liegt dabei auf Geschlechtsunterschieden zwischen Männern und Frauen, Fragen inwieweit diese Unterschiede auf Sozialisationsprozesse zurückzuführen sind und inwiefern das biologische und psychologische Geschlecht nicht notwendigerweise das gleiche Konstrukt repräsentieren.

Zieht man die Bedeutung der Geschlechterthematik in der heutigen Forschung in Betracht, überrascht es ein wenig, dass die Terminologie bezüglich der Schlagwörter (*sex, gender, stereotypes, roles, identity, etc.*) bisweilen uneindeutig und inkonsequent ist, wobei dies im stärkeren Maße für den englischen Sprachraum zutrifft. Die Begriffe *sex roles* und *gender roles* beispielsweise wurden über Disziplinen hinweg trotz unterschiedlicher Definition (Singleton, 1986) synonym eingesetzt, was eine Reflektion der thematisch relevanten Definitionen in den folgenden Abschnitten notwendig macht.

Gewöhnlich wird der Begriff *Geschlecht* im Sinne der biologischen Differenzierung von Männern und Frauen auf Grund von externen und internen physischen Charakteristiken verwendet. Im Zuge der theoretischen Entwicklung ist jedoch auch die Anerkennung des *psychologischen Geschlechts* von Bedeutung, das nicht gänzlich unabhängig von biologischen Voraussetzungen ist, und die Beziehung zur Männlichkeit und Weiblichkeit im sozialen Sinne herstellt (Archer & Lloyd, 2002; Havighurst, 1983). Beide Konzepte können zwar durchaus integriert werden und sollten daher nicht als zwei gänzlich unabhängige Aspekte betrachtet werden (Hines, 2004; Lippa, 2002), sie gleichzeitig jedoch als äquivalent zu behandeln und einzusetzen, wird den jeweiligen Bedeutungen nicht gerecht. Damit wird beispielsweise der Annahme Rechnung getragen, dass ein biologisch definierter Mann durchaus sehr weiblich in seinem Verhalten sein kann. Er bleibt physisch gesehen ein Mann (*biologisches Geschlecht*), in seinem Umfeld kann er jedoch als feminin (*soziales Geschlecht*) gelten. Weder Maskulinität noch Femininität sind von vornherein bestimmt oder stabil und können daher in beiden biologischen Geschlechtergruppen gleichermaßen entwickelt sein und das Verhalten sowie Kognitionen mitbestimmen.

Im englischen Sprachraum sollte die Konstruktunterscheidung verständlicher gelingen, da dort *sex* als Bezeichnung für das biologische Geschlecht oder Sexualität zum Einsatz kommt, während *gender* das psychologische Geschlecht betitelt. Archer und Lloyd (2002) wiesen darauf hin, dass *sex* entweder tatsächlich zur Bezeichnung von Sexualität verwendet

werden sollte oder zur Unterscheidung zwischen den binären biologischen Kategorien von *männlich* und *weiblich*.

Diese Äquivalenz im Gebrauch eigentlich unterschiedlich konnotierter Begrifflichkeiten zeigt sich auch bei den Begriffen Geschlechtsrolle und Geschlechtsrollenstereotype – verwandte, aber dennoch unterschiedliche Konstrukte. In Anlehnung an Singletons Definition (1986, S. 4) bezieht sich der Term *gender role* auf ein Set an Verhaltensweisen und Charakteristiken, die typisch für Männer und Frauen sind, während *gender-role stereotype* für konsensuelle Überzeugungen einer bestimmten kulturellen Gruppe bezüglich dieser Charakteristiken, Einstellungen, Werte und Überzeugungen als typisch für Männer und Frauen steht.

Williams und Best (1990, S. 15) nahmen eine Unterscheidung zwischen Geschlechtsrollenstereotype (*sex-role stereotypes*) und Stereotype bezüglich geschlechtsspezifischer Eigenschaften (*sex-trait stereotypes*) vor. Während Geschlechtsrollenstereotype sozial determinierte Überzeugungen bezüglich geschlechtsspezifischer Rollen sind, bezeichnen *sex-trait stereotypes* eine Konstellation psychologischer Eigenschaften, die als charakteristisch für das eine als für das andere Geschlecht gelten, z. B. Männer sind stark, widerstandsfähig und aggressiv. Stereotype werden dabei definiert als „*perceptions of persons, objects, activities, or concepts that are based on relatively rigid, oversimplified, and over-generalized beliefs or assumptions regarding the characteristics of males and females*“ (Beere, 1979, S. 164).

Das Geschlechterstereotyp beinhaltet sowohl kulturelle Vorstellungen über psychologische, personale oder auch physikalische Charakteristiken als auch geschlechterspezifische Verhaltensweisen, Aktivitäten, Präferenzen oder emotionale Dispositionen von Männern und Frauen (Alfermann, 1996; Ashmore & del Boca, 1979; Brannon, 2005; Deaux & LaFrance, 1998; Rudman & Glick, 2008) – es sind Einstellungen über Maskulinität und Femininität. Geschlechterstereotyp zeichnen sich durch Beständigkeit und einen starken, häufig impliziten und daher automatischen Einfluss auf kognitive Informationsverarbeitung aus. Sie existieren sowohl für die grobe Kategorisierung in Männer und Frauen als auch in Form spezifischer Subtypen (Deaux & LaFrance, 1998). Bezüglich der Richtigkeit von Geschlechterstereotypen verglich Swim (1994) die von Probanden durchschnittlich eingeschätzten Geschlechtsunterschiede hinsichtlich verschiedener Eigenschaften (wie mathematische Fähigkeiten) mit den in Meta-Analysen tatsächlich gefundenen Unterschieden und fand nur in 38% der Fälle eine Passung zwischen Stereotyp und Realität (verglichen mit 28% überschätzten und 34% unterschätzten Geschlechtsunterschieden). Unter diesem Aspekt sind Geschlechtsrollenstereotype Überzeugungen und Erwartungen bezüglich der Charakteristiken, des Verhaltens und des sozialen Status von Männern und Frauen, üblicherweise innerhalb einer Kultur. Ruble (1988) nannte dies auch das *public face* des Geschlechts. Wie oben bereits angeführt, ist aus der Gruppenforschung bekannt, dass der Kategorisierung nicht zuletzt auch eine motivationale Tendenz zugrunde liegt, die eigene Gruppe positiver zu bewerten. Dies trifft auch auf die Geschlechtergruppen zu. Männer stellen in den meisten Kulturen die dominante Gruppe dar.

Darüber hinaus ist über Länder hinweg nachgewiesen worden, dass das männliche Stereotyp im stärkeren Maße mit den Dimensionen Stärke und Aktivität assoziiert ist. Diese Bewertung kann u. a. dazu führen, dass geschlechterübergreifend in der Einschätzung das Männliche für wertvoller gehalten wird und das männliche Stereotyp mit mehr positiven Eigenschaften genannt wird (vgl. Alfermann, 1996). Diekman und Eagly (2000) führten verschiedenste Autoren auf, die den Nachweis erbracht haben, dass Geschlechterstereotype darüber hinaus verhältnismäßig akkurate Prädiktoren für durchschnittliche Geschlechtsunterschiede sind (Eagly & Diekman, 1997; Swim, 1994).

In Bezug auf die Geschlechtsrolle herrscht in der Literatur wenig Stringenz bezüglich der konsequenten Unterscheidung zwischen *gender role* und *sex role*, die ebenfalls häufig synonym verwendet werden. *Sex role* bezieht sich darauf, dass eine Person in einer – biologisch vorbestimmten – männlichen oder weiblichen Rolle lebt. Im Gegensatz dazu ist *gender role* eher Ausdruck dessen, wie die Person diese Rolle lebt, und wurde erweitert um den Begriff *gender role behaviour* als Verhaltensweisen, die typisch für diese Rolle sind. Geschlechtsrollen sind definiert durch das Verhalten und die Rollenausübung beider Geschlechter (Brannon, 2005).

Die Definition unterscheidet sich von der der Geschlechterstereotype insofern, da Geschlechtsrollen zusätzlich eine präskriptive Komponente beinhalten: sie integrieren über die reine Beschreibung der Unterschiedlichkeit auch normative Erwartungen hinsichtlich bestimmter Eigenschaften und insbesondere Handlungsweisen (Alfermann, 1996). Obwohl auch Geschlechterstereotype und Geschlechtsrollen definitorisch unterschiedliche Konzepte sind, stehen sie miteinander in Beziehung. Werden z. B. spezifische Verhaltensweisen entweder mit Männern oder Frauen assoziiert, erfolgt dies unter gleichzeitiger Ausblendung individueller Variationen und Erwartungen, wodurch die Vorstellung verstärkt wird, dass dieses Verhalten unweigerlich eher dem einen Geschlecht zugehörig ist als dem anderen. Brannon (2005) interpretierte daher die Geschlechtsrollen durchaus auch als eine der Grundlagen für Geschlechterstereotype.

Die Geschlechtsrollenorientierung (*gender role orientation*) ihrerseits entspricht der Selbstdefinition eines Individuums – feminin oder maskulin zu sein (vgl. Doyle & Paludi, 1995), wobei sich bezüglich der Formulierung erneut mangelnde Einigkeit andeutet, präferierte Hannover (1997) in diesem Fall doch von einem Geschlechtsrollenselbstkonzept zu sprechen.

Weitere wichtige zu definierende Begriffe sind die Geschlechtsidentität (*gender identity*) und Geschlechtsorientierung (*gender orientation*). Obwohl auch die Konstrukte *identity* und *orientation* ihrerseits unterschiedliche Bedeutungen haben, werden sie in Bezug auf das Geschlecht als Synonyme verwendet. Die Geschlechtsidentität speist sich aus der biologischen und sozialen Definition, was es heißt, Mann oder Frau zu sein, und wird auf der individuellen Ebene in das eigene Selbst verankert. Auf dieser Ebene nimmt die Geschlechtsidentität Einfluss auf kognitive, emotionale und verhaltensbezogene Aspekte (beispielsweise durch Normen der Geschlechtsrolle). Die Geschlechtsidentität, die aufgrund einer äußerst globalen Definition häufig auch als das individuelle Gefühl männlich oder weiblich zu sein beschrieben wird und

damit relativ konstant bleiben sollte, beinhaltet generell alle Aspekte des Selbst, die durch einen Geschlechterbezug gekennzeichnet sind, was sowohl Aktivitäten, Interessen, Wahrnehmungen oder Charakteristiken mit einschließt, und über Gruppen und Zeiten hinweg variieren kann (Deaux & Lafrance, 1998). Die Geschlechtsidentität repräsentiert die Übereinstimmung zwischen den kulturell vorherrschenden Geschlechtsrollenstereotypen und der Wahrnehmung bezüglich der eigenen Geschlechtsrolle (vgl. Kagan, 1964). Sie stellt das Verständnis dar, mit dem sich eine Person als typisch maskulin oder feminin einschätzt und damit die typischerweise geschlechtsspezifisch zugeschriebene Rolle akzeptiert und zu einem Teil ihrer Identität macht. Dabei gilt die Geschlechtsidentität gemeinhin als wichtiger für die subjektive Gesundheit als die eigentliche biologische Geschlechtszugehörigkeit.

Ein weiterer wiederholt auftauchender Begriff im Kontext der Geschlechterforschung sind *Geschlechterüberzeugungssysteme*, die laut Deaux und Lafrance (1998) als das vielfältige große Ganze, Stereotype, Geschlechtsrolleneinstellungen und Repräsentationen der eigenen Geschlechtsidentität beinhalten. Sie sind ein Komplex aller umfassender Ideen bezüglich des Geschlechts und übernehmen die Funktion eines Filters, der die individuelle Wahrnehmung der Geschlechter steuert und Einfluss auf Verhaltensweisen ausübt. Vor dem Hintergrund der Vielfältigkeit dieser Geschlechterüberzeugungssysteme muss zudem davon ausgegangen werden, dass individuelle Vorstellungen über das Geschlecht nicht ausschließlich aufgrund individueller Kognitionen analysiert werden können, sondern dass es vielmehr darum gehen sollte, auch soziale und strukturelle Aspekte sowohl in der Entwicklung als auch Aufrechterhaltung dieser Geschlechterüberzeugungssysteme zu beachten (Deaux & Lafrance, 1998).

In diesem Zusammenhang wird die gerechtfertigte Unterscheidung zwischen *sex* und *gender* offensichtlich, bedeutet *sexuelle Identität (sexual identity)* die objektive Kategorisierung des physiologischen Status als Mann oder Frau, während die Geschlechtsidentität die subjektive, aber doch recht stabile eigene Vorstellung als maskulin oder feminin repräsentiert.

Entscheidend ist, dass die biologischen Unterschiede die psychologischen zwar beeinflussen aber nicht unmittelbar bedingen (vgl. Maccoby & Jacklin, 1974). Es gibt keine direkte Beziehung derart, dass Männer zwingend maskulin sind und männliche Verantwortungen übernehmen, während Frauen die umsorgenden Aufgaben in einem Beziehungsgefüge übernehmen. So hatte Mead (1935) in ihren klassischen Studien dreier Stämme in Neuguinea schon gezeigt, dass Frauen sehr wohl als Familienoberhaupt fungieren können. Maskulinität und Femininität kommen demnach als Geschlechtsrolle für alle Menschen in Betracht.

Bei den meisten Menschen gibt es jedoch eine hohe Übereinstimmung zwischen biologischem, sozialem und dem eigenen akzeptierten Geschlecht, sie besitzen die Merkmale eines bestimmten Geschlechts, verhalten sich kulturabhängig diesem Geschlecht entsprechend und fühlen sich diesem Geschlecht auch zugehörig.

Die Geschlechtsrollenentwicklung (*gender role development*) bzw. die Geschlechtstypisierung (*gender typing*) wurde von Perry und Bussey (1984) als der Prozess

beschrieben, im Verlauf dessen Kinder alle Verhaltensweisen, Einstellungen Interessen, Motive und emotionale Reaktionen erlernen, die als kulturell angemessen für die Mitglieder des eigenen Geschlechts gelten. Damit nehmen die Autoren Bezug auf die gesellschaftlich unterschiedliche Behandlung von Personen in Abhängigkeit ihres biologischen Geschlechts.

Es lässt sich zusammenfassen, dass Geschlechtsrollen (maskulin vs. feminin) – als typische Charakteristiken von Männern und Frauen (Geschlechtsrollenstereotype) – im Prozess der Geschlechtsentwicklung geformt werden. Die Einflüsse auf die Entwicklung eben dieser Geschlechtsrolle umfassen neben biologischen, kognitiven und psychoanalytische Erklärungen, vor allem auch sozialpsychologische Theorien, die im nächsten Abschnitt diskutiert werden.

### **2.2.2 Biologische Grundlagen von Geschlechtsrollen und Geschlechterstereotypen**

Die Beziehung zwischen dem sozialen Geschlecht und biologischen Ursachen ist komplex. So können beispielsweise dem geschlechtstypischen Verhalten keine einfachen kausalen biologischen Ursachen zugrunde gelegt werden. Dennoch baut das soziale Geschlecht eines jeden Menschen zweifelsohne auf dem biologischen Geschlecht auf (siehe z. B. Muldoon & Reilly, 1998).

Biologische Unterschiede zwischen Jungen und Mädchen und deren Ursprünge sind in der Literatur klar belegt. Das Geschlechtschromosom des väterlichen Spermias bestimmt das genetische Geschlecht des Kindes. Das Durchsetzen des Y-Chromosoms in Kombination mit Hormonen stimuliert die Entwicklung der inneren männlichen Organe, des Gehirns und der äußeren männlichen Charakteristiken. Die Abwesenheit eben dieses Y-Chromosoms wiederum steuert die Entwicklung in die andere Richtung: Geschlechtsdrüsen werden zu Eierstöcken; das Baby wird ein Mädchen (Lewis & Weinraub, 1979; Ruble & Martin, 1998; Singleton, 1986).

Die Bedeutung der Geschlechtschromosomen und -hormonen wird durch die Betrachtung von Anomalitäten herausgestellt, die als Nachweis für einen andauernden biologischen Einfluss auch im fortgeschrittenen Alter stehen, z. B. Turner-Syndrom, Klinefelter-Doppel-Y-Syndrom oder Störungsbilder, die auf Androgenauffälligkeiten zurückgehen (Doyle & Paludi, 1995; Muldoon & Reilly, 1998). Innerhalb des Forschungsbereiches zu pränatalen Androgeneffekten werden unter anderem Störungsbilder untersucht, die mit einem erhöhten Androgenniveau beim XX-Karyotyp einhergehen (z. B. adrenogenitales Syndrom, AGS). Berenbaum (1998) konnte den Zusammenhang hormoneller Geschlechtsunterschiede zu geschlechtstypischen Verhaltensweisen nachweisen und kam zu dem Ergebnis, dass Mädchen mit dem AGS-Syndrom geschlechtsuntypische Präferenzen zeigen. So hatten sie nicht nur ein stärkeres Interesse an Jungenspielzeug und bevorzugten Jungen als Spielkameraden, sondern zeigten darüber hinaus auch häufiger aggressives Verhalten trotz vergleichbarer Sozialisationserfahrungen wie Mädchen ohne AGS.

Die biologischen Geschlechtsunterschiede sind von Geburt an augenscheinlich und resultieren in Verhaltensunterschieden, die zwischen den Geschlechtern unmittelbar in den ersten Tagen nach der Geburt auftreten, zu einer Zeit, in der Sozialisationsprozesse kaum hätten greifen können: männliche Babys sind größer, stärker und leichter reizbar als Mädchen (Maccoby & Jacklin, 1974; Ruble, 1988). Unabhängig von allen noch ausstehenden, den

Menschen formenden Prozessen, sind Jungen und Mädchen gewissermaßen für ihre sozialen Rollen „vordefiniert“. Jedoch sind die biologischen Voraussetzungen keineswegs der einzige und exklusive Faktor um unterschiedliches Geschlechtsrollenverhalten erklären zu können. Nach der Geburt wird dem Baby das Geschlecht aufgrund äußerer genitaler Charakteristiken zugeschrieben. Obwohl alsbald einsetzende soziale Einflüsse und Erwartungen zu unterschiedlichen Umgangsweisen mit den Kindern aufgrund ihres biologischen Geschlechts führen werden, hat ein jeder Mensch die Voraussetzungen, sowohl maskuline als auch feminine Verhaltensweisen entwickeln zu können. Hinzu kommt, dass Mädchen und Jungen biologisch gesehen durchaus mehr Ähnlichkeiten aufweisen als zunächst angenommen. So fassten Williams und Best (1990) Studien zusammen, die nur geringe Unterschiede in Morphologie und im hormonellen Status bis zum Alter von 10 Jahren fanden. Dennoch lassen sich auch genetische Einflüsse auf einen Teil der Geschlechtsrollenorientierung nachweisen (Lippa & Hershberger, 1999). Dazu hatten die Autoren in einer Zwillingsstudie Selbstbeschreibungsdaten analysiert sowie Angaben der Probanden hinsichtlich individueller Präferenzen für 160 Berufe und die Ausübung von 324 täglichen Aktivitäten und konnten zeigen, dass geteilte Umwelteinflüsse bisweilen weniger bedeutsam waren als ursprünglich angenommen.

### 2.2.3 Psychologische Theorien der Geschlechtsrollenstereotypisierung

Die Geschlechtsrollenstereotypisierung ist Teil des Sozialisationsprozesses, im Rahmen dessen die interpersonale und kulturelle Umgebung Einfluss auf die psychologische Entwicklung nimmt und ein Mensch zu einem Individuum in der Gesellschaft heranwächst (Archer & Lloyd, 2002). Allgemein existieren dabei bestimmte Ähnlichkeiten in der Sozialisation der Geschlechtsrolle über verschiedene Kulturen hinweg. Barry, Bacon und Child (1957) kamen zu dem Schluss, dass ein gewisser Sozialisationsdruck hin zur Entwicklung von Gehorsam, Verantwortung und Fürsorge stärker für Mädchen zu existieren scheint, während Jungen – in einem stärkeren Maße als Mädchen – in der Entwicklung von Leistungsstreben und Selbstvertrauen gefördert werden und dies über verschiedenste Kulturen hinweg.

Tabelle 1. Übersicht bedeutsamer Theorien bezüglich der sozialen Geschlechtsentwicklung

Theorie	Vertreter	Annahmen
Psychoanalyse	Freud	<ul style="list-style-type: none"> <li>▪ Libidoenergie - als Grundlage nahezu allen Verhaltens - erzeugt ein Verlangen, dass es zu befriedigen gilt</li> <li>▪ Entwicklung der oralen, analen und genitalen Phase</li> <li>▪ Ödipuskomplex</li> </ul>
Soziale Lerntheorie	Mischel, Bandura	<ul style="list-style-type: none"> <li>▪ Prinzipien der sozialen Umgebungsbedingungen, die die Persönlichkeit beeinflussen und formen</li> </ul>
Theorien der kognitiven Entwicklung	Kohlberg, Kagan, Ullian  Constantinople, Pleck, Bem	<ul style="list-style-type: none"> <li>▪ Entwicklungen von Geschlechtskonstanz und -identität das Kind als aktives Individuum</li> <li>▪ Geschlechterschema</li> </ul>

Bezüglich psychologischer Theorien und speziell derer, die die Sozialisation von Geschlechtsrollenstereotype thematisieren und ihre Entwicklung und Funktion erläutern, werden vor allem drei traditionelle Ansätze unterschieden: die Psychoanalyse, Lerntheorien und



kognitive Entwicklungstheorien (Hargreaves & Colley, 1986; Lewis & Weinraub, 1979), die im Folgenden näher dargestellt werden. Trotz der gänzlich unterschiedlichen Zugänge ist allen Theorien im gewissen Sinne die Anerkennung der Bedeutung vom Modelllernen und damit der entscheidenden Rolle der Eltern gemein (Tabelle 1).

### **2.2.3.1 Psychoanalytische Theorien (Freud, Erikson)**

Freuds (1927) Annahmen der Geschlechtsrollenstereotypisierung unter Berücksichtigung des Aspekts der psychosexuellen Motivation gelten auch heute noch als kontrovers. Sein Hauptaugenmerk legte Freud dabei sowohl vornehmlich auf die anatomischen Unterschiede zwischen Männern und Frauen sowie auch auf die starke Beziehung zwischen den gegebenen genitalen Unterschieden und den (daraus folgenden) psychologischen Charakteristiken von Männern und Frauen. Freuds' Ödipuskomplex besagt, dass die Hauptunterschiede zwischen Jungen und Mädchen während der ödipalen Entwicklungsphase entstehen, wenn sich das Kind mit dem gleichgeschlechtlichen Elternteil identifiziert. Im Falle von Jungen geht dem ein starkes Bedürfnis, ihre Mutter besitzen wollen, voraus, was die Jungen zu Konkurrenten ihres eigenen Vaters macht. Als Bestrafung fürchten Jungen den Verlust ihres Penis (die sog. *Kastrationsangst*). Anerkennend, dass sie ihre Mutter nie besitzen werden, unterdrücken Jungen ihr wahres Bedürfnis und beginnen sich stärker mit ihrem Vater zu identifizieren – ein Prozess, im Verlauf dessen sie maskuline(s) Verhalten und Eigenschaften erlernen. Im Vergleich dazu werden Mädchen dafür bestraft, dass sie keinen Penis haben, was zum sog. *Penisneid* führt. Obwohl ihre Mütter das Objekt ihrer Liebe sind, müssen Mädchen feststellen, dass auch ihre Mütter keinen Penis haben. Als Folge dessen richten sie die Liebe auf ihren Vater, eine Beziehung, die jedoch zum Scheitern verurteilt ist oder zumindest zeitlich begrenzt ist. Früher oder später erkennen Mädchen, dass sie den Besitz eines Penis nicht mit ihrem Vater teilen werden können, was sie an dem Punkt des Identifizierungsprozesses wieder zurück zu ihrer Mutter orientieren lässt. Das Verlangen, einen Penis zu besitzen, wird ersetzt mit dem Verlangen Kinder aufzuziehen und zu umsorgen. Mädchen entwickeln feminine Charakteristiken (Freud, 1927).

Ähnlich wie Freud, wenn auch nicht mit dem gleichen psychoanalytischen Schwerpunkt, postulierte Erikson (1950) eine Beziehung zwischen der biologischen Determination von Jungen und Mädchen und der späteren psychologischen Entwicklung. Eriksons Vorstellung bezüglich der Entwicklung eines Geschlechtskonzepts drehte sich dabei ebenfalls um die Sexualorgane, die ihrerseits die Ausrichtung eines Individuums im Leben mitbestimmen. In seiner Theorie korrespondierte das männliche Sexualorgan als externes und intrusives Merkmal mit der späteren aktiven Orientierung in der Welt. Im Gegensatz legte der weibliche Körper mit seinem internalen reproduktiven System den Grundstein für eine zarte und friedliche Einstellung.

In der Diskussion über Ursachen und Prozesse zur Geschlechtstypisierung sollten biologische und anatomische Unterschiede Berücksichtigung finden, da sie in der weiteren physiologischen, unterschiedlichen Entwicklung von Mädchen und Jungen zum Tragen kommen. Jedoch mögen gerade diese externen Unterschiede zwischen den Geschlechtern bedeutsamer für die Außenwelt und ihre Reaktion auf Jungen und Mädchen sein, als die von

Freud postulierte Funktion eines internen Antriebs im Prozess der Geschlechtstypisierung. Vielmehr scheint plausibel, dass die psychoanalytischen Theorien einen notwendigen – jedoch nicht erschöpfenden – Erklärungsansatz in Hinblick auf die Entwicklung des sozialen Geschlechts darstellen.

### **2.2.3.2 Soziale Lerntheorien (Mischel, Bandura)**

Während Evolutionstheoretiker davon ausgehen, dass psychologische Unterschiede zwischen den Geschlechtern das Ergebnis unterschiedlicher Reproduktionsrollen und -ansprüche an Männer und Frauen sind, führen Sozialisationsansätze die psychologischen Unterschiede zwischen den Geschlechtern nahezu ausschließlich auf unterschiedliche Sozialisationsverläufe und -einflüsse zurück (Looy, 2001).

Die soziale Lerntheorie, mit Mischel und Bandura als zwei ihrer sicherlich bedeutendsten Vertreter, betont alle sozialen Umgebungsbedingungen als Einfluss auf die Entwicklung der Persönlichkeit. Dabei leistet die Gesellschaft selbst den wohl größten Beitrag bei der Unterscheidung zwischen Maskulinität und Femininität und determiniert darüber schließlich die Geschlechtsrolle in einer Person. Die Entwicklung funktioniert dabei über Lernprozesse und der Anwendung von Verstärkern (Belohnung, Bestrafung), klassischer oder operanter Konditionierung sowie durch Lernen von Rollenmodellen (Mischel, 1966). Dieser Prozess nimmt seinen Beginn unmittelbar nach der Geburt des Kindes, wenn Eltern geschlechtsspezifisches und -angemessenes Verhalten interpretieren und verstärken (Lewis & Weinraub, 1979). Dieses Verstärkung erfahrende Verhalten wird dadurch in seiner Auftretenswahrscheinlichkeit zunehmen, wobei Verhalten, welches mit dem eigenen Geschlecht assoziiert ist, mit größerer Wahrscheinlichkeit imitiert wird, obwohl Verhalten des gegengeschlechtlichen Elternteils durchaus benannt und umgesetzt werden kann, wenn Kinder danach gefragt werden (Hargreaves, 1976).

Die Basis verschiedenster Eigenschaften und Verhaltensweisen wird demnach in den ersten Lebensjahren gelegt, in denen vor allem Eltern ihre Kinder darin unterstützen und animieren, geschlechtstypisches und -angemessenes Verhalten zu zeigen, z. B. in Bezug auf die Wahl des Spielzeugs (Idle, Wood, & Desmarais, 1993). So bekommen Jungen typischerweise Autos geschenkt, dürfen herumrennen und aktiv sein, während Mädchen bevorzugt mit Puppen spielen, kochen oder sich verkleiden. Einige Eltern zeigen großes Bestreben, auch geschlechtsuntypische Verhaltensweisen im Rahmen der elterlichen Sozialisation zu fördern. Vor allem kulturelle und gesellschaftliche Veränderungen haben in den letzten Jahrzehnten dazu geführt, dass diese Toleranz zunehmend gelebt wird und sich ursprünglich große Unterschiede zwischen Männern und Frauen als Resultat dessen verringern (Maccoby & Jacklin, 1974).

Viele Studien belegen, dass einige Eltern immer noch dazu tendieren, geschlechtstypisches Verhalten zu verstärken und geschlechtsuntypisches Verhalten zu missbilligen. Dieser Effekt ist umso stärker ausgeprägt, wenn das Kind ein Junge ist (Fling & Manosevitz, 1972). Mischel (1966) sah in den Eltern des Kindes die mächtigsten Verstärkungsagenten, die dabei auf eine Vielzahl von Strategien zur Belohnung / Bestrafung

zurückgreifen können, wie beispielsweise Lob, Ermutigung, Missbilligung oder Ärger. Trotz seines Schwerpunktes auf Beobachtungslernen hat Mischel die Existenz von mediierenden kognitiven Prozessen nie verleugnet, während er ihnen gleichzeitig jedoch weitaus weniger kausale Bedeutung in der Entwicklung zuschrieb, wie Kohlberg (1966) es tat.

Neben ihrer Verantwortung, Belohnung und Bestrafung einzusetzen, werden Eltern vor allem als Rollenmodelle betrachtet, die verfügbar, umsorgend und stark sind. Inwieweit dies geschieht, hängt von unterschiedlichsten Faktoren ab. So fasste Mischel (1966) diesbezüglich verschiedene Studien zusammen, die zeigen, dass der Beobachtende mit höherer Wahrscheinlichkeit Verhalten eines Modells kopiert, welches als ähnlich zu einem selbst wahrgenommen wird (z. B. Mädchen passen ihr Verhalten dem ihrer Mutter an, während Jungen sich stärker an ihren Vätern orientieren).

Darüber hinaus hatten Bandura, Ross und Ross (1963) jedoch demonstrieren können, dass der Sozialisationsprozess des sozialen Lernens weitaus komplexer ist als Mischels Ausführungen zunächst vermuten lassen. In ihrem klassischen Experiment von Aggression und Belohnung repräsentierten in einer 3-Personen-Gruppe ein Kind und zwei Erwachsene (jeweils eine Frau und ein Mann) gewissermaßen eine Familie, in der verschiedene Belohnungs- und Bestrafungssituationen bezüglich der Vergabe oder des Erhalts von Verstärkern gegeneinander getestet wurden. Im Gegensatz zu Mischels Erklärungen imitierten nicht nur die Jungen, sondern auch die Mädchen im gleichen Maße entweder die Person, die Verstärkung ausübte (dies galt für beide Elterngeschlechter) oder das belohnte männliche Modell. Zumindest die Mädchen wichen von der erwartungsgemäßen Linie des Zeigens geschlechtstypischen Verhaltens ab. Die Jungen, im Gegensatz dazu, orientierten sich selbst dann noch bevorzugt am Verhalten der männlichen Person, wenn diese keinerlei Belohnung erhalten hatte.

Maccoby und Jacklin (1974) betrachteten die Bedeutungszuschreibung des sozialen Lernens in der Geschlechtsrollenentwicklung kritisch und bezweifelten, dass die Geschlechtsrolle von Männern und Frauen direkt auf die geschlechtstypische Sozialisation von Mädchen und Jungen zurückgeführt werden kann, zumindest nicht in der Ausschließlichkeit mit der soziale Lerntheoretiker dies tun. Maccoby und Jacklin begründeten ihre Kritik vor allem anhand mangelnder Beweise aus der Forschung und stellten in Frage, inwieweit Jungen tatsächlich für aggressives oder Mädchen für abhängiges Verhalten verstärkt werden. Dafür erkannten Maccoby und Jacklin (1974) die Wirkung unbewusster und womöglich sogar unerwünschter Prozesse dahingehend an, dass die Kenntnis des kindlichen Geschlechts per se schon zu spezifischen und auch geschlechtstypischen Reaktionen seitens der Erwachsenen führt. Diese werden so zu sozial Agierenden, die – vielleicht sogar unbewusst – geschlechtstypisches Verhalten verstärken, z. B. sind Väter im Spiel mit Jungen etwas rauher und ungestümer als sie dies bei ihren Töchtern wären.

Neben sehr entscheidenden Jahren im frühen Kindesalter und in der Jugend erfahren geschlechtsbezogenes Verhalten, Einstellungen oder Geschlechtsrollen weiterhin andauernden Einfluss durch Sozialisationseinflüsse seitens Familie, Schule, Peers oder Print-, TV- und Filmmedien (Doyle & Paludi, 1995; Fagot, Rodgers, & Leinbach, 2000; Huston & Alvarez,

1990). Innerhalb der Familie sind es beispielsweise weniger die aus anderen Forschungssträngen bekannten beeinflussenden Faktoren wie Anzahl der Geschwister oder Geburtsfolge, die Unterschiede in den geschlechtstypischen Konzepten und Verhaltensweisen erklären können, sondern vor allem spätere die Erwartungen und Unterstützung seitens der Eltern sowie deren Arbeitsverhältnisse.

Obwohl der elterliche Einfluss im Entwicklungsverlauf anhält, werden dann jedoch in Phasen der Pubertät Peers zur Institution der sozialen Unterstützung oder (Miss-)Billigung von Verhalten oder auch Personen. Peers und deren sehr klare Vorstellungen von angemessenem und unangemessenem sozialen Verhalten kommen in dieser Entwicklungsphase der Entwicklung zum Tragen (vgl. Huston & Alvarez, 1990). Im Rahmen der Sozialisation machen sie ihren Einfluss geltend, indem sie zum Beispiel andere Personen aufgrund geschlechtstypischer Attribute akzeptieren oder auch ablehnen. Archer und Lloyd (2002) gingen sogar so weit anzunehmen, dass in dieser Phase der elterliche Einfluss hinter dem der Peers zurückfällt. So ist es dann auch möglich, dass zahlreiche und intensive Freundschaften mit Jungen zu Schulzeiten zu einer maskulinere Geschlechtsrolle beitragen können (Gurman, Wechselberg, & Long, 1993).

Die Rolle der Schule als Teil der Entwicklung von Geschlechtsrollenstereotypen erfährt ebenfalls mit den Jahren einen Bedeutungszuwachs. Mögliche Wechsel zwischen verschiedenen Schul(typ)en und die wiederkehrende Auseinandersetzung mit der Klassifikation von z. B. Mathematik und Naturwissenschaften als Männerdomänen nehmen Einfluss auf individuelle Geschlechterstereotype.

Schließlich leistet auch das Fernsehen mit seinen Shows, Nachrichten- oder Werbedarstellungen seinen Beitrag zu Geschlechterstereotypen und geschlechtsspezifischen Einstellungen und Verhaltensweisen. Wie verschiedene Studien zeigen, präsentieren die meisten Fernsehprogramme noch immer hoch stereotypisierte Bilder, indem sie Männer und Frauen in geschlechtstypischen beruflichen und familiären Rollen zeigen (Huston & Alvarez, 1990).

### **2.2.3.3 Kognitive Entwicklungstheorien (Kohlberg, Kagan, Ullian)**

In den bisher dargestellten Theorien nahm das Kind vor allem die Rolle eines passiven Rezipienten externer Faktoren ein. Diese Hypothesen liefern jedoch nicht alle möglichen Erklärungen im Sozialisationsprozess und resultieren in uneindeutigen Ergebnissen (vgl. Maccoby & Jacklin, 1974), die nicht ausreichend durch soziale Lerntheorien erklärt werden können. Bandura (1979) erkannte in seinen Überlegungen bereits die Bedeutung von kognitiven Faktoren und schrieb Individuen eine durchaus aktive Rolle außerhalb der *black box* zu.

Forscher wie Kohlberg (1966) und Ullian (Kohlberg & Ullian, 1974) führten dem entsprechend einen kognitiven Ansatz in der Erklärung zu Geschlechtsrollenentwicklung ein. Dieser Ansatz betrachtete vor allem interne und weniger externe Faktoren und konzentrierte sich auf die aktive Konstruktion der Welt seitens des Kindes während seines Interagierens in der sozialen Welt. Trotz der unterschiedlichen Überlegungen teilt die kognitive Entwicklungstheorie einige Annahmen der sozialen Lerntheorien, indem Modellprozesse in beiden Ansätzen

substantiellen Beitrag leisten. Die kognitive Theorie betrachtet ein Kind als weniger externen Einflüssen passiv ausgesetzt, sondern schreibt dem Kind eine bedürfnisorientierte, aktive Rolle sowie die Fähigkeit seine Umgebung zu verstehen zu. Statt lediglich verstärktes Verhalten zu zeigen, wie es die soziale Lerntheorie postulierte, kategorisieren sich Kinder eigenständig als entweder Junge oder Mädchen und werden folglich durch eine interne Geschlechtsrollenidentität „geleitet“. Die Identifikation des Jungen mit dem Vater ist nunmehr nicht mehr Ursache für die Geschlechtstypisierung sondern Folge derer.

Obwohl die theoretischen Erklärungen dies mitunter vermuten lassen, nimmt die Entwicklung der Geschlechtsidentität keinen geradlinigen Verlauf. Die Wahrnehmung der Konstanz des Geschlechts bildet sich graduell in drei Schritten: Geschlechtsidentität, Wahrnehmung der Geschlechtsstabilität sowie schließlich Geschlechtskonstanz (Eaton, 1983; Kohlberg, 1966; Slaby & Frey, 1976). Demnach kategorisieren sich Kinder eigenständig, lernen, dass ein Mädchen auch im späteren Verlauf des Lebens ein Mädchen / eine Frau bleiben wird und realisieren schließlich, dass das Geschlecht eine unveränderbare Konstante ist, unabhängig von möglichen externen Veränderungen. Die Wahrnehmung der Geschlechtskonstanz ist dabei notwendiger Entwicklungsschritt für den Identifikationsprozess mit gleichgeschlechtlichen Rollenmodellen. Geschlechtstypisches Verhalten wird quasi durch eine Art internalisierter Geschlechtsrolle initiiert. Dabei zeigen die Kinder individuelle Altersunterschiede im Erwerb der Geschlechtskonstanz, mit der Möglichkeit, dass einige Kinder diese Phase durchaus schon mit 18 Monaten erreichen können (Eaton, 1983; Lewis & Weinraub, 1979).

Kohlberg (1966) führte im Einzelnen aus, wie die kognitive Entwicklung zur eigenen Geschlechterstereotypisierung und der anderer beiträgt. Er ging davon aus, dass Kinder aktiv ihr eigenes Verhalten beeinflussen und hinterfragen und sich selbst durch das Ausüben geschlechtstypischen Verhaltens belohnen. Die Geschlechtsrollenidentität fungiert dabei gewissermaßen als Mediator zwischen dem Einfluss der Geschlechtsrollenstereotype und der eigenen individuellen Geschlechtsrollenentwicklung. Laut Kohlberg stellen Geschlechtsrollenattribute das Produkt der individuellen Geschlechtsrollenidentität und der Wahrnehmung der kulturell vorherrschenden Geschlechtsrollenstereotypen dar (Abbildung 1).

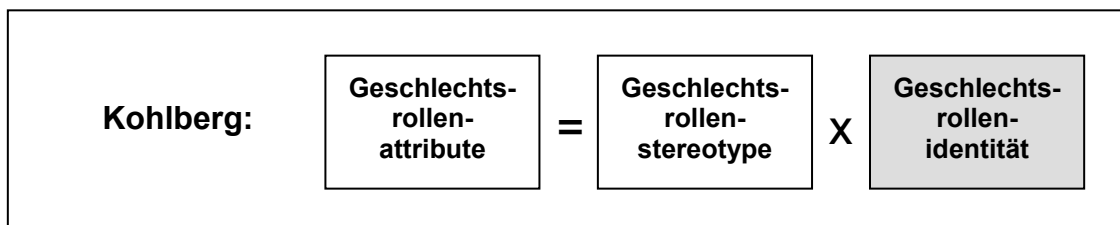


Abbildung 1. Kohlbergs (1966) Modell des Zusammenhangs von Geschlechtsrollenidentität, -attributen, und -stereotypen <sup>2</sup>

<sup>2</sup> adaptiert von Storms (1979), Kohlberg verwendet die Phrase 'sex' anstatt 'gender', z. B. 'sex role'

Individuen können demnach in Abhängigkeit der Veränderungen in den allgemeinen Geschlechtsrollenstereotypen aktiv ihre Geschlechtsrollenattribute beeinflussen (vgl. Kohlberg, 1966; Storms, 1979).

Im Gegensatz zu Kohlberg sah Kagan (1964) die Geschlechtsrollenidentität als Ergebnis eines Produktes aus den eigenen Geschlechtsrollenstereotypen und der Wahrnehmung von Geschlechtsrollenattributen (Abbildung 2).

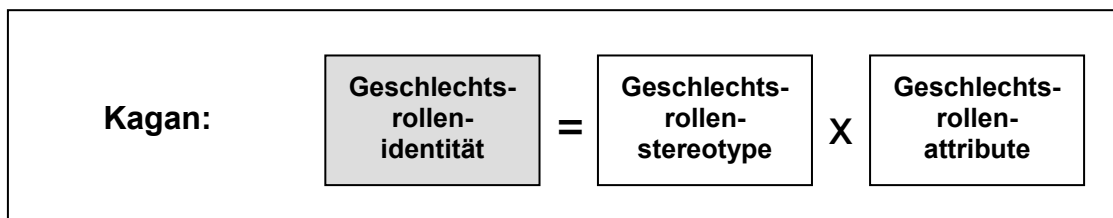


Abbildung 2. Kagans (1964) Modell<sup>3</sup> des Zusammenhangs von Geschlechtsrollenidentität, -attributen, und -stereotypen<sup>4</sup>

Demzufolge vergleicht eine Person die eigenen Attribute mit den allgemein vorherrschenden Geschlechtsrollenstereotypen, um relativ eigenständig zu einer realistischen Geschlechtsrollenidentität zu gelangen. Das Ausmaß, sich selbst als entweder maskulin oder feminin zu kategorisieren (*Geschlechtsrollenidentität*), hängt unter anderem vom Übereinstimmen der kulturellen Geschlechtsrollennormen (*Stereotype*) mit den eigenen Attributen ab.

Kagan betonte in ähnlicher Weise wie Kohlberg die aktive Rolle eines Individuums, dessen Motivation unter anderem darin liegt, sich mit einem Modell zu identifizieren. Dabei basiert diese Motivation vor allem auf dem Verlangen, die gleichen attraktiven Ziele (wie beispielsweise Macht, Liebe anderer) anzustreben, die das Modell bereits erreicht hat. Kagans Konzept zeigt dabei auch einleuchtend seine Beziehung zu den sozialen Lerntheorien (Abbildung 3), die ebenfalls die Geschlechtsrollenidentität zum Produkt verschiedener Einflüsse machen: die Bindung an und Identifikation mit der belohnenden Person.

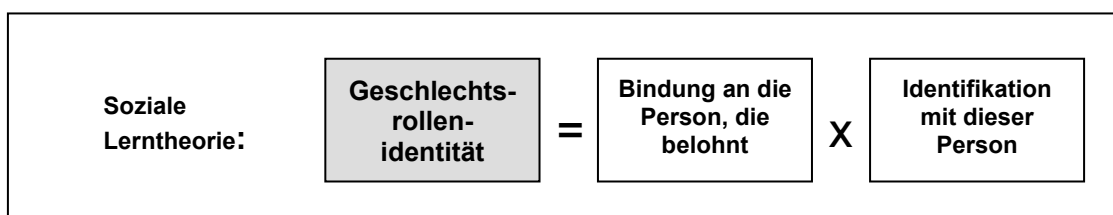


Abbildung 3. Das Modell der sozialen Lerntheorie über den Zusammenhang von Geschlechtsrollenidentität, -attributen, und -stereotypen

Kagans als auch Kohlbergs Theorie sagen beide eine enge Beziehung zwischen der Geschlechtsrollenidentität, -stereotypen und -attributen vorher, sind sich aber über die Kausalität dieser Beziehung uneinig. Verschiedenste Studien haben den Versuch unternommen haben, diese Kausalität näher zu untersuchen und beide Modelle gegeneinander zu testen (Spence, Helmreich, & Stapp, 1975; Storms, 1979). Dazu wurden beispielsweise in einer Studentenstichprobe sowohl das eigene Geschlechtsrollenselbstkonzept mittels des PAQs, die

<sup>3</sup> Kagan diskutiert *sex role standards*

<sup>4</sup> adaptiert von Storms (1979), Kagan verwendet die Phrase *sex* anstatt *gender*, z. B. *sex role*

Geschlechtsrollenidentität, als auch Stereotype hinsichtlich Männern und Frauen verglichen (Storms, 1979) mit leicht unterschiedlichen Ergebnissen und nachgewiesenen Plausibilitäten in beiden Theorien. Diese fehlende Übereinstimmung in der Präferenz des einen über das andere Modell wertet jedoch keinesfalls die Bedeutung der kognitiven Theorien ab, die zumindest unterschiedliche Einflüsse in der Entwicklung in Betracht ziehen.

Als Fazit sollten sowohl biologische als auch psychologische Faktoren in der Entwicklung und Erklärung der Geschlechtsrolle und den daraus folgenden Geschlechtsunterschieden Berücksichtigung finden (Abbildung 4). Während der Einfluss des biologischen Geschlechts durch genetische Determination und Stoffwechsel- sowie hormonelle Abläufen gegeben ist, erscheint die Rolle der kognitiven Faktoren jedoch weniger eindeutig geklärt. Im Zusammenwirken mit den biologischen Faktoren und den umgebenden sozialen Einflüssen, die ihrerseits in verschiedenen sich beeinflussenden Beziehungen stehen, sind sie jedoch für die Entwicklung der Geschlechtsidentität mitverantwortlich.

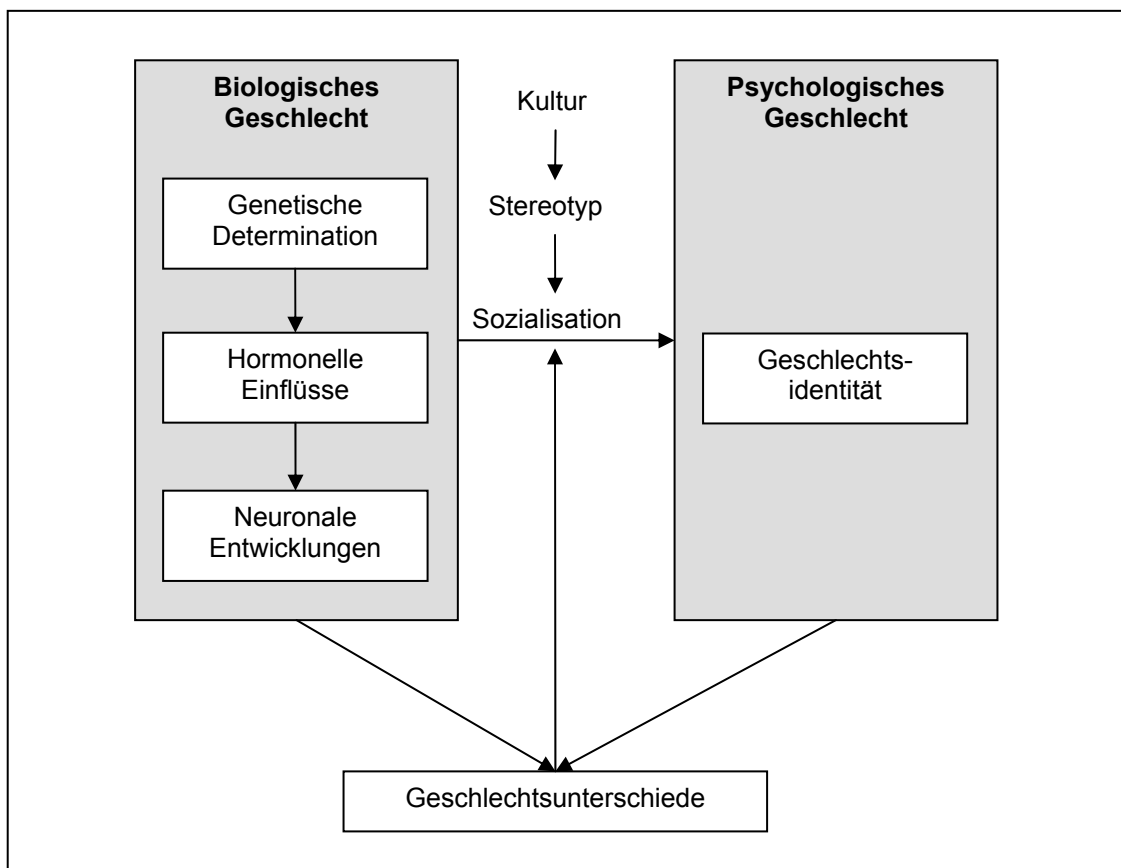


Abbildung 4. Übersicht des Zusammenhangs von biologischem und psychologischem Geschlecht<sup>5</sup>

#### 2.2.3.4 Geschlechterschematheorie / Geschlechtstypisierung (Bem)

Die Theorie der Geschlechtsschemata stellt einen kognitiven Ansatz zur Geschlechtsrollenstereotypisierung dar und basiert auf der Geschlechtstypisierung von

<sup>5</sup> Sofern Abbildungen in der Arbeit nicht mit einer Quelle versehen sind, stellen sie eine eigenständige Zusammenfassungen der berücksichtigten Variablen oder Konstrukte dar.

Männern und Frauen und wie Männer und Frauen zu maskulinen oder femininen Individuen mit geschlechtsentsprechenden Verhaltensweisen sozialisiert werden (Hargreaves, 1986). Die Betonung kognitiver Prozesse nimmt Bezug auf die Entwicklung eines Netzwerkes, das sich um die Konzeption von Maskulinität und Femininität organisiert. Dabei ist von individuellen Unterschieden in der Stärke und Verfügbarkeit dieser geschlechtsbezogenen Verknüpfungen auszugehen. Einige Individuen verfügen über ein großes, verzweigtes Netzwerk Stereotyp-assoziiierter Verknüpfungen, die sozial und kulturell geprägt wurden und auf zukünftige Informationsverarbeitung und Verhalten Einfluss nehmen – sog. Schemata (Bem, 1981b; Bem, 1984; Neisser, 1976; Pleck, 1975).

Bezogen auf die Geschlechterrollen bestehen diese kognitiven Strukturen aus geschlechtsspezifischen Assoziationen, die daher auch als Geschlechterschema bezeichnet werden; das Geschlecht fungiert als ein Organisationsprinzip, ein mentales Schema, welches Personen individuell in unterschiedlicher Art in ihre Informationsverarbeitung integrieren. Dabei ist laut Bem (1981b) die Art und Weise von Individuen geschlechtsspezifisch zu agieren oder andere Personen gemäß ihres Geschlechts zu betrachten und ihnen dementsprechend zu begegnen (*sex typing*) weniger das Resultat spezifischer individueller Persönlichkeitseigenschaften, sondern vielmehr der Tendenz, die Welt entsprechend der Kategorien *männlich* und *weiblich* zu kategorisieren. Geschlechtsschematische Personen wenden demzufolge die Kategorie Geschlecht im umfassenden Maße an: bezogen auf sich selbst, auf andere Personen und sogar auf eher abstrakte Konzepte. Lippa (2001) interpretierte Bems Postulierung als eine Verschiebung des Fokus von individuellen Charakteristiken zu der gesellschaftlichen Tendenz, das Geschlecht als zentrale und saliente Kategorie zu betrachten.

Bem (1981b) konnte zeigen, dass geschlechtstypische Personen (starke Beanspruchung geschlechtseigener Attribute) – im Vergleich zu Personen, die z. B. mehr geschlechtsuntypische Eigenschaften für sich beanspruchen – im größeren Maße Informationen geschlechtsspezifisch clusterten, schneller enkodierten und abriefen sowie schnellere schema-konsistente Entscheidungen das Selbst betreffend fällten. Der Geschlechtstypus einer Person bestimmt demnach die schematische, geschlechtsbasierte Verarbeitung sowie die individuelle *“generalized readiness to process information on the basis of the sex-linked associations that constitute the gender schema”* (Bem, 1981b, S. 354).

Es gilt als abgesichert, dass sich geschlechtstypisierte Individuen und nicht-geschlechtstypisierte Individuen darin unterscheiden, wie sie ihr Selbstkonzept um das Geschlechtsschema organisieren. Für eine geraume Zeit herrschte die generelle Annahme vor, dass ein hohes Maß an Geschlechtstypisierung in einem positiven Zusammenhang mit psychologischer Anpassung und Gesundheit steht. Im Verlauf der Zeit wurde diese Idee zunehmend kontrovers diskutiert und herausgefordert. Williams und Best (1990) kamen über eine Zusammenfassung mehrerer Studien zu dem Ergebnis, dass sich für alle Theorien Beweise finden lassen, keine jedoch allein die Unterschiede von Wirkung der Geschlechtstypizität erklären könne. Unabhängig davon scheint das Konzept der Androgynität bezüglich der Definition eines Standards psychologischer Gesundheit anderen Ansätzen



gegenüber im Vorteil zu sein. Androgyne Personen gelten als weitgehend besser angepasst und verfügen – im Gegensatz zu geschlechtstypisierten Peers in vergleichbaren Situation – über eine größere Flexibilität im Verhalten (Bem, 1974; 1977; 1984; Bem & Lenney, 1976; Bem, Martyna, & Watson, 1976); einen höheren Selbstwert (Spence & Helmreich, 1978); höhere Leistungsmotivation (Bierhoff-Alfermann, 1989); Führungsfähigkeiten (Kent & Moss, 1994), bessere persönliche und soziale Anpassung (Bem & Lenney, 1976; Harris & Schwab, 1979; 1990) oder das subjektive Gefühl des emotionalen Wohlbefindens (Lubinski, Tellegen, & Butcher, 1983) und größere Zufriedenheit in sozialen Beziehungen (Heilbrun, 1981). Neben gemischten Ergebnissen (Burchardt & Serbin, 1982; Cook, 1985; Lubinski, Tellegen, & Butcher, 1981; Lubinski et al., 1983; May & Spangenberg, 1997; O'Heron & Orlofsky, 1990; Paulhus & Martin, 1988; Skoe, 1995; Wiggins & Holzmuller, 1981; Wulff & Steitz, 1999) wurden insgesamt vielfältige Zusammenhänge von Androgynität zu positiven Konstrukten und Verhaltensflexibilität gefunden (Antill, 1983; Campbell, Steffen, & Langmeyer, 1981; Cheng, 1999; Green & Kenrick, 1994; Kirchmeyer, 1996; Rose & Montemayor, 1994; Sawrie, Watson, & Biderman, 1991; Shaver, Papalia, Clark, & Koski, 1996; Shimonaka, Nakazato, Kawaai, & Sato, 1997; Stake, 1997; Wubbenhorst, 1994).

Bereits Kinder unterscheiden demnach sich in der Art ihrer Schemata: einige verfügen über stärker entwickelte Schemata als andere, sie denken in stärkerem Maß geschlechtsschematisch, verstehen die Welt hinsichtlich geschlechtstypischer Themen besser und erinnern geschlechtskonsistente Informationen eher als geschlechtsinkonsistente (Stangor & Ruble, 1987). Es gilt jedoch zu beachten, dass Unterschiede zwischen den Gruppen der unterschiedlichen Verarbeitung geschlechtsrelevanter Informationen bisweilen nicht nur sehr gering ausfielen, sondern auch von anderen Studien nicht repliziert werden konnten, was ihren Wert einschränkt.

Bem (1981b) selbst machte zudem später noch einmal deutlich, dass ihr Messinstrument – Bems Sex Role Inventory (BSRI), welches an späterer Stelle in dieser Arbeit ausführlicher dargestellt wird – nicht per se die geschlechtsschematische Verarbeitung messen kann, sondern zur Identifizierung kognitiv unterschiedlich verarbeitender Individuen dient. Obwohl frühere kognitive Entwicklungstheorien wiederholt die Bedeutung der Geschlechtskonstanz thematisiert haben, wurde nun argumentiert, dass Kinder in der Lage seien, Geschlechtsinformationen zu kategorisieren (sie entwickeln Schemata), die ihre eigenen Präferenzen und Aktivitäten beeinflussen, sogar bevor sie das Geschlecht als Konstante akzeptieren.

Bems Theorie (1981b) zählt zwar zu einer der populärsten in der Geschlechterforschung und dennoch ist Bem selbst einen wirklich klaren Beweis schuldig geblieben. Andere empirische Nachweise in diesem Bereich gelten ebenfalls als unzureichend. So konnte bisher weder eindeutig gezeigt werden, dass Geschlechterschemata auch konstanten Einfluss auf die Verarbeitung geschlechtsspezifischer Informationen nehmen, noch dass entsprechende Messinstrumente des Geschlechtsrollenselbstkonzepts zu dieser Informationsverarbeitung einen adäquaten Zusammenhang herstellen können (Archer & Lloyd, 2002; Archer, Smith, &

Kilpatrick, 1995; Deaux, Kite, & Lewis, 1985; Edwards & Spence, 1987; Kite & Deaux, 1986; Payne, Connor, & Colletti, 1987; Ruble & Stangor, 1986).

Archer und Lloyd (2002), die wie Lippa (2001) die Nachweise für die Theorie bemängelten, stellten jedoch fest, dass die Messinstrumente womöglich zu einseitig in ihren Inhalten sind, um den Nachweis in dieser Form überhaupt erbringen zu können. So hatten bereits sowohl Spence und Helmreich (1981) als auch Deaux und Lewis (1984) angemerkt, dass der BSRI nur anteilig die Konzepte Maskulinität und Femininität abbildet. Dabei sind nicht zuletzt die Konstrukte und auch die kognitiven Assoziationen geschlechtsschematischer Personen um einiges komplexer. In Ermangelung umfassender Messinstrumente wiederum kann die Komplexität der kognitiven Informationsverarbeitung dadurch nur schwerlich eindeutig nachgewiesen werden.

Bems Theorie findet sich im weiteren Sinne jedoch auch in späteren Ansätzen wieder. So vermutete Hannover (2000) ähnliche Mechanismen des Geschlechtsbezuges von Informationen bei der Entwicklung des Selbst. Demnach ist das geschlechtsbezogene Selbstkonstrukt vor allem zwei permanenten Einflüssen ausgesetzt: zum einen den geschlechtsbezogenen Informationen der momentanen Situation und zum anderen der kollektiven geschlechtsbezogenen Erfahrungen aus sozialen Situationen aus dem Leben eines jeden Individuums. Hannover (2000) diskutierte beim Erwerb der Geschlechterstereotype ebenfalls die Differenzierung in geschlechtskongruentes und geschlechtsinkongruentes Geschlechterwissen. Dem Ansatz folgend, integrieren Kinder zu Beginn des Erwerbs von Geschlechterstereotypen vor allem geschlechtskongruentes Wissen in das eigene Selbst, was insgesamt betrachtet leichter kognitiv verfügbar ist. Mit fortlaufender Entwicklung wird schließlich auch geschlechtsinkongruentes Wissen in das Selbst aufgenommen und anerkannt, dass bei Frauen und Männern keine Ausschließlichkeit in der Verteilung hinsichtlich spezifischer Persönlichkeitseigenschaften, Interessen oder Aktivitäten existiert.

#### **2.2.3.5 Soziale Identität**

Im Hinblick auf das Selbst und die Definition des eigenen Ichs ist die soziale Identität eine bedeutende Theorie, die auch im Zusammenhang zum Geschlecht eine wichtige Rolle spielt und die Beziehung zwischen Geschlechtsidentifikation und sozialem Verhalten erklärt. Geschlecht fungiert dabei sowohl als soziale Kategorie als auch als persönliches Attribut und die Geschlechtsidentität beinhaltet nicht nur die eigene Klassifikation als männlich und weiblich, sondern darüber hinaus auch den Prozess der Identifizierung des Selbst mit denen, die derselben Kategorie angehörig sind (Trew, 1998). Ursprünglich als Theorie des Intergruppenverhaltens entwickelt und später durch die Selbstkategorisierungstheorie erweitert, kann Tajfels soziale Identitätstheorie (1969; 1978) ebenfalls bezüglich des Geschlechts angewendet werden. Demnach ist das Geschlecht lediglich nur eine von vielen sozialen Kategorien, die Macht- und Statusverhältnisse zwischen den Gruppen herstellt und zur individuellen Selbstdefinition herangezogen wird. Durch die eigene Kategorisierung als Mann oder Frau werden Ähnlichkeiten zwischen der eigenen Person und den anderen Mitgliedern der

Geschlechtergruppe akzentuiert, während gleichzeitig die Unterschiede zwischen der Eigen- zur Fremdgruppe betont und bisweilen sogar übertrieben werden (Trew, 1998).

Die Theorie der sozialen Identität kann durchaus zur Erklärung der Beziehungen zwischen Männern und Frauen und dem Einfluss auf die Selbststereotypisierung herangezogen werden – vor allem in solchen Kontexten, in denen das Geschlecht salient ist (Deaux & LaFrance, 1998; Swan & Wyer, 1997). Swan und Wyer (1997) hatten zeigen können, wie die soziale Identitätstheorie Auswirkungen der Salienz<sup>6</sup> des Geschlechts auf die eigene Identität erklären vermag. Dazu variierten sie die Experimentalsituationen für die Versuchspersonen derart, dass Minoritätenstatus über die Anzahl der Eigen- und Fremdgruppenmitglieder hergestellt wurde (Gruppe 1: ein Mann und mehrere Frauen; Gruppe 2: eine Frau und mehrere Männer; Gruppe 3: gleiche Anzahl Männer und Frauen). In der anschließenden Befragung sollten die Probanden sich hinsichtlich Geschlechterstereotype selbst einschätzen und Urteile über Personen abgeben, deren Geschlecht in gegebenen Vignetten uneindeutig gehalten wurde. Die Ergebnisse zeigen, dass der Minoritätenstatus die Wahrnehmung des Geschlechts als Kategorie aktiviert, die stereotypisch mit den Geschlechtskategorien verbunden sind und so auch die Interpretation des Verhaltens anderer beeinflussen kann. Die Resultate deuten an, dass sich Männer aufgrund der Wahrnehmung des eigenen Geschlechts im Vergleich zu Frauen ihres hohen sozialen Status' bewusst wurden, was die Konzepte und Attribute zur Selbstbeschreibung aktivierte, die die Mitgliedschaft in dieser Kategorie auch zu bestätigen vermögen. Frauen hingegen sehen der Herausforderung entgegen, der sozial niederen Gruppe anzugehören. Die Wahrnehmung dessen motiviert sie – im Vergleich zu Männern – die Konzepte zu bedienen und heranzuziehen, die sie maximal von Mitgliedern der eigenen Gruppe unterscheidet.

### **2.2.3.6 Rollentheorie des GSK**

Aufbauend auf dem biologischen Standpunkt, dass die elterliche „Investition“ (Stillen, Kindererziehung) für Frauen oftmals höher ist als für Männer, oder dass Männer in stärkerem Maße mit dem Umwerben zukünftiger Partner/innen und dem Bereitstellen von Essen beauftragt sind, argumentieren vor allem funktionale Ansätze, dass die stereotypen Unterschiede zwischen Männern und Frauen lediglich die unterschiedlichen Funktionen beider Geschlechter widerspiegeln (Buss & Kenrick, 1998; Parsons & Bales, 1955). Sozialpsychologische Geschlechterforschung betrachtet das Geschlecht als eine bedeutende soziale Rolle (Eagly, 1987) oder als eine soziale Kategorie (Deaux, 1984).

Soziale Rollen sind kulturell definierte soziale Positionen und können laut Gross, Mason und McEachern (1958) mit der Summe von Erwartungen gleichgesetzt werden, die andere Personen an den Inhaber einer Position herantragen. Diese Erwartungen haben normativen Charakter (Stryker & Statham, 1985). Auch *Geschlechtsrollen* sind mit Verhaltens- und Eigenschaftserwartungen gekoppelt (vgl. Alfermann, 1996). Unterschiedliche Erwartungen an

---

<sup>6</sup> Salienz bedeutet, dass ein Reiz (z. B. ein Objekt oder eine Person) aus seinem Kontext hervorgehoben und dadurch dem Bewusstsein leichter zugänglich ist als ein nicht-salienter Reiz. Geschlecht ist nicht per se ein salientes Merkmal, kann aber vor entsprechendem Hintergrund salient werden, wie beispielsweise ein einzelner Mann in einer Gruppe von Frauen.

Männer und Frauen resultieren nur zum Teil aus biologischen Gegebenheiten. Vorwiegend liegt die Ursache darin, dass in den meisten Gesellschaften eine mehr oder minder klare Arbeitsaufteilung zwischen den Geschlechtern besteht, auf die die gesamte Sozialisation abgestimmt wird (Eagly, 1987).

Soziale Rollentheorien führen Geschlechtsunterschiede in den Eigenschaften allgemein eher auf die geschlechtsspezifischen (Arbeits-)rollen zurück, die beide Gruppen ausfüllen (Eagly, 1987; Wood & Eagly, 2002): Eagly und Kollegen (Eagly, 1987; Eagly & Kite, 1987; Eagly & Steffen, 1984; 2000; Eagly et al., 2000) gingen davon aus, dass ganz allgemein gruppenstereotype Charakteristiken stark mit den Aktivitäten zusammenhängen, die mit den sozialen Rollen der Gruppe typischerweise assoziiert sind. Gerade bei Geschlechterstereotypen ist es demnach nicht das biologische Geschlecht per se, welches zu stereotypen Beschreibungen von Männern und Frauen führt, sondern die Rollen, die Frauen und Männer in der Gesellschaft füllen und ausüben sowie sich daraus ergebende Statusunterschiede zwischen beiden Geschlechtern. So konnten Eagly und Steffen (1984) auch demonstrieren, dass Geschlechterstereotype stark mit der individuellen Vorstellung der traditionellen sozialen Rollen von Männern und Frauen verknüpft sind. So führen Rollenunterschiede zwischen den Geschlechtern dazu, dass Frauen vor allem *communal* (ähnlich sozialer) und Männern *agentic* (ähnlich aktiver) Eigenschaften zugeschrieben werden.

Bemerkenswert ist dabei, dass Conway, Pizzamiglio und Mount (1996) zeigen konnten, wie wenig das eigentliche Geschlecht mit der Zuschreibung von weiblichen und männlichen Eigenschaften zu tun hat, und wie viel tatsächlich auf soziale Rollenunterschiede zurückgeführt werden kann. In einer Serie von Studien zu Berufen oder Beschreibungen fiktiver Kulturen fanden sie einen Zusammenhang zwischen Wahrnehmungen höheren vs. niederen Status der Zielgruppe und deren Bewertung hinsichtlich *communiality* und *agency*, sogar unabhängig des biologischen Geschlechts. Männern, die beispielsweise in hochdotierten und -angesehenen Positionen tätig sind, werden mit größerer Wahrscheinlichkeit *agentic* Eigenschaften, die in diesem Kontext relevant und notwendig sind, zugeschrieben. Das Ausüben einer statusniederen Anstellung wurde dementsprechend eher mit *kommunalen* Eigenschaften in Bezug gebracht.

Ähnliche Wirkungen auf die Geschlechtsunterschiede konnten in Forschungen bezüglich des sozialen Einflusses, sprachlicher Geschlechtsunterschiede oder Reaktionen auf geschlechtsuntypisches Verhalten ebenso gezeigt werden (Carli, 1990; Eagly, 1983; Eagly & Wood, 1982; Feinman, 1981; 1984; Lockheed, 1985; Meeker & Weitzel-O'Neill, 1977). Die soziale Identität eines Individuums, das Wissen darum, Mitglied einer bestimmten Kategorie oder Gruppe zu sein, sowie die Identifikation mit dieser Gruppe, beeinflussen die Wahrnehmung bezüglich des eigenen Selbst, Einstellungen und Verhalten (Abrams & Hogg, 1990; 2004; Hogg & Abrams, 1988). Dabei ist es möglich, dass jedes Individuum eine Vielzahl an sozialen Identitäten hat. Die Salienz einer spezifischen Kategorie sowie die individuelle Identifizierung mit eben dieser Kategorie aktiviert die Selbstkategorisierung und gruppennormatives Verhalten. Individuelle Bedürfnisse, Einstellungen und Verhaltensweisen werden zugunsten der

Gruppennorm zurückgestellt, in Folge dessen die Selbstbeschreibung dem Gruppenprototypen angeglichen wird (Hogg & Turner, 1987).

Parsons und Bales (1955) beschrieben die männliche Rolle noch als die des Familienernährers, der aus der Familie hinaus in den Beruf geht und Geld verdient. Die weibliche Rolle dagegen ist die der Familienbetreuerin, die zuständig für alle familiären Angelegenheiten ist, das Wohlergehen der Familie und das Funktionieren der Innenbeziehungen. Diese Arbeits- und Machtaufteilung zwischen Mann und Frau entspricht aus heutiger Sicht eher einem traditionellen Rollenverständnis. Im Laufe der letzten Jahrzehnte hat sich in westlichen Gesellschaften das Rollenbild in Richtung einer angestrebten Gleichstellung von Mann und Frau verändert (Deaux & LaFrance, 1998).

Die sozialen Rollen gehen nicht nur kulturell definiertem Ursprungs hervor, sondern bestimmen darüber hinaus die normativen Erwartungen, die andere Personen an die Rolleninhaber haben (Gross, Mason, & McEachern, 1958; Stryker & Statham, 1985) und die Männer und Frauen an sich selbst stellen. Dem Ansatz liegt die Annahme eines Sozialisationsprozesses zugrunde, der zur Entwicklung geschlechtsspezifischer Eigenschaften führt. Dabei lässt die Rollentheorie insgesamt offen, inwiefern ein Endstatus der Geschlechterrollenentwicklung erreicht wird und unterstreicht damit den dynamischen Status von Geschlechterrollenstereotypen.

Das Geschlecht gilt als eine der einflussreichsten Sozialisationsbedingungen: es steuert Erwartungen an das Verhalten eines Individuums, beeinflusst die Informationsverarbeitung und nimmt damit im Kontext zum Selbst und seiner Konstruktion eine bedeutende Rolle ein. Verschiedene Forschungsansätze haben sich in diesem Zusammenhang intensiv mit dem Einfluss des Geschlechts auf die Selbstkonstruktion beschäftigt und dabei jedoch leicht widersprüchliche Ergebnisse gefunden. Hannover (1997) konnte beispielsweise zeigen, dass das Ausüben geschlechtstypischer Tätigkeiten (maskulin: Nägel einschlagen; feminin: Puppen wickeln) geschlechtliches Selbstwissen aktiviert und leichter verfügbar macht (schnellere Selbsteinschätzung maskuliner und femininer Eigenschaften in einem Reaktionszeitexperiment).

Betrachtet man das geschlechtsspezifische Beziehungsverhalten und die Annahme, dass Männer als unabhängig, autonom und wenig beziehungsorientiert gelten (Cross & Madson, 1997), während Frauen eine interdependente Selbstkonstruktion entwickeln, lässt sich die Geschlechtertrennung in dieser Klarheit nicht immer bestätigen (Gabriel & Gardner, 1999; Li, 2002). Viel eher deutet sich an, dass sehr wohl beide Geschlechter beispielsweise soziale Beziehungen führen, sich jedoch in der Art ihrer Auslebung unterscheiden (Gabriel & Gardner, 1999). In diesem Zusammenhang hat sich gezeigt, dass das biologische Geschlecht allein zur Diskussion von unterschiedlichen Ergebnissen und Selbstkonstrukten nicht ausreicht, da es bisweilen eine klassische Geschlechtertrennung impliziert und von der Geschlechtsrolle abweichendes Verhalten nur bedingt erklären kann. Stattdessen müssen auch mögliche Einflüsse der Geschlechterrollenidentität (Alfermann, 1996; Bem, 1974), die unabhängig von dem biologischen Geschlecht ist, zur Diskussion herangezogen werden.

Zusammenfassend kann resümiert werden, dass zum einen Rollentheorien viele Ähnlichkeiten zu klassischen Lerntheorien und speziell dem Modelllernen aufweisen. Zum anderen wirkt der Einfluss der unterschiedlichen Rollen vor allem auf die Entwicklung der vorherrschenden Geschlechtsstereotype. Weiterhin kann davon ausgegangen werden, dass aufgrund sozialer Identitätskonstruktionen, positive (wie auch negative) geschlechtskongruente Eindrücke und Verhaltensweisen einen größeren Einfluss auf das Selbst ausüben als Einflüsse, die vom anderen Geschlecht ausgehen. Daher wäre für Messinstrumente zu erwarten, dass geschlechtskongruente Attributen eher für das eigene Selbst reklamiert werden sollten als geschlechtsinkongruente.

### **2.2.3.7 Psychologische Anpassung – Androgynität**

Eine besondere Stellung gerade auch im Rahmen situativer Flexibilität des Geschlechtsrollenselbstkonzepts nimmt die Androgynität ein. Williams und Best (1990) unterschieden ursprünglich drei Theorien bezüglich psychologischer Anpassung und mentaler Gesundheit: das Modell der Sozialisation (*socialisation model theory*), die Theorie der Androgynität (*androgyny theory*) sowie die männliche Anpassungstheorie (*male adjustment theory*). Abhängig von der Theorie gelten Personen dann als am besten an Lebensumstände angepasst, wenn sie entweder eher den männlichen stereotypischen Charakteristiken entsprechen, geschlechtstypisiert oder androgyn sind.

Insbesondere der Androgynitätsansatz (Bem, 1974; Spence et al., 1975) hat durch gesellschaftliche Veränderungen der letzten Dekaden, daraus resultierenden Rollen- und Aufgabenerweiterungen für beide Geschlechter sowie Veränderung in den Selbstbildern von Männern und Frauen an Popularität gewonnen. Bem (1981b) vertrat in ihrer Geschlechterschematheorie die Auffassung, Individuen lernen gesellschaftlich determinierte Informationen entsprechend ihres Geschlechterschemas wahrzunehmen, zu enkodieren und organisieren. Das eigene Selbstkonzept wird im Laufe der Entwicklung diesem Schema angepasst und ein Individuum ist entweder maskulin oder feminin. Das Auftreten einer Ausprägung balanciert die jeweils andere Tendenz aus. Demgegenüber stehen ausgehend vom Androgynitätsansatz *androgyn* Menschen, die nicht geschlechtstypisch maskulin oder feminin in ihrer Selbstwahrnehmung sind, sondern beide Ausprägungen gleichermaßen in ihr Selbst integriert haben.

Im Gegensatz zu Bems Theorie legten Spence, Helmreich und Stapp (1975) der Erklärung von Androgynität ein additives Eigenschaftsmodell mit Maskulinität und Femininität als zwei unterschiedliche und unabhängige Reaktions-Prädispositionen zugrunde. Die Addition beider Ausprägungen führe demnach durch mehr Flexibilität zum Verhaltensvorteil. Androgyn Menschen würden über gleichermaßen expressive wie instrumentelle<sup>7</sup> Eigenschaften verfügen, die gut mit anderen Merkmalen wie Fähigkeiten, Werten oder Motiven korrespondieren, was ihnen diese größere Verhaltensflexibilität ermögliche.

---

<sup>7</sup> Von Spence und Kollegen verwendeten Bezeichnungen für die unterschiedlichen Geschlechtsrollen. Instrumentalität weist Ähnlichkeiten zu Maskulinität, während Expressivität eher der Femininität zugeordnet werden kann.

Larsen und Seidman (1986) überprüften den BSRI hinsichtlich der Entsprechung der Geschlechterschematheorie (Bem, 1981b). Ergebnisse von Faktoranalysen, getrennt für geschlechtstypische und -untypische Individuen, bestätigen dessen theoretische Fundierung. Während sich die bipolare Faktorenstruktur für geschlechtstypische Probanden finden ließ, erwiesen sich Maskulinität und Femininität der geschlechtsuntypischen Gruppen als unterschiedliche Faktoren oder als zu demselben uni-dimensionalen Faktor gehörend – was die Autoren darauf zurückführten, dass die geschlechtsspezifische Selbstbeschreibung für geschlechtsuntypische Personen eine untergeordnete Rolle spielt. Geschlechtstypische Individuen dagegen scheinen bei der Beantwortung der BSRI-Items die Relevanz der bipolaren geschlechtsbezogenen Konzeption von Maskulinität und Femininität zu berücksichtigen – die Wahrnehmung dieser Eigenschaften ist beeinflusst durch ein kognitives Geschlechtsschema.

Bems Versuche (Bem & Lewis, 1975), den androgynen Typ sowohl theoretisch als auch experimentell ausführlicher zu definieren, führten zu weiteren Erkenntnissen derart, dass androgyne Individuen im Vergleich zu geschlechtstypischen Individuen in konformen Situationen sowohl eine ‚maskuline‘ Unabhängigkeit als auch ‚feminines‘ Umsorgen zeigten.

Insgesamt bleibt die Androgynitätsforschung einen eindeutigen kausalen Beweis des direkten Zusammenhangs gleichstarker Ausprägung von Maskulinität und Femininität mit anderen Konstrukten wie Interessen, Karriereplanung, Berufs- oder Familienrollen und einer globalen Verhaltensflexibilität schuldig (Bierhoff-Alfermann, 1989). Das soll den Erkenntniswert dieses Konzepts nicht einschränken, sondern verdeutlicht die Vielfältigkeit verhaltensaustösender Situationen, Erwartungen oder sozialer Einflüsse und den mitunter geringen Beitrag des Selbstkonzepts. Die Ausführungen zeigen die Flexibilität der Geschlechtsrolle, sowohl über die Lebensspanne als auch über sich verändernde Situationen hinweg.

#### **2.2.4 Stabilität des Geschlechtsrollenselbstkonzepts**

Vor dem Hintergrund der Veränderungen von Geschlechtsrollenselbstkonzepten können vor allem drei große Ansatzpunkte betrachtet werden: langfristig wirkende intra-individuelle Veränderungen einer Person in Abhängigkeit von Entwicklungsaufgaben und -einflüssen; kurzfristige Veränderungen als unmittelbare Anpassungsreaktionen an situative Anforderungen sowie globale Veränderungen z. B. innerhalb einer Gesellschaft: Rice und Coates (1995) untersuchten die Veränderungen der Geschlechtsrolleneinstellung von Südstaatlern und Nichtsüdstaatlern stellten eine Entwicklung weg von den traditionellen Rollenbildern fest, die sowohl auf intra-individuelle als auch situative Veränderungen nachhaltig Einfluss nehmen können.

Für die vorliegende Arbeit ist in diesem Zusammenhang vor allem wichtig, dass das Geschlechtsrollenselbstkonzept insgesamt eine stabile Repräsentation eines Individuums darstellt, welches von Aspekten der Situation nahezu unbeteiligt bleibt. Dazu ist es zunächst notwendig, individuelle Entwicklungen über die Zeit zu betrachten, im Verlaufe derer das Geschlechtsrollenselbstkonzept als ein Ergebnis hervorgeht. Im Anschluss wird die situationsspezifische Flexibilität im Verhalten und der Selbstbeschreibung diskutiert. In einem

letzten Abschnitt erfolgt schließlich die ausführliche Auseinandersetzung mit kulturellen Veränderungen und deren Einfluss auf das Geschlechtsrollenselbstkonzept im Wandel der Gesellschaft und vor allem vor dem Hintergrund unterschiedlicher Kulturen und daraus resultierender unterschiedlicher Repräsentation in Messinstrumenten.

#### **2.2.4.1 Individuelle Entwicklungsveränderungen über die Zeit**

Die Stabilität von Eigenschaften und dem Selbst beschäftigt klassischerweise nicht nur Persönlichkeitspsychologen (Asendorpf, 1992), sondern hat in allen psychologischen Teildisziplinen signifikante Bedeutung im Zuge der Betrachtung (intra-)individueller Unterschiede über Situationen und Zeiträume hinweg. Generell kann die Stabilität vor allem unter zweierlei Gesichtspunkten in Frage gestellt werden. Dabei spielen zum einen situative Anforderungen an Personen eine Rolle, die ursprünglich als stabil geltende Eigenschaften anders herausfordern und zu Veränderungen im Erleben und Verhalten führen können. Dies ist die eigentliche Flexibilität, von der im Zusammenhang mit Persönlichkeitseigenschaften mitunter gesprochen wird. Zum anderen gilt die Stabilität aber vor allem dann auch nur als bedingt, wenn man in Betracht zieht, dass das Selbst im Laufe der persönlichen Entwicklung an Komplexität zunimmt. Das Erleben neuer Situationen führt zu Wissenszuwachs und Ausdifferenzierung von vorhandenen Wissensstrukturen. So werden zum Beispiel während dieses Prozesses neue Selbst-Definitionen in das eigene Selbst integriert, wie beispielsweise von Hannover (2000) argumentiert. An dieser Stelle wird auch deutlich, dass gerade diese Flexibilität auch der Geschlechtsrollenkomponente zugrunde liegt und eine stabile Messung zur Herausforderung werden kann, zieht man Bems Geschlechterschematheorie oder die situations- oder rollenspezifischen Selbstzuschreibungen (Athenstaedt, Heinzle, & Lerchbaumer, 2008; Deaux, Winton, Crowley, & Lewis, 1985; Eckes, 1994a; Roberts & Donahue, 1994) in Betracht.

Verschiedenste Forschungsansätze (Costa & McCrae, 1988; Helson, Kwan, John, & Jones, 2002; McCrae & Costa, 1994) kamen zu dem Schluss, dass die Persönlichkeit bis zum 30. Lebensjahr entwickelt ist. So wird postuliert, dass Persönlichkeitseigenschaften das Ergebnis biologisch definierter Temperamentsunterschiede sein könnten, damit weniger den Umwelteinflüssen unterliegen und eine gewisse Stabilität erreichen. Kontrastierend dazu können Veränderungen jedoch auch im höheren Alter noch stattfinden, auch wenn diese Veränderungen immer weniger wahrscheinlich werden (Caspi & Roberts, 2001; Helson et al., 2002; Roberts & DeVecchio, 2000).

Caspi und Kollegen (Caspi & Roberts, 2001; Caspi, Roberts, & Shiner, 2005) diskutierten in diesem Zusammenhang Faktoren, die über die Kontinuität vs. Veränderungen der Persönlichkeit Einfluss nehmen können und betonten dabei die Relevanz von Veränderungen von Lebensumständen und sozialen Rollen und deren unterschiedlichen Einfluss auf die Persönlichkeit zu verschiedenen Punkten im Leben, was sich sowohl in Zu- als auch Abnahme von Eigenschaftszuschreibungen widerspiegeln kann.

Insgesamt zeigt sich wiederholt, dass Veränderungen vor allem im jungen Erwachsenenalter stattfinden und nahezu im gleichen Ausmaß für Männer und Frauen. So



bleiben Mittelwertsunterschiede zwischen den Geschlechtern erhalten (Caspi et al., 2005). Trotz der Anpassungsfähigkeit und Plastizität der Persönlichkeit bis ins hohe Alter überwiegt eine gewisse Konsistenz der Persönlichkeit, die auch die Interaktion mit der Umwelt beständiger macht.

Die Forschung bezüglich der Geschlechtstypisierung fokussiert vor allem das Säuglingsalter und die frühe Kindheit. Obwohl keine klare Einigkeit über das exakte Alter der Fähigkeit zur Geschlechterstereotypisierung besteht, ist allgemein akzeptiert, dass mit steigendem Lebensalter die Kenntnis stereotyper Eigenschaften wächst und mit dem Grundschulalter auch weitestgehend abgeschlossen ist (Spence, Deaux, & Helmreich, 1985). Bereits junge Kinder können weibliche Eigenschaften wie *schwach* oder *weichherzig* und als typische männliche Eigenschaften *stark* oder *aggressiv* nennen. Erstaunlich ist, dass diese Fähigkeiten und dieses Wissen – einmal entwickelt – über die Lebensspanne verhältnismäßig konstant sind und auch Gemeinsamkeiten über Nationen hinweg nachgewiesen werden können (Williams & Best, 1990).

Auch Maccoby und Jacklin (1974) argumentierten, dass die Entwicklung der Geschlechtsidentität früh im Verlauf des Lebenszyklus ihren Anfang nimmt und mit dem Erreichen der Adoleszenz als mehr oder weniger abgeschlossen gilt. Kinder erlernen bereits im frühen Alter den Unterschied zwischen Jungen und Mädchen und was es bedeutet, einer von beiden Kategorien anzugehören (Brannon, 2005; Martin, 1991; Martin, 1993; Martin, Wood, & Little, 1990). Alfermann (1996) setzte den Beginn kindlicher Fähigkeiten bedeutende Geschlechtsunterschiede wahrzunehmen bei etwa zwei Jahren an. In diesem Zeitraum lässt sich beobachten, dass und wie Kinder sich selbst und auch anderen die Etiketten männlich oder weiblich zuschreiben können. Dabei tauchen in kaum einem Bereich des Verhaltens so früh Geschlechterstereotype auf, wie man sie bei den Spielpräferenzen findet. Mit steigender Sprachflexibilität und dem allgemeinen Abstraktionsniveau lassen sich etwas später in der Entwicklung schließlich auch Vorstellungen über stereotype Eigenschaften nachweisen.

Bischof-Köhler (2006) stellte fest, dass bereits junge Kinder über Stereotype verfügen, die vor allem in Hinblick auf Unterschiede zum anderen Geschlecht betont werden, wenn die eigene Gruppe günstig abschneidet – während die Fremdgruppe diesen Unterschied oftmals noch nicht reflektieren kann. Im Verlauf der Entwicklung werden positive Stereotype einer Gruppe auch von der Fremdgruppe anerkannt, ebenso wie eigene negative Eigenschaften. Mit zunehmendem Alter werden die Einschätzungen zwar immer noch polarisiert, der zufolge das eigene Geschlecht positiver bewertet wird, während gleichzeitig jedoch durchaus eine Tendenz zu realistischeren Einschätzungen festzumachen ist. Auch Ullian (1976) nahm mit ihrem Modell der sechsstufigen Entwicklung der Geschlechtsrolle einen fortwährenden Prozess an und konnte zeigen, dass Überzeugungen und Ansichten über Maskulinität und Femininität auch mit zunehmendem Alter und fortschreitender kognitiver und sozialer Entwicklung noch veränderbar sind.

Studien unterschiedlicher Art zeigten ebenfalls, dass Kinder bereits im frühen Alter zwischen den biologischen Geschlechtern diskriminieren können (Brannon, 2005; Kite et al.,

2008; Martin, 1991; Martin, 1993; Martin et al., 1990), Wissen um die eigene Geschlechtskategorie besitzen und stereotype Einschätzungen hinsichtlich Zielpersonen mit gleichen geschlechtsspezifischen Spielinteressen im stärkeren Maße vornehmen (Martin, 1991; Martin, 1993; Martin et al., 1990). Martin und Kollegen gingen von einer Entwicklung in mehreren Stufen bis hin zur Adoleszenz aus: erst in Bezug auf das eigene Geschlecht und dann in Bezug auf das Andere, wodurch sich nicht nur die Art der Assoziation ändert, sondern auch die Qualität.

Andere Ansätze sehen keine Begrenzung in der geschlechtsbezogenen Entwicklung eines Menschen. So konnte nachgewiesen werden, dass sich stereotype Einstellungen und Geschlechtsrollen auch noch in der Adoleszenz und sogar darüber hinaus verändern, wie beispielsweise durch den Einfluss von Arbeitsbedingungen und -erfahrungen (Abele, 2007; Eagly, 1987). Die Bedeutung des Geschlechts – und dem, was Femininität und Maskulinität definiert und ausmacht – erfährt mit zunehmendem Alter Veränderungen (Huyck, 1996; Sedney, 1985; Strough, Leszczynski, Neely, Flinn, & Margrett, 2007).

Während eine andauernde Entwicklung von geschlechtstypischem Wissen mittlerweile als akzeptiert gilt, bleiben genauere Aussagen über die Spezifizierung und Richtung dieser Entwicklungen jedoch schwierig. Einige Studien präsentierten beispielsweise Resultate, die z. B. generell eine zunehmende Femininität mit steigendem Alter zeigen (Hyde, Krajinik, & Skoldt-Niederberger, 1991; Strong, 1943; Terman & Miles, 1936), während andere Studien zu dem Ergebnis kommen, dass vor allem typische Charakteristiken des anderen Geschlechts erst mit dem Alter in der Definition der eigenen Geschlechtsrolle von Männern und Frauen zunehmend bedeutender werden (Fischer & Leonard, 1981). Diesbezüglich lassen sich die Ergebnisse von Hyde et al. (1991) sehr gut einordnen, die in einer Längsschnittstudie eine Zunahme von androgynen Frauen verzeichnen konnten, was die Vermutung zulässt, dass Frauen zunehmend zu ihrer eigenen femininen Geschlechtsrolle womöglich auch vermehrt maskuline Eigenschaften in ihr Bild von sich selbst integrieren. Die Studien bezüglich der Geschlechtsrollenentwicklung lassen zwar eine eindeutige Linie in den Ergebnissen und der Argumentation vermissen, deuten aber an, dass eine kontinuierliche Entwicklung der eigenen Geschlechtsrolle weit über die Kindheit hinaus stattfindet.

Viele Forscher haben soziokulturelle Veränderungen und deren Einfluss auf Geschlechtsrollen untersucht, indem verschiedenste Geburtskohorten miteinander verglichen wurden. Helmreich (1982) und Spence (1985) führten diesbezüglich jedoch an, dass gesellschaftliche Veränderungen nicht notwendigerweise zu Veränderungen in den Stereotypen führen müssen, da die Eigenschaften stabile, internalisierte Eigenschaften sind, die verhältnismäßig unabhängig von allgemeinen sozialen Veränderungen beider Geschlechter sind („*quite stable clusters of internalized traits that appear to be relatively independent of social changes in the status of the two sexes*“ Helmreich, S. 161).

Strough, Leszczynski, Neely, Flinn und Margrett (2007) untersuchten altersbezogene Unterschiede in Selbstberichten von Geschlechterstereotypen vor dem Hintergrund möglicher einflussnehmender geschichtlicher Ereignisse. Sie konnten zeigen, dass die älteren Frauen

weniger wahrscheinlich maskuline und androgyne Eigenschaften für sich beanspruchten als dies z. B. jüngere Frauen taten. Demgegenüber zeigte sich bei den Männern ein leicht anderes Bild. Vor allem ältere Männer zogen im stärkeren Maße androgyne Eigenschaften zur Selbstbeschreibung heran als jüngere Mitglieder ihrer Geschlechtergruppe. Diese Studie zur Geschlechtsentwicklung über die Lebensspanne verdeutlicht den Einfluss historischer Entwicklungen (z. B. Feminismus-Bewegung) nicht nur auf allgemein veränderte Möglichkeiten im Leben, sondern darüber hinaus langfristig auch auf die individuelle Wahrnehmung und Entwicklung des eigenen Selbst und zeigt auf, wie schwierig ein systematisches Verständnis von altersbezogenen Unterschieden hinsichtlich des Geschlechts unter diesen Umständen ist.

Schwierigkeiten der Messung mit Selbstberichten sind durchaus bekannt und Bem (1974; Bem & Lewis, 1975) diskutierte eine gewisse Flexibilität in den Selbstberichten nicht nur vor dem Hintergrund des Geschlechtsrollenselbstkonzepts, sondern erweiterte methodisch den Rahmen, in dem sie zusätzlich eindeutige Verhaltensmaße ableitete. Wie bereits Hannover (2000) zur Ausdifferenzierung des Selbst in Abhängigkeit von neuen Situationen darstellte, bleibt auch in Hinblick auf das Geschlechtsrollenselbstkonzept schließlich die Frage offen, was mit dem Geschlechtsschema passiert (Martin, 2000 in Rustemeyer & Fischer, 2007). Durch biologische Determination und nach Abschluss der Sozialisationsphasen könnte auch hier eine gewisse Stabilität impliziert werden, um so mehr, wenn man in Betracht zieht, dass Personen die Tendenz haben, schemakonsistente Informationen eher wahrzunehmen und zu verarbeiten, was schlussendlich zu einem sich aufrecht erhaltenden System führt.

Insgesamt betrachtet lässt sich in Bezug auf Persönlichkeitseigenschaften und ihrer Stabilität nur schwerlich eine klare Antwort finden. Die Persönlichkeitspsychologie selbst befasst sich seit langem umfassend mit dem Konzept der Charaktereigenschaften als stabile individuelle Unterschiede und der Relevanz dieser Dispositionen für das Verhalten und persönliche Funktionieren (McCrae & Costa, 1991). Diese Perspektive geht davon aus, dass zugrundeliegende Eigenschaften ein Individuum prädisponieren, spezifische Verhaltensmuster konstant über eine Vielzahl von Situationen zu zeigen. Die Stabilität der Eigenschaften lässt sich jedoch nicht uneingeschränkt aufrecht erhalten, da beachtliche Nachweise erbracht wurden, die eine gewisse Variabilität von selbstberichteten Charakteristiken über die Zeit, Situationen, soziale Rollen oder Kontexte hinweg aufzeigen (Baird, Le, & Lucas, 2006; Funder & Colvin, 1991; Paulhus & Martin, 1988), und mit Anpassung und Wohlbefinden in Zusammenhang bringen.

Dennoch kann das Geschlechtsrollenselbstkonzept vor dem Hintergrund der dargestellten Studien als stabil genug und weitgehend entwickelt angesehen werden, um bei Erwachsenen mit Selbstberichtsmaßen zuverlässig erfasst werden zu können.

#### **2.2.4.2 Flexibilität über verschiedene Situationen**

Gerade wenn der Messung des Geschlechtsrollenselbstkonzepts in der Forschung solche Bedeutung beigemessen wird, muss die Stabilität über die Zeit hinweg berücksichtigt werden.

Die Annahme, das Geschlechterselbstkonzept sei über Kontexte hinweg statisch und konsistent, wurde bereits mehrfach angezweifelt. Verschiedene Studie wiesen die situative

Variabilität des Geschlechtsselbstkonzepts nach, z. B. durch eine experimentelle Aktivierung beeinflusste Flexibilität und Plastizität der Geschlechtsidentität (Hannover, 1997; Markus & Kitayama, 1991) und derart, dass eine aktivierte Expressivität und Instrumentalität die Zuschreibung der entsprechenden Eigenschaften im Selbstkonzept erhöhten (Hannover, 1997; Miller, Lewy, & Peckham, 1997; Smith, Noll, & Bryant, 1999).

Geschlechtsrollenstereotype haben sich in vielen Studien als weniger statisch herausgestellt als ursprünglich angenommen (Abele, 2003). In diesem Zusammenhang ist besonders das Prinzip der Bezugsgruppe relevant. Je nach Art der Fragestellung gerät einmal die eigene Gruppe, die eigene Subgruppe (mit eigenen Clustern von Geschlechterstereotypen, Athenstaedt et al., 2008) oder die eigene Person eher in den Vordergrund der Wahrnehmung und Bezugnahme – was die Einschätzung von allgemeinen Stereotypen für eine große, übergeordnete Gruppe erschwert. Die Existenz sowie Relevanz von Geschlechtssubgruppen ist in einer Reihe von Studien dokumentiert worden (Altermatt, DeWall, & Leskinen, 2003; Athenstaedt et al., 2008; Coats & Smith, 1999; Deaux et al., 1985; Eckes, 1994a; Eckes, 1994b; Eckes, Trautner, & Behrendt, 2005; Roberts & Donahue, 1994; Vonk & Olde-Monnikhof, 1998). Athenstaedt, Heinzle und Lerchbaumer (2008) beispielsweise untersuchten die Relevanz von Geschlechtersubgruppen und diesbezügliche Unterschiede für das eigene Geschlechtsrollenselbstkonzept. Dazu wurden die männlichen und weiblichen Probanden gebeten sich einer von vier beschriebenen Geschlechtersubgruppen zuzuordnen. Das Bemerkenswerte am Design bestand darin, dass die Versuchspersonen im lokalen Umfeld stereotyper Subgruppen rekrutiert wurden (z. B. „Spielplatz-Hausfrauen“; „Finanzbank-Yuppies“, etc.). Die Ergebnisse zeigten, dass das eigene Geschlechtsrollenselbstkonzept von der Kategorisierung zu einer bestimmten Subgruppe abhängt – jedoch nur entsprechend des jeweiligen Geschlechts (z. B. positive Femininität mit der Kategorisierung zur weiblichen Subgruppe). Ausgehend von der Selbstkategorisierungstheorie (Turner, Hogg, Oakes, Reicher, & Wetherell, 1987; Turner & Onorato, 1999) und der Annahme, dass Selbstkategorisierung im gewissen Ausmaß vom sozialen Kontext abhängt, werden wahrscheinlich solche Subgruppen zur Selbstkategorisierung und -beschreibung herangezogen, die vor dem Hintergrund stereotyper Umgebung salienter sind als andere (Hogg, 2003).

Moskowitz, Suh und Desaulniers (1994) konnten ebenfalls den Einfluss der Rollenausübung auf Geschlechtsrollenstereotype zeigen. Berichtetes dominantes Verhalten ist auf die Ausübung der momentanen Interaktionen (z. B. Führungspositionen) zurückzuführen. Martin (2000) demonstrierte die unterschiedliche Ausprägung traditioneller Geschlechterskripte und ihre Wirkung auf das Verhalten in Abhängigkeit der Situation und möglicher Interaktionspartner. Twenge (2001) beschrieb ähnliche Ergebnisse dahingehend, dass das Dominanzverhalten von Frauen durch den Status und die Rolle Variation erfährt. In Abhängigkeit situativer, kontextabhängiger Einflüsse entsteht eine Art dynamischer Interaktionismus: Aktivierung und Organisation unterschiedlicher Aspekte des Selbst in Abhängigkeit von dem jeweiligen sozialen Kontext, Salienz, Rollenausübung, Subgruppen und Interaktionspartner. Hannover (1997) ging in diesem Zusammenhang davon aus, dass mentale

Repräsentation geschlechtsbezogener Eigenschaften situational durch geschlechtsbezogene Tätigkeit aktiviert werden kann. Mädchen und Jungen ihrer Stichprobe bearbeiteten eine feminine oder maskuline Aufgabe (Babypuppe wickeln vs. Nägel einschlagen) und sollten sich anschließend hinsichtlich expressiver und instrumenteller Eigenschaften selbst einschätzen. Erwartungsgemäß schrieben sich beide Gruppen die geschlechtstypischen Eigenschaften der zu bearbeitenden Aufgabe (z. B. instrumentelle Eigenschaften nach der maskulinen Tätigkeit) im stärkeren und schnelleren Maß zu als andere Eigenschaften.

In diesem Zusammenhang kann auch die von Eagly und Kollegen (Eagly, 1987; Eagly & Wood, 1999; Eagly et al., 2000) – und bereits weiter oben diskutierte – aufgestellte soziale Rollentheorie zur Erklärung herangezogen werden. Demnach rühren Geschlechtsunterschiede in verschiedenen Situationen von der Tendenz her, sich den jeweils zutreffenden sozialen Rollen entsprechend zu verhalten. Roberts und Donahue (1994) merkten dazu an, dass die allgemeine Selbstbeschreibung den Individuen abverlangt, ihre Eigenschaften von verschiedenen spezifischen Rollenidentitäten in einer Einschätzung zusammenzufassen. Diese einzelne Einschätzung repräsentiert demzufolge eine Kombination verschiedenster rollenspezifischer Informationen und dies in Abhängigkeit von sowohl der Salienz, der Zufriedenheit mit als auch der Verbindlichkeit zu den jeweiligen Rollenidentitäten. Smith et al. (1999) kritisierten in diesem Zusammenhang, dass sich die Messung des allgemeinen Geschlechtsrollenselbstkonzepts mittels Selbstbeschreibung über Eigenschaftswörter stark von der Selbstbeschreibung in einem spezifischen Kontext unterscheiden kann. In ihrer Studie, die jedoch methodische Mängel (Stichprobenszusammensetzung, Vernachlässigung der Interaktionspartner in den Szenarien) aufweist, befragten Smith und ihre Kollegen Studierende mittels des BSRI in Bezug auf verschiedenste Situationen (z. B. das allgemeine Selbst; aus der Perspektive, Student zu sein; im sozialen Kontakt zu Freunden, etc.). Die Resultate zeigten deutlich, wie sowohl Femininitäts- als auch Maskulinitätswerte für beide Geschlechter über Situationen hinweg variieren.

Die zum Teil dargestellten Studien vermitteln bisweilen den Eindruck, dass das Geschlechtsrollenselbstkonzept nicht stabil zu erfassen sei. Zwei wesentliche Aspekte sollen daher in diesem Zusammenhang berücksichtigt werden.

Zum einen gilt das Selbst als strukturell multipel und prozedural flexibel (Athenstaedt et al., 2004; Deaux & Major, 1987; Hannover, 1997; Hannover, 2000; Markus & Wurf, 1987; Martin, 2000; Suh, Moskowitz, Fournier, & Zuroff, 2004). In einem späteren Abschnitt wird die Messung des Geschlechtsrollenselbstkonzepts betrachtet und in dem Zusammenhang auch die Bezugsgröße der jeweiligen Erfassungsart. So soll beispielsweise in der vorliegenden Arbeit nur der Teil des geschlechtsbezogenen Selbst erfasst werden, der durch eine Liste von Eigenschaftswörtern auch tatsächlich abgebildet werden kann und nicht den Anspruch erhebt, die Multidimensionalität des Gesamtkonstrukts Geschlecht zu repräsentieren. Zum anderen sei schließlich angemerkt, dass beispielsweise die durchschnittliche Korrelation aller eingeschätzten Rollen in der Stichprobe von Roberts und Donahue (1994) bei  $r = .81$  lag, was verdeutlicht, dass die Selbstwahrnehmung und daraus folgende Selbsteinschätzung des

Geschlechtsrollenselbstkonzepts – erhoben durch Eigenschaftswörter, so wie es auch in der vorliegenden Arbeit erfolgen soll – über Subgruppen hinweg verhältnismäßig konsistent und von situativen Einflüssen unbeeinflusst bleibt.

### **2.2.5 Kulturelle Veränderungen**

In den nächsten Abschnitten folgt die Diskussion kultureller Veränderungen und deren Einflüsse sowohl auf die Entwicklung des geschlechtsbezogenen Selbst als auch auf die Messung und dieses Konzepts. Dabei wird herausgearbeitet, wie gerade die Besonderheiten einzelner Kulturen unmöglich durch ein einzelnes Instrument (und seine Übersetzungen) abgebildet werden kann und warum – auch für den deutschsprachigen Raum trotz existierender Neuentwicklung in den 1980er Jahren (Schneider-Düker, 1978; Schneider-Düker & Kohler, 1988) – nunmehr eine Neukonstruktion unter Berücksichtigung aktueller Geschlechterstereotype notwendig ist.

#### *Der Anthropologische Ansatz*

Kultur spielt eine große Rolle im menschlichen Entwicklungsprozess und daher werden sowohl Studien präsentiert, die Ähnlichkeiten zwischen Kulturen nachwiesen, als auch solche, die Unterschiede fanden. Darüber hinaus werden kulturelle Veränderungen sowie Probleme von Untersuchungen im Kulturvergleich diskutiert.

Unabhängig von der oftmals allgemein vorherrschenden Annahme, die Beziehungen zwischen Männern und Frauen wären universell auf der ganzen Welt, konnten schon Anthropologen wie Margaret Mead (1935) zeigen, dass Geschlechtsrollen in großem Rahmen in verschiedenen Teilen der Welt variieren können (Doyle & Paludi, 1995). Dabei interagieren das Geschlecht und die ethnische Zugehörigkeit bezüglich der Ausbildung sozialer Normen und der Art und Weise, wie das Verhalten beeinflusst wird. Trotz dieser kulturellen Diversität gibt es gewisse universelle Muster in den Beziehungen zwischen Männern und Frauen (Arbeitsteilung, männliche Dominanz). Eine Befragung in 110 Kulturen (Barry, Bacon, & Child, 1957) konnte den Nachweis erbringen, dass existierende Geschlechtsunterschiede kulturell bedingt sind und man deshalb über Kulturen hinweg zu unterschiedlichen Aussagen kommt, was typisch männlich und typisch weiblich ist. Kulturen unterscheiden sich dahingehend, inwieweit traditionelle Geschlechtsrollenwerte gepflegt werden, was im Falle stark traditionell ausgeprägter Geschlechtsideologien bis zur Unterdrückung von geschlechtsuntypischen Verhaltensweisen führen kann (z. B. China, Chia, Moore, Lam, & Chuang, 1994). In einigen Kulturen sind bestimmte (berufliche) Rollen in der Gesellschaft, dem Militär oder der Politik ausschließlich den Männern vorbehalten, während in anderen Ländern diese Rollen durchaus auch von Frauen ausgefüllt werden, wie beispielsweise in Israel, wo auch Frauen einen Militärdienst zu absolvieren haben (Hyde & Rosenberg, 1980).

Kulturelle Unterschiede bezüglich der Erwartungen und Normen an die Geschlechter, wie sie sein sollen und welche Rollen sie ausfüllen, nehmen Einfluss auf die Entwicklung des Selbstkonzepts (Best, 2001). Diese kulturell definierte Art und Weise, Identitäten zu konstruieren, lenkt schließlich auch entscheidend subjektive Erfahrungen und Entwicklungen

und spiegelt sich in Selbstberichten wider. Markus und Kitayama (1991; 1998) erklärten gefundene Kulturunterschiede in den Selbstberichten damit, dass in den Kulturen sehr unterschiedliche Normen vorherrschen, die implizite Imperative an ihre Mitglieder stellen und so Einfluss auf die Entwicklung des Selbstkonzepts nehmen. Somit können Mitglieder verschiedener Kulturen ihr Selbst sehr unterschiedlich definieren. Dabei unterschieden die beiden Autoren zwischen zwei Arten der Selbstkonstruktion: independent (beschreibt die Besonderheit der eigenen Person gänzlich unabhängig von anderen) und interdependent (Selbstbeschreibung mittels Aspekten, die auch von anderen Personen geteilt werden).

Im Gleichklang mit der allgemeinen Entwicklung der Geschlechtsrollenforschung hat sich ein Teil der Studien speziell mit Kulturunterschieden beschäftigt, wobei im Wesentlichen vier Kategorien zu unterscheiden sind. Eine Forschungsrichtung verfolgte vor allem die Validierung und Beschreibung von Geschlechtsrollen in spezifischen Kulturen (Mexiko: Lara-Cantu & Navarro-Arias, 1986; Lara-Cantu & Navarro-Arias, 1987; Israel: Lobel, Slone, & Winch, 1997; Malaysia: Ward & Sethi, 1986). Eine weitere Richtung nahm Vergleiche zwischen Kulturen (USA mit China: Chia et al., 1994) vor, während die dritte Gruppe Unterschiede innerhalb von Kulturen betrachtete (Hogan, 1979: Highschool-Schüler und Studierende in den USA; Rice & Coates, 1995: Vergleich von Südstaatlern mit Personen aus anderen Regionen innerhalb der US). Eine letzte, vierte Gruppe schließlich verglich mindestens zwei verschiedene Kulturen um die Ausprägung traditioneller Geschlechtsrollen zu beleuchten (USA jeweils mit Fiji: Basow, 1984; Australien: Gackebach, 1981; Deutschland: Hogan, 1979; und dem damaligen Jugoslawien: Novakovic & Kidd, 1988; Mexiko: Schmitz & Diefentahler, 1998) oder den Zusammenhang zwischen Geschlechtsrollen zu anderen Variablen in den Zielkulturen herstellten (z. B. Zusammenhang zu Selbstwert und Popularität: Lobel et al., 1997).

Wegweisend für den Geschlechterstereotypenansatz im Bereich der kulturellen Forschung sind zweifelsohne die Arbeiten von Hofstede (1980; 1998). Seine über arbeitsbezogene Werte von IBM-Unternehmen in verschiedenen Nationen konzipierte Studie erbrachte fünf Kulturdimensionen, wovon für den Geschlechteransatz insbesondere die *Dimensionen* Maskulinität und Femininität von besonderer Bedeutung sind. Er wies den Kulturen einen MAS-Index zu: Kulturen mit einem hohen MAS-Wert sind durch überwiegend maskuline Normen und auch durch eine größere Distanz zwischen Männern, Frauen und deren Werte gekennzeichnet als feminine Kulturen. Maskulinität bezieht sich laut Hofstede auf Kulturen, in denen Männer durchsetzungsfähig und erfolgsorientiert sind, während sich Frauen durch Bescheidenheit auszeichnen und auf Beziehungsqualität fokussiert sind (z. B. USA, Deutschland, Japan, Mexiko). Im Vergleich dazu ist Femininität in Kulturen präsent, in denen sowohl Männer als auch Frauen bescheiden und beziehungs-fokussiert sind (Schweden, Finnland, Thailand, Costa Rica). Der Unterschied zu dem später von Bem (1974) entwickelten Messinstrument, welches Maskulinität und Femininität abbildet, liegt darin, dass Bem auf Individualebene zwischen diesen Dimensionen unterschied, während Hofstede dies auf Länderebene tat.

### *Ähnlichkeiten über Kulturen hinweg*

Trotz der methodischen Herausforderung, Stereotype, die bewiesenermaßen kultur- und bisweilen kohortenabhängig sind, mit Instrumenten zu erfassen, die hauptsächlich für eine einzige Kultur entwickelt und an dieser validiert wurden, gelingt es Studien wiederholt auch über Kulturgrenzen hinweg, Ähnlichkeiten zwischen Frauen und Männern festzustellen. So konnten Runge, Frey, Gollwitzer, Helmreich, und Spence (1981) in einer Studie, die westdeutsche und US-amerikanische Studierende mittels des Extended Personal Attribute Questionnaire (EPAQ)<sup>8</sup> verglich, auf allen Skalen signifikante Geschlechtsunterschiede für beide Nationen in ähnlicher Weise dokumentieren. Auch Sethi und Allen (1984) kommentierten anhand einer Vielzahl von Studien die überraschende Konsistenz bezüglich der unterschiedlichen Eigenschaften und Rollen zwischen Männern und Frauen, selbst über verschiedenste Kulturen hinweg und argumentieren darüber hinaus, inwieweit zum Beispiel der BSRI – wenn angemessen übersetzt – letztendlich doch als valides Instrument auch über Kulturgrenzen hinweg eine Diskriminierung zwischen den Geschlechtern aufzeigen kann.

Im Vergleich der methodisch anspruchsvollen Studien von beispielsweise Best und Williams (1997) oder Munroe und Munroe (1997a) – beide zitiert in Best (2001) – kam Best zu dem Fazit, dass unabhängig von biologischen Prädispositionen vor allem sozio-kulturelle Faktoren zur Entwicklung individueller Geschlechtskonzepte und der Geschlechtsrolle beitragen. Beide dort ausgewiesenen Studien kommen trotz unterschiedlicher Stichproben zu überraschend ähnlichen Ergebnissen, die die maskulinen Eigenschaften als das Ideal für bisweilen beide Geschlechtergruppen betonen.

### *Kulturelle zeitliche Veränderungen*

Die Einschätzungen und Ausprägungen des Geschlechtsrollenselbstkonzepts sind eng verzahnt mit den fortwährenden kulturellen Veränderungen und Anpassungen. Twenge (1997) fand in ihrer Meta-Analyse von 63 Studien, die zwischen 1973 und 1993 den BSRI<sup>9</sup> oder den Personal Attribute Questionnaire (PAQ)<sup>10</sup> in studentischen Stichproben einsetzten, eine Erhöhung der selbst zugeschriebenen Maskulinitätswerte bei Frauen, während die Femininitätswerte überraschend stabil blieben. Darüber hinaus verringerten sich die Unterschiede zwischen Männern und Frauen auf den beiden Teilskalen. Twenge führt diese Entwicklung bezüglich der Selbstzuschreibungen auf kulturelle Veränderungen zurück (was auch Stewart & Healy, 1989, schon diskutierten). So waren die letzten 50 Jahre beispielsweise von erheblichen Veränderungen und Entwicklungen sowohl in der Wirtschaft als auch der Politik gezeichnet und einer Verlagerung weg von den traditionellen Geschlechtsrollen mit dem Vater als zentraler und entscheidender Figur in einer Familie hin zu einer liberaleren, feministischen Kultur, mit gleichen Rechten und Möglichkeiten für Männer und Frauen stattfand – zumindest theoretisch.

---

<sup>8</sup> Der EPAQ ist eine Erweiterung des PAQ um sozial unerwünschte Maskulinität und Femininität

<sup>9</sup> Der BSRI ermöglicht Aussagen bezüglich der individuellen Ausprägung von Maskulinität und Femininität sowie eine Klassifikation in eine von vier möglichen Geschlechtertypen - die an späterer Stelle der Arbeit genauer ausgeführt werden. Es werden einfache Adjektive hinsichtlich ihrer Bedeutung für das Selbst eingeschätzt.

<sup>10</sup> Der PAQ fußt auf anderen Theorien als der BSRI und misst Instrumentalität sowie Expressivität, die jedoch inhaltliche Nähe zu Maskulinität und Femininität aufweisen. Die Einschätzung erfolgt mittels Itempaare.



In diesem Zusammenhang leistet die Forschung zu Reaktionen auf die Verletzung von Geschlechterstereotypen von Männern und Frauen, der inadäquaten Ausfüllung ihrer geschlechtstypischen Rolle, einen bedeutsamen Beitrag. Verschiedenste Ergebnisse zeigten, dass – obwohl ein gewisser gesellschaftlicher Wandel zu verzeichnen ist – eine Verletzung klassischer geschlechtstypischer Stereotype immer noch zu negativer Bewertung führen kann (Butler & Geis, 1990; Costrich, Feinstein, Kidder, Marecek, & Pascale, 1975; Rudman, 1998; Rudman & Glick, 1999; Tepper, Brown, & Hunt, 1993) – die Einschätzungen also weiterhin wenig liberal sind. So konnte Rudman (1998; Rudman & Glick, 1999; 2001) zur Wirkung impliziter Geschlechterstereotype auf die Bewertung von Bewerbern auf geschlechtstypisierte Anstellungen den sog. *Backlash* gegen instrumentelle Frauen demonstrieren. Bewältigen beispielsweise erfolgreiche Frauen den Aufbau instrumenteller Eigenschaften zusätzlich und in gleichem Ausmaß zu ihren traditionell weiblichen, expressiven Charakteristiken, fällt ihre Evaluation positiv aus. Waren in der Stellenausschreibung jedoch weibliche Qualitäten vorausgesetzt, wurden Frauen, die instrumentelle Eigenschaften in ihr Selbstbild integriert hatten, insgesamt negativer bewertet - im Gegensatz zu männlichen Bewerbern.

Dieses Beispiel verdeutlicht, wie langwierig solch ein Prozess des Wandels ist und wie – unabhängig von der scheinbaren allgemeinen gesellschaftlichen Akzeptanz des Wechsels zu mehr Übereinstimmung in den Eigenschaften – individuelle Interessen und Bedrohungen diesen Prozess auch verzögern können. Dennoch ließen sich Veränderungen in der Valenz von Eigenschaften und ihren ursprünglich exklusiven geschlechtsspezifischen Zuschreibungen feststellen. Swazina, Waldherr und Maier (2004) unterzogen 40 maskuline und feminine Eigenschaften (auch mit vorhandenen Werten aus dem Jahr 1978 von Schneider-Düker, 1978) einer Überprüfung der sozialen Erwünschtheit um der von Hager, Schmuck und Mecklenbräuer (1996) gefundenen Veränderung in der Valenz der Zuschreibung nachzugehen. Dabei konnte ein signifikanter Wandel der sozialen Erwünschtheit bestätigt werden, sowohl vor allem für ursprünglich als geschlechtsinkongruent deklarierte (einige maskuline Eigenschaften sind nunmehr für Frauen sozial erwünschter als noch 1978) als auch für geschlechtstypische Eigenschaften (maskuline Eigenschaften für Männer stiegen beispielsweise in der sozialen Erwünschtheit).

Das Bundesministerium für Familie, Senioren, Frauen und Jugend (2007) hatte in einer umfangreichen Bestandsaufnahme 20-jährige Frauen und Männer zu ihren Rollenbildern und Vorstellungen zu Beruf, Partnerschaft, Familie, Kinder und Zukunft in der heutigen Gesellschaft befragt. Die qualitative Analyse erbrachte den Befund, dass Perspektiven und Vorstellungen in Abhängigkeit vom Geschlecht, Bildungsgrad und dem sozialen Milieu mitunter stark differenzieren. So demonstrieren Frauen mit Abitur eine selbstbewusste Geschlechtsidentität. Frau sein bedeutet für sie neben äußeren Attributen wie Schönheitsideale vor allem auch die Kombination aus Eigenschaften, die sich sowohl typisch femininen als auch typisch maskulinen Beschreibungen zuordnen lassen: Intelligenz, Selbstbewusstsein, Durchhaltevermögen, etc. Besonders auffällig ist hierbei, dass ursprünglich männliche Eigenschaften nicht mehr als exklusiv „männertypisch“ gelten. Sie wünschen sich für Männer neben den klassischen

männlichen Attributen auch feminine Rollenbeschreibungen als Ideal: Emotionalität, Sinnlichkeit oder Sanftmut. Männliche Befragte zeigen eine traditionellere Einstellung. Zumindest für die Gruppe mit Bildungshintergrund deutet sich eine wichtige Tendenz an. Frauen und Männer erkennen die gesellschaftlichen Veränderungen gleichermaßen an, scheinen dies jedoch mit unterschiedlicher Motivation zu tun.

Bezüglich der Konstruktion ihres Messinstruments BSRI suchte Bem (1974) ursprünglich nach Eigenschaften, die sich als maskulin bzw. feminin qualifizierten, wenn sie von Männern und Frauen unabhängig voneinander als signifikant erwünschter für jeweils das eine Geschlecht beurteilt wurden als für das andere. Bem merkte dazu an:

*“... the BSRI was founded on a conception of the sex-typed person as someone who has internalized society's sex-typed standards of desirable behaviour for men and women ... personality characteristics were selected as masculine or feminine on the basis of sex-typed social desirability”* (Bem, 1974, S. 155-156).

Obwohl Bem wiederholt damit zitiert wird, dass sie nach Eigenschaften in Abhängigkeit derer sozialen Erwünschtheit suchte, mischte sie in dieser Ausführung soziale Erwünschtheit und Typizität. Wie von Swazina et al. (2004) angemerkt, sind diese beiden Aspekte nicht identisch, wenn auch nicht gänzlich unabhängig voneinander (Mecklenbräuker, Hager, & Möller, 1994). Unter Bezugnahme des Gender-Belief-Systems (Deaux & LaFrance, 1998) und der Vorstellung, wie Frauen und Männer typischerweise und idealerweise sein sollten, erklärten Swazina und Kollegen, dass geschlechtstypische Eigenschaften eine deskriptive Komponente haben, während geschlechtsspezifisch sozial erwünschte Eigenschaften sogenannte *prescriptive gender stereotypes* darstellen – dem Ideal entsprechen. Diese inhaltliche Unterscheidung zwischen präskriptiven und deskriptiven Eigenschaften ist schlussendlich auch theoretische Grundlage für die Entwicklung von Messinstrumenten, die sich je nach Schwerpunkt der Definition auch genau darin unterscheiden. Im Gegensatz zum BSRI, der vor allem sozial erwünschte Eigenschaften berücksichtigte, versucht der Personality Attributes Questionnaire, PAQ (Spence et al., 1974; Spence & Helmreich, 1978) „to tap descriptive and prescriptive gender stereotypes“ (Spence, 1991, S. 145), demnach beiden Eigenschaftsaspekten gerecht zu werden.

Die gesellschaftlichen Wertvorstellungen, deren Wurzel soziale Repräsentationen über die Geschlechter sind, erlebten seit den 1970er Jahren einen stetigen Wandel bezüglich ihrer sozialen Erwünschtheit (Deaux & LaFrance, 1998). Dabei muss jedoch generell die Frage gestellt werden, inwieweit sich nur die Ideale geändert haben oder ob die Selbstbeschreibung auch tatsächlich einen Wandel dahingehend erlebt, dass das entsprechende Geschlecht die nunmehr neu erwünschte Eigenschaft wahrnimmt und für sich in der Selbstbeschreibung beansprucht (z. B. ehrgeizige Frauen und warmherzige Männer). In Bezug auf die soziale Rollentheorie konnten Diekman und Eagly (2000) in mehreren Experimenten aufzeigen, dass die Rollen von Männern und Frauen zunehmend ähnlicher werden. So glaubten z. B. die Probanden, dass Frauen der Gegenwart maskuliner sind als Frauen der Vergangenheit, und dass Frauen in der Zukunft hinsichtlich der Persönlichkeitseigenschaften sogar noch maskuliner sein werden. Dazu baten sie Probanden, sich eine durchschnittliche Frau sowie einen Mann zu

einer bestimmten, vorgegebenen Zeit zwischen 1959 und 2050 vorzustellen. Die Autorinnen führten diese Entwicklung unter anderem auf die Veränderung der Arbeitsmarktstruktur zurück, die einen Anstieg an arbeitnehmenden Frauen verzeichnet (z. B. in den USA: von 34% auf 60% zwischen 1950 und 1998, zitiert in Diekman & Eagly, 2000), während die Veränderungen männlicher Rollen weniger stark ausfallen. Wilde & Diekman (2005) hatten in einem ähnlichen Versuch dynamische Stereotype an der Zeitachse 1950-1975 und 2000 (ohne den Zukunftsaspekt) in einem kulturellen Vergleich von Deutschen und US-Amerikanern untersucht und konnten dabei in Übereinstimmung zu der früheren Studie zeigen, dass Frauen eine Zunahme maskuliner Eigenschaften über die Zeit bescheinigt wird, während männliche Charakteristiken als stabiler eingeschätzt wurden. Bezüglich der Vorgehensweise von beiden Studien (Diekman & Eagly, 2000; Wilde & Diekman, 2005) bleibt jedoch offen, inwieweit das Erfassen der zu den entsprechenden Zeiten vorherrschenden Rolleneinstellung über die hier eingesetzte Methode möglich ist. Inwieweit hierbei tatsächlich Wissen über die damalige Zeit geprüft wird, welches unabhängig vom Wissen über bereits stattgefundenen Veränderungen, z. B. Veränderungen der sozialen Rollen, Zunahme der Maskulinität bei Frauen (siehe Twenge, 1997) ist, bleibt ungeklärt.

Zahlreiche Studien belegten einen Wandel der Geschlechtsrollen über die Zeit auch für den deutschsprachigen Raum (z. B. Hager, Schmuck, & Mecklenbräuer, 1996; Swazina, Waldherr, & Maier, 2004). Hinzu kommt, wie bereits oben angeführt, dass selbst für Gruppen innerhalb der USA, dem Land, welches den kulturellen Rahmen für die Erstellung des BSRI darstellte, als auch für andere Nationen Konzepte von Maskulinität und Femininität mitunter variieren und ganz allgemein die Anwendung des BSRI in Frage stellen.

Auf der Ebene der empirischen Erfassung des Geschlechtsrollenselbstkonzepts sind daher Instrumente erforderlich, die den aktuellen innerkulturellen Konsens in der betreffenden Zielgruppe hinsichtlich der Geschlechtstypizität von Eigenschaften sowie ihrer Bewertung widerspiegeln, was durch eine Neuentwicklung, wie diese Arbeit sie anstrebt, gewährleistet werden soll.

### **2.2.6 Korrelate des GSK auf der Ebene von Affekten, Kognitionen und Verhalten**

Seit jeher beschäftigt sich psychologische und soziologische Forschung mit der Frage nach dem biologischen Geschlecht und dessen Auswirkungen auf das Leben. Viele Studien haben sich dabei besonders der Frage nach Geschlechtsunterschieden angenommen und diese Unterschiede in Bezug auf die Sozialisation von Mädchen und Jungen über verschiedenste Kulturen hinweg zeigen können (vgl. Rosenblatt & Cunningham, 1976). Für die vorliegende Arbeit sind diese Resultate vor allem für den Aspekt der Validierung des neuen Instruments relevant.

Die Ergebnisse früherer Studien zeigen, dass Mädchen zu Verantwortung und Gehorsam erzogen wurden, während Jungen Leistung und Selbstbewusstsein demonstrieren sollten (Barry et al., 1957). Bezüglich aktuell bestehender und nachgewiesener Geschlechtsunterschiede differenziert man zwischen denen, die faktisch existieren und nachgewiesen werden können, und denen, die lediglich eine Widerspiegelung bestehender Stereotype sind.

So konnten Unterschiede in Persönlichkeitseigenschaften (Temperament), Fähigkeiten (vgl. Archer & Lloyd, 2002; Maccoby & Jacklin, 1974, für eine Zusammenfassung) oder Rollenverteilungen gefunden werden, wobei verschiedene Autoren anführen, dass zum einen diese Unterschiede eher geringe Effektgrößen aufweisen und zum anderen im Verlauf der vergangenen Jahre geringer geworden sind (McGuinness, 1998).

Die Annahme, dass beispielsweise Jungen aggressiv und laut sind, besser in Mathematik und Sport, während Mädchen im Vergleich vor allem sprachlich brillieren und über ein besseres soziales Geschick verfügen, lässt sich nur teilweise durch die Forschung bestätigen, da die Unterschiede nur geringen diagnostischen Wert haben (vgl. Hogg & Vaughan, 2008) oder gar nicht erst systematisch gefunden werden können (Alfermann, 1996). So fasste Alfermann (1996) zusammen, dass nachgewiesene, stabile Unterschiede, wie in der Aggressivität oder im prosozialen Verhalten, mit lediglich 10% Varianzaufklärung demnach zu 90% auf andere interindividuelle Ursachen zurückgeführt werden müssen als auf das Geschlecht. Darüber hinaus ist ebenfalls der Zeitpunkt der Differenzierung zwischen den Geschlechtern bedeutsam; so ist beispielsweise bezüglich physischer Aggressivität ein ursprünglich mittelgroßer Geschlechtsunterschied aus dem Kindesalter im Erwachsenenalter um ein Erhebliches reduziert, wie von Meta-Analysen aufgezeigt worden ist (vgl. ebenfalls Alfermann, 1996). Derartige Meta-Analysen offenbaren auch, dass eine Vielzahl von Geschlechtsunterschieden mitunter kleiner sind als bisher angenommen. So zeigte Hyde (2005) in 46 Studien, dass in 78% der betrachteten Domänen die Unterschiede klein bis nicht existent sind.

Es wird gern argumentiert, dass Männer und Frauen gleich(-wertig) sind und dementsprechend das Recht auf gleiche Anstellungen, Möglichkeiten und Entlohnungen haben sollten. Die Realität zeigt, dass einige berufliche Tätigkeiten jedoch eindeutig geschlechtsspezifisch sind. Krankenpflegerinnen und Frauen in sozialen oder artistischen Berufen sind ihren männlichen Kollegen in Zahlen immer noch weit überlegen, während handwerkliche Berufe wie Ingenieur oder Mechaniker mit größerer Wahrscheinlichkeit von Männern ausgeübt werden (Lippa, 1998; 2005). Diese Geschlechtsspezifität innerhalb beruflicher Rollen zeigt sich auch in Löhnen oder Aufstiegsmöglichkeiten. Somit scheint das Geschlecht eines Bewerbers immer noch entscheidender für die Feststellung der beruflichen Eignung als die eigentlichen Qualifikationen (Williams & Best, 1990). Unterschiedliche Ausprägungen bestimmter Persönlichkeitseigenschaften legen den Schluss nahe, dass die Sozialisation mehr Varianz in dieser Entwicklung erklären kann als genetische Prägung allein (vgl. Archer & Lloyd, 2002; Maccoby & Jacklin, 1974).

Trotz unterschiedlicher theoretischer Erklärungen und methodischer Zugänge zu Geschlechtsunterschieden gibt es verschiedene Bereiche, in denen Männer oder Frauen die jeweilige andere Gruppe nachweislich übertreffen. In Tabelle 2 sind einige Bereiche den entsprechenden Geschlechtern mit der jeweils stärkeren Ausprägung zugeordnet.

Tabelle 2. Zusammenfassung der berichteten Geschlechtsunterschiede (in Auszügen)

Jungen / Männer	Mädchen / Frauen
<ul style="list-style-type: none"> <li>• handwerkliche Berufe</li> <li>• visuelle u. räumliche Wahrnehmung</li> <li>• mentaler Rotationstest</li> <li>• höhere Risikofreude / Neugier</li> <li>• abweichendes Sexualverhalten</li> <li>• höherer Substanzmissbrauch</li> </ul>	<ul style="list-style-type: none"> <li>• soziale / künstlerische Berufe</li> <li>• sprachliche Leistungen</li> <li>• Konzentrationsfähigkeit</li> <li>• Depressionen, Angststörungen</li> <li>• Neurotizismus, Gewissenhaftigkeit</li> <li>• höhere Prävalenz für Alzheimer</li> </ul>

### *Leistungsbezogene Kognitionen und kognitive Fähigkeiten*

Im Bereich des Selbstkonzepts wurde für Frauen in diesem Zusammenhang nachgewiesen, dass sie ein geringeres Selbstbewusstsein haben, ihre Fähigkeiten unterschätzen (Lenney, 1977), negative Ereignisse bevorzugt internal attribuieren (Boggiano & Barrett, 1991, Hanson & O'Leary, 1985) und sich zusätzlich eher selbst die Schuld in stressigen Situationen zuweisen (Korabik & VanKampen, 1995<sup>6</sup>, Wingerhoets & VanHeck, 1990; alle zitiert in Helgeson & Fritz, 1998).

Innerhalb des allgemeinen kognitiven Bereichs ist vor allem die Intelligenz eine der prominentesten Fragen im Rahmen der Geschlechterforschung. Im Bereich der Intelligenz wurden wiederholt Leistungsunterschiede in den verschiedensten Teilgebieten untersucht: Höhe und Struktur der Intelligenz, visuelle vs. sprachliche Intelligenz oder auch weiterführende Konstrukte wie Problemlösen und Konzentration.

Dabei zeigten sich insgesamt betrachtet oftmals keine oder nur geringe Unterschiede, die in manchen Fällen sogar lediglich aufgrund der untersuchten Stichprobengrößen Signifikanz erreichten, wie die Analyse durchschnittlicher Effektgrößen offenbarte (Hyde, 2005). Dies gilt beispielsweise für die allgemeine Höhe der Intelligenz (Eysenck, 1988; Kimura, 1992; Srocke, 1989). Bezüglich der Intelligenzstruktur zeichnen sich dagegen etwas deutlicher ausgeprägte Geschlechterdifferenzen in der visuellen Wahrnehmung mit besseren Leistungen männlicher Individuen (vor allem im Bereich der räumlichen Wahrnehmung, oder bei mathematischen und mentalen Rotationstest) und in der sprachlichen Leistung mit besseren Leistungen der Mädchen und Frauen ab (Deaux & LaFrance, 1998; Eagly, 1987; Gage & Berliner, 1992; Kimura, 1992; Maccoby & Jacklin, 1974; Sabini, 1992). Generell schränken uneindeutige Definitionen und Operationalisierungen der Konstrukte die Generalisierbarkeit der Ergebnisse erheblich ein. Maccoby und Jacklin (1974) fanden beispielsweise solche Geschlechtsunterschiede erst ab dem 10. Lebensjahr, was auf geschlechtsspezifische Sozialisationseffekte zurückgeführt werden kann. Geschlechterdifferenzen innerhalb des Problemlösens deuten eine größere Fähigkeit männlicher Personen an, bedingt durch ihre etwas höhere Risikofreudigkeit und Neugierde (Gage & Berliner, 1986; Maccoby, 1966). Demgegenüber verfügen Frauen wiederum über eine bessere Konzentrationsfähigkeit, die sich im besseren Abschneiden bei Merkfähigkeits- und Konzentrationstests widerspiegelt (Brickenkamp, 1962; Trost, 1992).

### *Unterschiede im Verhalten*

Selbst über kognitive und verbale Fähigkeiten hinaus, werden biologische und hormonelle Besonderheiten von Männern und Frauen für unterschiedliche Ausprägungen in anderen

Bereichen verantwortlich gemacht. Geschlechtsunterschiede sind auch im Gesundheits-Krankheitsbereich gut belegt. So wurde ein höheres Auftreten von Alzheimer bei Frauen nachgewiesen. Ein weiteres Beispiel ist die Männer-Frauen-Sterbe-Ratio: wie Selbstmord (viermal so viele Männer), Mord (dreimal so viele Männer), Unfälle (zweimal so viele Männer) und chronische Leberkrankheiten (zweimal so viele Männer) (Doyle & Paludi, 1995; van Wersch, 1998). Einige Forscher haben in Bezug auf der Verhaltensebene höhere Ausprägungen von Männern in den Bereichen wie Aggression; abweichendes Sexualverhalten oder Substanzmissbrauch nachgewiesen (Eagly & Steffen, 1986; LaBrie, Shaffer, LaPlante, & Wechsler, 2003; McCabe, 2002; Rhoads, 2004a; van Wersch, 1998; Wechsler, Kuo, Lee, & Dowdall, 2000; Wechsler & Nelson, 2001). Darüber hinaus lassen sich im Bereich der psychischen Störungen (Angststörungen und Depressionen häufiger bei Frauen, antisoziales Verhalten häufiger bei Männern) eindeutige geschlechtsspezifische Inzidenz- und Prävalenzraten nachweisen (Hartung & Widiger, 1998). Selbst über verschiedene Kulturen hinweg finden sich Geschlechtsunterschiede in diesen Bereichen. In einer Studie konnten in der Mehrzahl der 55 untersuchten Nationen höhere Werte bei Frauen bezüglich Neurotizismus, Extraversion, Verträglichkeit und Gewissenhaftigkeit aufgezeigt werden (Schmitt, Realo, Voracek, & Allik, 2008).

#### *Methodenkritik und Erklärungen der Geschlechtsunterschiede*

Zusammenfassend lässt sich festhalten, dass die Untersuchungsergebnisse bezüglich der Geschlechtsunterschiede keinesfalls so eindeutig sind, wie bisweilen postuliert wurde, und oftmals nur ein geringer Anteil der Varianz in der Gesamtbevölkerung durch das Geschlecht erklärt wird. Hinzu kommen nicht selten methodische Mängel wie in der Definition beteiligter Konstrukte, der Operationalisierung oder der Generalisierbarkeit von Ergebnissen.

Einige Forscher (Archer, 2006; Davies & Shackelford, 2006; Lippa, 2006) kritisieren das bloße Betrachten von Effektgrößen bei der Einschätzung der Bedeutsamkeit gefundener Geschlechtsunterschiede (wie z. B. durch Hyde, 2005). Sie erklären dies vor dem Hintergrund, dass analysierte abhängige Maße im Bereich der Verhaltensmessung, wo häufig nur ein Maß erhoben wird, selbst bisweilen wenig zuverlässig sind (Lippa, 2006). Des Weiteren führen sie als Argument an, dass verschiedene Theorien zur Erklärung der Geschlechtsunterschiede herangezogen werden müssen – vor allem aus dem Bereich der Evolutionstheorie (Archer & Lloyd, 2002). Dabei wiesen sowohl Archer (2006) als auch Davies und Shackelford (2006) wiederholt auf die Rolle der sexuellen Selektion zur Erläuterung der Unterschiede zwischen Männern und Frauen hin in Abhängigkeit davon, ob sich Männer und Frauen in der Evolution z. B. ähnlichen Herausforderungen gegenüber sahen oder ob die Evolution ihnen geschlechtsspezifische Aufgaben stellte. Folgt man Archers (2006) Argumentation lassen sich gerade Bereiche, in denen sich eine männliche Überlegenheit fand, nahezu eindeutig evolutionäre Entwicklungen und daraus resultierende Anforderungen an die Geschlechter zuordnen (z. B. motorische Fähigkeiten, physische Aggression).

Eine Ergänzung zu dieser evolutionären Perspektive sah Archer (2006) vor allem auch in der sozialen Rollentheorie (Eagly, 1987; Eagly, Wood, & Johannesen-Schmidt, 2004) als eine

weitere wichtige Grundlage zur Erklärung von Geschlechtsunterschieden. Ähnlich der Evolutionstheorie erwartet die soziale Rollentheorie gerade bei Attributen, die entweder spezifisch mit der männlichen (z. B. Arbeit, Sport) oder der weiblichen (z. B. Familie, Haushalt) sozialen Rolle und ihren Anforderungen an diese Geschlechtergruppe assoziiert sind, Geschlechtsunterschiede.

Lippa (2006) schloss sich im Wesentlichen Archers (2006) Argumentation an und betonte ebenso die Relevanz verschiedener sozialer und biologischer Faktoren, die zur Aufschlüsselung des komplexen Gebildes psychologischer Geschlechtsunterschiede herangezogen werden sollten. Einige Unterschiede werden vor allem auf die Physiologie zurückgeführt: so verfügen Männer über größere Gehirne oder einen größeren Hypothalamus (Lippa, 2002) und schneiden vermutlich daher insgesamt betrachtet beispielsweise besser bei Intelligenztests im Bereich der mentalen und räumlichen Verarbeitung ab (Jackson & Rushton, 2006). Darüber hinaus werden neben biologischen vor allem hormonelle Variationen für Unterschiede zwischen und innerhalb der Geschlechtergruppen verantwortlich gemacht. Lippa (2002) argumentierte diesbezüglich, dass beispielsweise der Einfluss pränataler Geschlechtshormone und genetischer Variationen zu individuellen Unterschieden in Maskulinität und Femininität führt – sowohl bei Männern als auch bei Frauen. Mädchen, die in der pränatalen Phase hohen Mengen an Testosteron ausgesetzt waren, zeigten in der späteren Entwicklung stärker ausgeprägte maskuline Eigenschaften, wie Aggressivität und Wettbewerbsorientierung, und dies nicht nur zusätzlich zu ihren femininen Eigenschaften, sondern auch unabhängig vom Sozialisationsbemühen um eine höhere Femininität (Rhoads, 2004b).

### **2.3 Empirische Erfassung des Geschlechtsrollenselbstkonzepts**

Die bisherige Forschung bezüglich des Erwerbs der Geschlechtsrolle fokussierte bevorzugt kulturelle und soziale Einflüsse ohne jedoch die Bedeutung der Kognitionen dabei gänzlich zu vernachlässigen. Nach der umfangreichen Reflektion der relevanten Theorien zur Geschlechtstypisierung sowie der Indikation, dass dieser Prozess durch weitere kognitive Entwicklungsschritte und dem Wechsel der einflussnehmenden sozialen Partner bis ins Erwachsenenalter andauert, wird im folgenden Schritt eine Diskussion der Instrumente zur Messung des Geschlechtsrollenselbstkonzepts notwendig. Die Herleitung und theoretische Fundierung vieler dieser Instrumente leisten einen wichtigen Beitrag zum Verständnis von Geschlechterstereotypen, ihrer Entstehung und Funktionsweisen.

Im Hinblick auf die Erfassung des GSK sind in Abbildung 5 alle beteiligten Komponenten als Übersicht dargestellt. So werden die Beziehungen einiger Konstrukte bereits hergestellt, was zur Strukturierung der folgenden Abschnitte beitragen soll. Zusammengefasst bedeutet dies: aus der Theorie sind positive und negative Aspekte der Selbst- und Fremdbeschreibung bekannt. Dieser Einfluss der Valenz von Eigenschaften gilt in der Form auch für das Geschlechtsrollenselbstkonzept. Bisherige Messinstrumente beschränkten sich ausschließlich wie im Falle des BSRI auf die Erfassung positiver Maskulinität und Femininität

(die in anderen Messinstrumenten mitunter anders benannt sind, wie beispielsweise die vergleichbare Instrumentalität beim PAQ). Theoretische Überlegungen und Forschungsergebnisse bedingten sich weitgehend gegenseitig und führten schließlich zur allgemeinen Anerkennung auch negativer Maskulinität und Femininität des GSK (und ihren entsprechenden Teilskalen in anderen Instrumenten wie beispielsweise Unmitigated Agency).

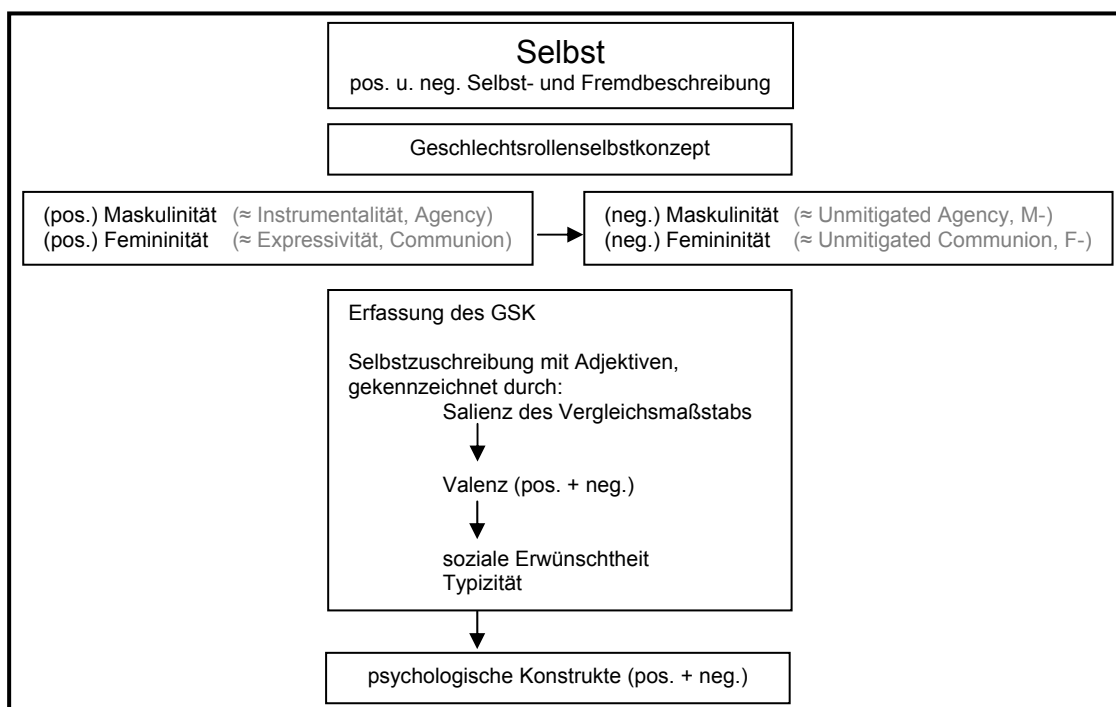


Abbildung 5. Darstellung wichtiger Komponenten bei der Erfassung des GSK

Bezüglich der Erfassung kommen vor allem attributbasierte Messinstrumente zum Einsatz. Diesbezüglich gilt es in Betracht zu ziehen, dass die Beschreibung im von der Salienz einer Art Vergleichsmaßstab abhängt. Beschreibt sich beispielsweise eine Person als Mann unter Frauen oder als Mann unter erfolgreichen Männern. Des Weiteren sind Attribute per se durch eine Valenz gekennzeichnet sind, die diese Attribute allgemein positiv oder negativ wertet. Im Zuge der Selbstbeschreibung beeinflusst die Valenz u. a. das Merkmal der sozialen Erwünschtheit der stereotypen Eigenschaften für die Gruppen der Männer und Frauen, welches das Ausmaß der Selbstbeschreibung mit bestimmt. Neben der sozialen Erwünschtheit ist die Typizität von Eigenschaften für Männer und Frauen ebenfalls entscheidend für die Selbstbeschreibung. Instrumente zur Messung des GSK, die diese beeinflussenden Variablen und sowohl positive als auch negative Eigenschaften berücksichtigen, sollten sich sehr gut zur Vorhersage von psychologischen Konstrukten eignen, die ihrerseits positiv oder negativ sein können.

In den folgenden Abschnitten wird nun ausgehend von der bisher dargestellten Theorie zunächst die Entwicklung des Verständnisses des GSK nachgezeichnet, was sich in der Konstruktion verschiedener Messinstrumente widerspiegelte. Dabei wird vor allem die Implementierung der negativen Selbstbeschreibung in die Instrumente betrachtet, wie diese erhoben wird und warum. Aus den Ergebnissen des Einsatzes dieser Messinstrumente wird gefolgert, welchen Einschränkungen diese Verfahren bisweilen unterliegen, und was generell



bei der Erhebung des GSK zu beachten ist. In diesem Zusammenhang werden Aspekte der Valenz, der sozialen Erwünschtheit sowie Typizität diskutiert, um abschließend zur Konstruktion eines neuen Inventars überzuleiten, im Zuge derer diese Erkenntnisse berücksichtigt werden sollen.

### **2.3.1 Übersicht der Entwicklung der Definition und Messinstrumente**

#### **2.3.1.1 Maskulinität und Femininität als bipolare Dimensionen**

Lippa (2001; 2005) sah in dem 456 Items umfassenden Fragebogen *Attitude Interest Analysis Survey (AIAS)* von Terman und Miles (1936) den ersten in einer langen Reihe von Instrumenten zur Messung des Geschlechtsrollenselbstkonzepts, die Maskulinität und Femininität als unipolare Skalen und damit sich einander ausschließender Gegensätze eines Kontinuums definierten. Dabei ergab sich zwangsweise, dass eine hohe Maskulinität (positive Punktwerte) mit gleichzeitiger geringer Femininität (negative Punktwerte) einher ging. Terman und Miles (1936), die Wort- und Bildassoziationen, allgemeines Wissen, Interessen, Einstellungen, Persönlichkeit und die emotionalen Reaktionen in zahlreichen Situationen erfassten, sahen eine gewisse Analogie zwischen Intelligenz und Geschlechtsrollenstereotypen dahingehend, dass ihr Fragebogen einem Abbild der Geschlechtsunterschiede in verschiedenen Bereichen wie Intelligenz oder beruflichen Interessen der damaligen Zeit entspricht. Obwohl methodisch nicht frei von Kritik, stellte der Fragebogen von Terman und Miles mit seiner theoretischen bipolaren Grundannahme den konzeptuellen Rahmen für alle zunächst folgenden Messinstrumente in diesem Bereich dar.

So selektierte Strong (1943) für die M-F-Skala (Maskulinität und Femininität) seines Berufsfragebogens (*Strong Vocational Interest Blank*) in ähnlicher Weise wie seine Vorgänger Items vor allem basierend auf statistisch signifikanten Geschlechtsunterschieden. Generell setzte sich in den 1940er und 1950er Jahren die Einführung einer M-F-Skala in den allgemein gängigen Persönlichkeitsinventaren durch: Guilford-Zimmerman Temperament Inventory (Guilford & Zimmerman, 1956), California Psychological Inventory (CPI; Gough, 1957), Minnesota Multiphasic Personality Inventory (MMPI; Hathaway & McKinley, 1951), die sich in Abhängigkeit der postulierten Theorie in den Inhalten leicht unterscheiden. Maskulinität und Femininität wurden dabei in ähnlicher Weise definiert: Maskulinität als instrumentelle (durchsetzungsfähig, dominant, unabhängig) und Femininität als expressive Persönlichkeitseigenschaften (einfühlsam, fürsorglich; sensibel) (Bakan, 1966; Parsons & Bales, 1955), wobei die Diskussion um die Bezeichnung der zwei Skalen noch Dekaden andauern sollte. Parsons und Bales (1955), die in einer soziologischen Untersuchung gesellschaftliche Funktionen im System amerikanischer Kleinfamilien betrachteten, führten beispielsweise die Begriffe instrumentell und expressiv ein:

*"The area of instrumental function concerns relations of the system to its situation outside the system, to meeting the adaptive conditions of its maintenance of equilibrium, and 'instrumentally' establishing the desired relations to external goal-objects. The expressive area concerns the 'internal' affairs of the system, the maintenance of integrative relations between the members, and regulation of the patterns and tension levels of its component units."* (Parsons & Bales, 1955, S. 47).

Die instrumentelle Funktion / Orientierung ist das Streben nach externen Zielobjekten, verbunden mit Macht und als exekutive Rolle durch den Vater ausgefüllt, während die Mutter in einer expressiven Orientierung Dinge innerhalb der Familie fokussiert. Beide Konzepte sind durchaus denen der Maskulinität (Instrumentalität) und Femininität (Expressivität) ähnlich.

Bakan (1966), dem ein großer Teil der begrifflichen Grundlagen zugeschrieben wurde, führte schließlich etwas später in seinem Essay über Psychologie und Religion („*The duality of human existence*“) noch zwei weitere Begriffe ein: *Agency* und *Communion* ein:

„I have adopted the terms "agency" and "communion" to characterize two fundamental modalities in the existence of living forms, agency for the existence of an organism as an individual, and communion for the participation of the individual in some larger organism of which the individual is a part.“ (S. 14-15).

Mit *Agency* als eine Form des individuellen Bestrebens nach Effektivität, Beherrschen, Selbstschutz und Selbstwirksamkeit in ständiger Auseinandersetzung mit der sozialen Umwelt und *Communion* als dem Streben nach Bildung und Aufrechterhaltung zwischenmenschlicher Beziehung unterschied Bakan zwei grundlegende Formen der menschlichen Existenz, die zunächst einmal von ihm nicht in direkten Zusammenhang mit den Geschlechtern gebracht wurden. Die inhaltliche Nähe zu den später als geschlechtstypisch identifizierten Eigenschaften führte dann zu der kontrastierten Einteilung der Geschlechterstereotype in *Agency* und *Communion*. Auch in diesem Zusammenhang können Beziehungen zu Teilskalen anderer Messinstrumente herangezogen werden. So würde *Agency* am ehesten mit Maskulinität und *Communion* mit Femininität in Beziehung gesetzt werden. Während sowohl *Agency* / *Communion* als auch *Instrumentalität* / *Expressivität* Bestandteile von Theorien und Instrumenten zur Messung des GSK sind, repräsentieren heute *Maskulinität* und *Femininität* das GSK im klassischen Sinne.

Eine der ersten wichtigen Arbeiten zu Geschlechterstereotypen stellt die von Broverman (1970) aus den 1960er Jahren dar. Maskulinität und Femininität, immer noch als zwei Enden eines einzelnen bipolaren Geschlechterrollenkontinuums betrachtet (Abbildung 6), wurden jeweils als stereotypisch definiert, wenn mindestens 75% der Probanden es typischer für das eine Geschlecht als für das andere hielten.

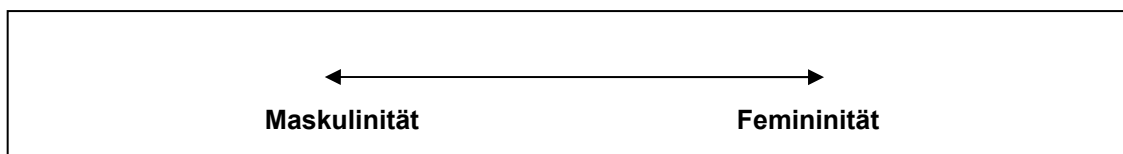


Abbildung 6. Maskulinität und Femininität als bipolare Dimensionen

Diese klare Zweiteilung spiegelte sich auch in der gesellschaftlichen Akzeptanz des geschlechtstypischen Verhaltens wider: so galt es normal und angemessen, für Männer typisch *maskuline* (z.B. Aggression, Dominanz, Ambitionen) und für Frauen typisch *feminine* (z.B. Leidenschaft, Empathie, Fürsorge) Verhaltensweisen zu zeigen. Die Geschlechterrollen waren einander ausschließend, maskuliner zu sein bedeutete gleichermaßen weniger feminin zu sein und umgedreht.

### 2.3.1.2 Maskulinität und Femininität als unabhängige Dimensionen

Diese Konzeption von Maskulinität und Femininität als bipolare Enden eines Kontinuums wurde in Frage gestellt: die Forschungsergebnisse bezüglich einer möglichen Multidimensionalität, einer bipolaren Konzeption oder der Itemauswahl konnten nur unzureichend erklärt werden. Vor dem Hintergrund der fortschreitenden gesellschaftlichen Entwicklung und damit nicht zuletzt auch der modernen Frauenbewegung ändert sich der Blick auf Maskulinität und Femininität und Constantinople (1973) diskutierte beide Begriffe als zu den theoretisch und empirisch schwammigsten Konzepten in der Psychologie zählend: “[...] *both theoretically and empirically [masculinity and femininity] seem to be among the muddiest concepts in the psychologist’s vocabulary*” (S. 390).

Aufgrund einer Vielzahl von faktoranalytischen Studien und theoretischer Verfeinerung der Definition erfolgte in den 1970er Jahren schließlich die notwendige Überarbeitung von Maskulinität und Femininität in eine zwei-dimensionale Konzeption, wonach Maskulinität und Femininität voneinander unabhängige und orthogonale Dimensionen darstellen mit der für eine Person daraus resultierenden Möglichkeit auf beiden Dimensionen unabhängig ihres biologischen Geschlechts hohe Ausprägungen aufzuweisen (Constantinople, 1973). Diese theoretische Fundierung zugunsten zweier orthogonaler unabhängiger Faktoren (Bem, 1974; Block, 1973; Constantinople, 1973; Deaux & Kite, 1987; Helmreich, Spence, & Wilhelm, 1981; Lippa, 1991; Lippa, 2002; Lippa & Tan, 2001; Lubinski et al., 1983; Spence et al., 1975; Spence & Helmreich, 1978) sollte für die folgenden 30 Jahre die Grundlage für Forschungen im Bereich der Geschlechtsrollenstereotypen bilden.

Vor allem Bem (1974; 1977; 1979) hatte großen Anteil an dieser neuen Entwicklung. Sie ging in ihren Ausführungen davon aus, dass Maskulinität und Femininität innerhalb der individuellen Persönlichkeit gleichberechtigt koexistieren können. Diese Unabhängigkeit der Dimensionen und die flexiblen Ausprägungen lassen eine neue Geschlechtsrolle zu, die Bem als *Androgynität*<sup>11</sup> bezeichnete.

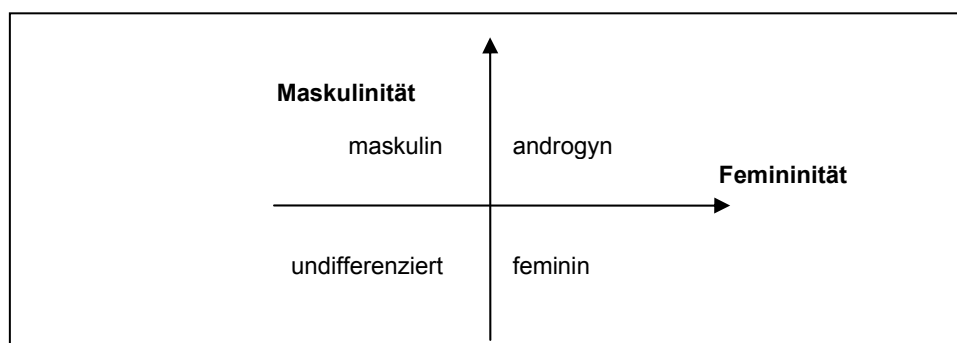


Abbildung 7. Maskulinität und Femininität als zwei unabhängige Dimensionen

Ursprünglich unspezifisch definiert als Balance zwischen Maskulinität und Femininität und unabhängig von der absoluten Höhe der Ausprägungen, waren es nicht zuletzt Spence, Helmreich und Stapp (1975), die Bem (1977; Bem et al., 1976) wie auch andere später davon überzeugten, den Begriff *androgyn* auf Individuen zu beschränken, die tatsächlich hohe

<sup>11</sup> Die aufgrund der Androgynität existierende situative Flexibilität wurde bereits weiter oben in der Theorie diskutiert.

Ausprägungen auf Maskulinität und Femininität aufweisen und damit einen klaren Unterschied zu Personen aufmachen, die auf beiden Dimensionen niedrige Ausprägungen haben und von nun an als *undifferenziert* klassifiziert wurden (Abbildung 7). Androgynität hat dabei ihren sprachlichen Ursprung im Griechischen *άνήρ* (*androgynos*), setzt sich zusammen aus *άνδρός* (Mann) und *γυνή* (Frau) und meint die Vereinigung von männlichen und weiblichen Eigenschaften.

Durch diese Unabhängigkeit von Maskulinität und Femininität sowie der Einführung des neuen Konstruktes Androgynität, wurden die bis dahin entwickelten Geschlechtsrolleninstrumente gegenstandslos und neue Verfahren zur Erfassung des Geschlechtsrollenelbstkonzepts notwendig. Verschiedene Forschungsgruppen haben zu diesem Vorhaben und der Diskussion um theoretische Fundierung sowie der methodischen Umsetzung beigetragen: z. B. Bem Sex Role Inventory (BSRI, Bem, 1974; 1981a), der Personal Attribute Questionnaire (PAQ; Spence et al., 1974), die Adjective Check List (Heilbrun, 1976), die PRF ANDRO Scale (Berzins, Welling, & Wetter, 1978) oder die Multicultural Masculinity Ideology Scale (Doss & Hopkins, 1998). Bisweilen liegen den Instrumenten sehr unterschiedliche konzeptuelle Ansätze zugrunde, so basierte Bem (1974; 1981a) ihren BSRI

*“on a theory about both the cognitive processing and the motivational dynamics of sex-typed and androgynous individuals”* (Bem, 1979, S. 1048).

Die Items des BSRI wurden aufgrund konsensueller Zustimmung (1 = *nicht erwünscht* bis 7 = *sehr erwünscht*) von amerikanischen Studierenden aus sozial erwünschten Persönlichkeitseigenschaften für Männer und Frauen ausgewählt (z. B.: *“In American society, how desirable is it for a man / woman to be ...?”*) (Bem, 1974, S. 157). Eine Eigenschaft galt demzufolge dann als feminin, wenn sowohl männliche als auch weibliche Studierende diese Eigenschaften den Frauen als stärker sozial erwünscht zuschrieben als sie dies für Männer taten und umgekehrt. Der BSRI besteht aus 60 Persönlichkeitseigenschaften, von denen 20 stereotype weibliche (*gentle, sensitive to others*) und 20 stereotype männliche Charakteristiken (*assertive, self-reliant*) repräsentieren. Die verbliebenen 20 Attribute sollten als Maß der sozialen Erwünschtheit fungieren, gelten mittlerweile jedoch als geschlechtsneutrale Items (*conventional, happy*). Die dargebotenen Eigenschaften sind zum größten Teil als Adjektiv (*dominant*) oder länger gefasste Umschreibungen (*sensitive to the needs of others*) formuliert.

Mittels der Selbstbeschreibung erlaubt der BSRI zusätzlich zu Aussagen bezüglich der individuellen Ausprägung von Maskulinität und Femininität eine Klassifikation in eine von vier möglichen Geschlechtertypen:

- maskulin (hohe Maskulinitätswerte und niedrige Ausprägung auf Femininität)
- feminin (geringe Maskulinitätswerte und hohe Ausprägung auf Femininität)
- androgyn (hohe Maskulinitätswerte und hohe Ausprägung auf Femininität) sowie
- undifferenziert (niedrige Maskulinitätswerte und niedrige Ausprägung auf Femininität).

Neben Bems (1974) *Sex Role Inventory* bediente sich die Erforschung des psychologischen Geschlechts seit den 1970er Jahren vor allem auch noch des von Spence et al. (1974; 1975) entwickelten Personal Attribute Questionnaires (PAQ). Auf Grundlage der erweiterten Fassung

des *Sex Role Stereotype Questionnaires* (SRSQ) von Rosenkrantz (1968) fußend, entschieden Studierende anhand bipolarer Eigenschaften, welche Eigenschaften typisch oder ideal für Männer und Frauen waren. Ähnlich wie bei Bem bestanden die Items aus Adjektiven (*emotional*, *independent*) oder Umschreibungen (*helpful to others*). Für den Fragebogen wurden solche Items ausgewählt, die über mehrere Beurteilergruppen als jeweils typischer oder charakteristischer für das eine als das andere Geschlecht, aber gleichermaßen sozial erwünscht für beide Geschlechter (anders als beispielsweise bei Bem) beurteilt wurden.

Beide Instrumente unterscheiden sich in der theoretischen Fundierung (Geschlechterschematheorie vs. Geschlechtsidentitätstheorie, vgl. Spence, 1991), der Skalenbenennung [Maskulinität / Femininität (Bem, 1974) vs. Instrumentalität / Expressivität (Spence & Buckner, 1995; Spence & Helmreich, 1980)] der Konstruktion der Fragebögen oder der operationalen Vorschriften zur Typenklassifikation und Androgynitätsbestimmung [t-Test-Methode (Bem, 1974) vs. Median-Split-Methode (Spence & Helmreich, 1978)].

Obwohl der Messbereich der Instrumente wiederholt auf das instrumentelle und expressive Selbstkonzept eingegrenzt wird (Sieverding & Alfermann, 1992), gilt das additive Vier-Typen-Modells mittlerweile als anerkannt (Strauss, Köller, & Möller, 1996).

### **2.3.1.3 Kritik am BSRI**

Der BSRI gehört zu dem mit Abstand am weitesten verbreiteten Instrument zur Messung des Geschlechtsrollenelbstkonzepts Erwachsener (Beere, 1990). Eine Vielzahl von Ergebnissen bescheinigten dem Fragebogen zufriedenstellende Reliabilitäten und Validitäten der Skalen, z. B. Cronbach Alpha Koeffizienten von .86 bis .87 für Maskulinität und .75 bis .78 für Femininität (Ballard-Reisch & Elton, 1992; Bem, 1974; Holt & Ellis, 1998; Taylor & Hall, 1982). Dennoch sahen sich Bems Annahmen insgesamt sehr schnell einer scharfen theoretischen und empirischen Kritik ausgesetzt (Ballard-Reisch & Elton, 1992; Blanchard-Fields, Suhrer-Roussel, & Hertzog, 1994; Locksley & Colten, 1979; Marsh, 1985; Myers & Gonda, 1982a; Pedhazur & Tetenbaum, 1979; Spence & Buckner, 2000; Spence & Helmreich, 1980; Strauss et al., 1996).

Im Zuge der Entwicklungen der letzten Jahre haben verschiedene Forscher die Itemauswahl repliziert (Edwards & Ashworth, 1977; Walkup & Abbott, 1978), die Faktorenstruktur untersucht (Feather, 1978; Gaudreau, 1977; Moreland, Gulanick, Montague, & Harren, 1978; Waters, Waters, & Pincus, 1977) oder den Zusammenhang zu Verhaltensvariablen hergestellt (Bem & Lenney, 1976; Bem & Lewis, 1975; Bem et al., 1976). Die bestehenden, unterschiedlich sozial erwünschten Teilskalen Maskulinität und Femininität wurden einer umfangreichen Prüfung unterzogen. Dabei fanden sich uneinheitliche Resultate, die teilweise den BSRI in seiner Struktur bestätigten und teilweise scharfe Kritik übten an der Skalenstruktur, den spezifischen Itemkennwerten oder gewählter Methoden für seine Konstruktion (Antill & Russell, 1982; Marsh & Myers, 1986; Ricciardelli & Williams, 1995; Russell & Antill, 1984; Spence, Helmreich, & Holahan, 1979), die – berücksichtigt man die enge inhaltliche und theoretische Nähe zur Originalkonzeption des Messinstrumentes von Bem – in gleicher Weise für die deutsche Version von Schneider-Düker und Kohler (1988) gelten dürfte.

#### *Skalenbezeichnung*

Schon Beere (1990) stellte in ihrer Übersicht fest, dass bestehende Instrumente bisweilen sehr inkonsequent in ihrer Namensgebung seien und nahm in diesem Zusammenhang Bezug auf Bems Sex Role Inventory, welches aufgrund der theoretischen Definition der Begrifflichkeiten korrekterweise Bem Gender Role Inventory heißen müsste - misst es doch das Geschlechtsrollenelbstkonzept (vgl. Abschnitt 2.2.1 Definition von Geschlechtsrollen und Geschlechterstereotypen, S. 9).

Lippa (1991; 1995) stellte des Weiteren eine Ähnlichkeit zu den Big Five<sup>12</sup> fest: maskuline Instrumentalität weist Gemeinsamkeiten zu Extraversion und zu Neurotizismus sowie feminine Expressivität zu Verträglichkeit auf. Die folgende Tabelle präsentiert Beispiele bevorzugter Skalenbezeichnungen, die in direkter oder indirekter Weise mit Maskulinität oder Femininität eine theoretische und empirische Ähnlichkeit aufweisen.

*Tabelle 3. Bezeichnungen von Maskulinität und Femininität in verschiedenen Studien*

	<b>Maskulinität</b>	<b>Femininität</b>
Bakan (1966)	Agency	Communal aspects
McClelland, Atkinson, Clark & Lowell (1976) <sup>a</sup> McAdams & Vaillant (1982) <sup>b</sup>	Bedürfnis nach Leistung <sup>a</sup>	Bedürfnis nach Intimität <sup>b</sup>
Parsons & Bales (1955) Locksley & Colten (1979) Spence & Helmreich (1978)	Instrumental traits	Expressive traits
Berzins et al. (1978)	z. B. Autonomy	z. B. Self-subordination
Pedhazur & Tetenbaum (1979)	Assertiveness / Instrumentality	Interpersonal sensitivity
Blatt & Shichman (1983)	Anaclitic Orientation	Introjective Orientation
Lubinski, Tellegen & Butcher (1983)	Assimilative style	Accommodative style
Wiggins & Trapnell (1996)	Dominance Dimension	Nurturance Dimension

Trotz unterschiedlicher Benennungen für Maskulinität und Femininität in Tabelle 3, Kritik an der Bezeichnung von Bems Skalen und der Zweifaktorenstruktur als inadäquate Lösung für die Repräsentation von Maskulinität und Femininität konnten weitere Faktoren wiederholt den beiden ursprünglich postulierten Kategorien zugeordnet werden und somit Bems Zweifaktorenlösung unterstützen (Brems & Johnson, 1990). Dennoch gilt es, die jeweils verwendeten Kategorientitel mit Sorgfalt zu beachten, da sie nicht nur unterschiedliche theoretische Fundierungen repräsentieren, sondern darüber hinaus auch den Einsatz unterschiedlicher Messinstrumente implizieren, die nicht zuletzt bei der Interpretation und dem Vergleich von Ergebnissen berücksichtigt werden sollten.

#### *Validität und Reliabilität*

Die Hauptkritik des BSRI bezog sich zwar bevorzugt auf die Bezeichnung und Struktur der zwei definierten Kriterien, jedoch gab es zusätzliche Aspekte, die eine fortwährende methodische

<sup>12</sup> Ein Fünf-Faktoren-Modell (FFM) der Persönlichkeitspsychologie, das Neurotizismus, Extraversion, Offenheit für Erfahrungen, Verträglichkeit und Gewissenhaftigkeit erfasst.

Diskussion nicht abreißen ließen. So gab es unter anderem Kritik bezüglich der Studentenpopulationen als Konstruktions- und Validierungsgruppe (Locksley & Colten, 1979) oder der Tatsache, dass der BSRI an und für die amerikanische Gesellschaft konzipiert wurde (Pedhazur & Tetenbaum, 1979), was nicht nur die Generalisierbarkeit der Ergebnisse einschränkt, sondern auch einen Einsatz über die amerikanischen Landesgrenzen hinaus in Frage stellt.

Pedhazur und Tetenbaum (1979) attackierten Bems empirischen Ansatz zur Auswahl der maskulinen und femininen Items. Die Autoren argumentieren, dass durch die Selektion der Items zu den Skalen aufgrund der Tatsache, eine Eigenschaft müsse sozial erwünscht für das eine Geschlecht jedoch nicht für das andere sein, Bem schließlich doch eine Bipolarität der Eigenschaften voraussetzt, was wiederum dem theoretischen Konstrukt widerspricht.

Ursprünglich ist der BSRI aufgrund sozial erwünschter maskuliner und femininer Eigenschaften erstellt worden. In den vergangenen Jahren gab es jedoch eine Verschiebung des evaluativen Inhaltes einiger Items, wonach einige BSRI-Items nun eher zu den sozial unerwünschten zu zählen sind, was jedoch wiederum vor allem für feminine Items gilt (*gullible*). Demnach wären Ergebnisse dieser Skala allgemein schwierig zu interpretieren (Grimmel & Stern, 1992; Spence, 1993).

Strauß et al. (1996) zeigten beim Einsatz des BSRI, dass einige Items beider Skalen sehr geringe Trennschärfen von unter .30 aufwiesen (*liebe Sicherheit, verführerisch, benutzt keine barschen Worte, glücklich, abhängig, etc.*) und lieferten einen Hinweis darauf, dass die Skalen auf 18 Items für die M-Skala und 10 Items für die F-Skala verkürzt werden können. Einschränkend muss man jedoch anmerken, dass die Stichprobe aus Schülern und Schülerinnen bestand.

In zwei Studien mit Einschätzungen bezüglich einer unspezifizierten Zielperson, Beurteilung der sozialen Erwünschtheit der Items sowie Selbst-, Typizitäts- und Idealeinschätzung zweifelten Myers und Gonda (1982a) ebenfalls die Validität des BSRI an. So fanden die Autoren signifikante Unterschiede in Abhängigkeit der verschiedenen Instruktionen, fehlende Übereinstimmung in den Einschätzungen von Männern und Frauen sowie ein inkonsistentes Muster in den itemspezifischen Geschlechtsunterschieden.

Darüber hinaus adressierte eine Vielzahl von Studien die Validität (Ballard-Reisch & Elton, 1992; Thompson, 1989; Wong, McCreary, & Duffy, 1990) sowie wiederholt Bems postulierte Zwei-Faktorenstruktur. In Anwendung verschiedenster Methoden (Campbell, Gillaspay, & Thompson, 1997; Choi, Fuqua, & Newman, 2007; Marsh & Myers, 1986; Martin & Ramanaiah, 1988), konfirmatorischer Faktorenanalysen oder Strukturgleichungsmodelle (Blanchard-Fields et al., 1994) kamen die Forschungsgruppen beim BSRI zu einer Zwei-, Drei- oder sogar Vier- und Mehr-Faktorenlösung. Dabei ist es vor allem die Vierfaktorenlösung, die sich in der Diskussion und Anwendung durchzusetzen scheint (Brems & Johnson, 1990; Gross, Batlis, Small, & Erdwins, 1979; Pedhazur & Tetenbaum, 1979; Waters et al., 1977).

Auch der Personal Attribute Questionnaire (Spence & Helmreich, 1978) ist nicht gänzlich frei von Kritik. Abgesehen von gemeinsamen Debatten über Theorien, Konzeptionen

oder rechtfertigender Empirie liefert beispielsweise der Aufbau des PAQ einen Anstoß zur Skepsis. Während das Design des BSRI einzelne Einschätzungen von Eigenschaften (*helpful, yielding*) ermöglicht, sind es im PAQ Eigenschaftspaare. In einigen Fällen verwenden die bipolaren Items beispielsweise bei der Kontrastierung der Extreme den gleichen Wortstamm/-laut wie z. B. *very little need for security* vs. *very strong need for security*, während in anderen Fällen unterschiedliche Inhalte zum Einsatz kommen (*very rough* vs. *very gentle*) und jeweils auf einer Fünf-Punkte Skala bewertet werden sollen. Die Präsentation dieser gepaarten Items bleibt etwas fragwürdig, da in der Konstruktion nicht sichergestellt wurde, dass jedes Paar auch tatsächlich Gegensätze des gleichen Konstrukts darstellt – ein Problem, das nahezu allen Instrumenten inne liegt, die sich eines Aufbaus ähnlich dem des semantischen Differentials bedienen.

In den letzten Jahren konzentrierte sich die Forschung bezüglich des BSRI vor allem darauf inwieweit die Items eine Differenzierung zwischen Männern und Frauen in der heutigen Gesellschaft noch gewährleisten können (Auster & Ohm, 2000; Harris, 1993; Holt & Ellis, 1998; Walkup & Abbott, 1978). Auster und Ohm (2000) fanden in ihren Analysen der Erwünschtheitsratings, dass immer noch 18 von 20 Originaleigenschaften als feminine Items klassifiziert werden können, während dies nur auf acht von 20 maskulinen Charakteristiken zutraf. Zudem zeigte eine Rangordnung der Eigenschaften nach Erwünschtheit, dass beispielsweise die femininen Eigenschaften als erwünschter für Frauen gelten, was die Persistenz traditioneller Geschlechterbeschreibungen demonstriert. Im Vergleich dazu konnten Holt und Ellis (1998) ebenfalls die soziale Erwünschtheit von 18 der ursprünglich 20 femininen Items für Frauen bestätigen, ebenso wie die aller 20 maskulinen Eigenschaften für die Männer. Die Studien zeigen, wie variabel und unzuverlässig die Ergebnisse mit dem BSRI sind.

#### *Kulturunterschiede*

Obwohl der Vergleich von Geschlechtsrollenselbstkonzepten über mehrere Kulturen hinweg sehr populär ist, gestaltet sich dieser methodisch schwierig, da in der Mehrheit der Studien auf Messinstrumente zurückgegriffen wurde, deren Ursprung in westlich geprägten Kulturen liegt und somit kulturspezifische Aspekte vernachlässigt. Dementsprechend wurden im Bereich der kulturvergleichenden Forschung Ähnlichkeiten zwischen Kulturen vor allem in westlich geprägten Ländern nachgewiesen (Runge, 1981; Williams, 1977; 1979; 1981). Dabei hatte selbst Mead (1935) in ihren Arbeiten gezeigt, dass einige Geschlechtsunterschiede universell sind und damit durchaus in allen Nationen vorkommen müssten. Wie zu recht auch von Bozionelos und Bozionelos (2003) angemerkt, sind die meisten Instrumente auf ein angelsächsisches Cluster aufgebaut – was erfordern würde, die Modelle für andere Cluster zu überprüfen, da jeweils kulturspezifische Erwartungen an die Geschlechter gestellt werden.

Dies bestätigen auch Studien mit dem BSRI, die wiederholt demonstrieren, dass die Konzeptualisierung dessen, was als männlich bzw. weiblich gilt, kulturspezifisch für den US-amerikanischen Raum und die weiße Bevölkerungsgruppe ist. Williams und Best (1990) konnten in ihrem Vergleich von 30 verschiedenen Ländern (aus Amerika, Europa, Asien, Afrika und Ozeanien) zwar mehr Ähnlichkeiten als Unterschiede aufzeigen, die Anzahl der



kulturübergreifenden allgemein gültigen Attribute für beide Geschlechter bleibt jedoch gering. So wurden lediglich sechs Attribute identifiziert, die als typisch für Männer in allen Kulturen anerkannt sind (dominant, unabhängig, stark) sowie drei für Frauen (sentimental, gehorsam, abergläubisch). Insgesamt verzeichneten Williams und Best (1990) in allen Ländern eine stärkere Zuschreibung maskuliner Eigenschaften für beide Geschlechter, die die Autoren auf den besseren Status des maskulinen Stereotyp zurückführten, welches gerade für die so häufig getesteten Studierenden mit einem eher ausgeprägten Leistungsmotiv ansprechender ist als das feminine Stereotyp. Sowohl für andere Gruppen innerhalb der USA als auch für andere Nationen variieren die Konzepte von Maskulinität und Femininität bisweilen erheblich und stellen die dortige Anwendung des BSRI in Frage. Studien mit spanisch-stämmigen Amerikanern (Harris, 1994; Sugihara & Warner, 1999), Afro-Amerikanern (Harris, 1994; Konrad & Harris, 2002) und weiblichen indianischen Ureinwohnern (Portman, 2001) zeigen, dass Frauen und Männer in unterschiedlichem Maße nicht den Normen des BSRI aus der Validierungsstudie von 1978 entsprechen.

Die Geschlechtsschematheorie blieb überraschend offen bezüglich der inhaltlichen Ausprägung der Eigenschaften, die als typisch männlich bzw. weiblich aufgefasst werden. Bem (1981b) hält diesbezüglich fest:

*"Because sex-typed individuals are seen as processing information in terms of and conforming to whatever definitions of masculinity and femininity the culture happens to provide, it is the process of partitioning the world into two equivalence classes on the basis of the gender schema, not the contents of the equivalence classes, that is central to the theory." (S. 356)*

Gerade Bems Auffassung der ursprünglichen Kategorisierung der Welt in zwei Klassen sowie die Anerkennung der kulturellen Spezifikation von Maskulinität und Femininität unterstreichen die „Anfälligkeit“ der Geschlechterstereotypen für Veränderungsprozesse.

Troche, Weber, Hennigs, Andresen und Rammsayer (2007) fanden bei ihrer Anwendung des BSRI einen Hinweis auf Unterschiede in der femininen Selbstbeschreibung in Abhängigkeit der Nationalitätzugehörigkeit der Probanden (z. B. italienische Probanden schrieben sich die femininen Eigenschaften im geringeren Maße zu als schwedische Probanden). Der BSRI, ursprünglich für die amerikanische Gesellschaft konzipiert, wird mittlerweile weltweit eingesetzt, häufig durch eine einfache Übersetzung der Eigenschaften. In überraschend wenig Fällen kommt es zu einer Neukonstruktion vor dem Hintergrund spezifischer vorherrschender Rollen und Vorstellungen von Männern und Frauen, wie z. B. im Falle einer japanischen (Katsurada & Sugihara, 1999; Sugihara & Katsurada, 2000) oder der deutschen Version (Schneider-Düker & Kohler, 1988). Obwohl Bem (1981a) selbst die Kulturspezifität ihres Instrumentes thematisierte:

*"[...] the BSRI is based on a conception of the traditionally sex-typed person as someone who is highly attuned to cultural definitions of sex-appropriate behavior [...]" (S.4)*

*„In order to ensure that each judge would rate the various characteristics from the perspective of the culture at last [...]" (S. 11),*

überrascht es, dass beispielsweise keine klassische BSRI Validierung für Großbritannien existiert, die über die Einschätzungen von sehr jungen Studierenden hinaus geht<sup>13</sup>.

Moore und Rosenthal (1980) fanden in einer australischen Stichprobe zufriedenstellende Ergebnisse hinsichtlich der Reliabilität, was wiederum die Stärke des Instruments für den englischsprachigen Raum aufwertet, da Australien und Großbritannien beispielsweise als ähnlich in Organisationsstrukturen gelten (Brown, 1998). Die westlichen, englischsprachigen Nationen teilen zweifelsohne gesellschaftliche Normen, dennoch kann nicht automatisch postuliert werden, dass kulturspezifische Unterschiede nicht auch zu anderen Ergebnissen mittels des BSRI führen können.

Tabelle 4. Einsatz des BSRI und weiterer Instrumente in verschiedenen Nationen

Instrument	Auswahlkriterien	Sprache / Land / Übersetzung	Besonderheiten
<b>BSRI</b>			
Bem (1974)	Soziale Erwünschtheit	Englisch / USA	
Wilson, McMaster, Greenspan, & Mboyi (1990)		Englisch / ZIM	
Persson (1999)		Englisch / SWE	70% der Attribute trennen zwischen M und F in Schweden; nicht-zutreffende Worte: shy, yielding, gullible, childlike
Schneider-Düker & Kohler (1988)		Deutsch / GER	
Gana (1995)		Französisch / FRA	
Fontayne, Sarrazin, & Famose (2000)		Französisch / FRA	
Sebastián (1990)		Spanisch / ESP	
De Leo, Villa, Magni, Andreatta, & Gagliardi (1986)		Italienisch / ITA	
Katsurada & Sugihara (1999); Sugihara & Katsurada (2000)	Typizität, Soz. Erwünschtheit <sup>14</sup> , Vor- u. Rückübers., Selbsteinsch.	Japanisch / JAP	Insgesamt höhere Werte der F-Skala. Geschlechtsunterschiede nur bei der M-Skala (Männer höhere Werte)
Peng (2006)		Chinesisch / TPE	
<b>BSRI - Überprüfung der Validität</b>			
Wilcox & Francis (1997b)		English / GBR	
Acuña & Bruner (1986)		Spanisch / MEX	
Qin & Yanjie (2003)		Chinesisch / CHN	
Sethi & Allen (1984)	Soziale Erwünschtheit	Hindu / IND	Replikation Bems Auswahlprozesses, 60 BSRI + 44 neue; Ähnlichkeit der Itemzuordnung
Maznah & Choo (1986)	keine Angabe (lediglich Bezug zur Übersetzung)	Malaiisch / MAS	Bestätigung der Multidimensionalität des BSRI - neben M und F weitere Unterfaktoren
<b>PAQ</b>			
Spence, et al. (1974)	Typizität	Englisch / USA	
Rinfret & Lévesque (2000)		Französisch / CAN	PAQ übersetzt, an Franz., Eng.-sprechender, kanad. Manager (öffentl. Dienst) validiert. Test-Retest-Reliabilität überprüft.
Sebastián (1990)		Spanisch / ESP	
Runge (1981)		Deutsch / GER	d. Stichprobe. Ergebnisse stützen konzept. Model von Spence & Helmreich, 1978
<b>weitere Instrumente</b>			
Antill, Cunningham, Russell, & Thompson (1981)		Englisch / AUS	

Wilcox und Francis (1997a; 1997b), die den BSRI in England einsetzten, ließen weibliche Studenten (16-19 Jahre) die Erwünschtheit der Items für die britische Bevölkerung einschätzen,

<sup>13</sup> E-Mailkonversation mit S. Bem (26. Juni 2001)

<sup>14</sup> In der Studie von 2000 wurden die Probanden gebeten die Typizität im Ausschlussprinzip einzuschätzen: das Item als maskulin, feminin oder beides zu identifizieren. Alle maskulinen Eigenschaften, jedoch nur 12 der femininen Eigenschaften wurden als geschlechtstypisch eingeschätzt, wobei die Definition der sozialen Erwünschtheit bezüglich dieser Form der Auswahl zu hinterfragen wäre.

wonach nur drei aus den beiden jeweils 20 Items umfassenden Skalen als valide Messung für die maskuline und feminine Geschlechtsrolle galten.

Für den deutschsprachigen Raum kam die Studie von Swazina, Waldherr und Maier (2004) zu dem Ergebnis, dass sich die Soziale Erwünschtheit von Eigenschaften in Deutschland und Österreich nicht allzu sehr unterscheidet und einen grenzübergreifenden Einsatz daher ermöglicht. Eine Legitimation für einen weltweiten deutschsprachigen Einsatz bedeutet dies jedoch noch nicht, da nationenspezifische und kulturelle Unterschiede immer Berücksichtigung finden sollten. Tabelle 4 (S. 56) präsentiert einen Überblick – welcher lediglich als Auszug verstanden werden kann – über den Einsatz des BSRI in verschiedenen Nationen und inwieweit kulturspezifische Validierungen durchgeführt wurden oder eine einfache Übersetzung zum Einsatz kam.

#### **2.3.1.4 GSK als multidimensionales Konstrukt**

Schon im Zusammenhang mit den zwei klassischen Messinstrumenten BSRI und PAQ wurde die Dimensionalität von Maskulinität und Femininität diskutiert. Die Anerkennung bis dato unberücksichtigter Aspekte und eine erste Erweiterung der Items um zumindest negative Geschlechterstereotype erfolgte bereits vom PAQ zum EPAQ durch Spence, Helmreich, und Holahan (1979), die negative Maskulinität durch Eigenschaften wie *egoistisch*, *arrogant* oder *diktatorisch* und negative Femininität mittels Items wie *weinerlich* oder *nörgelnd* abbildeten.

Auch Constantinople (1973) argumentierte bereits, dass sich das Geschlechtsrollenselbstkonzept durch gewisse Komplexität und Multidimensionalität auszeichnet, denen bisherige Instrumente zur Messung von Maskulinität und Femininität nicht gerecht werden können, zumindest nicht, wenn sie der ausschließlichen Unterscheidung in maskuline und feminine Eigenschaften folgen.

Eine Vielzahl Studien wies darüber hinaus nach, dass Geschlechterstereotype, -rollen wie auch -identitäten hierarchischer und vielfältiger sind als ursprünglich mit Maskulinität und Femininität angenommen (Ashmore, 1990; Athenstaedt, 2003; Baber & Tucker, 2006; Connell & Messerschmidt, 2005; Deaux & LaFrance, 1998; Deaux & Lewis, 1984; Deaux & Major, 1987; Eagly & Kite, 1987; Feather, 1984; Freeman, 1987; Helgeson, 1994a; Huston, 1983; Jackson & Cash, 1985; Koestner & Aube, 1995; Marsh & Myers, 1986; Myers & Gonda, 1982a; Oswald & Lindstedt, 2006; Ricciardelli & Williams, 1995; Ruble & Martin, 1998; Spence, 1993; Spence, 1999; Spence & Buckner, 2000; Spence & Sawin, 1985; Taylor & Hall, 1982; Twenge, 1999).

Spence und Kollegen (Spence, 1993; Spence & Sawin, 1985) führten in diesem Zusammenhang die *Multifactorial Gender Identity Theory* an, der die Annahme zugrunde liegt, die Geschlechtsidentität sei durch eine Kombination und Verknüpfung mehrerer geschlechtsbezogener Phänomene (Einstellungen, Interessen, Rollenverhalten oder eben Persönlichkeitseigenschaften) definiert. Diese verschiedenartigen, geschlechterdifferenzierenden Faktoren durchlaufen ihrerseits individuell unterschiedliche Entwicklungen unter dem Einfluss multipler Faktoren, die nicht notwendigerweise geschlechtsbezogen sein müssen; womit sich Ausprägungsunterschiede in den Charakteristiken auch innerhalb eines Geschlechtes erklären ließe. Konsequenterweise resümierte Spence, dass damit weder der BSRI noch ihr eigenes

Messinstrument, der PAQ (1984; 1991; 1993), geeignete Messinstrumente für globale geschlechtsbezogene Konstrukte sind, sondern lediglich relevante Variablen der Instrumentalität und Expressivität abbilden.

In Bezug auf die zwei Teilskalen Maskulinität und Femininität sprechen viele Analysen für ein inhaltliches Vierfaktorenmodell (Ricciardelli & Williams, 1995), was die Komplexität des Geschlechterrollenselbstkonzepts andeutet. Selbst dieses Vierfaktorenmodell ist womöglich nicht komplex genug, um die umfangreichen Rahmen der Geschlechtseigenschaften adäquat abbilden zu können (Blanchard-Fields et al., 1994; Spence, 1993; Thompson, Pleck, & Ferrera, 1992). Trotz der Popularität der Maskulinitäts- und Femininitätsmessung erfahren demnach die Skalen eine andauernde psychometrische und konzeptuelle Kritik dahingehend, dass es womöglich gar nicht Maskulinität und Femininität sind, was gemessen wird (Sieverding & Alfermann, 1992; Spence & Buckner, 1995; Spence & Helmreich, 1980). Andere kritische Stimmen merken an, dass die Teilskalen nur geschlechtsspezifische stereotypische Persönlichkeitseigenschaften abbilden, während ebenfalls relevante und gleichermaßen bedeutsame geschlechtsbezogene Charakteristiken wie beispielsweise Verhalten, Interessen, Fähigkeiten, Attraktivität oder beruflicher Status (Deaux & Lewis, 1984; Orlofsky, Ramsden, & Cohen, 1982; Spence & Helmreich, 1980) gänzlich unbeachtet bleiben. Geschlechterstereotype lassen sich inhaltlich in verschiedenen Bereichen beschreiben lassen. Myers und Gonda (1982a) fanden bei der von Probanden gegebenen Definition von Maskulinität und Femininität ebenfalls weniger spontan genannte instrumentelle oder expressive Charakteristiken, sondern vor allem Eigenschaften aus den Bereichen der physischen Erscheinung, Sexualität, sozialer Rollen; oder biologischer Aufgaben. Ähnlich sahen das auch Eagly und Kollegen (Cejka & Eagly, 1999; Diekman & Eagly, 2000; Oswald & Lindstedt, 2006), die in Hinblick auf Maskulinität und Femininität zwischen Persönlichkeit, physischen und kognitiven Eigenschaften unterschieden, während Athenstaedt (2003) beispielsweise bei ihrer Argumentation des geschlechtsbezogenen Selbstkonzepts die Unabhängigkeit positiver und negativer Eigenschaften sowie das soziale Rollenverhalten anführte. Das männliche Stereotyp steht für Aktivität, Stärke, während das weibliche Stereotyp wiederum Eigenschaften von Emotionalität, Soziabilität, Passivität und praktischer Intelligenz enthält (vgl. Alfermann, 1996).

Die Beschränkung auf geschlechterstereotype Eigenschaftswörter kann demnach unmöglich der Komplexität von Geschlechterrollenstereotypen gerecht werden (Fiske & Stevens, 1993). Im Prinzip sind der PAQ und der BSRI auf sozial erwünschte instrumentelle und expressive Persönlichkeitseigenschaften reduziert, während das Geschlechterrollenselbstkonzept darüber hinaus durch eine Vielzahl weiterer Konstrukte konstituiert wird. Unabhängig von den nachgewiesenen Charakteristiken der Geschlechterstereotype als multidimensional, idiosynkratisch oder multifaktoriell ist eines allen kritischen Stimmen gemeinsam: die Tatsache, dass die bis dato existierenden Skalen, selbst in ihrer bipolaren, orthogonalen Definition, eben lediglich einen Aspekt von Maskulinität und Femininität abbilden können (Spence, 1984; Spence, 1993; Twenge, 1999).

Sieverding und Alfermann (1992) empfahlen daher beim Einsatz von Eigenschaftslisten eben nicht von der *Maskulinität* oder *Femininität* zu sprechen, da diese nur einen Teilaspekt abdecken. Baber und Tucker (2006) stellten darüber hinaus in Frage, inwieweit gefundene Unterschiede zwischen den Geschlechtern tatsächlich Unterschiede in der Identifikation mit maskulinen und femininen Eigenschaften widerspiegeln oder sie lediglich ein Abbild ihrer Zustimmung dieser Eigenschaften sind.

### **2.3.1.5 Weitere Messinstrumente des GSK**

An dieser Stelle sollen weitere Instrumente zur Erfassung des Geschlechtsrollenselbstkonzepts aufgeführt werden. So existieren eine Vielzahl von aktuellen Messinstrumenten darüber, wie sich Personen im Hinblick auf geschlechts(rollen)bezogene Inhalte positionieren, z. B. *Social Role Questionnaire* (Baber & Tucker, 2006), zur Erfassung internalisierter, persönlicher Geschlechtsrollennormen, *Normative Geschlechtsrollenorientierung*, NGRO (Athenstaedt, 2000) oder der wahrgenommenen Gleichstellung von Männern und Frauen (Katenbrink, 2006).

Gerade die vorherrschenden gesellschaftlichen Normen und deren Auswirkungen stellen mittlerweile einen großen Teil der geschlechtsbezogenen Forschung dar. Demnach unterscheiden Personen sich dahingehend, inwieweit sie mit dieser Art von Normen konform gehen. In der Literatur wird in diesem Zusammenhang oft von Geschlechtsrolleneinstellungen gesprochen (Deaux, 1985; Deaux & LaFrance, 1998; Spence & Helmreich, 1972), wodurch der bewertende Aspekt in den Vordergrund gerückt wird. Andere Autoren wiederum sprechen von Geschlechtsrollen-Ideologie (Williams & Best, 1990). Im deutschen Sprachraum ist es eher üblich, von normativer Geschlechtsrollenorientierung zu sprechen (Krampen, 1979; 1983), während sich der amerikanische Sprachraum auf die Geschlechtsrollenorientierung bezieht, wie sie beispielsweise mittels der *Attitude towards Women Scale* (AWS) von Spence und Helmreich (1972) gemessen wird. Andere Maße wie der EPAQ (Spence et al., 1979) nahmen ersten Bezug zu den im BSRI vernachlässigten negativen Aspekten des Geschlechtsrollenselbstkonzepts, dies jedoch ohne genauen Bezug zu den theoretischen Grundlagen und der Relevanz dieser negativen Aspekte für die Geschlechtsrollenidentität.

Aufgrund eines Teils der Kritik zu attributbasierten GSK-Instrumenten erarbeitete Lippa (1991; 1995; Lippa & Connelly, 1990) ein alternatives Messinstrument, welches vor allem auf Verhaltensunterschiede aufbaut. Lippa (1991; 1995; Lippa & Connelly, 1990) betrachtete Maskulinität und Femininität als historisch und kulturell variierende Dispositionen, die durch Verhaltensunterschiede von Männern und Frauen spezifischer Populationen erfasst werden können. Auf Grundlage eben dieser Verhaltensunterschiede entwickelte Lippa daher das auf Diskriminanzanalysen basierende Verfahren *Gender Diagnosticity* (GD). Dieses Instrument misst geschlechtsbezogenes Verhalten, indem die Probanden auf fünfstufigen Skalen angeben, wie gern sie bestimmte Hobbies (Golf, Malerei), Tätigkeiten (reparieren, kochen) oder Berufe (Anwalt, Lehrer) ausführen würden. Lippas GD Verfahren bestimmt dann über alle Einschätzungen mithilfe der Diskriminanzanalyse die Vorhersagewahrscheinlichkeit, mit der ein Individuum als Mann oder Frau klassifiziert wird. Eine maskuline Person ist demnach durch

maskulines Verhalten im Vergleich zu einer unmittelbaren Referenzgruppe bestimmt. Durch dieses Vorgehen konnten Lippa und Connelly (1990) 92% der Männer und 88% der Frauen korrekt aus deren geschlechtsrollenkonsistenten Berufswünschen vorhersagen.

Dieses Verfahren bietet mehrere Vorteile: Der GD ist sowohl konzeptuell als auch empirisch distinkt von Maskulinität und Femininität, wie Bem (1974) sie definierte, sowie deren Kritik und lässt sich leichter mit Verhaltenskorrelaten in Beziehung setzen. Jedoch werden im Unterschied zur klassischen Berechnung von Maskulinität und Femininität die GD Werte für jede Gruppe neu kalkuliert. Während Lippa selbst (Lippa & Connelly, 1990; Lippa & Tan, 2001) dies als Vorteil gegenüber dem fixierten Eigenschaftsansatz von Bem betrachtet, lässt die unvermeidliche Situations- als auch Stichprobenabhängigkeit Lippas Instrument etwas fragwürdig erscheinen. Selbst ein verhaltensbasierter Ansatz ist also nicht notwendigerweise eine gelungene Alternative zu klassischen Selbstbeschreibungsmaßen bei der Messung der GSK - so kritikbehaftet diese Instrumente auch sein mögen.

Zusammenfassend kann festgestellt werden, dass der BSRI das am weitesten verbreitete Instrument zur Messung des GSK ist, gleichzeitig jedoch unter verschiedenen theoretischen und methodischen Aspekten kritisiert wurde, die den uneingeschränkten Einsatz dieser Skala mehr als in Frage stellen. Darüber hinaus ist der BSRI auf positive Aspekte des GSK beschränkt. Es ist allgemein anzuerkennen, dass Messinstrumente, die auf Persönlichkeitseigenschaften beruhen, lediglich einen Teil des GSK abbilden können.

Im folgenden Abschnitt soll die Bedeutung negativer Aspekte für das Geschlechtsrollenselbstkonzept deutlicher herausgearbeitet werden. Dazu wird zunächst die Rolle der positiven wie auch negativen Stereotypenzuschreibung theoretisch aus der Intergruppenforschung hergeleitet. Im Anschluss erfolgt die Diskussion von Aspekten der Typizität, Valenz sowie sozialen Erwünschtheit im Zusammenhang zu Eigenschaftsattributen, ehe der Bogen zu der konkreten Diskussion negativer Aspekte im Bereich der Messung zum GSK geschlagen wird. Diesbezüglich erfolgt danach der Bezug zu den Konzepten *Unmitigated Agency* und *Unmitigated Communion*, die die theoretischen Voraussetzungen negativer geschlechtsbezogener Zuschreibungen berücksichtigen und konkrete Vorhersagen der Beziehungen dieser Teilskalen zu einander implizieren.

### 2.3.2 Die Bedeutung negativer Aspekte der Geschlechtsrollenidentität

*„There are a number of ‚masculine‘ and ‚feminine‘ characteristics that are socially undesirable and have consequences for their possessors or those around them that may be deleterious“ (S. 1674).*

(Spence et al., 1979)

Das Geschlechtsrollenselbstkonzept wird allgemein definiert als die Menge der Geschlechterstereotype, die sich eine Person selbst zuschreibt (Athenstaedt, 2003; Bem, 1974; Hannover, 2000; Hoffman, 2001; Spence et al., 1974). Während die Soziale Identitätstheorie (Tajfel & Turner, 1986) postuliert, dass Menschen nach einer positiven sozialen Identität streben und daher mit größerer Wahrscheinlichkeit positive Attribute in ihre Selbstkonzept integrieren, werden negative Stereotype unter dieser Annahme zu einer Herausforderung für das Individuum.

Die Bedeutung negativer Eigenschaften für das Geschlechtsrollenselbstkonzept einer Person ist bislang weder theoretisch noch methodologisch hinreichend geklärt. Ursprünglich wird ein maskulines Geschlechtsrollenselbstkonzept dann angenommen, wenn sich eine Person maskuline Eigenschaften in stärkerer Ausprägung zuschreibt als feminine – dies bisher nahezu ausschließlich beschränkt auf positive Eigenschaften, nicht zuletzt durch die Art der Messinstrumente. Dabei wird jedoch ignoriert, dass Geschlechterstereotype durchaus auch negative geschlechtstypische Attribute beinhalten (Alfermann, 1996; Kelly & Worell, 1977; Prentice & Carranza, 2002; Ricciardelli & Williams, 1995), die nicht nur ad hoc produziert werden können (Ferring, Boll, & Neumann, 2002; Fiebert & Meyer, 1997; Oswald & Lindstedt, 2006), sondern auch sehr wohl zwischen den Geschlechtern differenzieren können (Kelly, Caudill, Hathorn, & O'Brien, 1977; Woodhill & Samuels, 2003). Williams und Best (1990) zeigten, dass Kinder Vorstellungen von sowohl positiven als auch negativen Stereotypen für beide Geschlechter haben.

Diese negativen Beschreibungen können auf die aus der Gruppen- und Selbstforschung bekannten negativen Aspekte der Identität einer Person zurückgeführt werden. Zusammengefasst bedeutet dies, dass beispielsweise bei einem negativen Vergleich zwischen der Eigen- und Fremdgruppe, das eigene Prestige sich verringert und damit eine negative soziale Identität gefördert wird. In diesem Zusammenhang diskutierten Tajfel und Turner (1979) zwar den Mechanismus, dass alles, was die Eigengruppe gegenüber der Fremdgruppe im positiven Sinne hervorhebt und abgrenzt, automatisch positiver bewertet wird, gleichzeitig erkennen sie jedoch die Relevanz eines verlorenen Gruppenvergleichs und der daraus resultierenden negativen Identität an. Diese negative Identität spiegelt sich auch im Geschlechtsrollenselbstkonzept und Ergebnissen mit anderen Konzepten wider.

Eine Vielzahl von Studien zu Depressionen, einem Krankheitsbild mit höherer Prävalenzrate für Frauen als für Männer, hat bezüglich möglicher psychologischer Geschlechtsunterschiede nahezu ausschließlich *Agency* und *Communion* (beide positiv in ihrer Valenz) untersucht und dabei einen Zusammenhang zwischen Depression und geringer Agency, nicht aber (hoher) Communion gefunden (für einen Vergleich siehe Helgeson & Fritz, 1999). Der Bezug zu geschlechtsspezifischen Persönlichkeitseigenschaften scheint zwar damit hergestellt, dennoch korrelierten traditionelle Messinstrumente von weiblichen geschlechtsbezogenen Eigenschaften nicht mit Depressions- oder Stresssymptomen.

Dieser fehlende Zusammenhang zu der den Frauen zugeschriebenen Geschlechtsrolle überrascht und stellt die Qualität derartiger Messinstrumente in Frage. Das Ziel kann nicht sein, die Abstinenz von Eigenschaften, gemessen mit einer (positiven) Teilskala (in dem Beispiel: Agency), die dem eigenen Geschlecht weniger entspricht, zur Vorhersage des geschlechtsspezifischen Verhaltens (Depression) heranzuziehen. Das Problem liegt bei der Abstinenz selbst. Die Feststellung einer geringen Ausprägung eignet sich nur unzureichend zur Vorhersage von Verhalten. Eine Abstinenz von etwas kann zwar tatsächlich dafür stehen, dass entsprechende Eigenschaften explizit gering ausgeprägt sind. Sie kann aber auch dadurch bedingt sein, dass die Eigenschaften eben nur weniger motiviert zur Selbstbeschreibung

herangezogen wurden – weil Personen sich nicht mit ihnen identifizieren wollen. Die Selbstzuschreibung auf diesen Eigenschaften liefert ein unzulängliches Bild von einem selbst und kann daher keine stabile Verhaltensvorhersage ermöglichen. Darüber hinaus lässt sich bei dem dargestellten Beispiel diskutieren, inwieweit *Communion* womöglich die feminine Geschlechtsrolle nicht ausreichend abzubilden vermag. Der Weg, einen verlässlichen Zusammenhang zu Verhaltenskonstrukten nachzuweisen, kann nur über die Gesamtheit von personenbeschreibenden Adjektiven – positiver und negativer Valenz – gehen. Erst die Betrachtung sozial unerwünschter Eigenschaften kann Forschern nicht nur einen neuen Zugang zur Erklärung von Problemverhalten wie Substanzmissbrauch ermöglichen, sondern auch die Ableitung angemessener Präventions- und Interventionsmaßnahmen. Hier greift auch Helgesons (Helgeson & Fritz, 1998; Helgeson & Fritz, 1999) Argumentation, die die Bedeutung der zwei zusätzlichen geschlechtsbezogenen Eigenschaftskomponenten *Unmitigated Agency* (UA) und *Unmitigated Communion* (UC)<sup>15</sup> herausarbeitet, die die Vorhersageleistung in Bezug auf andere psychologische Konzepte maximieren können.

Der folgende Abschnitt demonstriert in einem Überblick die allgemeine Bedeutsamkeit eben solcher negativer Eigenschaften in der Selbstbeschreibung. Dazu werden sowohl theoretische Grundlagen aus unterschiedlichen Theorien herangezogen als auch allgemein die Inhalte der Geschlechterstereotype diskutiert. In einer abschließenden Darstellung von Ergebnissen empirischer Forschung mit negativen Eigenschaften soll deren Bedeutung für das Geschlechtsrollenselbstkonzept beleuchtet werden. Durch diese Argumentation soll schließlich die Notwendigkeit unterstrichen werden, mittels einer Neuentwicklung eines Messinstrumentes diese unerwünschten Eigenschaften angemessen zu berücksichtigen.

### **2.3.2.1 Selbstzuschreibungen: Theoretische Perspektiven**

Einige der theoretischen Grundlagen finden sich in den kognitiven Geschlechtsschemata (Bem, 1981b; 1984; 1987a; 1987b), der Vorurteilsforschung sowie der Selbstkategorisierungstheorie (Turner, 1981; Turner, 1982; Turner et al., 1987). Unabhängig davon, dass Bem und Spence (1991) unterschiedliche Annahmen bezüglich der Faktorenstruktur und der theoretischen Untermauerung der Geschlechtsrollenitems getroffen hatten, erkennen beide die Rolle der kognitiven Voraussetzungen an. Wie bereits angeführt, geht Bem von einer kognitiven Struktur bzw. einem Netzwerk von Assoziationen aus, welches die Wahrnehmung eines Individuums organisiert. Es existieren Netzwerke und Vorstellungen bezüglich dessen, was gesellschaftlich determiniert, typisch maskulin oder feminin ist. An keiner Stelle erfolgt dabei eine Einschränkung auf nur positive Eigenschaften, so dass davon auszugehen ist, dass diese Netzwerke aus sowohl positiven als auch negativen Eigenschaften bestehen. Studien aus der Stereotypenforschung vermuten, dass beispielsweise Mitglieder von stigmatisierten Gruppen sehr wohl über die negativen Beschreibungen ihrer eigenen Gruppe Bescheid wissen, sich jedoch weigern, diese zu internalisieren, da sie unter anderem eine Bedrohung für den

---

<sup>15</sup> UC ist definiert durch den Fokus auf andere unter Ausschluss des eigenen Selbst, während UA das Selbst in den Vordergrund stellt unter gleichzeitigem Ausschluss der anderen (beide Aspekte werden in einem späteren Teil der Arbeit ausführlicher vorgestellt).



Selbstwert und das Verhalten darstellen (Crocker & Major, 1989; Sinclair, Hardin, & Lowery, 2006; Steele, 1997). Folgt man darauf aufbauend den Ausführungen der Stereotypenforschung, soll die Art der Zuschreibung der Eigenschaften die Ähnlichkeit zur eigenen Gruppe erhöhen und gleichzeitig die Differenzierung zur anderen Gruppe (Turner et al., 1987), trotz möglicher Überlappungen und eingeschränkter Zuschreibungen von Eigenschaften der anderen Gruppe, maximieren. Um die Differenzierung zur Fremdgruppe möglichst zu erhöhen, sollten sich Gruppenmitglieder im stärkeren Maße die stereotypen Eigenschaften der eigenen Gruppe zuschreiben als die der anderen (Pickett et al., 2002). Auch in diesem Zusammenhang erfolgt keine Beschränkung auf ausschließlich positive Eigenschaften.

Smith und Mackie (2000) gingen in ihrer Definition des Stereotyps sogar einen Schritt weiter und erkennen an, dass Stereotype sowohl positive (die ihrerseits wiederum durchaus negative Konsequenzen haben können) als auch negative Charakteristiken beinhalten können. Gerade im direkten Vergleich mit anderen Gruppen und umso mehr mit Minoritäten zeigte sich, dass die dominante Gruppe als besser angepasst und leistungsorientierter gilt sowie insgesamt mehr vorteilhafte Eigenschaften auf sich vereint, während die Minorität beispielsweise als wesentlich aggressiver wahrgenommen wird (Zimet & Zimet, 1978). Oswald und Lindstedt (2006) hatten sich ebenfalls des Inhalts und der Funktion von Geschlechterstereotypen angenommen und studentische Probanden hinsichtlich ihres Wissens und Geschlechterstereotype getestet. Insgesamt schätzten die Studierenden in der Studie von Oswald und Lindstedt (2006) positive Stereotype als charakteristischer für die Selbst- als die Gruppenbeschreibung ein, während die negativen Stereotype – in Einklang mit der *Sozialen Identitätstheorie* (Tajfel & Turner, 1986), der *Geschlechtsidentitätstheorie* (Spence, 1993) sowie der *Selektiven Selbststereotypisierung* (Biernat et al., 1996) – geeigneter für die Gruppen- als für die Selbstbeschreibung eingeschätzt wurden. Die Tendenz der individuellen selektiven Selbstbeschreibung stand dabei im Zusammenhang mit persönlichem und kollektivem Selbstwertgefühl. Hogg und Vaughan (2008) argumentierten in ihrer Darstellung von Vorurteilen und deren kognitive, affektive und konnotative Komponenten, dass gerade die Wertkomponente durchaus negativ sein kann. Ebenfalls im Zusammenhang mit Stigmatisierung, im Sinne negativ sozialer Gruppenattribute, wird die Daseinsberechtigung negativer Eigenschaften deutlich. Stigmatisierte Gruppen wissen um ihren negativen Status und erleben dadurch den *Stereotype Threat* (Steele, 1997; Steele & Aronson, 1995; 2004), die Sorge diesen negativen Stereotypen zu entsprechen und sie zu bestätigen, was langfristige Auswirkungen auf Angsterleben und Leistung haben kann.

Aus der Selbst- und Gruppenforschung kann demnach zunächst abgeleitet werden, dass auch negative geschlechtsbezogene Eigenschaften bei der Zuschreibung von Eigenschaften bekannt sind und berücksichtigt werden.

#### 2.3.2.1.1 Die Perspektive der Stereotypenforschung: *Stereotype Threat*

Eng verknüpft mit der Stereotypisierung von Gruppen ist die Wahrnehmung einer Bedrohung, verursacht durch eben diese stereotype Zuschreibung von Eigenschaften – *Stereotype Threat*

(vgl. Steele & Aronson, 1995; 2004). In Abhängigkeit der Salienz von Kategorien, denen man angehören kann, wissen Personen nicht nur um eine aufgemachte Gruppierung, sondern sind sich in den meisten Fällen durchaus bewusst, dass sie stereotypisiert werden. Die daraus resultierende Evaluation hinsichtlich dieser Stereotype kann dazu führen, dass – gerade im Zusammenhang mit negativen Stereotypen – die Befürchtung entsteht, diese eigentlich überzeichneten Verallgemeinerungen schlussendlich doch zu erfüllen und damit zu bestätigen. Diese wahrgenommene Bedrohung durch Stereotype kann wiederum Einfluss auf Verhaltens- und Leistungsmaße haben und zur Minderung von Leistungen führen. Steele und Aronson (1995) konnten diesen Stereotype Threat in einer Gruppe Studierender nachweisen. So erzielten afro-amerikanische Probanden nur in der Bedingung, die ihre ethnische Gruppe salient machte, schlechtere Ergebnisse bei einem kognitiven Leistungstest als Versuchspersonen in anderen Bedingungen (Weiße u. Nicht-Salienzmachung der Kategorie) und sie waren zusätzlich davon überzeugt, dass diese schwächeren Leistungen das ethnische Stereotyp, Afro-Amerikaner seien insgesamt weniger intelligent als Weiße, bestätigen.

Ein nicht unwesentlicher Teil der Studien zum Stereotype Threat befasst sich mit Unterschieden zwischen Männern und Frauen. Cadinu, Maass, Rosabianca und Kiesner (2005) testeten in ihrer Studie zum Einfluss von negativen Gedanken auf Leistungsdefizite 60 weibliche Probanden in zwei Bedingungen, entweder mit oder ohne stereotypische Bedrohung. Neben der Bewältigung eines mathematischen Tests sollten alle Frauen zusätzlich ihre Gedanken hinsichtlich des Tests an sich und der Mathematik im Allgemeinen angeben. Die Ergebnisse zeigten, dass die Frauen unter der Stereotype Threat Bedingung nicht nur schlechter in dem Test abschnitten, sondern darüber hinaus auch noch mehr negative Gedanken äußerten. Obwohl in dieser Studie ausschließlich Frauen einer Überprüfung unterzogen wurden, zeigt deren unterschiedliches Abschneiden in einer Leistungsaufgabe unter einer stereotypen Bedrohungssituation bereits auf, dass auch die saliente Kategorisierung durch das Geschlecht die jeweiligen Gruppen für diesen Effekt anfällig machen könnte. Nguyen und Ryan (2008) hatten in ihrer Meta-Analyse nachgewiesen, dass die Leistungseinbußen bei schwierigen Tests im stärkeren Maße für Minderheiten als für Frauen nachgewiesen werden konnten. Es zeigte sich auch, dass die größten leistungsmindernden Effekte für Frauen unter den Bedingungen zustande kamen, in denen die Bedrohungsaktivierung eher subtil erfolgte, oder auch, dass vor allem Frauen, die sich im mittleren Grade mit Mathematik identifizierten, stärker von der Stereotypenbedrohung beeinflusst waren als Frauen, die sich im geringen oder sehr hohen Maße mit Mathematik identifizierten.

Die Stereotypenbedrohung für das einzelne Individuum ist abhängig von Faktoren wie die Stärke der eigenen Identifikation mit der Gruppe oder mit der betrachteten Leistungsdomäne. Beim Geschlecht als Kategorie können Individuen dahingehend unterschieden werden, in welchem Ausmaß sie sich selbst der Gruppe des Geschlechts zugehörig fühlen und auch wie stark die Identifizierung mit der Eigengruppe über verschiedene Kontexte hinweg ist (Marx, Stapel, & Muller, 2005; Schmader, 2002).

Der Stereotype-Threat hat nicht nur Einfluss auf die soziale Identität, sondern darüber hinaus auch auf die Identifizierung mit Aktivitäten, die mit der eigenen sozialen, bedrohten Gruppe assoziiert sind und somit unter bestimmten Umständen sogar zur Distanzierung von der stereotypisierten Gruppe (z. B. um den Fokus des Stereotyps zu verlassen) oder zu einer Art Teildistanzierung von der Eigengruppe führen kann. Eine Studie diesbezüglich haben Pronin, Steele, und Ross (2004) vorgelegt, die demonstrierten, dass Frauen ihre Selbstbeschreibung unter Bedingungen des Stereotype-Threat modifizieren; ein Ergebnis, welches vor allem in der Diskussion über die Messung von Geschlechtsrollenselbstkonzepten Berücksichtigung finden sollte. Die Autoren konnten in ihren Studien zeigen, dass Frauen unter Umständen des Stereotype-Threat feminine Eigenschaften, die mit diesem Stereotyp assoziiert sind, im Zuge der Selbstbeschreibung stärker ablehnten als feminine Eigenschaften, die nur eine geringe Assoziation mit dem Stereotyp aufwiesen. Dies konnte umso stärker den Frauen nachgewiesen werden, die sich mit der Domäne Mathematik im stärkeren Maße identifizierten.

#### *2.3.2.1.2 Die Perspektive der Intergruppenforschung: Eigengruppenbevorzugung und Selbstkategorisierung*

Zieht man in Betracht, dass viele GSK Instrumente auch zur Einschätzung anderer Gruppen verwendet werden, ergibt sich nahezu fordernd die Notwendigkeit, auch negative Eigenschaften einzubeziehen. Soziale Vergleiche stellen eine wichtige Informations- und Orientierungsfunktion dar und es lassen sich eine Vielzahl von kognitiven (Gebrauch von Stereotypen) oder zugrundeliegenden motivationalen Erklärungen unterscheiden: das Bedürfnis nach einer exakten Selbsteinschätzung (*self-evaluation*), nach Kompetenzsteigerung (*self-improvement*) und das Bedürfnis nach einer Stabilisierung oder gar Erhöhung des Selbstwertgefühls (*self-enhancement*); das Herunterspielen möglicher persönlicher Diskriminierung; (vgl. Petersen, Stahlberg, & Dauenheimer, 2000; oder Wood, 1989, für motivationale Grundlagen); sowie der *Better Than Average Effect* (Alicke, Klotz, Breitenbecher, Yurak, & Vredenburg, 1995); oder das *Selective Self-Stereotyping* (Biernat et al., 1996).

Im Rahmen des *Selective Self-Stereotyping* erkannten auch Biernat et al. (1996) durchaus an, dass negative Eigenschaften einen Beschreibungswert für die übergeordnete soziale Gruppe haben, dass diese aber zur Beschreibung des Selbst nicht selektiv und freiwillig herangezogen werden. Die Eigenzuschreibung von negativen Eigenschaften bleibt problematisch und paradox, weil dadurch eine Reduktion / ein Verlust des Selbstwerts riskiert wird. Gleichzeitig sind jedoch die negativen Eigenschaften der eigenen als auch der anderen Gruppe sehr wohl bekannt und eine Ablehnung eben dieser negativen Ingroup-Eigenschaften repräsentiert im weitesten Sinne eine Verleugnung der eigenen Identität (Lorenzi-Cioldi, 1991) oder der sozialen Realität (Oakes, Haslam, & Turner, 1994).

Demnach scheint plausibel, dass sich Männer und Frauen diese Eigenschaften – aufgrund der negativen Valenz – durchaus weniger stark zuschreiben (um unter anderem den Selbstwert zu schützen) als die positiven Eigenschaften, jedoch dies immer noch in der geschlechtsentsprechenden Form vornehmen, und dass damit ein Informationszugewinn durch negative Eigenschaften zu erreichen ist. Pickett et al. (2002) konnten dieser Argumentation

folgend dann auch eine signifikante Rolle der Identifizierung mit positiven und negativen Eigenschaften aufweisen: in ihren Studien zum Einfluss des Bedürfnisses nach Annäherung und Differenzierung in und zwischen Gruppen konnten sie zeigen, dass nur Probanden, die sich stark mit der eigenen Gruppe identifizierten, gewillt waren, sich auch negative Eigenschaften zuzuschreiben.

Auch Biernat et al. (1996) stützten indirekt dieses Bild. Sie ließen Studierende aus Verbindungen und Nicht-Verbindungsmitglieder in der Selbstbeschreibung Eigenschaften einschätzen, die sowohl positiv als auch negativ waren. Erwartungsgemäß wurden negative Eigenschaften weniger zur Selbstbeschreibung herangezogen. So schrieben sich die Verbindungsmitglieder die positiven Eigenschaften stärker zu als die negativen typischen Eigenschaften der Eigengruppe, beide im stärkeren Maße als die jeweils positiven und negativen Eigenschaften, die typischer für die Nichtverbindungsmitglieder waren. Übertragen auf Geschlechtsrollensereotype entspricht dies zum Beispiel dem hypothetischen Konstrukt der maskulinen Skala. Männer sollten die höchsten Werte auf den positiven Eigenschaften in der Selbstbeschreibung erreichen, vor ihrer negativen Selbsteinschätzung, gefolgt von Frauen in Bezug auf die positiven maskulinen Eigenschaften und auf den negativen Eigenschaften.

Morris, Kowk und Iyengar (2004) hatten die Zuschreibung negativer Eigenschaften auf Fremdgruppen in Hinblick auf daraus resultierendes Konfliktverhalten (formell vs. informell) unter situativen und kulturellen Gesichtspunkten untersucht. Konfliktsituationen sind häufig durch negative attribuierte Eigenschaften bezüglich der Gegenpartei bestimmt. In der kulturvergleichenden Studie mit Versuchspersonen chinesischer und US-amerikanischer Abstammung konnten die Autoren die Bedeutung negativer Eigenschaftszuschreibungen für nachfolgende Verhaltensweisen zeigen und darüber hinaus demonstrieren, dass die Art der negativen Zuschreibung kulturabhängig komplex ist. So nahmen Chinesen in dieser Studie insgesamt weniger Eigenschaftszuschreibungen vor, diese dann aber in einer negativeren Ausprägung.

In ähnlicher Art und Weise spielen negative Eigenschaften sowie insbesondere deren Akzeptanz als Teil der eigenen Identität eine wichtige Rolle in der allgemeinen Selbstbewertung. So konnten unterschiedliche Studien einen Zusammenhang zwischen der Akzeptanz negativer Eigenschaften einer stigmatisierten Gruppe und einem verringerten Selbstwert herstellen. Chassin und Stager (1984) wiesen diesen reduzierten Selbstwert bei delinquenten Jugendlichen nach, die der negativen Bewertung ihrer eigenen Gruppe zustimmen, im Vergleich zu solchen Jugendlichen, die diese negative Bewertung als wenig bedeutsam und richtig ansehen. Das könnte eventuell erklären, warum als für eine Gruppe typisch klassifizierte Eigenschaften nicht zu unterschiedlichen Einschätzungen bezüglich der sozialen Erwünschtheit führen – Männer und Frauen sind sich der Unterschiede womöglich sehr wohl bewusst, wissen im gleichen Moment aber auch um die Relevanz und gesellschaftliche Bedeutsamkeit und Bewertung dieser Eigenschaften für ihr Selbstkonzept.

### 2.3.2.1.3 Valenz und Salienz sozialer Kategorien

Ein weiterer bedeutsamer Aspekt im Rahmen der Gruppenkategorisierung und Selbstzuschreibung von Eigenschaften betrifft die Salienz der existierenden Gruppen, die auf die Selbstzuschreibung mit Attributen unterschiedlicher Valenz wirken kann.

Bezüglich des Einflusses von Salienz der Eigengruppe Geschlecht fanden Hogg & Turner (1987), dass das Selbststereotypisieren akzentuiert wird und zu stärkeren Geschlechtsunterschieden in der Einschätzung führt. Generell konnte aufgezeigt werden, dass die Selbststereotypisierung von Eigenschaften besonders robust bei salienter Kategorienmitgliedschaft ist (Ashmore, 1981). Wird Frauen vor der Beantwortung eines Geschlechtsrollenfragebogens die eigene Geschlechtsgruppe salient gemacht, würde die Theorie der sozialen Identität vorhersagen, dass die Antworten im stärkeren Maße prototypisch weiblich ausfallen als ohne den vorher gesetzten Fokus auf geschlechtliche Eigen- und Fremdgruppe (diskutiert von Hardie & McMurray, 1992).

Hinsichtlich der Intergruppendifferenzierung und der dabei auffälligen positiv-negativ Asymmetrie, die so weder von der Theorie der sozialen Identität noch der Selbstkategorisierung vorhergesagt wurde, testeten Mummendey, Otten, Berger und Kessler (2000) in zwei Studien den Einfluss der Valenz (positiv / negativ) auf die Salienz (niedrig / hoch) der Gruppenkategorisierung. Die Autoren haben Gruppendiskussionen zur Geldverteilung durchgeführt und die Salienz der Gruppenkategorien in unterschiedlichem Maße manipuliert. Besonders in der zweiten Studie wurde dabei im stärkeren Maße darauf geachtet, dass es durch Besetzung der Gruppen mit Studierenden der jeweiligen Fakultäten einen stärkeren persönlichen Bezug gab. Im Ergebnis bestätigte sich die positiv-negativ Asymmetrie nur unter Bedingungen geringer Salienz, während sich mit zunehmender Salienz dieser Effekt reduzierte und zur Favorisierung der Eigengruppe unter beiden Valenzbedingungen führte. Eigengruppeneinschätzungen bezüglich negativer Dimensionen blieben von der Salienzmanipulation unbeeinflusst. Einschätzungen der Fremdgruppe fielen dagegen negativer in Bedingungen hoher Salienz aus. Mummendey und Kollegen argumentierten in ihrer Zusammenfassung den manipulativen Einfluss der Valenzvariation auf die Kategoriensalienz dahingehend, dass negative Bewertungsdimensionen die Salienz der Intergruppenunterscheidung verringern.

Darüber hinaus verdeutlichen die Studien vor allem den allgemeinen Zusammenhang von Valenz und Salienz im Rahmen des Gruppenkontextes und die Besonderheiten im Gruppenvergleich bezüglich negativer Dimensionen. Andere Studien haben ebenfalls die Valenz von Ereignissen thematisiert und in Zusammenhang mit Regulationsmechanismen gebracht (Sassenberg, Kessler, & Mummendey, 2003). Dabei zeigte sich, dass soziale Diskriminierung bei salienter sozialer Identität höher war und ein Regulationsfit (*promotion focus*: Aufsuchen positiver Ereignisse, vs. *prevention focus*: Vermeiden negativer Ereignisse) vorlag.

Die einzuschätzenden Eigenschaftswörter beinhalten an sich bereits einen Valenzanteil, durch den sie in positive und negative Eigenschaften klassifiziert werden können. Das Konzept

der Valenz und seine Einbettung in die psychologische Forschung soll im folgenden Abschnitt näher betrachtet werden.

#### 2.3.2.1.4 Valenz als Charakteristik von Attributen

Der Valenz-Aspekt wurde vor allem durch die umfangreiche Review von Baumeister, Bratslavsky, Finkenauer, und Vohs (2001) ausführlich diskutiert. Analysen zeigten, dass insgesamt mehr negative Phänomene wissenschaftlich untersucht werden. Baumeister und Kollegen präsentierten eine Vielzahl von Ergebnissen, die einen größeren Einfluss von negativen (wie beispielsweise Verlust von Werten, Personen, Anerkennung) im Vergleich zu positiven Ereignissen in alltäglichen Situationen, Attributionen, Beziehungen, Stereotypen, neuronalen Prozessen und Entwicklungen zeigen. Die Autoren diskutierten diese Tendenz vor dem Hintergrund der Salienz dieser Ereignisse sowie motivationaler und kognitiver Gegebenheiten: negative Informationen und Eindrücke bekommen mehr Aufmerksamkeit, werden stärker verarbeitet, schneller evoziert, sind persistenter, resistenter und werden ebenso leichter und schneller erinnert.

Diesen Einfluss negativer Informationen konnten Baumeister und Kollegen ebenfalls sehr gut im Bereich des Selbst und der Eigenschaftszuschreibung finden und zitierten Studien, in denen beispielsweise Personen bei Stroop-Aufgaben größere Schwierigkeiten hatten, die Farbe bei negativen Wörtern zu benennen, oder Studien, die den erheblich größeren Einfluss von negativen Eigenschaften bei der Eindrucksbildung von Personen demonstrieren konnten. Insgesamt gelten nur wenige Adjektive als *wert-neutral*, da die meisten als *positiv* oder *negativ* bewertet werden (Mecklenbräuker et al., 1994).

Wiederholt wurde eine mögliche Ergebniskonfundierung von Inhalt und Valenz diskutiert und kritisiert (Pahl & Eiser, 2005; Perkins & Forehand, 2006). Wenn beispielsweise Versuchspersonen Attribute wie *faul* und *nett* vor dem Hintergrund der Valenzeffekte beurteilen sollten, waren die Ergebnisse uneindeutig, wenn nicht auch die Unterschiedlichkeit des Inhalts kontrolliert wird. Die Unterscheidung von denotativer (Semantik) und konnotativer (Valenz) Bedeutung hat gerade im Bereich sozialer Ratingskalen von Persönlichkeit, Einstellung oder Gruppen große Auswirkungen.

Heyman und Giles (2004) hatten in ihrer Studie über evaluative Eigenschaften (*nice*, *mean*, *smart*, oder *not smart*) die Bedeutung der Valenz als Informationsgehalt nachgewiesen. In drei unterschiedlich alten Stichproben konnten sie zeigen, dass bereits Kinder im Alter von 7 Jahren Vorstellungen über Valenzaspekte besitzen, die die eigenen Überlegungen beeinflussen. Bei der Einschätzung sozialer und akademischer Fähigkeiten von Zielpersonen waren es vor allem deren positive Eigenschaften, die als stabiler galten und eher zum Rückschluss auf Kompetenz herangezogen wurden als negative Eigenschaften. Darüber hinaus bestand die Annahme, dass positive Einflüsse von biologischen und Sozialisationsfaktoren mögliche negative Einflüsse ausbalancieren können und generell Verhaltensweisen wahrscheinlicher in eine positive Richtung verändert werden können als in eine negative. Obwohl diese allgemeinen Überzeugungen mit dem Alter geringfügig abnahmen, demonstrierte

die Arbeit von Heyman und Giles (2004) eine optimistische Tendenz von Versuchspersonen in Bezug auf Verhaltenseigenschaften und den daraus gezogenen Rückschlüssen.

Auch aus anderen Forschungsbereichen ist bekannt, welche Bedeutung der Valenz und vor allem der negativen Wertigkeit von Attributen und Ereignissen zukommt. So zeigte sich, dass, obwohl Personen glauben, ihre eigene Zukunft sei besser als die anderer, negative Ereignisse im stärkeren Maße das eigene positive Selbst beeinflussen als dies positive Ereignisse vermögen (unrealistischer Optimismus; z. B. Hoorens, 1995). Darüber hinaus ist ebenfalls akzeptiert, dass negative Ereignisse in ihrem Einfluss stärker und wirkungsvoller sind als positive Ereignisse, berücksichtigt man beispielsweise den Einfluss schlechter Eltern, das Bestreben, ein negatives Selbstbild zu vermeiden, oder die Resistenz von negativen Stereotypen gegenüber Veränderungen (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001).

Fiske (1980) konnte in ihrer Studie zur Aufmerksamkeit und Gewichtung von Verhaltenswahrnehmung zeigen, dass die Versuchspersonen bevorzugt solches Verhalten aufmerksamer betrachten und gewichten, welches extrem oder negativ ist. Fiske diskutiert die Wirkungsweise der negativen Ereignisse vor dem Hintergrund ihres auffälligeren Charakters im Vergleich zu positiven Verhaltensweisen, die die Basis darstellen, ebenfalls wie deren geringe Vorkommensrate. Es scheint, dass Versuchspersonen den Hinweisreizen mehr Bedeutung beimessen, die auffälliger und seltener sind, da deren Informationsgehalt größer ist.

So ist es beispielsweise auf lange Sicht hin wenig ratsam, trotz des Bestrebens um ein positives Selbstbild, negative Charakteristiken gänzlich zu ignorieren, da die Wahrnehmung und Akzeptanz eigener Grenzen und Unzulänglichkeiten eine Person davor bewahrt, womöglich gefährliche oder wenig lohnende Situationen aufzusuchen (Alicke, 1985).

#### 2.3.2.1.5 Valenz vs. Typizität

Eine Unterscheidung zwischen Typizität und Valenz kann im weiteren Sinne auf die Stereotypenforschung und im Speziellen auf die Differenzierung von *stereotypical knowledge* vs. *personal beliefs* zurückgeführt werden (Ashmore & Del Boca, 1981; Augoustinos & Rosewarne, 2001; Devine, 1989; Devine & Elliot, 2000; Karlins, Coffman, & Walters, 1969). Demnach sind Stereotype im Allgemeinen vielfältig strukturiert und beinhalten physikalische und soziale Beschreibungen, genauso wie persönliche Präferenzen. Während generell argumentiert wird, dass Vorurteile ein bedingtes Ergebnis des Kategorisierungsprozesses sind (Allport, 1954; Tajfel, 1979; Tajfel, 2001; Tajfel & Forgas, 2000), kann jedoch zwischen den tatsächlichen Stereotypen, dem was man als allgemeine Überzeugungen aller annimmt, und dem was man persönlich als wahr in Bezug auf eine bestimmte Gruppe (*personal beliefs*) annimmt unterschieden werden. Der tatsächlichen Kenntnis typischer Eigenschaften (*Stereotype*) stehen die eigene Überzeugung und damit die Einschätzung der persönlichen Relevanz entgegen, die nicht kongruent mit eben jenen, eher überdauernden Stereotypen sein muss. Indem man diese Trennung von Stereotypen und persönlichen Überzeugungen vornahm, wurde auch der Tatsache Rechnung getragen, dass die Existenz von Stereotypen nicht notwendigerweise auch das Vorhandensein von Vorurteilen erklärt. Diese unterschiedlichen Aspekte von

Stereotypwissen und persönlichen Überzeugungen im Bereich der Vorurteilsforschung ließ sich sowohl bei Erwachsenen (Augoustinos, Ahrens, & Innes, 1994; Devine, 1989) als auch bereits bei älteren Kindern sehr gut aufzeigen (Augoustinos & Rosewarne, 2001).

Devine (1989) lieferte in ihrer Arbeit zu automatischen und kontrollierten Prozessen weitere Spezifikationen beider Aspekte. So teilen diese Prozesse zwar gewisse Eigenschaften, können grundlegend jedoch als distinkte kognitive Konzepte gelten, die dadurch auch unterschiedliche Implikationen für Evaluations- und Verhaltensprozesse bezüglich bestimmter Gruppen haben können. Bedeutsam in diesem Zusammenhang ist Devines (1989) Argumentation der zeitlichen Entwicklungsabfolge: da persönliche Überzeugungen als Entscheidungen über die (situative) Angemessenheit stereotypischer Zuschreibungen verstanden werden können, stellen demzufolge Stereotype die älteren, verfügbaren kognitiven Strukturen dar, die stärker verknüpft sind und somit automatischer aktiviert und abgerufen werden können. Dieser Ansatz der Vorurteilsforschung würde auch den Unterschied zwischen Personen mit starken und schwachen Vorurteilen erklären. Während beide Ausprägungen grundsätzlich ähnliches Wissen über die einzuschätzende Zielgruppe (stereotype knowledge) besitzen, sind es die persönlichen Überzeugungen, die schließlich zur Abgrenzung führen. Demzufolge akzeptieren Personen mit starken Vorurteilen die kulturellen Stereotype nahezu vollständig als ihre persönlichen Überzeugungen, während Personen mit schwachen Vorurteilen die Stereotype in der entsprechenden Situation als unangemessen einschätzen und deren automatische Aktivierung kontrollieren. In dem sie in drei Studien die Bedingungen manipulierte, unter denen automatische Prozesse mehr oder weniger gut kontrollierbar sind, konnte Devine (1989) diese kognitiven Prozesse nachweisen.

Diese Differenzierung kann nun in ähnlicher Form auch auf das Geschlechtsrollenselbstkonzept übertragen werden – durch die Unterscheidung zwischen der Kenntnis der typischen geschlechtsspezifischen Eigenschaften für eine Gruppe (auch der eigenen) und der eingeschätzten Wichtigkeit eben dieser Eigenschaften für das eigene GSK. Hort, Fagot und Leinbach (1990) hatten dies indirekt nachgewiesen, indem sie ihre Probanden baten, Frauen und Männer aus der eigenen Gesellschaft auf dem PAQ zum einen dahingehend einzuschätzen, wie die Gesellschaft sie sieht (*Stereotype*), und zum anderen, wie sie sie persönlich sehen (*personal belief*). Auch Hort und Rothbart (1989, zitiert in Hort, Fagot, & Leinbach, 1990) fanden unterschiedliche kognitive Konstruktionen in Hinblick auf bestimmte Zielgruppen dahingehend, ob sie soziale Einschätzungen und persönliche Wahrnehmungen vornehmen sollten. So fielen beispielsweise die sozialen Beschreibungen für Männer und Frauen stereotypischer (und mit mehr Konsens seitens der Probanden) aus als die persönlichen Wahrnehmungen, die ihrerseits stereotypischer waren als tatsächliche individuelle Präferenzen.

In ähnlicher Art und Weise ordnet sich hier die Diskussion um Typizität und Valenz von Geschlechtsrollenstereotypen ein. Aus der Geschlechterforschung ist bekannt (Deaux & Lewis, 1984), dass Geschlechterstereotypen multidimensional sind. Die Typizität bezogen auf Eigenschaftswörter beispielsweise würde dem klassischen stereotypen Wissen entsprechen: sowohl Frauen und Männer besitzen eine klare Vorstellung davon, was typisch Mann und



typisch Frau ist. Im Bereich der Entwicklung von Messinstrumenten (Bem, 1974; Spence et al., 1974) macht man sich beispielsweise diesen Umstand bei der Auswahl der Items zunutze. Auch von Kindern weiß man, dass sie geteiltes, stabiles Wissen über typische Eigenschaften von Männern und Frauen besitzen (Williams & Best, 1990). Diese Typizität stellt jedoch nur einen Aspekt der Geschlechterstereotype dar, denn erst die individuellen, persönlichen Überzeugungen sollten in den Einschätzungen des eigenen Geschlechtsrollenselbstkonzepts den Unterschied machen.

Um den inhaltlichen Bezug zu dem *personal belief* aus der allgemeinen Stereotypenforschung herzustellen, würde sich im Rahmen der Geschlechtsrollenforschung die Frage nach der Wichtigkeit geschlechtsbezogener Eigenschaften für das eigene Selbstkonzept anbieten. Der bedeutsame Aspekt in diesem Zusammenhang ist die Tatsache, dass mit der Implementierung neuer, auch negativer Aspekte dieser Wichtigkeit insofern eine besondere Rolle zuteil wird, da die Wichtigkeit nicht nur per se von der eigenen Einschätzung, sondern zudem auch von den Inhalten der Stereotypen abhängt und deren Valenz, wie positiv und negativ diese im Allgemeinen eingeschätzt werden. Im Abschnitt 2.4 *GSK als Erwartungs-mal-Wert-Konstrukt* auf S.98 wird dieser Aspekt der Wichtigkeit noch einmal aufgegriffen.

#### 2.3.2.1.6 Das Konzept der sozialen Erwünschtheit

Bei der Erfassung von Einstellungen und persönlicher Einschätzung hinsichtlich unterschiedlicher Persönlichkeitseigenschaften oder Verhaltensweisen spielt, wie im vorherigen Abschnitt gezeigt, die Valenz eine große Rolle. Die Wahrnehmung der Wertigkeit von beschreibenden Adjektiven oder Aussagen ist dabei unweigerlich verknüpft mit der sozialen Erwünschtheit, dem, was die Gesellschaft als Norm und verhaltensweisende Richtlinie akzeptiert. Daher soll im nachstehenden Abschnitt eine kurze Ausführung zu Funktionen und Wirkungsweise der sozialen Erwünschtheit folgen.

In der heutigen Gesellschaft ist es für Männer sozial erwünscht maskulin zu sein, ebenso wie Femininität als gesellschaftlich bevorzugt für Frauen gilt – unabhängig davon, ob dem im Einzelfall nach objektiv gemessenen Standards wirklich so ist. Männer, die sich dagegen als wenig maskulin beschreiben, riskieren den Verlust des Selbstwerts, was im Umkehrschluss die Gefahr sozial erwünschten Antwortens erhöht. Diese komplexe Beziehung zwischen Selbstbeschreibung, Selbstwertbezug, gesellschaftlichen Standards sowie Möglichkeiten der bewussten Täuschung wird bisweilen durch bestehende Geschlechtsrolleninstrumente zusätzlich verkompliziert, da große Unterschiede in der sozialen Erwünschtheit zwischen den Attributen bestehen, mit dem Vorteil auf Seiten maskuliner Eigenschaften, die als sozial erwünschter gelten als die femininen (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972). Im Bereich der Forschung, die sich mit Valenz und Erwünschtheit beschäftigt, gelten wiederum jedoch die kommunalen, femininen, Eigenschaften als anfälliger für sozial erwünschtes Antworten (Schwartz, Verkasalo, Antonovsky, & Sagiv, 1997). Individuen, die beispielsweise eine starke Tendenz haben, sozial erwünscht zu antworten, äußern in ihren generellen Werten den Wunsch, soziale Missbilligung zu vermeiden,

was sich in hohen Werten auf Skalen wie Konformität, Sicherheit, Tradition und Wohlwollen widerspiegelt.

Hinsichtlich der Geschlechtsrollenattribute kann das geteilte Wissen um gesellschaftliche Standards / Erwartungen für Männer und Frauen über Kohorten und Kulturen hinweg Verzerrungen möglich machen. Dies zu verhindern ist im Falle von Adjektiv-basierten Selbstbeschreibungen auch nahezu unmöglich; dem zu entsprechen durch Gewährleistung annähernd vergleichbarer sozialer Erwünschtheit der typisch maskulinen und femininen Attributen, ist jedoch die Mindestvoraussetzung an ein Instrument.

In der psychologischen oder auch soziologischen Forschung werden Persönlichkeitseigenschaften und Verhaltensweisen mittels Frequenz- oder Ratingskalen in Selbstbeschreibungen erfasst, die nicht selten bis zu einem gewissen Grad durchschaubar und für Verfälschungstendenzen anfällig sind. Daher beschäftigt sich die Forschung intensiv mit möglichen gezielten Antwortverzerrungen solcher Selbstbeschreibungen (*faking good*) aufgrund des Bedürfnisses der Probanden sich in einem positiven Licht darzustellen (Amelang & Borkenau, 1981; Converse, Peterson, & Griffith, 2009; Edwards, 1953; Edwards, 1957; Hager et al., 1996; Krahe & Herrmann, 2003; Lobel & Teiber, 1994; Lönnqvist, Verkasalo, & Bezmenova, 2007; Millham & Jacobson, 1978; Mummendey, 1981; Paulhus, 1986; Robson, Jones, & Abraham, 2008; Schahn, 2002; Schmitt & Steyer, 1993; Stöber, 1999; Stocké, 2004; Strickland, 1977; Twenge & Im, 2007). Vorherrschende kulturelle Normen und die damit einhergehende Bewertung von Eigenschaften und Verhaltensweisen machen dabei einige von ihnen sozial erwünschter als andere (Anderson, 1968; Lück & Timaeus, 1969), was nicht nur zu einem Bias in der Selbstbeschreibung führen kann, sondern darüber hinaus die Validität der gemessenen Konstrukte (Persönlichkeitskonstrukte, Einstellungen oder Werte) erheblich einschränkt.

Es gilt mittlerweile anzuerkennen, dass sich Personen in ihrem Bedürfnis nach Anerkennung und somit einer möglichen sozial erwünschten Verschiebung ihrer Selbsteinschätzungen unterscheiden – in Abhängigkeit der verschiedenen Erhebungssituationen wie auch in der Interaktion von Person-Situation. Dieser Ansatz könnte auch erklären, warum Unterschiede in den Ergebnissen zu zwei Messzeitpunkten nicht gänzlich durch Messfehler zu erklären sind, ungeachtet der Tatsache, dass die Eigenschaft als konstant angesehen wird. Dem nachgehend testeten Schmitt und Steyer (1993) die deutsche Version (Lück & Timaeus, 1969) der Marlowe-Crowne-Skala (Crowne & Marlowe, 1960) in Bezug auf soziale Erwünschtheit, indem sie systematische Effekte der Erhebungssituation als auch der Personen-Situation-Interaktion in zwei Studien mit mehreren Messzeitpunkten berücksichtigten. In ihrer Studie konnten Schmitt und Steyer zeigen, dass ein – wenn auch kleiner – Anteil der interindividuellen Antwortvarianz signifikant systematisch sowohl von der psychologischen Erhebungssituation als auch der Interaktion zwischen Person und Situation abhängt. Schmitt und Steyer (1993) argumentierten, dass die Methodenspezifität einer Skala davon determiniert ist, wie stark individuelle Items von verschiedenen Faktoren abhängen und wie viele Items in einer Skala zusammengefasst sind.

Die soziale Erwünschtheit als Auswahlkriterium für Items zur Bestimmung des Geschlechtsrollenselbstkonzepts spielt insofern eine Rolle, da sie in bereits existierenden Instrumenten – wenn auch in unterschiedlicher Ausrichtung (BSRI: „*how socially desirable*“, PAQ: „*which where ideal for each sex*“) – zugrunde gelegt wurde, während gleichzeitig in verschiedenen Studien gezeigt werden konnte, dass sie einen nicht unerheblichen Einfluss auf die Faktorenstruktur der angestrebten Dimensionen haben kann (Altstötter-Gleich, 2004; Altstötter-Gleich, Eglau, & Kramer, 2000; Pedhazur & Tetenbaum, 1979).

Mögliche Zielitems dahingehend einzuschätzen, wie positiv / negativ bzw. sozial erwünscht diese sind, kann politisch oder kulturell erwünschtes Verhalten im Rahmen der Bewertung und möglichen Bedrohung der eigenen Gruppe(nidentität) induzieren. Aus der Forschung um geschlechtsbezogene Normen und Geschlechtsrolleneinstellungen ist hinreichend bekannt, dass Personen sich sehr wohl darin unterscheiden, inwieweit sie mit existierenden Normen konform gehen (Deaux, 1985; Deaux & LaFrance, 1998; Krampen, 1979; Spence & Helmreich, 1972; Williams & Best, 1990) und damit eine eher traditionelle oder egalitäre Geschlechtsrollenorientierung aufweisen (Athenstaedt, 2000).

Mögliche erklärende Konzepte sind in diesem Zusammenhang *reverse discrimination*<sup>16</sup> (Branscombe & Smith, 1990) oder *women are wonderful effect*<sup>17</sup> (Eagly & Mladinic, 1989; Eagly, Mladinic, & Otto, 1994). Diese Konzepte können auf der Bewertungsebene einem Muster der möglichen bewussten Aufwertung der eigenen Gruppe oder Abwertung / Ablehnung der anderen Gruppe entsprechen, basierend auf der Sozialen Identitäts- und Selbstkategorisierungstheorie und in Abhängigkeit vom Grad der Identifizierung mit der Eigengruppe und dem Selbstwert (Tajfel, 1969; Tajfel & Turner, 1979; Turner, 1978a; 1978b). Dies setzt voraus, dass die Bewertungsgrundlage und deren Inhalte zu einem gewissen Grad hin durchschaubar sind.

Jackson (1985) ließ Versuchspersonen in offenen Fragen Stellung zu sich selbst nehmen („*I am*“) und untersuchte diese Angaben hinsichtlich maskuliner und femininer Bezüge und der Übereinstimmung zur eigenen Geschlechtsrolle sowie der gegebenen Definition beider Geschlechtsrollenselbstkonzepte. Die Ergebnisse zeigten, dass die Mehrheit der gegebenen offenen Antworten sozial erwünscht, aber ohne direkten Geschlechtsbezug war.

Pierce und Kollegen (Pierce, Sydie, Stratkotter, & Krull, 2003) ließen zur Überprüfung von Rollenstereotypen Versuchspersonen die Konzepte *Feminist*, *Mann* und *Frau* auf bipolaren Skalen in Form von semantischen Differentialen einschätzen. Mittels dieser Form der Erfassung, die positiv gewertete Items mit ihren negativen Gegensätzen paart (z. B. *pleasant - unpleasant*, *forgiving - unforgiving*), konnten die Autoren in einer Faktorenanalyse auch einen evaluativen Faktor identifizieren, der immerhin 28% der Varianz aufklärte und für die Geschlechterstereotypenforschung von großer Relevanz ist. Pierce et al. (2003) konnten nachweisen, dass die individuelle Wahrnehmung geschlechtsbasierter Konzepte in ihrer Studie

---

<sup>16</sup> = Favorisieren der Minderheit

<sup>17</sup> = Obwohl sich Frauen fortgesetzter Diskriminierung ausgesetzt sehen, erfahren sie mitunter eine positivere Gesamtbewertung als Männer

zum größten Anteil aufgrund evaluativer Komponenten erfolgte, und weniger aufgrund anderer Faktoren.

Es ist nachgewiesen, dass Frauen und Männer Wissen über Geschlechterstereotype und -normen besitzen, dieses jedoch in unterschiedlichem Ausmaß in Situationen zur Selbstbeschreibung berücksichtigen. In Abhängigkeit von einer möglichen Bedrohung für den eigenen Selbstwert fällt demnach das sozial erwünschte Antworten unterschiedlich stark aus. In einer von Abele (1997) durchgeführten Studie wurden beispielsweise die Geschlechterstereotype eindeutig repliziert, während es jedoch im Selbstbild keinerlei Unterschiede zwischen Männern und Frauen gab. Dazu, dass Ergebnisse generell motivierte Einschätzungen reflektieren können (controlled information processing, z. B. Petty & Wegener, 1993), kommt zusätzlich erschwerend hinzu, dass kulturelle Veränderungen tatsächliche Veränderungen in den Normen nach sich ziehen (Abele, 1997; Hofstede, 1994; Jansen, 1997; Kanter, 1997), über die Männer und Frauen für Bewertungen dieser Art Wissen besitzen müssen..

Die Berücksichtigung der Prinzipien und der beeinflussenden Umstände des sozial erwünschten Antwortverhaltens ist nicht nur für die Selbstbeschreibung evident und relevant, sondern umso mehr im Rahmen der Instrumentenerstellung. Die Vorstellung bezüglich der Typizität oder sozialen Erwünschtheit von Frauen und Männern fällt mitunter sehr klar und eindeutig stereotyp aus – sowohl für positive als auch negative Attribute. Probanden fällt es nicht schwer, ad hoc Eigenschaften zu produzieren oder dahingehend zu bewerten, inwieweit diese typischer für Männer oder eben für Frauen sind. Bemerkenswert ist jedoch das sich ergebende Bild, wenn nicht nur nach den Typizitätseinschätzungen, sondern zusätzlich gleichzeitig auch nach der Selbstbeschreibung gefragt wird, wodurch ein Vergleichsrahmen aufgemacht wird.

#### *2.3.2.1.7 Soziale Erwünschtheit (negativer) Eigenschaften*

Die Tendenz einer positiven Selbstdarstellung scheint durch zwei Strategien möglich: der Zuschreibung sozial erwünschter und dem Ablehnen unerwünschter Eigenschaften, deren Einzigartigkeit bei durchaus vorhandener Ähnlichkeit als bereits bestätigt gilt (Hoorens, 1995; Jacobson, Kellogg, Cauce, & Slavin, 1977; Roth, Snyder, & Pace, 1986). Es wurde nachgewiesen, dass Aspekte des Verleugnens (das Ablehnen sozial unerwünschter Eigenschaften für das eigene Selbst) und der Zuschreibung (das Beanspruchen sozial erwünschter Eigenschaften für das Selbst) bei einigen Instrumenten zur Erfassung der sozialen Erwünschtheit eine Rolle spielen und in unterschiedlichem Maße mit verfälschendem Verhalten in Beziehung gebracht werden (siehe für einen Überblick: Paulhus, 2002; S. 53 ff.). Dies wurde beispielsweise in Zusammenhang mit Sexualität nachgewiesen, bezüglich derer bekannt ist, dass Individuen eine Tendenz haben, bestimmte Erfahrungen zu über- oder unterschätzen sowie bestimmte Erfahrungen auszulassen, durch die eine negative soziale Bewertung erfolgen könnte.

Paulhus hat in seinen Untersuchungen das Konzept der sozialen Erwünschtheit einer umfangreichen Analyse unterzogen und zwei bis dato konkurrierende Faktorenmodelle gegeneinander getestet (Paulhus, 1984; Paulhus, 2002). Dabei unterschied Paulhus generell

die soziale Erwünschtheit hinsichtlich kurzlebig motivierten Antwortverschiebungen (*response sets*) und andauernder überhöhter positiver Selbstbeschreibung über Zeiten und Fragebögen hinweg (*response styles*). Demgegenüber steht ein Modell, in dem *Selbsttäuschung* (Personen schenken der eigenen positiven Beschreibung Glauben) und *Impressionsmanagement* (ursprünglich: *other deception*; Versuchsperson verstellen sich bewusst) kontrastiert werden (vgl. Paulhus, 1984). Paulhus konnte zeigen, dass das Modell *Selbsttäuschung / Impressionsmanagement* durch die Daten am besten repräsentiert wurde. Paulhus (2002) unterteilte den Aspekt der Selbsttäuschung schließlich noch einmal in *Selbsterhöhung*, dem Herausstellen positiver Eigenschaften („*My first impression always turn out to be right*“) sowie dem *Verleugnen negativer Eigenschaften* („*I have never thought about killing someone*“), die in unterschiedlichem Maße mit den Big Five Faktoren in Beziehung stehen. Darüber hinaus zergliederte er die soziale Erwünschtheit in egoistische (selbsttäuschende Erhöhung; *agentic management*) und moralische (selbsttäuschende Verleugnung; *communion management*) Aspekte und unterschied dabei zwischen Inhaltslevel (*agency vs. communion*) und Verarbeitungslevel (bewusst vs. unbewusst).

Es bleibt festzuhalten, dass es verschiedene Mechanismen gibt, die das Antwortverhalten beeinflussen können, insbesondere wenn es um die Selbstbeschreibung mit negativen Eigenschaften geht. In dem Fall müssen diese Antworttendenzen sowohl bei der Erstellung als auch beim Einsatz der Messinstrumente zu berücksichtigt werden.

### **2.3.2.2 Positive und negative Inhalte der Geschlechterstereotype**

Am Ende des letzten Jahrtausends wurden die Limitationen in ein- oder zwei-dimensionalen Ansätzen offensichtlich (Korabik, 1999). Kritiken haben wiederholt festgestellt, dass die ausschließliche Betrachtung zweidimensionaler Geschlechtsrollentheorien häufig eine eingeschränkte Vorhersagevalidität demonstrierte (Cook, 1985). McCreary und Korabik (1994) mahnten an, dass durch Untersuchungen von ausschließlich *Agency* und *Communion* zu Korrelaten wie Gesundheitsverhalten und Wohlbefinden eine Reduzierung auf favorisierte, sozial erwünschte Eigenschaften stattfindet, die zu einem Konfundieren der Ergebnisse mit einer dritten Variable führt: der sozialen Erwünschtheit von Persönlichkeitseigenschaften. Diese Reduzierung der Geschlechtsrolleneigenschaften auf sozial erwünschte Beschreibungen hat mögliche unerwünschte Geschlechtsrollenaspekte in den Hintergrund rücken lassen (Aubé & Koestner, 1992; Marsh & Byrne, 1991), wodurch langfristig die Generalisierbarkeit der Ergebnisse einschränkt wird.

Eine Reihe von Forschungsgruppen haben die Inhalte von Geschlechterstereotype diskutiert und festgestellt, dass diese nicht nur sozial erwünschte Aspekte von Femininität und Maskulinität beinhalten (Albert & Porter, 1983; Basow, 1992; Broverman et al., 1972; Ellis & Bentler, 1973; Kelly & Worell, 1977; Ricciardelli & Williams, 1995; Rosenkrantz, Bee, Vogel, & Broverman, 1968; Williams & Bennett, 1975; Williams & Best, 1990). Im Rahmen dieser Studien wurden unter anderem von Gruppen unterschiedlichen Alters verschiedenste Stereotype selbst produziert oder gegebene Attribute eingeschätzt. Dabei wurde nicht deutlich, dass

beispielsweise weibliche Stereotype weniger positiv bewertet werden als männliche, sondern auch, dass negative Eigenschaften ungern für die eigene Gruppe akzeptiert werden sowie dass negative Eigenschaften durchaus selbständig produziert werden können. Nirgendwo in der Literatur wird beschrieben, dass sich Individuen niemals in sozial unerwünschter Weise beschreiben. Stattdessen wurde bereits festgestellt, dass in einigen sozialen Situationen negative Aspekte sogar dominanter sein können.

Williams und Bennett (1975) baten beispielsweise Studierende die Typizität von 200 Eigenschaften für Männer und Frauen einzuschätzen. Ähnlich wie in späteren Studien zeigte sich eine zufriedenstellende Übereinstimmung zwischen Frauen und Männern in der Beurteilung, die zu 30 typischen Adjektiven für Frauen und 33 für Männer führte. Auffällig ist, dass diese Listen auch negative Eigenschaften enthielten. Letztendlich waren fünf der Eigenschaften für Frauen positiv und 10 negativ, während es für Männer umgekehrt war: 10 positive und fünf negative Items, was dahingehend interpretiert wurde, dass Einschätzungen dieser Art häufig aufzeigen, dass männlichen Charakteristiken mehr Wert zugeschrieben wird. Auch Long (1991) konnte zeigen, dass die US-amerikanische Kultur kompetenzorientierte maskuline Eigenschaften positiver einschätzt und gegenüber den eher „passiven“ weiblichen Eigenschaften favorisiert. Maskuline Eigenschaften wie *Unabhängigkeit* oder *Rationalität* wurden positiv eingeschätzt, während feminine Eigenschaften wie *Intuition* oder *Emotionalität* negativ eingeschätzt wurden.

Im Gegensatz dazu hatten Hosoda und Stone (2000) an einer Stichprobe von Studierenden Attribute hinsichtlich ihrer stereotypischen Aktualität sowie des evaluativen Inhaltes untersucht. Über das Auszählen von Häufigkeiten zeigte sich, dass das maskuline Stereotyp insgesamt 78 Attribute umfasste (32 erwünscht, 37 unerwünscht, 9 neutral), während das feminine Stereotype 62 Attribute enthielt (26 erwünscht, 21 unerwünscht, 15 neutral). Das Ergebnis, dass Männer insgesamt mehr unerwünschte Attribute zugeschrieben bekamen als Frauen, interpretierten die Autoren als negativeres, maskulines Stereotyp.

Insgesamt betrachtet, existiert mittlerweile eine Vielzahl von Überlegungen und Befunden, die sich sehr wohl mit negativen Aspekten des Geschlechtsrollenselbstkonzepts und deren Relevanz befasst haben, zum Teil aus Bakans ursprünglicher Unterscheidung von *agency* und *communion* herrührend (Fritz & Helgeson, 1998; Helgeson & Fritz, 1998; Helgeson & Fritz, 1999; Runge, Frey, Gollwitzer, Helmreich, & Spence, 1981; Spence et al., 1979) und zum Teil datenbasiert aus Untersuchungsergebnissen bezüglich des Aufklärungswerts anderer Konstrukte (z. B. Stress, Gesundheit, antisoziales Verhalten, etc.) hergeleitet (Athenstaedt, 2003; Bruch, 2002; Danoff-Burg, Revenson, Trudeau, & Paget, 2004; Fritz & Helgeson, 1998; McCreary & Korabik, 1994; Sieverding, 1999).

#### 2.3.2.2.1 Valenz bei den Geschlechterstereotypen

Obwohl die meisten bisherigen Messinstrumente auf positive Eigenschaften beschränkt sind, wird allein durch diese Beschränkung bereits die Bedeutung der Valenz bei Geschlechterstereotypen deutlich.

Hinsichtlich des stereotypen Frauenbildes gelten Eigenschaften wie *empfindsam*, *passiv* oder *warm* in den westlichen Kulturen als typisch (Eckes, 1994b; Williams & Best, 1990) und werden generell als positiv in der Valenz eingeschätzt, was sich unter anderem in einer positiven Einstellung gegenüber Frauen niederschlägt (Eagly & Mladinic, 1994). Dieses homogene feminine Stereotyp enthält zwar überwiegende sympathische Eigenschaften, die jedoch – gerade im Vergleich zum Männerstereotyp – eine deutliche Limitierung offenbaren: es sind allesamt Eigenschaften, die wenig mit Attributen von beruflichem Erfolg, Leistungsfähigkeit, Dominanz und Kompetenz zu tun haben. Studien zu Rollentheorien und Gruppenkategorien stellen zwar durchaus Veränderungen in den letzten Jahren fest, aber ebenso eine recht stabile Überschneidung des allgemeinen Frauenstereotyps mit dem Stereotyp von Hausfrauen, jedoch nicht mit dem einer Karrierefrau (Diekmann & Eagly, 2000; Locksley & Colten, 1979). In Bezug auf die Beschreibung einer Frau ist dabei das aktivierte Stereotyp von entscheidender Bedeutung. So konnten Wänke, Bless und Wortberg (2003) zeigen, dass das Beispiel einer stereotyp inkonsequenten Frau in Kombination mit Informationen, die die Inklusion dieser wenig prototypischen Vertreterin in die Gruppe der Frauen nahelegen, die stereotypische feminine Beurteilung von Frauen im Allgemeinen reduziert.

Bezüglich des Geschlechtsrollenselbstkonzepts ist die Valenz der einzuschätzenden Eigenschaften – wie in vorherigen Abschnitten dargestellt – zweifelsohne sehr stark verknüpft mit der resultierenden Valenz des Selbstkonzepts. Die Voraussetzungen durch die zur Selbstbeschreibung verfügbaren maskulinen und femininen Eigenschaften sind jedoch ungleich verteilt, betrachtet man die die Valenz der Stereotype. So gelten einige eindeutig positive Eigenschaften als eher typisch für Maskulinität, während feminine Eigenschaften zwar durchaus auch als positiv angesehen werden, in ihrem Wert für das Selbstkonzept aber weniger bedeutsam sind (Sieverding & Alfermann, 1992). Obwohl die Stereotype der Geschlechter unterschiedliche Inhalte von möglichen Persönlichkeits- und Verhaltenseigenschaften widerspiegeln, wird deutlich, dass hinsichtlich der sozialen Erwünschtheit maskuline Attribute im Vergleich zu den femininen als wertvoller und positiver eingeschätzt werden (Broverman, Broverman, Clarkson, Rosenkrantz, & Vogel, 1970; Broverman et al., 1972) und somit das Geschlecht von vornherein ungleich soziale Valenz-Kategorien herstellt. Der positive Effekt von Maskulinität geht einher mit dem größeren Wert, den die heutige Gesellschaft Männern und ihren Eigenschaften zuschreibt (Broverman et al., 1972; 1994). So ist es wenig überraschend, dass die Entwicklung von Maskulinität in dieser Gesellschaft immer noch die notwendige Strategie zu sein scheint, um zu einer positiven Einschätzung des eigenen Selbst zu gelangen, unter anderem gezeigt in Studien von Klingenspor (1994a; 1994b; 2002; mit Rastetter, 2004), die den Einfluss des Geschlechtsrollenselbstkonzepts mediiert durch Selbstwert auf Essstörungen zeigen konnte. Klingenspor hatte unter anderem hypothetisiert, dass das Risiko für eine Bulimie in Teilen durch die weibliche Geschlechtsidentität bestimmt wird. Nachdem erste Ergebnisse auf eine geschlechtstypische Identität bei bulimischen Frauen hindeuten (62% der bulimischen Frauen in der Stichprobe wurden als geschlechtstypisch klassifiziert), konnte Klingenspor (1994a) in weiteren Studien zeigen, dass Frauen, die eher maskuline

Eigenschaften in ihr Selbstkonzept integrieren, erfolgreicher darin sind, das eigene Selbst positiv zu bewerten und daher weniger anfällig sind eine Essstörung zu entwickeln.

Die Valenz der Eigenschaften wird nicht nur durch kulturell vorherrschende Geschlechterstereotype bestimmt, sondern liegt bisweilen in den zur Selbstbeschreibung herangezogenen Worten selbst. In einer Studie zu Wortnormierungen, die die semantische Bedeutung von Wörtern in den Vordergrund stellt, haben Leibetseder und Kollegen (Leibetseder, Mackinger, Hammerer, Staudinger, & Vogl, 1996) geschlechtsspezifische Einschätzungen von insgesamt 36 Substantiven, Adjektiven und Verben hinsichtlich ihrer Valenz (Verursachung eines angenehmen oder unangenehmen Gefühls), Aktivierung und Symbolik untersucht. Davon abgesehen, dass die Ergebnisse abhängig sind von Wortrubrik und Geschlecht der Versuchsperson, verdeutlicht diese Studie in erster Linie, dass Wörter per se eine Valenz besitzen, was gerade im Bereich der adjektiv-basierten Stereotypenforschung von essentieller Bedeutung ist. Zusätzlich offenbarte sich, dass das Wort *abhängig* den geringsten Valenzwert besitzt und Männer und Frauen sich diesbezüglich einig sind (ein Aspekt, der an späterer Stelle der Arbeit noch einmal aufgegriffen wird).

Ein weiterer wichtiger Ansatz in diesem Zusammenhang kommt von Perkins und Forehand (2006), die überzeugt waren, dass Attribute nicht aufgrund ihrer positiven Valenz eher zur Selbstbeschreibung herangezogen werden, sondern auch in Abhängigkeit davon wie stark sie mit der bestehenden Idee des eigenen Selbst übereinstimmen, dadurch eher wahrgenommen und in die mentale Repräsentation des Selbst integriert werden. Bezüglich der Bedeutsamkeit von Attributen für das eigene Selbst vermuteten die Autoren vor allem zwei grundlegende Prinzipien: Positivität und Übereinstimmung. Perkins und Forehand (2006) überprüften demnach einen möglichen Valenzwert in der mentalen Repräsentation mit dem Selbst. Die Autoren wollten auf Arbeiten von Markus (1977; 1983) aufbauend mithilfe eines IATs sowohl die valenzbezogenen als auch die semantisch bedeutenden Prozesse identifizieren, die eine bestimmte Eigenschaft selbst-relevant machen, die dadurch als stärker verknüpft mit dem Selbst gilt und als ein Resultat dessen vom Individuum stärker zur Selbstbeschreibung herangezogen wird. In vier Studien, in denen Perkins und Forehand (2006) den Einfluss von Valenz und semantischer Bedeutung bezüglich der Selbstbezogenheit variierten, zeigte sich, dass zwar sowohl Valenz als auch die semantische Bedeutung eine Rolle bei den kognitiven Assoziationen spielen, dass jedoch überraschend die semantische Bedeutung die stärkere Ursache für die Herstellung der Assoziationen ist. Die Art der Zuschreibung ging sogar so weit, dass Personen eine größere Assoziation mit Attributen demonstrieren, die semantisch anderen Eigenschaften ihres Selbstkonzept entsprachen und dabei jedoch von negativer Valenz waren (z. B. *halsabschneiderisch*) als jenen, die semantisch von gegensätzlicher, aber positiver Wertigkeit waren (z. B. *gelassen*). Obwohl diese Arbeit die Bedeutung der semantischen Verknüpfung im Selbstkonzept herausstellte, berücksichtigte sie ebenfalls den Einfluss der Valenz im Prozess der Selbstbeschreibung.

Pahl und Eiser (2005) zielten darauf ab die reinen Valenzeffekte zu isolieren, indem sie linguistische Antonyme der gleichen Dimension paarten (z. B. *faul* vs. *hart arbeitend*). Inhalt und



Valenz von Attributen wurden kontrolliert, erwünschte und unerwünschte Items noch einmal in zwei Kategorien bezüglich (mittel, extrem) unterteilt und Versuchspersonen dahingehend befragt, sich entweder selbst mit anderen (Fokus Selbst-Andere) oder andere mit dem eigenen Selbst zu vergleichen (Fokus Selbst-Andere). Die Autoren fanden in ihrer Studie, dass der Positivierungsbias bei der eigenen Einschätzung stärker im Selbst-Andere Fokus auftrat und dass weniger die Valenz selbst, sondern vor allem die Typizität und dabei insbesondere das Ausmaß der Extreme in der Itemwertigkeit (sowohl für sozial erwünschte als auch unerwünschte) diesen Positivierungsbias beeinflusste. Je erwünschter oder unerwünschter bestimmte Eigenschaften waren, desto stärker fiel der Bias aus.

In Bezug auf die Einschätzung zur Typizität und sozialen Erwünschtheit könnte dies folgendes bedeuten: bei der Typizitätsbestimmung liegt der Fokus eindeutig auf den Unterschieden zwischen den jeweiligen Gruppen. Um die Differenzierung zur Fremdgruppe möglichst hoch zu halten und aufgrund des Wissens um kulturell vorherrschende Geschlechterstereotype, können die Probanden klare stereotype Eigenschaften den jeweiligen Gruppen zuschreiben. In dem Zusammenhang fällt es nicht schwer, auch negative Eigenschaften als geschlechtsspezifisch auszugeben.

Ein anderes Bild ergibt sich bei der sozialen Erwünschtheit. Bezugspunkt ist hierbei im größeren Maße das eigene Ich – die eigene Identität – womit die Einschätzung als solche eine größere Gefahrenquelle für die Herabsetzung des Selbstwertes darstellt. Frauen – im Minoritätenstatus und von der Rollenverteilung immer noch klassisch auf die familiäre Verantwortung reduziert – sind sich sehr klar darüber, dass maskuline Eigenschaften – positive als auch negative – vonnöten sind, um als gleichberechtigt angesehen zu werden. Dementsprechend könnten die Einschätzungen der Valenz womöglich schon eine sozial erwünschte Verteilung der Geschlechtsrollenstereotype in der Gesellschaft widerspiegeln und demnach als ausschließliches Auswahlkriterium in der heutigen Zeit nicht mehr geeignet sein. Der Einfluss dieser Tendenzen des sozial erwünschten Antwortverhaltens bezüglich der Bewertung von Geschlechtereigenschaften, die sehr wohl sozial unterschiedlich erwünscht sind, ist durch bereits vorliegende methodische Ergebnisse wie der Korrelation von negativer maskuliner und negativer femininer Teilskalen indiziert (Athenstaedt, 2003; Athenstaedt & Altstötter-Gleich, 2009).

#### 2.3.2.2.2 Soziale Erwünschtheit (negativer) geschlechtsbezogener Eigenschaften

Swazina und Kollegen (2004) fassten in ihrer Studie zu geschlechtsspezifischen Idealen im Wandel der Zeit, verschiedene Arbeiten zusammen, die ein Plädoyer dafür halten, in auf Adjektiven basierenden Arbeiten nicht nur Repräsentativität oder Bedeutungshaltigkeit zu überprüfen, sondern auch die soziale Erwünschtheit. Ricciardelli und Williams (1995) konnten die Bedeutung sozial unerwünschter geschlechtsbezogener Eigenschaften herausarbeiten und zeigen, dass unerwünschte Aspekte der Maskulinität mit häufigerem und stärkerem Trinkverhalten sowohl bei Frauen als auch Männern einhergehen.

Prentice und Carranza (2002) initiierten eine Diskussion hinsichtlich der sozialen Erwünschtheit im Rahmen der Auswahl der Items am Beispiel des BSRI (*“defends own beliefs”*

gilt als erwünschter für Männer als für Frauen). Diese Art der Auswahl lässt zwei Interpretationen zu, dahingehend, dass die soziale Norm, für seine eigene Meinung einzustehen, für Männer stärker ist, während die andere Erklärung impliziert, dass die Umsetzung dieser Norm für Frauen nicht so bedeutend ist. Beide Auswahlkriterien haben einen gemeinsamen Nenner, zeigen aber umso mehr die Bedeutung der Valenz und der Perspektive in diesem Zusammenhang auf. Um dieser Sichtweise auf den Grund zu gehen, machten Prentice und Carranza (2002) einen für sie passenderen Vergleichsrahmen auf, indem sie das jeweilige Geschlecht<sup>18</sup> mit dem einer allgemeinen Person<sup>19</sup> vergleichen ließen und nur dann von der Geschlechtstypizität ausgingen, wenn die soziale Erwünschtheit eines Geschlechts die einer Person übersteigt. Sollte eine Eigenschaft als gleichermaßen sozial erwünscht für Männer und eine Person im Allgemeinen gelten, stellt diese Eigenschaft kein präskriptives Stereotyp für Männer dar. Ausgehend von dieser Sichtweise hypothetisieren Prentice und Carranza (2002) ein Kategorisierungsprinzip (Tabelle 5), definiert durch generelle soziale Erwünschtheit (hoch vs. niedrig) und der Richtung, in welche das Zielgeschlecht von der sozialen Norm abweicht (höher oder niedriger).

Tabelle 5. Kategorien der Eigenschaften, die sich in ihrer sozialen Erwünschtheit für Männer und Frauen unterscheiden (entnommen aus: Prentice & Carranza, 2002)

Valenz der Eigenschaften	Erwünschter für das Zielgeschlecht verglichen mit einer Person	Weniger erwünscht für das Zielgeschlecht verglichen mit einer Person
Sozial erwünscht	Gender-intensified prescriptions	Gender-relaxed prescriptions
Sozial unerwünscht	Gender-relaxed proscriptions	Gender-intensified proscriptions

*Gender-intensified prescriptions* sind demnach Eigenschaften, die hoch allgemein sozial erwünscht eingeschätzt wurden und umso höher für das Zielgeschlecht (präskriptive Eigenschaften). *Gender-relaxed prescriptions* sind ebenfalls hoch in allgemeiner sozialer Erwünschtheit, jedoch signifikant geringer für das Zielgeschlecht – und ebenso geschlechtstypisch, aber im geringeren Maße einem sozialen Vergleichs- und Abrechnungsmaßstab ausgesetzt. *Gender-relaxed proscriptions* sind sozial unerwünscht, jedoch immer noch charakteristischer für das eine als für das andere Geschlecht (erlaubte „Fehler“ oder „Unzulänglichkeiten“ der Geschlechter). *Gender-intensified proscriptions* sind schließlich sozial unerwünschte Eigenschaften, die sogar noch geringere Werte der sozialen Erwünschtheit für das Zielgeschlecht aufweisen.

Nach dieser Definition würde dies für Frauen folgendes bedeuten: Eigenschaften, die generell hoch sozial erwünscht und höher für Frauen im Vergleich zu einer Person sind, entsprechen *intensified prescriptions*; Eigenschaften – hoch in allgemeiner sozialer Erwünschtheit, aber geringer für Frauen = *relaxed prescriptions*; Eigenschaften – gering in allgemeiner sozialer Erwünschtheit, aber höher für Frauen = *relaxed proscriptions*; und

<sup>18</sup> „how desirable it is in American society for a man to possess each of these characteristics“ (S. 271)

<sup>19</sup> „how desirable it is in American society for a person to possess each of these characteristics“ (S. 271)

schließlich Eigenschaften, sowohl gering in allgemeiner sozialer Erwünschtheit als auch eingeschätzt für Frauen = *intensified proscriptions*. Zur Überprüfung dieser Kategorisierung befragten Prentice und Carranza (2002) männliche und weibliche Studierende, in einem ersten Teil nach der Selbstbeschreibung und anschließend im zweiten Teil nach der Einschätzung einer von fünf Varianten (1. soz. erwünscht für Frauen, 2. soz. erwünscht für Männer, 3. soz. erwünscht für eine Person, 4. typisch für Männer, 5. typisch für Frauen). Während ein möglicher Bias (durch die Art der Erfassung und des Triggers der Valenz) problematisch ist, konnten die Ergebnisse generell die postulierten Annahmen bestätigen. So gab es beispielsweise ein Cluster positiver femininer Eigenschaften (*intensified prescriptions* für Frauen; und *relaxed prescriptions* für Männer) oder ein Cluster negativer, maskuliner Eigenschaften (*relaxed proscriptions* für Männer, und *intensified proscriptions* für Frauen).

Ein Aspekt schränkt die Ergebnisse weiter ein. Prentice und Carranza (2002) hatten ausführlich die Auswahl ihrer Items und die Kombination von Synonymen beschrieben – ein Vorgehen, welches einer ökonomischen Plausibilität folgt, jedoch auch Gefahr läuft, das Kategoriensystem in Frage zu stellen, da umgekehrt gepolte Items, die womöglich in unterschiedlichen Zellen des Kategorisierung landen würden, schon von vornherein reduziert wurden. Dadurch wurden nicht alle inhaltlichen Möglichkeiten und Flexibilität seitens der Versuchspersonen ausgeschöpft. Darüber hinaus gab es eine Vielzahl von Charakteristiken in jeder Kategorie, die nicht in der komplementären Kategorie des anderen Geschlechts auftauchten. Bedeutsam waren in diesem Zusammenhang vor allem Items bei denen es keine Überlappung für Geschlechter gab, wie beispielsweise *intelligent* oder *humorvoll*, die als weniger sozial erwünscht für Frauen galten, aber nicht erwünschter für Männer, ebenso wie *happy* und *optimistisch* weniger erwünscht für Männer galten aber nicht erwünschter für Frauen. Diese Ergebnisse zeigen vor allem auf, dass soziale Präskriptionen für Frauen und Männer nicht einfach spiegelbildlich sind, in denen die Standards des einen Geschlechts als *relaxed* und weniger wichtig für das andere gelten. Darüber hinaus gab es für eine Vielzahl von Items eine hohe Korrespondenz von sozialer Erwünschtheit und Typizität.

Die Differenz der sozialen Erwünschtheit als Marker für die Beschreibung der Eigenschaft eines Geschlechts ist laut Prentice und Carranza (2002) kritisch und ignoriert zudem die sozial unerwünschten Eigenschaften, die erst in einer zweiten Entwicklungsstufe betrachtet wurden (Antill, Cunningham, Russell, & Thompson, 1981; Bryson & Corey, 1977; Kelly et al., 1977; Spence et al., 1979; Stoppard & Kalin, 1978). undefiniert blieb zunächst, wie diese sozial unerwünschten Eigenschaften die Geschlechtsbeschreibungen beeinflussen. So ist zum einen möglich die unerwünschten geschlechtskongruenten Eigenschaften im gleichen Maße zu demonstrieren wie die erwünschten. Zum anderen könnten gesellschaftlich negativ eingeschätzten Eigenschaften vermieden werden (Broverman et al., 1972; Stoppard & Kalin, 1978). Fest zu stehen scheint, dass es in erster Linie darum gehen muss, die geschlechtsinkongruenten sozial unerwünschten Eigenschaften für sich selbst abzulehnen, wie Stoppard und Kalin (1978) zeigen konnten. So erhielten geschlechtskongruente Eigenschaften die höchsten Bewertungseinschätzungen im Vergleich zu den geschlechtsinkongruenten

Eigenschaften. Das gleiche Bild ergab sich – wenn auch in abgeschwächter Form – für sozial unerwünschte Eigenschaften. Beide Eigenschaftsgruppen sind präskriptiver für das eine Geschlecht als das andere – über die Valenzen hinweg – soziale Standards wirken aber unterschiedlich bestimmend und wertend.

Wenn es darum geht, einen Gruppenvergleich vornehmen zu müssen und dabei gleichzeitig den eigenen Selbstwert zu erhalten oder sogar zu erhöhen, stellt die Bewertung von negativen Eigenschaften in diesem Zusammenhang eine Herausforderung dar.

#### 2.3.2.2.3 Zuschreibung negativer geschlechtsbezogener Eigenschaften

Insgesamt stellt sich die Frage, wie sich die negativen Eigenschaften ins Gesamtkonzept des Geschlechtsrollenselbstkonzepts theoretisch und empirisch eingliedern lassen und was für die Zuschreibung von Eigenschaften von größerer Relevanz ist: die Typizität oder die Valenz von Eigenschaften. Phoenix (1982) hat bei der Überprüfung des Zusammenhangs positiver und negativer maskuliner sowie femininer Eigenschaften mit Selbstverwirklichung festgestellt, dass Personen, die die meisten positiven maskulinen sowie negativen femininen Eigenschaften für sich beanspruchen, die niedrigsten Werte bei der Selbstverwirklichung erreichten. Dies ist ein Indiz dafür, dass diese beiden Skalen einen großen gemeinsamen inhaltlichen Bezug haben.

Powlishta (1995) konnte zeigen, dass schon 8- bis 10-jährige Kinder eine generelle Tendenz haben, vorgegebene positive Eigenschaften dem eigenen Geschlecht zuzuschreiben und negative Eigenschaften dem anderen, zunächst einmal unabhängig von der Geschlechtstypizität der Eigenschaften. Stoppard & Kalin (1978), ließen ihre studentischen Versuchspersonen feminine und maskuline Eigenschaften auf 4 Dimensionen einschätzen: wie verbindlich (*obligatory*), verboten (*prohibited*) und charakteristisch diese Eigenschaften jeweils für Männer und Frauen sind (*characteristic*) sowie inwieweit sie auf Zustimmung oder Ablehnung für die jeweiligen Geschlechter stoßen (*approval / disapproval*). Damit stellten die Autoren einen Bezug von Messungen der Geschlechterstereotype (*characteristic*) zu denen der Geschlechtsrollenwahrnehmung (*obligatory, prohibited, approval / disapproval*) her. Dabei zeigte sich, dass sozial erwünschte geschlechtstypische Eigenschaften die höchsten Werte für Zustimmung und Verbindlichkeit erhielten. Im Vergleich dazu erreichten sozial unerwünschte geschlechtstypische Eigenschaften durchschnittliche Einschätzungen auf allen Dimensionen und die geschlechtsuntypischen die negativsten Werte. Stoppard & Kalin (1978) sahen dies als Hinweis darauf, dass unerwünschte Charakteristiken weniger relevant für Beschreibungen sind.

Coleman und Hong (2008) hatten sich ebenfalls der Bedeutung negativer Eigenschaften angenommen und dies mit dem individuellen Glauben an die Ursächlichkeit von Geschlecht in Bezug gesetzt. In ihren Studien unterschieden sie zwischen (ausschließlich) Frauen, die an eine biologische Determination des Geschlechts glauben, und denen, die die Sozialisationstheorie favorisieren und betrachteten dies in Zusammenhang mit der individuellen Selbststereotypisierung, der Zuschreibung von geschlechtsspezifischen Eigenschaften. Neben dem Haupteffekt der Valenz stellte sich zudem heraus, dass der Glaube an die biologische Theorie des Geschlechts mit einer größeren Zuschreibung femininer (auch negativer)

Eigenschaften einher ging und die Reaktionszeiten beim Ablehnen stereotyper femininer Eigenschaften zurückgingen. Coleman und Hong (2008) erklärten ihr Ergebnis damit, dass bei negativen Eigenschaften zunächst die Tendenz besteht, diese im geringeren Maße zur Selbstbeschreibung heranzuziehen. Unter den Bedingungen jedoch, in denen Frauen an eine biologische Bestimmung glauben, gehören negative Charakteristiken zum Status Quo, und die Zuschreibung derer einfacher. Auch die Diskussion der Selbststereotypisierung (Biernat et al., 1996) erkennt an, dass mit ausreichender Motivation sich Individuen durchaus auch negative Eigenschaften zuschreiben können.

Aus dem bisherigen Stand der Forschung kann man schlussfolgern, dass negative Eigenschaften nicht nur für die Beschreibung von Fremdgruppen, sondern auch für die eigene Gruppe herangezogen werden – schon von frühester Kindheit an. Die genaue Vorgehensweise und Gründe dafür – vielfältig und noch nicht endgültig geklärt – reichen von Aufrechterhaltung und Anerkennung des bestehenden Status Quo bis zur Abwertung der Fremdgruppe zum Schutz des eigenen Status. Im Zusammenhang der Zuschreibung von positiven und negativen Attributen wird wiederholt diskutiert, welche Rolle der semantischen Bedeutung, der Wertigkeit im Sinne einer Valenz (Greenwald et al., 2002) oder auch der Kontrollierbarkeit (Alicke, 1985) von diesen Attributen und Ereignissen im Rahmen der Selbstzuschreibung zukommt. Darüber hinaus wird ebenfalls untersucht inwieweit der Fokus und die Individualisierung der Zuschreibung (Selbst vs. Andere) (Alicke et al., 1995) sowie kognitive als auch motivationale (Better-Than-Average-Effekt; Alicke, 1985) Ursachen dafür verantwortlich gemacht werden können. Deutlich wird jedoch bereits, dass negative Attribute ein wichtiger Bestandteil in der Beschreibung von Individuen und Gruppen sind und ihnen auch im Bereich des Geschlechtsrollenselbstkonzepts die entsprechende Rolle zuteilwerden sollte.

In Studien jedoch offenbaren negativ bewertete Eigenschaften wiederholt Schwierigkeiten, dahingehend, dass sie beispielsweise eine wesentlich höhere Varianz in einigen Gruppen zeigen, was gerade die Voraussetzungen für bestimmte statistische Verfahren verletzt. Im Vergleich zu den positiven zeigen die negativen Eigenschaften über Gruppen hinweg zusätzlich kein konsistentes Bild, wie und in welchem Ausmaß sie zur Beschreibung herangezogen werden (Fiske, Xu, Cuddy, & Glick, 1999). Während bisweilen vermutet wird, dass einige der negativen Eigenschaften weniger häufig im Sprachgebrauch sind und oft extremer in ihren Formulierungen, bleibt die Frage offen, in welcher Form der Erhebung die negativen Eigenschaften ohne die bekannten Schwierigkeiten durch beispielsweise ein extra Set an Verhaltens- und Persönlichkeitseigenschaften optimal abgebildet werden können oder inwieweit geringe Zuschreibungen der positiven Variablen einfacher vorzunehmen sind und als Richtungsgeber ausreichen.

Fest steht jedoch, dass ein Instrument zur Erfassung des GSK negative Attribute berücksichtigen muss, sowohl im Rahmen des Konstruktionsprozesses der Skalen als auch im späteren Einsatz. In den folgenden Abschnitten wird die bisherige Messung negativer Eigenschaften des GSK vorgestellt. Dabei wird vor allem auch auf ein Instrument von Helgeson und Fritz (Fritz & Helgeson, 1998; Helgeson, 1993; 1994b) eingegangen. Zunächst wird der

Beitrag dieses Instrumentes zur Forschung und zur Herstellung eines Bezugs zu Erlebens- und Verhaltenskonzepten diskutiert, ehe die Korrelate der in dieser Arbeit geplanten Teilskalen vorgestellt werden.

### **2.3.2.3 Berücksichtigung negativer Eigenschaften bei der Messung des GSK**

Aus den bisher berichteten Studien lässt sich ableiten, dass eine Person sich auch zu den negativen Elementen des männlichen bzw. weiblichen Geschlechterstereotyps in seiner geschlechtsbezogenen Identität positionieren können sollte. Ähnlich wie bei den positiven Eigenschaften müsste sich demnach das GSK auch geschlechterabhängig in der Selbstbeschreibung negativer Eigenschaften zeigen.

In Bezug auf Messinstrumente wurde dies in Ansätzen bereits in der *Adjective Check List* (Williams & Best, 1990), der *Australian Sex Role Scale*, ASRS (Antill et al., 1981) und dem *Extended Personal Attributes Questionnaire*, EPAQ (Spence et al., 1979) umgesetzt. Die ASRS, als Weiterentwicklung des BSRI, enthält neben der sozialen Erwünschtheitskala vier Skalen mit jeweils 10 positiven männlichen und weiblichen sowie 10 negativen männlichen und weiblichen Eigenschaften. Der EPAQ mit sieben Skalen (M+, F+, M-F, M-, Fva- und Fc-<sup>20</sup>) erfasst als Erweiterung seines Vorgängers PAQ zwei zusätzliche, sozial unerwünschte Aspekte im Geschlechtsrollenelbstkonzept, die extreme Ausprägungen von Maskulinität und Femininität reflektieren. Dabei ähnelt diese negative Maskulinität der von Bakans (1966) festgestellten *unmitigated agency* (UA) und die negative Femininität der *unmitigated communion* (UC)<sup>21</sup>. Laut Bierhoff-Alfermann (1989) sollen diese negativen Attribute „*der zusätzlichen Charakterisierung der Probanden dienen [...]*“ (S.33) und nicht zur Typisierung bezüglich des Geschlechtsrollenelbstbildes berücksichtigt werden. Der EPAQ wurde jedoch für die deutsche Fassung lediglich übersetzt, ohne jedoch die Typizitäts- und Erwünschtheitswerte der Attribute für diesen gesellschaftlichen Kontext zu validieren (GEPAQ, Runge, Frey, Gollwitzer, Helmreich, & Spence, 1981). Athenstaedt (2003) konnte dennoch für den GEPAQ zeigen, dass sich Frauen sozial unerwünschte feminine Eigenschaften stärker zuschrieben als sozial unerwünschte maskuline Eigenschaften. So wiesen Athenstaedt, Heinzle und Lerchbaumer (2008) mittels des GEPAQ ebenfalls nach, dass die Selbstkategorisierung und Zuordnung zu vorgegebenen Subgruppen von den stereotypen Umgebungsmerkmalen der Umgebung abhängt. Befragte Frauen auf Spielplätzen beispielsweise ordneten sich demzufolge eher der Gruppe der Hausfrauen zu, während Männer an Schulen und Universitäten die Gruppe der Professoren bevorzugten und sich gleichzeitig mehr negative feminine Geschlechtereigenschaften zuschrieben als die anderen männlichen Untergruppen (z. B. Machos, Hippies).

<sup>20</sup> M+ (sozial erwünschte Maskulinität), F+ (sozial erwünschte Femininität), MF (bipolare Maskulinität und Femininität), M- (sozial unerwünschte Maskulinität), Fva- (feminine verbale Aggressivität), Fc- (feminine *Unmitigated Communion*)

<sup>21</sup> Communion und Agency sind im Deutschen sprachlich eher unzulänglich repräsentiert. Die Übersetzung „Gemeinschaftsorientierung“ (Brunstein, Lautenschlager, Nawroth, & Pöhlmann, 1995) beispielsweise deckt das Konstrukt Communion eher unzureichend ab. Die reine Übersetzung von *unmitigated* würde zu *ungemildert* führen, repräsentiert die Teilskalen jedoch nur unzureichend, zumal dadurch die Beziehungen zu den jeweils anderen (Agency und Communion) unklar bleiben und zu wenig Abgrenzung erfahren. Aufgrund der Übersetzungsprobleme sollen daher beide Termini in der Originalbenennung verbleiben.

Athenstaedt und Altstötter-Gleich (2009) beschäftigen sich mit einer neuen Skala zur Messung des Geschlechtsrollenselbstkonzepts, die auch negative Aspekte in der Selbstbeschreibung berücksichtigt und haben dem neuen Instrument im Gegensatz zum bisherigen Vorgehen nicht etwa eine Unterscheidung der sozialen Erwünschtheit der Eigenschaften zugrunde gelegt, sondern lediglich die Typizität als verschieden zwischen den Geschlechtern vorausgesetzt. Bezüglich der sozialen Erwünschtheit sollte diese für beide Geschlechter gleich sein. Leider blieben sie eine Erklärung für dieses Vorgehen schuldig, ähneln in ihrem Vorgehen aber der ursprünglichen Itemauswahl des PAQ, für den solche Adjektive als geeignet erachtet wurden, die sich nicht in der sozialen Erwünschtheit für beide Geschlechter unterschieden, wohl aber als typischer für das eine als für das andere Geschlecht eingeschätzt wurden. Altstötter-Gleich (2004) – wenn auch nur in Bezug auf positive Items – setzte ebenfalls die soziale Erwünschtheit als Auswahlkriterium ein, berichtete jedoch, dass die Items hohe Erwünschtheitswerte aufweisen, was die Vermutung nahelegt, dass in diesem Fall ebenfalls keine unterschiedliche Bewertung der Items für das eine im Vergleich zum anderen Geschlecht zugrunde gelegt wurde. Orlofsky (1982) setzte bei der Entwicklung der *Sex Role Behavior Scale* ebenfalls eine Gleichheit der sozialen Erwünschtheit für beide Geschlechter voraus, während Typizitätsratings die Zugehörigkeit zu den jeweiligen Geschlechtern und damit auch Teilskalen *male-valued* (Segeln) und *female-valued* (Bridge spielen) bestimmten. Dennoch existieren in diesem Messinstrument zusätzlich auch Eigenschaften, die sehr wohl sozial erwünschter für das jeweils eine oder andere Geschlecht sind und als *sex-specific* Aktivitäten (Fußball und Stricken) definiert werden.

#### 2.3.2.3.1 *Unmitigated Agency und Unmitigated Communion*

Die Einführung der negativen, sozial unerwünschten Aspekte des Geschlechtsrollenselbstkonzepts stellte berechnete Fragen nach dem Zusammenhang der Teilskalen dahingehend, ob auch die sozial unerwünschten Teilaspekte ein voneinander unabhängiges zweidimensionales System darstellen (wie von den positiven Aspekten bereits bekannt, Bem, 1974; Spence et al., 1975), oder ob die jeweils inhaltlich verknüpften, unterschiedlich valenzierten Konzepte womöglich Pole eines Kontinuums abbilden.

Ein Ansatz, der sich dieser Beziehungen annahm, war im Weiteren der von *Agency* (AG) und *Communion* (CM) (Spence et al., 1974; Spence et al., 1979). *Agency* (Instrumentalität) als repräsentativer Aspekt von Maskulinität mit einem Fokus auf das eigene Selbst in Abgrenzung zu anderen und dem Erreichen persönlicher Ziele manifestiert sich vor allem in Persönlichkeitseigenschaften wie Durchsetzungsfähigkeit (*assertiveness*) und Selbstvertrauen (*confidence*). Kontrastierend dazu repräsentiert *Communion* (Expressivität) Aspekte der Femininität mit dem überwiegenden Fokus auf andere, Unterstützung derer beim Erreichen ihrer Ziele und genereller interpersönlicher Beziehungen und zeigt sich in Eigenschaften wie mitfühlend und aufmerksam (Bakan, 1966; Ghaed & Gallo, 2006; Helgeson, 1993; Helgeson & Fritz, 1998; 1999; 2000; Spence et al., 1979). Ganz ursprünglich wurden *Communion* und *Agency* von Bakan (1966) als umfassende, divergierende Persönlichkeitskonstrukte eingeführt, wobei *Communion* – die Betonung des Sozialen

repräsentierend – typischer für Frauen ist, während *Agency* – Betonung der Individualität – vor allem in Männern determiniert ist, obwohl Bakan schon damals anmerkte, dass es sich um umfassende Organisationsprinzipien der menschlichen Existenz handelt, die nur schwerlich mittels eines einzelnen Instruments abgebildet werden können.

Beide Aspekte erfuhren schließlich eine Re-Konzeptualisierung durch Fritz und Helgeson (Fritz & Helgeson, 1998; Helgeson, 1993; 1994b) derart, dass nun vor allem die extremen, weniger angepassten Ausprägungen beider Eigenschaften: *Unmitigated Agency* (UA) und *Unmitigated Communion* (UC) mitberücksichtigt wurden. UC ist definiert durch den Fokus auf andere unter Ausschluss des eigenen Selbst (dementsprechend wahrnehmbarer Abwesenheit von Agency, Helgeson & Fritz, 1998), während UA das Selbst in den Vordergrund stellt unter gleichzeitigem Ausschluss der anderen. In Selbstbeschreibungsinstrumenten wird UA durch die entsprechende Subskala des EPAQ (Spence et al., 1979) erfasst und UC durch eine von Helgeson entwickelte Skala (Fritz & Helgeson, 1998; Helgeson, 1993; Helgeson & Fritz, 1998), die das Reagieren auf Bedürfnisse anderer in den Vordergrund stellt: z. B. „/ *always place the needs of others above my own*“.

Das Bedeutsame an diesen Konzepten ist, dass sie vor allem weniger angepasste und damit auch sozial unerwünschte Aspekte enthalten: *Unmitigated Agency* beispielsweise repräsentiert den weniger erwünschten Teil des gesellschaftlichen maskulinen Ideals und gilt daher als separates Konstrukt sowohl zu *Agency* als auch zu den sozial unerwünschten Aspekten der femininen Orientierung. Ähnlich zu *Femininität* und *Maskulinität* (Bem, 1974) betonte Helgeson (1993) die Unabhängigkeit der vier Konstrukte (wie beispielsweise durch die Stärke der Korrelationen zueinander bestimmt) und bezeichnete sie als relativ stabile interpersonale Persönlichkeitsdimensionen. Ebenfalls entsprechend der Konzeption von Bem wiesen die nun vier Konstrukte eindeutige Geschlechtstypizitäten auf: gewöhnlich tendieren Frauen sozialisationsbedingt eher dazu, sich jeweils Eigenschaften von (*Unmitigated*) *Communion* zuzuschreiben (Fritz & Helgeson, 1998), während Männer typischerweise höhere Werte auf Instrumenten zur Messung von (*Unmitigated*) *Agency* erreichen (Helgeson, 1994b; Helgeson & Fritz, 1999; Helgeson & Fritz, 2000).

Der Beitrag Helgesons und Kollegen liegt in der ausführlichen Diskussion der Beziehungen dieser vier Teilskalen zueinander mit dem Resultat, dass die negativen Teilskalen nicht einfach eine Umkehr oder Erweiterung der bereits bestehenden Konzepte *Agency* und *Communion* darstellen, sondern sehr wohl ihren eigenen Beitrag im Rahmen einer individuellen Selbstbeschreibung leisten. So stellte Helgeson (Helgeson & Fritz, 1998; 1999; 2000) beispielsweise klar, dass hohe *Communion* und niedrige *Agency* nicht einfach gleichzusetzen sind mit *Unmitigated Communion*. Ein hoher Fokus auf andere und eine geringe Berücksichtigung des eigenen Selbst implizieren dementsprechend nicht automatisch, dass der hohe Bezug auf andere mit der Vernachlässigung eigener Bedürfnisse einhergeht. Helgeson & Fritz (1998) bewiesen die Unabhängigkeit von *Unmitigated Communion* und *Communion* empirisch mit unterschiedlichen kognitiven und Verhaltens-Manifestationen.



### 2.3.2.3.2 Beziehungen positiver und negativer Teilskalen

Der vorherige Abschnitt zeigte, dass die Konzepte von negativer Maskulinität und Femininität nicht gänzlich neu sind, wenn man berücksichtigt, dass die Arbeiten von Spence und Helgeson Konstrukte einführten, die zwar einen eigenen theoretischen Hintergrund haben, aber dennoch inhaltlich sehr viel Ähnlichkeit zu möglichen negativen Teilskalen des Geschlechtsrollenselbstkonzepts mit Femininität und Maskulinität aufweisen.

Vor diesem Hintergrund sollen in den folgenden Abschnitten zunächst bereits festgestellte Beziehungen der Konzepte dargestellt werden, deren Rückschlüsse für die Hypothesen der hier vorliegenden Arbeit herangezogen werden. Im Anschluss werden schließlich bereits gefundene Erlebens- und Verhaltenskorrelate eben jener eher sozial unerwünschten Eigenschaften des geschlechtsbezogenen Selbstkonzepts diskutiert, die für die Durchführung der vorliegenden Arbeit einen Teil der Grundlage stellen sollen.

In Bezug auf die Teilskalen zueinander haben beispielsweise Helgeson & Fritz (1999; 2000) einen positiven Zusammenhang von Unmitigated Communion zu Communion gefunden bei gleichzeitiger negativer Korrelation zu Agency, was der Darstellung unkorrelierter Agency und Communion von Spence et al. (1979) widersprach. Individuen, mit hohen Werten der Unmitigated Communion, bringen ein ungünstiges Set an Persönlichkeitsausprägungen mit, die im direkten Zusammenhang zu persönlichem Stress (ungünstiger Selbstwert; Vernachlässigung eigener Bedürfnisse), situationsspezifischem Stress (größere Auseinandersetzung mit den Problemen anderer Menschen) und Gesundheitsproblemen als Resultat dessen stehen – sowohl bei Erwachsenen als auch schon bei Jugendlichen (Helgeson, Escobar, Siminerio, & Becker, 2007).

Gleichermaßen konnten Helgeson und Fritz (1999) zeigen, dass Unmitigated Agency in keiner oder positiver Beziehung zu Agency selbst steht und konsistenterweise in inversivem Zusammenhang zu Communion. Unmitigated Agency, im EPAQ (Spence et al., 1979) als *cynical* oder *hostile* erfasst, ist nicht gleichbedeutend mit einer hohen Ausprägung auf Agency und geringer auf Communion. Es ist nachgewiesen, dass die negativen Teilskalen ihren eigenen Beitrag leisten. So fanden Helgeson und Fritz (Helgeson & Fritz, 1998; 1999; 2000) beispielsweise ein inkonsistentes Bild vom Zusammenhang Communion zu psychologischen Gesundheitsaspekten, welches sich nahezu sofort lösen lässt, wenn man auch sozial eher unerwünschte Aspekte berücksichtigt und unmitigated Aspekte aus Communion separat betrachtet. Demnach lässt sich kein Zusammenhang feststellen von Communion zu Stress, wohl aber von Unmitigated Communion.

Ein anderer Ansatz betrachtete ebenfalls das Zueinander möglicher vier Teilskalen: das Circumplexmodell von Wiggins (1979; 1991), welches sowohl eine Orthogonalität von M+ / F+ und M- / F- als auch bipolare Beziehungen zwischen M+ / M- und F+ / F- impliziert, die empirisch bestätigt werden konnten (Bem, 1974; Helgeson, 1994b; McCreary & Korabik, 1994; Spence et al., 1975). So gingen Wiggins und Holzmueller (1978; 1981) beispielsweise davon aus, dass jeweils die Maskulinitätsskalen (M+ und M-) sowie die Femininitätsskalen (F+ und F-) stark negativ miteinander korreliert sein sollten und konnten dies auch nachweisen.

Durch Korrelationen konnten in anderen Studien die möglichen Beziehungen aller Teilskalen untereinander genauer analysiert werden. So zeigten sich beispielsweise negative Korrelationen von jeweils Unmitigated Agency mit negativer Femininität sowie Unmitigated Communion mit positiver Maskulinität (Helmreich et al., 1981; McCreary & Korabik, 1994; Runge et al., 1981; Saragovi, Koestner, Di Dio, & Aubé, 1997). Bei Spence et al. (1979) waren die Korrelationen zwischen den beiden maskulinen Teilskalen, den beiden femininen Teilskalen wie auch zwischen negativer Maskulinität und negativer Femininität klein oder nicht signifikant. Athenstaedt (2003) fand ebenfalls geringe bis nicht signifikante Zusammenhänge zwischen den entsprechenden positiven und negativen Teilskalen (F- und F+) oder auch zwischen den gepaarten Valenz-basierten Skalen (z. B. F+ und M+). Es lässt sich also ableiten, dass die negativen Teilskalen in einigen Studien geringe Zusammenhänge zu jeweils anderen, positiven Teilskalen aufweisen, insgesamt jedoch als eigenständige Aspekte des GSK betrachtet werden können.

Spence und Helmreich (1978) konnten durch Zusammenhänge zu anderen Konstrukten die unabhängige Existenz positiver und negativer Geschlechtsrollenkonzepte bestätigen. Alle einzelnen Konstrukte können wichtige und einzigartige Informationen über relevante sozial-emotionale Aspekte für Gesundheit und das Wohlbefinden liefern. Spence und Helmreich (1978) fanden in ihrer Studie Hinweise, dass Agency erwartungsgemäß positiv mit verbaler Aggression assoziiert war und negativ mit Depression und Angst, UA in positivem Zusammenhang zu physischer Aggression stand und UC mit feindseligen Gedanken und erhöhtem emotionalen Stress einher ging. AG wies – entgegen der Ergebnisse von Helgeson und Fritz (z. B. 1999), aber in Übereinstimmung mit Saragovi, Koestner, Di Dio und Aubé (1997) – Zusammenhänge mit wahrgenommener sozialer Unterstützung auf – im Sinne der besseren sozialen Anpassung und Ausgeglichenheit von Individuen hoher AG Ausprägung. Wobei Saragovi et al. (1997) diskutierten, inwieweit die Assoziation zwischen AG und emotionalen Konstrukten vor allem durch ähnliche Iteminhalte überschätzt wird.

Gerade diese postulierte Unabhängigkeit von Unmitigated Agency und positiver Maskulinität sowie Unmitigated Communion und positiver Femininität differenziert die ursprünglich als hoch geschlechtstypisch (maskulin, feminin, Bem, 1974) definierten Individuen, die sich in ihren Ausprägungen der zwei zusätzlichen Teilskalen unterscheiden können. Die komplexen Zusammenhänge, der aus doch sehr unterschiedlichen Forschungsansätzen stammenden negativen Teilskalen des Geschlechtsrollenselbstkonzepts, verdeutlichen Möglichkeit und Notwendigkeit, diese vier Konzepte auch tatsächlich getrennt voneinander zu betrachten und durch eigene Messinstrumente abzubilden um Vorhersagefähigkeiten zu maximieren.

#### *2.3.2.3.3 Erlebens- und Verhaltenskorrelate von Unmitigated Agency und Communion*

Gerade Unmitigated Communion hat viel Aufmerksamkeit durch die Forschung erhalten, da wiederholt ein Zusammenhang zu negativen Effekten mentaler und physischer Gesundheit aufgezeigt werden konnte. So argumentierte Helgeson (1993; 1994b) dies nicht unplausibel im

Hinblick auf die extreme Ausprägung der Fokussierung auf andere und die nach außen gerichtete Prioritätensetzung, die eventuell zu Lasten der Aufmerksamkeit auf die eigenen physischen und psychologischen Bedürfnisse geht – soweit, dass diese Personen äußerst zurückhaltend sind was die Offenbarung eigener Gefühle angeht. Sie bieten gern Unterstützung an, fühlen sich aber unwohl diese Unterstützung im Umkehrschluss anzunehmen (Fritz & Helgeson, 1998).

Obwohl sowohl geringe Ausprägungen von Agency als auch hohe von Unmitigated Communion in Beziehung zu psychologischem Stress zu stehen scheinen, konnten Helgeson und Fritz (1996) demonstrieren, dass es sich hierbei nicht um die gleichen Wirkungsmechanismen handelt und die Ursachen des Stresserlebens nicht die gleichen sind, derart, dass Beziehungsstressoren zwar als Mediatoren der Beziehung von UC zu Gesundheitsverhalten fungieren, nicht jedoch in der von AG und Gesundheitsverhalten. Die anzunehmende Unabhängigkeit beider Ausprägungen konnten die Autoren durch eine weitere Studie bestätigen, in der sie AG und UC unterschiedliche Zusammenhänge mit Depression oder interpersönlichen Problemen nachweisen konnten (Helgeson & Fritz, 1999).

Eine Betrachtung der sowohl positiven als auch negativen Aspekte der Geschlechtsrollenorientierung erfolgte ebenfalls durch Woodhill und Samuels (2003), die sich dabei des durch Bem initialisierten Androgynitätstypen, gebildet aus den nun vier Teilskalen, annahmen und einen Zusammenhang zum Wohlbefinden und psychologischer Gesundheit herstellten. Dazu nahmen die beiden Autoren auf Grundlage der Messung des GSK mittels des EPAQ bei ihren Probanden eine Einteilung in einen von sieben Typen vor und konnten demonstrieren, dass sich z. B. positiv und negativ androgyne Personen signifikant voneinander unterscheiden, was die Bedeutung negativer Teilskalen untermauert. Insgesamt gesehen zeigten positiv androgyne, positiv maskuline und positiv feminine Individuen die besten Ausprägungen bezüglich des Gesundheitsverhaltens und Wohlbefindens.

Helgeson (1994b) schließlich gab einen Überblick zu Zusammenhängen zwischen spezifischen Persönlichkeitseigenschaften und der Gesundheit. So kamen unterschiedliche Studien zum Ergebnis, dass beispielsweise Agency eng verknüpft ist mit einer besseren Gesundheit, geringerer Ängstlichkeit und Depression sowie einem höheren Selbstwert, jedoch auch mit Aggression und Delinquenz. Communion wirkt sich positiv auf soziale Fähigkeiten und prosozialeres Verhalten aus. In den negativen Ausprägungen spiegelt sich dementsprechend negativeres Gesundheitsverhalten wider, in Form der jeweils anderen Extremen (Unmitigated Agency – körperliche Beschwerden; Unmitigated Communion – Burnout). Shapiro (1994) beschäftigte sich ausführlich mit der gefühlten Kontrolle im Leben eines Individuums und ordnete diese Kontrollwahrnehmung geschlechtsrollentypisch zu: so geht positive Maskulinität beispielsweise mit erhöhter Durchsetzungsfähigkeit im Sinne einer aktiven Veränderungsherbeiführung einher, während eine hohe Ausprägung der negativen Femininität mit nachgiebiger Kontrolle assoziiert wird. Demnach haben Personen mit negativer Femininität ein Gefühl von geringer bis zu gar keiner Kontrolle über ihre Situation, was sich in entschlassarmem und abhängigem Verhalten widerspiegelt.

Tabelle 6 stellt eine Übersicht verschiedener Studien dar, die die Auswirkungen der Persönlichkeitsausprägungen auf Verhaltensweisen untersucht haben. Dazu sind jeweils den GSK-Teilskalen Verhaltenskonstrukte zugewiesen, sowie mit den entsprechenden Literaturabgaben belegt.

Tabelle 6. Zusammenhang von GSK-Aspekten mit Verhaltensweisen

Agency	Unmitigated Agency	Communion	Unmitigated Communion
Beziehung zu Selbstwert, Selbstbewusstsein begünstigendes Gesundheitsverhalten Typ A Verhalten Gesundes Essverhalten	negativ assoziiert mit Schwierigkeiten in persönlichen Beziehungen und Empfinden von Emotionen, „acting out“ normabweichendes Verhalten erhöhter Alkoholkonsum geringes Wissen um Krankheitsaspekte Substanzmissbrauch	positiver Zusammenhang zu sozialen Beziehungen, z. B. eheliche Zufriedenheit, Erhalt von sozialer Unterstützung, Geben sozialer Unterstützung, Empathie, gezügeltes Essverhalten <sup>22</sup>	negative Zusammenhänge mit mentaler und physischer Gesundheit, z. B. psychologischen Stress, negatives Gesundheitsverhalten, höhere Depressionswahrscheinlichkeit, schlechte Anpassung an Krankheiten (Herz, Diabetis, Brustkrebs) ungünstiges Essverhalten negatives Schulverhalten cognitive restraint, relationship events, Co-Abhängigkeit, Vernachlässigung seiner selbst, gezügeltes Essverhalten <sup>23</sup>
(Danoff-Burg, Mosher, & Grant, 2006; Ghaed & Gallo, 2006; Helgeson, 1993; Helgeson & Lepore, 1997)	(Danoff-Burg et al., 2006; Helgeson & Fritz, 1999; Helgeson & Fritz, 2000; Helgeson & Lepore, 1997; Ricciardelli & Williams, 1995; Snell, Belk, & Hawkins, 1987; Spence et al., 1979; Williams & Ricciardelli, 2001)	(Helgeson & Fritz, 1996; 1998; 2000)  (Williams & Ricciardelli, 2000) <sup>23</sup>	(Danoff-Burg et al., 2004; Fritz, 2000; Helgeson, 1993; Helgeson et al., 2007; Helgeson & Fritz, 1996; 1998; 1999; 2000; Nagurney, 2008; Ricciardelli & Williams, 1995) (Dear & Roberts, 2002) <sup>23</sup> (Williams & Ricciardelli, 2000) <sup>23</sup>

Es lässt sich zusammenfassend die Notwendigkeit der Erfassung aller vier Facetten des GSK feststellen. Um der Komplexität des Geschlechtsrollenselbstkonzepts über attributenbasierte Messverfahren maximal gerecht zu werden, müssen sowohl jeweils positive als auch negative Aspekte der Femininität und Maskulinität berücksichtigt werden. Daraus ergeben sich Vorhersagen derart, dass die Beziehung zwischen positiven und negativen Merkmalen innerhalb des femininen bzw. maskulinen GSK für die Geschlechter stärker sein sollte als die Beziehung zwischen den negativen Merkmalen des männlichen und weiblichen GSK. Neben den Vorteilen der umfassenderen Messung der nunmehr vier Facetten des GSK bietet sich darüber hinaus die Möglichkeit, das GSK mit weiteren als bisher berücksichtigten psychologischen Konstrukten in Beziehung zu bringen.

### 2.3.3 Korrelate positiver und negativer Maskulinität / Femininität

Maskulinität und Femininität, die sich in unterschiedlichen Variationen von Verhaltensweisen und Kompetenzen äußern können, sollten demnach auch sehr unterschiedliche und eindeutig identifizierbare Zusammenhänge zu anderen Verhaltenskonstrukten haben. Wie von

<sup>22</sup> Daten erhoben mit der *Australian Sex-Role Scale*; berichtet wurden Korrelationsergebnisse; ausschließlich weibliche Stichprobe

<sup>23</sup> Die Geschlechtsrolle wurde in dieser Studie mit der dem Personal Description Questionnaire (Antill et al., 1981) erhoben. Dieser Fragebogen ist dem PAQ konzeptionell ähnlich – entwickelt jedoch für eine australische Stichprobe. Bemerkenswert bei Dear und Roberts (2002) ist, dass Co-Abhängigkeit sowohl mit hoher Identifikation sozial unerwünschter femininer Charakteristiken einhergeht, wie auch mit geringen Ausprägung positiver maskuliner Aspekte, was auf eine Ähnlichkeit der beiden Teilskalen hinweist.

unterschiedlichen Autoren wiederholt festgestellt, kann ein gewisser Kern der Maskulinität verhältnismäßig eindeutig Persönlichkeitseigenschaften wie Aggression, Leistung und Erfolg, Kraft, Kontrolle oder auch Ambitionen zugeordnet werden sowie gleichzeitig einer geringeren emotionalen Fähigkeit (für einen Überblick siehe Helgeson, 1994a). Obwohl diese Maskulinität häufig nur als positiv konzipierte Skala erfasst wurde, fällt auf, dass die in Beziehung gebrachten Konstrukte sowohl positive Eigenschaften (Erfolg, Kraft, Kontrolle, Ambitionen) als auch negative Merkmale (Aggressionen, reduzierte Emotionalität) betreffen.

Während die Dichotomisierung der Geschlechtsrollen in Maskulinität und Femininität nachweislich spezifischen Verhaltensweisen zugeordnet werden konnte, ist vor allem die zusätzliche Aufnahme eben jener neuen negativen Teilskalen eine Strategie, die vermuteten stärkeren Beziehungen zu anderen Konstrukten zu identifizieren, und die vier Teilskalen für zukünftige Forschungen attraktiv zu machen. Da der Einsatz auch negativer Geschlechtsrollenteilskalen nicht gänzlich neu ist und Erlebens- und Verhaltenskorrelate bestehender negativer Teilskalen bereits diskutiert wurden, lassen sich nun verschiedene Konstrukte in Betracht ziehen, die spezifisch einer der vier neuen Teilskalen zur Validierung zugeordnet werden sollen.

Im Rahmen der Validierung werden verschiedene Konstrukte zum Einsatz kommen, die sowohl die Konstrukt-, konvergente als auch divergente Validität des neuen Messinstrumentes unter Beweis stellen werden. So sollte bezüglich der Konvergenzvalidität, die den Zusammenhang zu Instrumenten ähnlichen Inhalts und Intention des zu messenden Konstrukts herstellt, der BSRI vergleichsweise erhoben werden. Darüber hinaus wurden bezüglich der vier Teilskalen verschiedene Kriterien angelegt, die beispielsweise die Relevanz der zusätzlichen Erfassung negativer Selbstbeschreibung belegen sollten. So wird unter anderem erwartet, dass bezüglich der Konstruktvalidität positive Vorzeichen der negativen Skalen mit den Problem-bereichen bei gleichzeitigem Fehlen negativer Vorzeichen der positiven Skalen einhergehen.

Eagly (1987) merkte bereits an, dass Forscher eben diesen Bezug zu reliablen, aggregierten Messungen des geschlechtsbezogenen Verhaltens vernachlässigt haben. Lippa und Connelly (1990) brachten daher eine Vielzahl von Verhaltensmessung in ihrer Studie zum Einsatz: mathematische, visuelle Fähigkeiten, Berufspräferenztest oder Hobbies.

In ähnlicher Form soll dies auch in der vorliegenden Studie geschehen. Im Folgenden werden die verwendeten Validierungskonstrukte theoretisch hergeleitet sowie methodisch beschrieben. Zur besseren Strukturierung geschieht dies jeweils getrennt für die vier Teilskalen. In Tabelle 7 sind alle in der vorliegenden Arbeit verwendeten Validierungskonstrukte und ihre Quellen aufgeführt. Der Methodik vorweggreifend sei an dieser Stelle bereits angekündigt, dass es zwei Validierungsstudien (Studie 5 und Replikationsstudie 6) geben wird, in denen unterschiedliche psychologische Konzepte zu den vier Teilskalen in Bezug gesetzt werden. Für die maskuline positive Teilskala sind dies beispielsweise beruflicher Ehrgeiz und Hoffen auf Erfolg in Studie 5 sowie Leistungsorientierung in Studie 6.

Tabelle 7. Quellen und Zuordnung der Validierungskonstrukte

<b>Konstrukte</b>	<b>Quelle</b>
<b>Studie 5</b>	
<b>Prädiktoren für MAS+</b>	
berufl. Ehrgeiz (AVEM)	Schaarschmidt & Fischer (1996)
Hoffen auf Erfolg (AMS)	Göttert & Kuhl (1980)
<b>Prädiktoren für MAS-</b>	
Aggressivität (AQ)	Möller (2006)
Werteskala Macht	Schröder (2005)
<b>Prädiktoren für FEM+</b>	
soziale Werte	Schröder (2005)
Empathie (Sorge)	Maes, Schmitt & Schmal (1995)
<b>Prädiktoren für FEM-</b>	
ESWS	Schütz & Sellin (2006)
Ängstl. Depressivität (BSF)	Hörhold & Klapp (1993)
<b>Studie 6</b>	
<b>Prädiktoren für MAS+</b>	
Leistungsorientierung (FPI)	Fahrenberg, Hampel, & Selg (2001)
<b>Prädiktoren für MAS-</b>	
Aggressivität (FPI)	Fahrenberg, Hampel, & Selg (2001)
<b>Prädiktoren für FEM+</b>	
Fürsorglichkeit (PRF)	Stumpf, Angleitner, Wieck, Jackson, & Beloch-Till (1985)
<b>Prädiktoren für FEM-</b>	
STAI	Laux, Glanzmann, Schaffner, & Spielberger (1981)

In den folgenden Abschnitten werden die Korrelate der einzelnen Teilskalen vorgestellt und deren Auswahl für die vorliegende Arbeit begründet.

### 2.3.3.1 Maskulinität Positiv

Bezüglich der (positiven) Maskulinität wurde in der Vergangenheit – neben klassischen Konstrukten wie kognitiven Fähigkeiten – unter anderem wiederholt die Beziehung zur Leistungsmotivation geknüpft. Obwohl Orlofsky und Stake (1981) den Personal Attribute Questionnaire als Messinstrument für Geschlechtsrollen einsetzten, konnten sie nachweisen, dass beispielsweise Maskulinität – und nicht etwa das biologische Geschlecht – zu Unterschieden in der Leistungsmotivation führte. Sie fanden, dass höhere berichtete Maskulinität einher ging mit größerer Leistungsmotivation und geringerer Furcht vor Misserfolg. Maskulinität konnte somit zwischen Leistung und wahrgenommenen kognitiven Fähigkeiten diskriminieren (Orlofsky & Stake, 1981).

Sieverding (1999) kam zu ähnlichen Schlussfolgerungen, derart, dass bei maskulinen Personen – unabhängig vom Geschlecht – Anpassungsfähigkeit, Kompetenz, Kreativität, Flexibilität und Unkonventionalität am stärksten ausgeprägt sind. Spence (1980) griff das Leistungskonzept in größerem Umfang auf und konnte nicht nur zeigen, dass Männer wettbewerbsorientierter sind, sondern zudem auch, dass die maskuline Teilskala die stärkeren signifikanten Beziehungen zu diesen Leistungsaspekten aufwies (z. B. *mastery*, *competitiveness*). Andere Studien fanden selbst bei ausschließlicher Betrachtung von Frauen als Versuchspersonen ebenfalls den größeren Bezug von Maskulinität und Leistungsmotiven (Marshall & Wijting, 1980).

Die Beziehung der Maskulinität zu leistungsbezogenen Aspekten wie Orientierung und Motivation impliziert auch eine Korrelation von Maskulinität zu Komponenten des AVEMs (Schaarschmidt & Fischer, 1996). Alle drei zum Einsatz gekommenen Instrumente bilden in ihrem Kern im weitesten Sinne Komponenten des Leistungsmotivs und der -orientierung ab.

Alle diese Konzepte werden Männern im stärkeren Maße zugeschrieben. Da es sich um positive Aspekte der Selbstbeschreibungen handelt, werden sie mit positiver Maskulinität in Beziehung gebracht, was sie für den Einsatz in der vorliegenden Studie prädestiniert.

#### *Beruflicher Ehrgeiz (AVEM)*

Beim AVEM (Arbeitsbezogenes Verhaltens- und Erlebensmuster) handelt es sich um ein Instrument zur interventionsbezogenen Diagnostik beruflichen Bewältigungsverhaltens (Schaarschmidt & Fischer, 1996). *Beruflicher Ehrgeiz* ist eine der 11 Teilskalen und nachweislich signifikant bedeutsamer für Männer als für Frauen, wodurch ebenfalls ein Zusammenhang zu positiver Maskulinität angenommen werden kann.

#### *Hoffen auf Erfolg (AMS)*

Die Achievement Motive Scale (AMS), ursprünglich von Gjesme und Nygard (1970) entwickelt und eine Dekade später von Göttert und Kuhl (1980) ins Deutsche übersetzt, misst neben der Hoffnungskomponente (HE), die Furchtkomponente (FM) als einen Kern des Leistungsmotivs.

#### *Leistungsorientierung (FPI)*

Im Zuge der Validitätsüberprüfung beim AVEM haben Schaarschmidt und Fischer (Schaarschmidt & Fischer, 1996) *beruflichen Ehrgeiz* mit dem Konzept der Leistungsorientierung, gemessen mit dem FPI (Fahrenberg, Hampel, & Selg, 2001), in Beziehung gesetzt und eine positive Korrelation ( $r = .35$ , S. 18) nachgewiesen. Die gemeinsame Varianz beider Konzepte impliziert eine ähnliche Beziehung beider Skalen zur positiven Maskulinität. Der Bereich der Leistungsorientierung erfasst individuelle Unterschiede der Leistungsorientierung im beruflichen und außerberuflichen Bereich; macht dabei aber keine Aussagen über direkte, aufgabenbezogene Leistungsmotivation oder tatsächliche Leistungsfähigkeit. Die Leistungsorientierung kontrastiert aktives, schnelles und ehrgeiziges Handeln mit weniger energischer und konkurrierender Ausprägung. Probanden mit hohem Skalenwert gelten als insgesamt leistungsorientierter, denen Konkurrenz nichts ausmacht und die Aufgaben energisch angehen und effizient bewältigen. Demgegenüber präsentieren sich Probanden mit geringer Ausprägung auf diesem Konzept als weniger ehrgeizig und eher passiv, denen beruflicher Erfolg und Anerkennung nicht derart bedeutsam sind. Die Skala ist geringfügig vom Geschlecht abhängig mit der Tendenz einer höheren Leistungsorientierung bei Männern.

#### **2.3.3.2 Maskulinität Negativ**

Negative Maskulinität weist Parallelen auf zur Unmitigated Agency, einem Konstrukt, das durch den Fokus auf das Selbst bis hin zum Ausschluss anderer charakterisiert ist (Bakan, 1966; Helgeson, 1994b). Personen mit hoher Ausprägung negativer Maskulinität legen mit höherer Wahrscheinlichkeit weniger Wert auf soziale Beziehungen und sind durch eine reduzierte Bereitschaft zu unterstützen und zu helfen charakterisiert. Eigene Bedürfnisse werden über die

anderer gestellt und das Durchsetzen dieser erfolgt um nahezu jeden Preis. Bereitwilligeres aggressives Agieren und das Verlangen, nach vorn zu streben und Macht über Situation und Personen zu haben, kennzeichnet diese Menschen.

#### *Aggressivität (AQ) / Aggressivität (FPI)*

Bezüglich möglicher Validierungskonstrukte für negative Maskulinität soll auch in dieser Studie die Trait-Aggressivität herangezogen werden, von der aus früheren Studien bekannt ist, dass männliche Probanden eine höhere Ausprägung aufweisen als Frauen (Möller, 2006). Es gibt zahlreiche Studien, die höhere Aggressionswerte (Brigham, 1986; Eagly & Steffen, 1986; Maccoby & Jacklin, 1974) oder Ärgerwerte (Holahan & Spence, 1980) bei Männern berichten, dies vor allem stärker bei physischer als bei verbaler Aggression. Evans und Dinning (1982) kamen beim Einsatz von MMPI-Korrelaten mit sowohl BSRI als auch dem PAQ zu ähnlichen Schlussfolgerungen. Während sich in ihren Untersuchungen mit ausschließlich männlichen psychiatrischen Patienten hoch maskuline Männer als frei von Neurotizismus und geringem Selbstwert beschrieben, wiesen sie gleichzeitig eine erhöhte Feindseligkeit mit aggressiven Tendenzen auf.

In einem ähnlich aktuellen Skalenkonstruktionsvorgang von Geschlechtsrollen für den Einsatz bei Kindern und Jugendlichen konnte ebenfalls der stärkere Zusammenhang der bereits neu entwickelten negativen maskulinen Teilskala (im Vergleich zu beiden femininen Teilskalen und vor allem auch der positiven maskulinen Teilskala) zu Trait-Aggressivität nachgewiesen werden (Krahé, Berger, & Möller, 2007). In dieser Studie haben die Autoren eine veränderte, hoch reliable Version des Aggression Questionnaire (Buss & Warren, 2000), der zwischen verbaler, relationaler und physischer Aggression sowie Ärger unterscheidet, zum Einsatz gebracht.

Die nachweislich bei Männern höher ausgeprägte Aggressivität wird – aufgrund ihrer negativen Valenz – mit der negativen maskulinen Teilskala in Beziehung gesetzt und neben der von Möller (2006) entwickelten Aggressivitätsskala auch durch die entsprechende Teilskala des FPI erfasst. Der Bereich der Aggressivität (FPI) (Fahrenberg et al., 2001) misst Aspekte des aggressiven Verhaltens und kontrastiert dabei spontanes und reaktives Verhalten, sich durchsetzen mit wenig aggressiven, kontrollierenden und zurückhaltenden Verhaltensweisen. Dementsprechend zeigen Probanden mit hohem Skalenwert eine Bereitschaft zu aggressiver Durchsetzung ihrer Interessen, was sich u.a. darin äußern kann, dass sie andere eher bloßstellen, sich schneller aktiv zur Wehr zu setzen und sogar körperliche Gewalt einsetzen. Obwohl der Skalenwert nur geringfügig mit dem Geschlecht zusammenhängt, zeigte sich, dass vor allem jüngere Männer die höchsten Werte erzielten, was die Skala für den Einsatz in dieser Studie prädestiniert.

#### *Werteskala Macht*

In früheren Studien zeigten Männer höhere Werte als Frauen im Bereich Rache / Vergeltung / Macht (Stuckless & Goranson, 1992), einem Konstrukt, das verwandt ist mit Aggressivität.



Diese Ergebnisse sind in Übereinstimmung mit Untersuchungen, die einen Schritt weitergingen und über den klassischen Geschlechtsunterschied hinweg nicht nur den Zusammenhang zu Geschlechtsrollen herstellten, sondern darüber hinaus die unterschiedliche Valenz möglicher Teilskalen in Betracht zogen und zeigen konnten, dass beispielsweise Ärger positiv mit negativer Maskulinität korreliert ist (Holahan & Spence, 1980). Im Zusammenhang mit Macht ist eine hohe Merkmalsausprägung Ausdruck einer Führungspersönlichkeit, die andere zu beeinflussen versucht. Dem gegenüber neigen Personen mit niedrigerer Merkmalsausprägung zu weniger Kontrolle und Beeinflussung anderer. Die Wichtigkeit bestimmter machtbezogener Werte (und zur Validierung von Mas-) für die Probanden wird mittels eines an den Schwartz Value Survey (Schwartz, 1992) angelehnten Fragebogens von Schröder (2005) erfasst.

### **2.3.3.3 Femininität Positiv**

Femininität gilt als eng assoziiert mit sowohl positiven als auch negativen emotionalen Inhalten (Hunt, 1993; Sieverding, 1999). Feminine Eigenschaften gehen mit höherem sozialen Selbstwert sowie größerer Zufriedenheit in und mit interpersonalen Beziehungen, dem Bedürfnis nach Bestätigung, Liebe oder auch Furcht vor Misserfolg einher (Orlofsky & Stake, 1981).

#### *Soziale Werte*

Da positive Femininität starke Ähnlichkeit mit Communion hat, kann man auch diesem Konstrukt einen Fokus auf andere (Bakan, 1966; Helgeson, 1994b) unterstellen, der dem Wohl anderer und der Zeit im sozialen Miteinander einen großen Stellenwert einräumt. Individuen mit hoher Ausprägung werden mit höherer Wahrscheinlichkeit Eigenschaften zeigen, die sie im sozialen Leben Bedürfnisse ihrer Mitmenschen erkennen und angemessen begegnen lassen. Da das Gestalten interpersonalen Beziehungen mit positiven Werten belegt ist, wird bei den berücksichtigten Konzepten der sozialen Werte, Empathie und Fürsorglichkeit jeweils auch die stärkste Beziehung zu der positiven femininen Teilskala erwartet. Eine hohe Ausprägung positiver femininer Eigenschaften sollte dementsprechend einhergehen mit hoher Fürsorglichkeit und Empathie sowie stark ausgeprägten sozialen Werten. Die Ausprägung sozialer Werte und damit die Berücksichtigung der Bedürfnisse anderer wird durch den an den Schwartz Value Survey (Schwartz, 1992) angelehnten Fragebogen von Schröder (2005) erfasst.

#### *Empathie (Sorge)*

Zahlreiche Studien konnten höhere Empathiewerte für Frauen belegen (Eisenberg & Lennon, 1983). Zusätzlich konnte in Bezug auf Empathie und Fürsorglichkeit in verschiedenen Studien bereits die im Vergleich zum Geschlecht stärkere Beziehung zur femininen Geschlechtsrolle (Karniol, Gabay, Ochion, & Harari, 1998; Krahe et al., 2007) hergestellt werden.

Trotz der Popularität des Konzepts Empathie und seiner Bedeutung für das soziale Verhalten stellt die Erfassung aufgrund der Komplexität dieser Komponente eine Herausforderung dar.

Daher kann es bei der Messung von Empathie weniger um ein Gesamtkonstrukt gehen, sondern lediglich um die Darstellung eines Teilkonzepts. Ein häufiger methodischer Ansatzpunkt in der Empathiemessung ist der Einsatz sogenannter Fallvignetten, die jedoch in der vorliegenden Studie aufgrund von Befragungsökonomie und der möglichst gleichbleibenden Itemdarstellung ausgeschlossen werden sollten. In der vorliegenden Studie wird daher ein von Maes, Schmitt und Schmal (1995) entwickelter Fragebogen zum Einsatz kommen.

#### *Fürsorglichkeit (PRF)*

Bem (Bem et al., 1976) selbst hatte Fürsorglichkeit mit Femininität ihres Instrumentes in Beziehung gebracht und auch andere Studien konnten diese positive Korrelation nachweisen (Lippa, 1997); je höher die Femininität, desto höher die Fürsorglichkeit. Diese positive Korrelation sollte gerade auch erhalten bleiben bei der Beschränkung der Vorhersage auf positive Femininität. In der vorliegenden Studie wird Fürsorglichkeit mit der Teilskala aus der Personality Research Form (PRF) von (Jackson, 1974; Stumpf, Angleitner, Wieck, Jackson, & Beloch-Till, 1985) erfasst.

#### **2.3.3.4 Femininität Negativ**

Sieverding (1999) stellte über mehrere Studien hinweg nicht nur Geschlechterunterschiede im Gesundheitsbereich fest, sondern vor allem auch, dass Frauen aus westlichen Industrieländern häufiger an psychischen Krankheiten leiden oder auch häufiger therapeutische Hilfe in Anspruch nehmen. Studien im Gesundheitsbereich, in denen Probanden einen gesunden Mann oder eine gesunde Frau beschreiben sollten (Beckwith, 1993; Broverman et al., 1970), konnten beispielsweise zeigen, dass Unterschiede zwischen Männern und Frauen im Gesundheitsverhalten bestehen, die teilweise durch Geschlechterstereotype erklärt werden können. So fasste Helgeson (Helgeson, 1994a) beispielsweise Studien zusammen, die einen Zusammenhang zwischen Femininität und Konzepten wie Stress oder Einsamkeit nachweisen konnten.

Während die Ergebnisse hinsichtlich der Frage, ob nun ein größeres Ausmaß an Maskulinität oder das an Femininität für die Wahrnehmung einer gesunden Person ausschlaggebend sind oder welche Ausprägung eher mit ungesunden Verhaltensweisen in Bezug zu bringen sind, nicht eindeutig sind (Sieverding, 2002), zeigen sie die Relevanz der Geschlechtsrolle (für eine Übersicht siehe Danoff-Burg et al., 2004) bei der Beantwortung eben dieser Frage. Ausgehend von den Androgynitätsansätzen der 1970er Jahre ist die bedeutende Rolle der maskulinen und femininen Geschlechtsrolle in diesem Zusammenhang belegt (Bem, 1974; Spence et al., 1975) und bedarf lediglich einer Anpassung der Erwartungen an negative Teilskalen.

Depression gehört zu einem populären Erkrankungsmuster in diesem Zusammenhang, da in der Literatur bereits konsistent und vielfach ein Geschlechtsunterschied nachgewiesen werden konnte, im Sinne einer durchschnittlich stärkeren Symptombelastung von Frauen (Beck, Steer, & Garbin, 1988; Dalgard et al., 2006; Schmitt et al., 2003; Schmitt, Altstötter-Gleich, Hinz, Maes, & Brähler, 2006; Schmitt & Maes, 2000; Sieverding, 1999). Grundsätzlich kann davon

ausgegangen werden, dass Frauen im Schnitt höhere Depressionswerte als Männer aufweisen; das Geschlecht selbst konnte jedoch lediglich 1.5% der Varianz des Depressionswertes erklären (Schmitt et al., 2003; Schmitt et al., 2006; Schmitt & Maes, 2000), was die Bedeutsamkeit der Geschlechtsrollen unterstreicht.

Da allgemein die Gefahr einer Über- vs. Unterschätzung durch Selbstberichte wie dem BDI diskutiert wird (Schmitt et al., 2006), soll kein klassisches Depressionsinventar zum Einsatz kommen, sondern inhaltlich verwandte Konstrukte.

Helgesons (1994b) Nachweis, dass Unmitigated Communion eine Eigenschaft ist, die als typischer für Frauen gilt und nachweislich einen wichtigen Einfluss auf die (psychische) Gesundheit hat, lässt die Schlussfolgerung zu, dass mit einer neuen negativen femininen Teilskala eine positive Beziehung zu Konstrukten wie depressive Verstimmung zu erwarten ist. Die negative Valenz der Teilskala ist der der Depressivität ähnlich und sollte den Zusammenhang stärken. Unterstützt wird diese Annahme durch Ergebnisse von Craighead und Green (1989), die in ihrer Studie zur Beziehung von depressiver Stimmung, Selbstbewusstsein und geschlechtstypisierten Persönlichkeitseigenschaften bei Kindern zeigen konnten, dass negative feminine Eigenschaften hoch mit Depression korreliert war.

#### *Ängstliche Depressivität (BSF)*

Ängstliche Depressivität als ein Aspekt aus der Weiterentwicklung (Hörhold & Klapp, 1993) des Mehrdimensionalen Stimmungsfragebogens (BSF) von Hetcheltjen und Mertesdorf (1973) wird zur Validierung der femininen negativen Teilskala eingesetzt.

#### *Trait-Ängstlichkeit*

Der wiederholt hergestellte Bezug von Depression und (negativer) Femininität und den Versuchen, derartige Ergebnisse u. a. mit Ausprägung von Angst zu erklären, macht die Trait-Ängstlichkeit zu einem sehr guten Validierungskonstrukt für negative Femininität. So ist beispielsweise aus den Normdaten des STAI, in der deutschen Fassung von Laux, Glanzmann, Schaffner und Spielberger (1981), bekannt, dass Frauen sowohl höhere Trait- als auch State-Ängstlichkeit berichten als Männer dies tun. In Bezug auf die Geschlechtsrolle konnte zudem eine signifikante Korrelation zwischen Femininität und Angst nachgewiesen werden (Gall, 1969). Erneut würde man diesbezüglich erwarten, dass die jeweils negativen Aspekte der Angst und der Femininität zu positiven signifikanten Korrelationen in der vorliegenden Studie führen würden. Zur Erfassung der Trait-Ängstlichkeit wurde die Trait-Skala des State-Trait-Angstinventar (STAI) in der deutschen Fassung von Laux, Glanzmann, Schaffner und Spielberger (1981) eingesetzt.

#### *ESWS*

In Bezug auf Selbstwert und Geschlechtsrollen konnten Spence et al. (1975) eine signifikante (positive) Korrelation zwischen Femininität und Selbstwert (*self-esteem*) nachweisen – sowohl für Männer als auch für Frauen. Damit lieferten Spence und Kollegen den Beweis, dass Femininität und Selbstwert ganz allgemein eine Beziehung zueinander aufweisen. Marsh, Antill

und Cunningham (1987) fanden dagegen eine negative Korrelation zwischen Femininität und Selbstwert: je höher Femininität ausgeprägt ist, desto geringer fällt der berichtete Selbstwert aus. Somit eröffnen die Autoren auch die Möglichkeit der Vorhersage eines Zusammenhangs von negativer Femininität und Selbstwertschätzung, derart, dass hohe negative Femininität einher geht mit geringer Selbstwertschätzung.

Diese von Schütz und Sellin (2006) entwickelte Skala der emotionalen Selbstwertschätzung umfasst den evaluativen Aspekt selbstbezogener Einstellungen und orientiert sich grundsätzlich am international gebräuchlichen Begriff des *self-esteem*s.

## **2.4 GSK als Erwartungs-mal-Wert-Konstrukt**

In Bezug auf das Geschlechtsrollenselbstkonzept operationalisierten bisherige Arbeiten die Selbstbeschreibung entweder ausschließlich auf Grundlage der Typizität oder der sozialen Erwünschtheit. Frauen und Männer schreiben sich maskuline und feminine Eigenschaften zu, die als typisch oder sozial erwünscht für ihre kulturelle Gruppe gelten. Dabei wurde in unzähligen Arbeiten diskutiert, dass sich Geschlechtsattribute (unabhängig von einer sorgfältigen Skalenkonstruktion) in ihrer kulturbasierten Valenz unterscheiden. Gerade das Wissen um kulturell geformte und geteilte Stereotype kann nicht nur zu sozial erwünschtem Verhalten führen, sondern unterstreicht die mögliche Bedeutsamkeit einer zusätzlich erfassten, individuellen Wertekomponente, wie sie im Falle des Erwartungs-mal-Wert-Modells in der Einstellungsforschung bereits Berücksichtigung fand.

### **2.4.1 Ursprung aus der Motivations- und Einstellungsforschung**

Seit den 1940er Jahren haben sich Erwartungs-mal-Wert-Theorien in der Motivationsforschung entwickelt, deren grundlegende Idee das Prinzip der Nutzenmaximierung darstellt. Dabei ist der Nutzen repräsentiert durch ein Produkt aus erzielbarem Wert (Anreiz) und der Wahrscheinlichkeit, ein bestimmtes Ergebnis zu erzielen (Erwartung). Die Kombination aus hohem Anreiz und hoher Erwartung das Ziel durch eigenes Handeln zu erreichen maximiert die Handlungsintention (Heckhausen, 1989). Die Erfolgswahrscheinlichkeit (= die Einschätzung der Wahrscheinlichkeit) eine bestehende Situation durch eigenes Handeln so zu verändern, dass ein erwünschtes Ergebnis eintritt, findet sich in den Arbeiten zum erweiterten kognitiven Motivationsmodell von Heckhausen (Heckhausen, 1977a; Heckhausen, 1977b). Auch Eccles et al. (1983) berücksichtigten in einem Modell zur Vorhersage leistungsbezogener Verhaltensweisen eine Erwartungs- (Erfolgserwartung) und eine Wertekomponente (objektive und subjektive Wertschätzungen).

Die aus der Motivationspsychologie etablierten Konzepte von Erwartung und Wert wurden schließlich von Fishbein und anderen in die Einstellungsforschung integriert (Ajzen, 2001; Ajzen & Fishbein, 2005; 2008; Feather, 1982; Feather & O'Brien, 1987; Fishbein, 1963; 2007; Fishbein & Ajzen, 1975; LaPiere, 1934; Mazis, Ahtola, & Klippel, 1975; Rosenberg, 1956; Rotter, 1954; Tolman, 1932; Viklund & Sjöberg, 2008; Vroom, 1964). Die Theorie von Fishbein und Ajzen (1975) definierte die Einstellung zu einem Objekt über die Erwartung (subjektive Wahrscheinlichkeit), dass ein Einstellungsobjekt eine bestimmte Eigenschaft besitzt, und dem

Wert, dem das Individuum dieser Eigenschaft zumisst. So lässt sich eine positive Einstellung zum Zähneputzen vorhersagen durch das Produkt (a) der Erwartung, dass regelmäßiges Zähneputzen die Zahngesundheit fördert und (b) der Stärke der positiven Bewertung der Zahngesundheit im Einstellungssystem des Individuums. Abbildung 8 stellt diese Beziehung dar.

$$A = \sum_{i=1}^n b_i e_i$$

Abbildung 8. Formel des Zusammenhangs von Einstellung, Bewertung und Erwartung nach Ajzen und Fishbein (Ajzen, 1985; Fishbein, 1967)

Das Erwartungs- x Wert Modell gehört zu den wohl wichtigsten kognitiven Ansätzen in der Motivationsforschung, welche auf individuellen Glaubenssystemen basieren und Erwartungen und Bewertungen seitens des Individuums berücksichtigen. Dennoch ist es in seiner Funktionsweise und Anwendung nicht unumstritten, da der Zugewinn durch die multiplikative Verknüpfung (Interaktionsterm) bisweilen unter den von Fishbein und Ajzen ursprünglich hypothetisierten Möglichkeiten bleibt (Feather & O'Brien, 1987; Viklund & Sjöberg, 2008). Viklund und Sjöberg (2008) stellten in ihrer Studie fest, dass eine multiplikative Verknüpfung von Überzeugungen und Wertzuschreibung im Vergleich zur separaten Verwendung beider Summenwerte wenig zur Erklärung beiträgt.

#### 2.4.2 Anwendung des Erwartungs-mal-Wert-Modells auf das GSK

Die Idee, das Erwartungs-mal-Wert-Modell im Rahmen der Geschlechterforschung einzusetzen, ist nicht gänzlich neu. Eagly und Mladinic (1989) ließen Versuchspersonen in einem ersten Schritt fünf Eigenschaften einer Zielperson auflisten. Im Anschluss sollten die Probanden dann nicht nur die Anzahl der Personen in der Zielgruppe (Frauen, Männer, politische Gruppen) schätzen, die diese Eigenschaft besitzen, sondern auch diese Eigenschaft auf einer Skala von gut bis schlecht bewerten. Die Autoren konnten zeigen, dass die Einstellung zu Frauen und Männern sehr stark mit der Bewertung der Geschlechterstereotype zusammenhing und Stereotype von und Einstellungen zu Frauen insgesamt positiver waren.

Dementsprechend kann das Geschlechtsrollenselbstkonzept durch eine Summe aus Produkten der geschlechtsbezogenen Selbstbeschreibung sowie der auf Itemebene zugehörigen Wichtigkeit eben dieser Zuschreibung repräsentiert werden. Da über die Valenz der einzelnen Attribute übereinstimmendes, geteiltes Wissen existiert, welches lediglich bekannte Werte aus der Messung der sozialen Erwünschtheit zur Itemkonstruktion widerspiegeln würde, wird das neue Produkt durch die Multiplikation der Selbstzuschreibung (GR = gender role) mit der individuellen Wichtigkeit der zuvor vorgenommenen Selbstzuschreibung realisiert ( $I$  = importance;  $a$  = die einzelnen Aspekte / Adjektive, Abbildung 9).

$$GRI = \sum_{a=1}^n GR_a I_a$$

Abbildung 9. Mögliche Formel der Geschlechterrolle als Funktion der eigenen Selbstzuschreibung in Abhängigkeit der persönlichen Bedeutsamkeit dieser Zuschreibung

Die Adaption der Einstellungsformel für die Messung des Geschlechtsrollenelbstkonzepts ähnelt dem *Adequacy-Importance Model* von Cohen, Fishbein, und Ahtola (1972) aus der Konsumentenforschung, welches beispielsweise fragt: "*indicate the degree of importance you attach to each of these attributes when buying a particular product*" (Mazis et al., 1975; S. 43).

Mit dem erweiterten Erwartungs-mal-Wichtigkeit Ansatz soll nun ein zusätzlicher Aspekt der persönlichen Wertkomponente implementiert werden. Der Logik des ursprünglichen Ansatzes folgend ist beispielsweise die Stärke der Zuschreibung der femininen Geschlechtsrolle als multiplikative Verknüpfung von Selbstzuschreibung und Wichtigkeit definiert: die feminine Geschlechtsidentität ist demnach um so stärker ausgeprägt, je stärker eine Person sich weibliche Attribute zuschreibt und je wichtiger (im Falle positiver Attribute) und weniger wichtig (im Falle negativer Attribute) sie diese für sich selbst einschätzt. Anders ausgedrückt ist das positive feminine GSK (und entsprechend das männliche GSK) umso stärker ausgeprägt, je höher das Produkt aus Selbstzuschreibung und Wichtigkeit bei den positiven femininen Items und je niedriger das Produkt aus Selbstzuschreibung und Wichtigkeit bei den positiven maskulinen Items ist. Analog ist das negative feminine GSK umso stärker ausgeprägt, je höher das Produkt aus Selbstzuschreibung und Wichtigkeit der negativen femininen im Verhältnis zu den maskulinen negativen Attributen ist. Wenn schon die Zuschreibung negativer Items erfolgt, dann werden die für das eigene Geschlecht typischen "schlechten Eigenschaften" eher in das GSK integriert als die für das andere Geschlecht typischen. In Bezug auf den Erwartungs-mal-Wert-Ansatz sollte die Geschlechtsidentität als Produkt aus Selbsteinschätzung und individueller Wichtigkeit ausgewählte Validierungskonstrukte besser vorhersagen als die einfache Selbstzuschreibung.

### **2.4.3 Bedeutung des Produktes**

#### *Erfassung der Wichtigkeit*

Außer der Definition und Interpretation von Wichtigkeit gibt es verschiedene Möglichkeiten diese zu operationalisieren. Jaccards, Brinberg und Acherman (1986), mit einem Fokus auf Konsumentenforschung, verglichen sechs elaborierte Methoden der Erfassung von Wichtigkeit in einer multi-traitmethod Matrix und identifizierten den Einsatz direkter Einschätzung der Wichtigkeit im Vergleich als insgesamt bevorzugt gegenüber allen anderen Methoden.

Es existiert eine Reihe von weiteren Studien, die ebenfalls die Bedeutung der Produktkalkulation in den Zusammenhang mit ursprünglich hypothetisierten Modellen bringen und das Plädoyer für die kombinierte Verwendung von bipolaren (Evaluation) und unipolaren (Überzeugung) Skalen halten, da sie die theoretischen Modelle am besten repräsentieren (Ajzen, 1991; Ajzen & Fishbein, 2008; Hewstone & Young, 1988; Smith, 1996; Sparks, Hedderley, & Shepherd, 1991). Sparks et al. (1991) diskutierten darüber hinaus generell die Berücksichtigung der Beziehung von Zustimmung und Ablehnung und von Ergebnis und Nicht-Ergebnis bei den Kalkulationen und zukünftigen Forschungen.

Ajzen und Fishbein (2008) beteiligten sich an der Debatte um die Methodik hinter der Erwartungs-mal-Wertverknüpfung und unterstrichen den Zugewinn an Varianzaufklärung durch das Produkt bei der Vorhersage von Verhalten mittels einer multiplen Regression (Blanton &

Jaccard, 2006); zumindest in den Fällen, in denen die Messungen aller beteiligten Variablen alle Ausprägungskombinationen ermöglichen und erfassen (und negative Werte der bipolaren Skalen nicht wie in einigen anderen Studien vor der Produktbildung rekodiert werden). Dieses Vorgehen, möglicherweise in Kombination mit der nicht unumstrittenen Ausrichtung der Prädiktoren am Mittelwert (Aiken & West, 1991), würde zusätzlich das Problem der Multikollinearität reduzieren.

Die Überlegungen zu dem möglichen Einfluss einer zusätzlich erfassten individuellen Wichtigkeit von geschlechtsbezogenen Eigenschaften im Zuge der Selbstbeschreibung sollen im Rahmen der vorliegenden Arbeit in einer abschließenden Studie explorativ geprüft werden.

## 2.5 Zusammenfassung

Unabhängig von ungenauer Definition und Diskussion um mögliche unzureichende methodische Abbildung der zugrunde liegenden Multidimensionalität, gelten Geschlechterstereotype als akzeptiert, beständig und traditionell. Aus der diskutierten Theorie kann die Bedeutung der Geschlechtsrollen für das Selbstkonzept hergeleitet werden:

- (1) Demnach haben Menschen grundsätzlich ein Selbstkonzept, das relativ stabil, aber anpassungsfähig ist.
- (2) Die Kategorie Geschlecht stellt in diesem Zusammenhang einen zentralen Bestandteil der individuellen Selbstkonstruktion dar.
- (3) Das Geschlechtsrollenselbstkonzept entwickelt sich im Laufe der Sozialisation in Auseinandersetzung mit den vorherrschenden Vorstellungen der umgebenden Kultur.
- (4) Das GSK steht in Beziehung zu Affekten, Kognitionen und Verhaltensweisen in einer Vielzahl von Bereichen.
- (5) Das GSK umfasst nicht nur ausschließlich positive Aspekte, sondern beinhaltet auch die Auseinandersetzung mit negativen Attributen.
- (6) Dabei lässt sich auf der Basis identitätstheoretischer Modelle aus dem Intergruppenkontext herleiten, dass den geschlechtskongruenten Merkmalen mehr Gewicht beigemessen wird als den geschlechtsinkongruenten.

Die bisher vorliegenden Instrumente zur Messung des GSK weisen eine Reihe von Problemen auf: Zum einen ist in den häufigsten Fällen die Aktualität der Instrumente nicht geprüft. Zum anderen berücksichtigen viele Maße die negativen Attribute des Geschlechtsrollenselbstkonzepts entweder nicht oder ohne ausreichende theoretische Fundierung.

Obwohl die beiden positiven Aspekte des GSK wiederholt mit anderen Verhaltens- oder Persönlichkeitsmerkmalen in Zusammenhang gebracht werden konnten, besteht Konsens darüber, dass die positiv valenzierten Attribute bestehender Messinstrumente die Beziehungen zu anderen Konstrukten oft jedoch nur unbefriedigend zu erklären vermögen. Während dafür bisweilen die fehlende Multidimensionalität der Messinstrumente verantwortlich gemacht wurde, was unter Umständen sogar zur vorgeschlagenen Abkehr von den Konzepten der Femininität und Maskulinität führte, basiert die vorliegende Arbeit auf der Überzeugung, dass eine theoretische Fundierung der Einbeziehung negativer Attribute des GSK sowie die Entwicklung

eines entsprechenden Messinstruments einen wesentlichen Beitrag zur Weiterentwicklung der Forschung im Bereich des GSK leisten kann.

Aus den dargestellten Forschungsergebnissen der Bereiche Selbst und Gruppe weiß man um den komplexen Identitätsbildungsprozess und die Bedeutung auch negativer Charakteristiken. Sowohl der Bereich des Stereotype-Threats als auch Studien zum allgemeinen Selbstkonzept arbeiteten die Relevanz negativer Aspekte heraus. Aktivierte Stereotype, die sich auf die eigene Gruppe beziehen, entfalten eine stärkere Wirkung als die Aktivierung von Stereotypen, die sich auf Fremdgruppen beziehen. Daraus ergibt sich, dass auch negative Eigenschaften, die mit dem eigenen Geschlecht über Typizität oder Valenz assoziiert sind, für die Selbstkonstruktion bedeutsamer sind als Eigenschaften, die mit dem anderen Geschlecht assoziiert sind. Das Geschlechtsrollenselbstkonzept stellt in diesem Zusammenhang keine Ausnahme dar.

Während die theoretische Verknüpfung der Ergebnisse aus der Gruppenforschung zu Geschlechtsrollenstereotypen noch nicht ausreichend erfolgt ist, hat die Forschung zur Messung des GSK einige dieser negativen Attribute in Bezug auf das soziale Geschlecht bereits in ihren Instrumenten berücksichtigt. Vor dem Hintergrund der kulturellen Veränderungen in den letzten Dekaden und der dadurch eingeschränkten Stabilität des zu erfassenden Geschlechtsrollenselbstkonzepts über die Zeit ist eine Überarbeitung der theoretischen Herleitung ebenso wie die Erstellung eines neuen deutschsprachigen Instruments notwendig. Dabei gilt es nicht nur, Theorien aus der Selbst-, Identitäts-, Geschlechts- und Gruppenforschung zu berücksichtigen, sondern auch verschiedene Ansätze zur Erstellung von GSK-Instrumenten kritisch in Betracht zu ziehen und daraus abgeleitete Verbesserungsansätze im Zuge der Neu-Entwicklung umzusetzen.

## 2.6 Fragestellung und Hypothesen

In der vorliegenden Arbeit wird eine Reihe von aufeinander aufbauenden Studien präsentiert. Zunächst wird in einem ersten Schritt der deutsche BSRI einer Überprüfung unterzogen, um Schwächen der Items ausschließlich positiver Valenz aufzudecken und die Aktualität der unterstellten geschlechtstypischen Valenz der verwendeten Attribute empirisch zu überprüfen (Studie 1 und 2).

In einem zweiten Schritt wird eine neue Skala entwickelt, die diese Schwächen durch einen ausführlichen Konstruktionsprozess adressiert und vor allem von Beginn an auch negative Selbstbeschreibungen berücksichtigt, um der Komplexität des geschlechtlichen Selbst gerecht zu werden (Studie 3 und 4). Dabei soll Studie 3 im Konstruktionsprozess der Generierung einer Ausgangsstichprobe von Eigenschaften für Männer und Frauen dienen, die jeweils typischer für das eine Geschlecht sind als für das andere und sich zusätzlich hinsichtlich ihrer Valenz unterscheiden (positiv, negativ). In Studie 4 wird aufgrund der Einschätzungen zur Typizität und sozialen Erwünschtheit (wie es in anderen Konstruktionsstudien umgesetzt wurde) sowie mit ersten Resultaten aus der Selbstbeschreibung die Auswahl der Items für die Teilskalen vorgenommen. Die Einschätzung der Typizität wird zusätzlich gewählt, weil die



ausschließliche Bewertung der sozialen Erwünschtheit den stärkeren Bezug zur eigenen Identität herstellt, wodurch vorgenommene Einschätzungen eine größere Gefahrenquelle für die Herabsetzung des eigenen Selbstwertes darstellen und die Urteile beeinflussen können.

Schließlich werden in zwei weiteren Studien (Studie 5 und 6) die vier entwickelten Teilskalen des neuen GSK-Inventars einer Validierung unterzogen. Jeder der Teilskalen werden theoriegeleitet spezifische Konstrukte zugeordnet. Im Ansatz des destruktiven Testens (Anderson & Anderson, 1996) soll der Nachweis angetreten werden, dass alle Teilskalen ihren eigenen Beitrag zur Vorhersage psychologischer Konzepte leisten können.

Der Schwerpunkt dieser Arbeit liegt damit auf dem Entwicklungsprozess des neuen Inventars. In bisherigen Studien wurde die Auswahl von möglichen Items für ein GSK Instrument verschiedenartig vorgenommen: so gibt es Beispiele, die sich an schon existierenden Skalen orientierten und sich bewährter Items bedienten, während andere Studien Übersetzungen ursprünglich englisch publizierter Skalen vornahmen. Gerade der Übersetzungsprozess vernachlässigt dabei mitunter die in der Zielgruppe vorherrschenden kulturellen Besonderheiten, deren relevanter Einfluss in der Geschlechtsrollenforschung jedoch bekannt ist. Die theoretische Fundierung von GSK Instrumenten auf Stereotype bezüglich Männern und Frauen in der zu untersuchenden Gruppe macht daher den Rückgriff auf Eigenschaften, die in dieser Gruppe als aktuell geschlechtstypisch gelten, für die Itemgenerierung erforderlich. Dieses Argument gewinnt um so mehr an Plausibilität, betrachtet man die Dauer des Bestehens einiger Instrumente. So können die Veränderungen von Femininität und Maskulinität nur noch einschränkt durch GSK Instrumente, deren Entwicklung auf stereotype Vorstellungen von Männer und Frauen von vor 30 Jahren basieren, abgebildet werden. Aus diesem Grund wird eine grundlegende Neugenerierung von geschlechtsspezifischen Eigenschaften angestrebt, die sowohl den aktuellen Vorstellungen davon, was typisch männlich und weiblich ist entsprechen, als auch den aktiven Wortschatz in der heutigen Zeit repräsentieren. Klapprott (1994) wies in seinen Ausführungen bereits auf die Bedeutung der Alltagssprachlichkeit zu verwendender Eigenschaftsbegriffe hin sowie auf die Notwendigkeit, Repräsentativität, Soziale Erwünschtheit und Bedeutungshaltigkeit (Vergleich von Synonymen und Antonymen) der Adjektive zu überprüfen. Gerade veraltete Wörter gegebener Instrumente oder unübliche Wörter (z. B. *yielding*) erschweren die individuelle Charakterisierung. Ebenso tragen Modewörter wie *interessant* oder *liberal* wenig zur Differenzierung bei, da sie auf sehr viele Personen angewendet werden können (Klapprott, 1994).

Das in der vorliegenden Arbeit eingesetzte Vorgehen orientiert sich weitgehend an Studien, die sich mit dem sogenannten lexikalischen Ansatz, der Taxonomie von Eigenschaftswörtern und der Einschätzung von Erwünschtheit beschäftigten, wie das Handbuch deutschsprachiger Wortnormen (Hager & Hasselhorn, 1994).

Generell zeigte sich wiederholt, dass in der freien Personenbeschreibung die Versuchspersonen am häufigsten auf Kategorien zum Erscheinungsbild und äußeren Charakteristiken (Körper-, Körperbau- und Gesichtsmerkmale) zurückgreifen (Angleitner &

Ostendorf, 1994; Helgeson, 1994a; Myers & Gonda, 1982b), was auch bereits im Rahmen der Multidimensionalität des Geschlechts diskutiert wurde. Für die Erstellung des neuen Instruments ist jedoch essentiell, dass die Zielitems vor allem aus dem Bereich überdauernder Dispositionen und Charaktereigenschaften<sup>24</sup> stammen. Bezugnehmend auf die angeführten Kriterien zur Auswahl der relevanten Adjektive, sollten sich die gefundenen Adjektive in einen der folgenden Sätze einfügen lassen: „Wie [Adjektiv] bin ich eigentlich?“ (Angleitner & Ostendorf, 1994).

## 2.7 Validierungskonstrukte und Hypothesen

### 2.7.1 Hypothesen zu den Teilskalen des GSK

In Bezug auf die vorgeschlagene Konzeptualisierung des GSK über die beiden Dimensionen Typizität und Valenz werden zunächst folgende Hypothesen postuliert:

- (H1) Positive bzw. negative Maskulinität bzw. Femininität stellen unterschiedliche Facetten des Geschlechtsrollenselbstkonzepts dar. (Vier-Faktoren-Struktur)
- (H2) Männer schreiben sich die maskulinen Eigenschaften im stärkeren Maße zu als Frauen, während Frauen sich die femininen Eigenschaften im stärkeren Maße zuschreiben als Männer dies tun (HE: Typizität).
- (H3) Männer und Frauen schreiben sich die positiven Merkmale des GSK stärker zu als negative (HE: Valenz).
- (H4) Frauen schreiben sich negative Attribute eher zu, wenn sie für Frauen typischer sind als für Männer, Männer schreiben sich umgekehrt negative Attribute eher zu, wenn sie typisch männlich sind (Interaktion vom Valenz & Typizität).

### 2.7.2 Hypothesen zu den Validierungskonstrukten

Zur Abschätzung der Konstruktvalidität wurden verschiedene Konstrukte herangezogen, die in Tabelle 8 den entsprechenden Teilskalen zugeordnet sind, und für die folgende Zusammenhänge postuliert werden:

- (H5) Die positive Maskulinitätsskala ist spezifischer Prädiktor und klärt mehr Varianz bezüglich der Konstrukte beruflicher Ehrgeiz, Hoffen auf Erfolg sowie Leistungsorientierung auf, indem sie eine bessere Vorhersageleistung erzielt als die drei übrigen Konstrukte.
- (H6) Die negative Maskulinitätsskala sagt zwei unterschiedlichen Konstrukte der Aggressivität sowie das Konzept der machtbezogenen Werte besser vorher als Mas+ und die beiden Femininitätsskalen.

---

<sup>24</sup> Angleitner und Ostendorf (1994) kategorisierten personenbeschreibende Adjektive in Bezug auf: Temperament, überdauernde Dispositionen, Charaktereigenschaften, intrapsychische oder körperliche Zustände, kurzfristige Aktivitäten, Effekte auf andere, soziale Bewertungen, soziale Rollen, Beziehungen, Herkunft, Merkmale der äußeren Erscheinung

- (H7) Die positive feminine Teilskala hat spezifischen Vorhersagewert für die Konstrukte der sozialen Werte, Empathie sowie Fürsorglichkeit und ist für diese Konzepte ein besserer Prädiktor als die drei übrigen Konstrukte
- (H8) Die negative Femininitätsskala sagt die Konstrukte emotionale Selbstwertschätzung, Ängstliche Depressivität und Trait-Angst besser vorher als die maskulinen Teilskalen und Fem+ und klärt mehr Varianz auf.

Tabelle 8. Validierungskonstrukte der vier Teilskalen der Geschlechtsrollen

Mas+	Mas-	Fem+	Fem-
beru. Ehrgeiz (AVEM)	Aggressivität (AQ)	Empathie (Sorge)	Emot. Selbstwertschätzung
Hoffen auf Erfolg (AMS)	Werteskala Macht	soziale Werte	Ängstliche Depressivität (BSF).
Leistungsorientierung (FPI)	Aggressivität (FPI)	Fürsorglichkeit (PRF)	Trait-Angst (STAI)

## 2.8 Realisierung als Onlinestudien

Die einzelnen Studien zu den in dieser Arbeit aufgestellten Hypothesen wurden mittels mehrerer Onlinebefragungen realisiert. Aus diesem Grund werden im Folgenden allgemeine Informationen zu dem Medium Internet sowie Vor- und Nachteile von online umgesetzten Datenerhebungen diskutiert.

### 2.8.1 Datensammlung im Internet

Auf der gesamten Welt steigt die Internet-Nutzung – mit jedem Jahr werden zunehmend mehr Menschen mit diesem Datennetz verbunden. So berichtete der Bundesverband Informationswirtschaft Telekommunikation und neue Medien e.V. (BITKOM., 2008) über das Jahr 2007 von einem weltweiten Zuwachs an Internetnutzern von hundert Millionen Menschen auf insgesamt rund 1.23 Milliarden Internetnutzern, womit sich die Online-Community innerhalb von fünf Jahren verdoppelt hat. Nahezu jeder 5. Mensch verfügt mittlerweile über einen Onlinezugang. In Europa ist mehr als die Hälfte der Bevölkerung regelmäßig im Internet (i2010, 2008), in Deutschland nutzen fast zwei Drittel aller Personen über 18 Jahre das Internet (Forschungsgruppe Wahlen Online, 2008). Bis 2010 wird ein weltweites durchschnittliches Wachstum von 7.7 % pro Jahr erwartet. Zweifelsohne bleiben einige Regionen unerreicht, der repräsentative Ausschnitt der Gesamtbevölkerung online erfährt jedoch täglich Zuwachs.

### 2.8.2 Vergleich von Offline und Onlineforschung

Onlinestudien sind gewöhnlich in drei grundlegenden Arten möglich: E-Mailbefragungen, HTML Formen und herunterladbare interaktive Befragungsmodule (siehe Council of American Survey Research Organizations., 1998). Fragebogenuntersuchungen jeglicher Form stellen dabei die populärste Datenerhebungsanwendung im Netz (Henning-Thurau & Dallwitz-Wegner, 2002) dar, nicht zuletzt aufgrund der unkomplizierten Umsetzung.

Die Fragestellung, inwieweit online erworbene Daten vergleichbar zu denen von Offlinestudien sind, hat eine Vielzahl von wissenschaftlichen Arbeiten beschäftigt (Buchanan, 2001; Buchanan, 2002; Knapp & Kirk, 2003; Rietz & Wahl, 2002; Tuten, Urban, & Bosnjak, 2002). Dazu haben verschiedene Studien über die letzten Jahre hinweg einen extensiven

Methodenvergleich von Paper-Pencil, Telefon- oder computervermittelten Kommunikationsformen in verschiedenen Forschungsbereichen vorgenommen und keine nennenswerten Unterschiede finden können (Erdman, Klein, & Greist, 1983; Knapp & Kirk, 2003; Morrison, Leigh, & Rogers Gillmore, 1999; Rietz & Wahl, 2002; Schillewaert & Meulemeester, 2005; Skinner & Allen, 1983). Auch Gosling, Vazire und John (2004) konnten zeigen, dass Versuchspersonen aus dem Internet nicht unmotivierter, sozial isolierter oder depressiver sind. Die Onlinestichproben sind zwar demografisch mannigfaltig und arbeiten unter größerer Anonymität, dennoch können Ergebnisse als vergleichbar betrachtet werden.

Knapp und Kirk (2003) konnten nachweisen, dass webbasierte Umfragen zu vergleichbaren Ergebnissen führen wie Befragungen im Paper-Pencil-Format (sowohl mit als auch ohne Versuchsleiter), online oder via Telefon. Im Antwortverhalten bezüglich sensibler Fragen über alle drei verwendeten Medien hinweg fanden sich keine signifikanten Unterschiede.

Darüber hinaus stützen zahlreiche weitere Studien, die gezielt verschiedene Persönlichkeitsskalen untersuchten, den positiven Eindruck der Online-Forschung. Gefundene Ergebnisse erwiesen sich in ihren statistischen Kennwerten als vergleichbar und zufriedenstellend, sowohl bezogen auf interne Konsistenzen oder Faktorenstrukturen (Pasveer & Ellard, 1998), Ähnlichkeiten in den gefundenen Konstrukten (Schwarzer, Mueller, & Greenglass, 1999) als auch Reliabilitätsindizes (Davis, 1999). Während Krantz und Dalal (2000) vergleichbare Teilnehmerzahlen, Mittelwerte und Varianzen der zum Einsatz gekommenen Likert-Skalen in ihren on- und offline Erhebungen berichteten, fanden Buchanan and Smith (1999) in ihren Untersuchungen zur Selbstüberwachung sogar eine höhere Reliabilität sowie einen besseren Fit des hypothetisierten theoretischen Modells in ihren Onlinedaten.

Methodisch gesehen sind vor allem zwei Unterschiede zwischen online und offline erhobenen Daten relevant: die größere Varianz in den Onlinedaten aufgrund der größeren Versuchsgruppe und die Selbstselektion der Probanden (Bandilla, 1999; Reips, 1999; Reips, 2000). Gerade diese Selbstselektion ruft kritische Stimmen hervor, da bei dieser passiven Auswahl von Personen durch überrepräsentierte geübte und häufig teilnehmende Probanden eine größere Gefahr der Verzerrung von Daten besteht. Die Kritik bezüglich der Validität und Selbstselektion kann nicht ignoriert werden, ihretwegen die Onlineforschung abzuwerten ist jedoch ebenfalls unnötig. Zum einen entstehen Kosten und ein Mehraufwand an Zeit bei Offline-Studien und zum anderen ist die Umsetzung einer Studie offline nicht automatisch gleichbedeutend mit einem Zuwachs an Qualität bezüglich der Stichprobe, denn auch diese Methode muss sich mit dem Problem der Selbstselektion auseinandersetzen. Daher erfordert es vor allem die Anerkennung der besonderen Umstände in der Form der Datenerhebung sowie Berücksichtigung dieser bei Analysen und Interpretationen der Ergebnisse.

Im Großen und Ganzen sind online und offline erhobene Daten in ihrer Qualität äquivalent in dem, was sie messen und wie gut sie dies tun. Zusammenfassend bleibt zu sagen, dass online umgesetzte Forschung nicht nur eine gute Alternative zu klassischen Paper-

Pencil Befragungen darstellt, sondern darüber hinaus noch über eine Vielzahl von Vorteilen verfügt, derer sich die vorliegende Arbeit bedient.

### 2.8.3 Merkmale der webbasierten Forschung

Obwohl das World Wide Web (WWW) in seiner heutigen Form von der technischen Umsetzung als jung gilt, verfügt die Onlineforschung über vielversprechendes Potential und eine Reihe von Vorteilen verglichen mit traditionellen Datenerhebungsmethoden (Batinic, 2001; Batinic & Bosnjak, 1997; Reips, 1997; Schmidt, 1997). Dabei ist vor allem die schnelle und einfache Erreichbarkeit einer größeren, demografisch heterogenen Online-Stichprobe (Alter, Bildung, Beruf) von großer Bedeutung. Das Verlegen der Datenerhebung in das WWW ermöglicht sowohl den zeitlich als auch lokal unabhängigen Zugang zu geografisch beinahe unbegrenzten Stichproben als auch zu spezifischen Randgruppen von potentiellen Versuchspersonen (*Asynchronität* und *Allokation*) (Marcell & Falls, 2001, die z. B. das visuelle und auditive Gedächtnis bei Familienmitgliedern mit Down- oder Williams-Syndrom testeten; Reips, 2000). Durch die verbesserte Erreichbarkeit von vielen Gruppen sind kulturvergleichende Studien besser zu realisieren als über die offline Methode (Hewson, Laurent, & Vogel, 1996; Reips, 1997).

Der vereinfachte Zugang zu Versuchspersonen sowie die Art der Umsetzung eines Fragebogens reduzieren die Kosten an Zeit und finanziellen Aufwendungen (z. B. Kosten für die Nutzung von Laboren, Versuchsleitern, Administration, etc.).

Auf der technischen Seite erlaubt die Onlineforschung eine viel stärkere *Automatisierung* von Prozeduren und (Auswertungs-)Analysen, beispielsweise durch die Programmierung von Prozessserien, Filtersetzung oder automatischer Erfassung und Analyse von Antworten. Durch die *Flexibilität der Darstellung* können interaktive und multimediale Inhalte wie Audio, Video etc. integriert werden. Darüber hinaus können *Meta-Daten* (Antwortzeiten, Browser) aufgezeichnet werden, die zusätzliche Rückschlüsse auf den Antwortprozess (Reihenfolge, Vollständigkeit der beantworteten Fragen, Dauer der Befragung) erlauben.

Einer der wichtigsten Diskussionspunkte der Onlineforschung betrifft zweifelsohne die *Repräsentativität* des Internet-Benutzers. Es gilt anzuerkennen, dass – zumindest zurzeit – die Gesamtpopulation noch nicht vollständig im Internet repräsentiert ist. Obwohl in Offline-Bedingungen zunächst keine technischen Voraussetzungen Menschen per se für die Stichprobe ausschließen, wird die Erreichbarkeit der Offline-Population in diesem Zusammenhang überbewertet; dies jedoch als Argument gegen die Forschung im WWW geltend machen zu wollen, erscheint kurzsichtig. Umso mehr, wenn man psychologische Studien offline genauer betrachtet, und feststellt, dass 80% auf Studierende zurückgreifen, obwohl diese Subgruppe lediglich 3% der Gesamtbevölkerung ausmacht (Schultz, 1972; Smart, 1966).

Zusammenfassend lässt sich festhalten, dass insbesondere die gute Erreichbarkeit großer heterogener Stichproben, die Effektivität in der Durchführung (bezogen auf Geschwindigkeit, Kosten, Unabhängigkeit von Ort und Zeit) sowie vergleichbare Ergebnisse zu Offlinestudien zu den großen Vorteilen der Onlineforschung zählen, die sich dadurch als geeignete Methode zur Datenerhebung qualifiziert.

Verschiedene Studien haben darüber hinaus Aspekte diskutiert, die die Onlineforschung für den Einsatz in bestimmten Studien vor Herausforderungen stellt: Abhängigkeit von der technischen Ausstattung, Abwesenheit eines Versuchsleiters, Datenqualität, mehrfaches Absenden des Datensatzes, Sicherheit und Datenintegrität, Antwortverhalten, Drop-Out, sozial erwünschtes Antworten, Selbstselektion und Stichprobendemografie der Probanden (Bailey, Foote, & Throckmorton, 2000; Bandilla & Hauptmanns, 1998; Bason, 2000; Batinic & Bosnjak, 1997; Bosnjak, 2001; Bosnjak & Tuten, 2001; Dillman, 2000; Dillman & Bowker, 2001; Dillman, Tortora, Conradt, & Bowker, 1998; Joinson, 2001; Knapp & Heidlingsfelder, 2001; Krantz & Dalal, 2000; Reips, 1997; Reips, 1999).

Eine Vielzahl der Nachteile von Studien im WWW müssen durchaus auch bezüglich von Offline-Datenerhebungen diskutiert werden. Grundsätzlich kann durch sorgfältiges Arbeiten sicher gestellt werden, dass elektronische Studien den traditionellen Offlinestudien in nichts nachstehen. Beachtet man zusätzlich die nicht unerheblichen Vorteile, ist die Onlineforschung der Umsetzung offline in einigen Belangen sogar überlegen.

Daher erfordern die Nachteile weniger eine Abkehr von Onlinestudien als Methode, sondern legen vielmehr eine Berücksichtigung derer in Planung und Programmierung nahe. So gilt es bei der Arbeit mit IP-Adressen den Richtlinien der Landesbeauftragten für den Datenschutz und für das Recht auf Akteneinsicht Brandenburg (2005) zu folgen, gemäß denen Identifizierungsinformationen und Daten voneinander getrennt aufbewahrt und nur unter hohem Aufwand wieder zusammengeführt werden sollten. Durch die Erfassung der demografischen Aspekte zu Beginn der Befragung wird die Verbindlichkeit der Probanden erhöht und zusätzlich persönliche Informationen sicher gestellt, die – in Fällen, in denen die Versuchspersonen die Studie doch noch vorzeitig verlassen – dann zur weiteren Analyse (bis zum Drop-out) und Interpretation herangezogen werden können.

## **2.8.4 Technische Realisation der Fragebögen**

### **2.8.4.1 Design & Programmierung der Fragebögen**

Die Datenerhebungen der verschiedenen Studien wurden (mit Ausnahme eines Teils der Pilotstudie) mit Hilfe von Onlinebefragungen in den Jahren 2006-2009 durchgeführt. Jeder Proband wurde über einen Link zur ersten Seite geleitet, auf der allgemeine Informationen angegeben waren. Reips (1997), Schmidt (1997), Dillman, Tortora und Bowker (1999) sowie Dillman (2000) einigten sich auf grundlegende Prinzipien von Webfragebögen und schlugen u. a. vor eine einladende erste Seite zu schalten, die zunächst alle wichtigen Informationen verständlich aufführt (Name des Versuchsleiters, Titel der Umfrage, Zweck der Studie, wissenschaftlicher Hintergrund, möglicherweise Link zum Fragebogen in anderen Sprachen, Versicherung der Anonymität und des Feedbacks; Batinic & Bosnjak, 1997). So wurde in allen Studien darauf hingewiesen, dass die Daten ausschließlich für wissenschaftliche Zwecke gespeichert und ausgewertet werden. Die Anonymität wurde versichert sowie das technische Vorgehen dahingehend erklärt, dass die Antworten einer einzelnen Person mit denen vieler anderer gemischt werden, was eine Identifizierung des Einzelnen verhindert. Anlehnend an Reips (1999) Technik der Hohen Hürde wurden die demografischen Fragen ebenfalls am

Anfang des Fragebogens erfasst, um die Verbindlichkeit der Probanden zu erhöhen und die Abbruchquoten zu verringern.

Alle einzelnen Fragebogenteile wurden mit einer eigenen kurzen Instruktion eingeleitet, in der unter anderem spezifische Antwortskalen oder Aufgaben genau erläutert wurden. Zusätzlich zu der ohnehin erklärten Antwortskala, wurde zwischen den einzelnen Fragebogenblöcken die Antwortskala erneut grafisch abgebildet, um das Scrollen innerhalb einer Seite unnötig zu machen.

Die Fragebögen waren in HTML (Hypertext Markup Language; version 3.2) mithilfe von DreamWeaver programmiert worden, wobei das Prinzip des Least Compliant Browser (LCB<sup>25</sup>), wie von Bowker (2000) empfohlen, umgesetzt wurde. Zusätzlich wurde mit PHP<sup>26</sup> programmiert, wodurch die von den Probanden bereitgestellten Informationen seitenweise verwaltet und in verschiedene Textdateien abgespeichert wurden. Die abgespeicherten Variablen konnten in dem vorliegenden Format direkt in SPSS-Dateien importiert werden. Die Fragebögen wurden allesamt auf dem Server der Homepage <http://www.w-lab.de> abgespeichert und konnten mittels jedes gängigen Internetbrowser (Mozilla Firefox, Internet Explorer, etc.) aufgerufen werden.

#### **2.8.4.2 Bewerbung der Fragebögen**

Die Bewerbung der verschiedenen Studien erfolgte auf unterschiedlichen Wegen – jedoch in gleicher konsequenter Weise für alle unterschiedlichen Fragebögen. So wurden die Links mit einer Beschreibung der Untersuchung auf verschiedenen, gängigen Testportalen gepostet. Die Hauptquelle der Rekrutierung möglicher Versuchspersonen stellte jedoch das studiVZ dar. Bei studiVZ handelt es sich um eine Web-2.0-Plattform, welche mittlerweile dem Holtzbrinck Networks zugehörig ist. Ähnlich dem angloamerikanischen Online-Netzwerk Facebook zählt studiVZ zur sog. sozialen Software<sup>27</sup>, die menschliche Kommunikation und soziales Zusammenarbeiten unterstützt und fördert. StudiVZ verfolgt dabei das Ziel, Studierenden die Möglichkeit zu geben, ihre sozialen Kontakte online zu verwalten und alte Kontakte wiederzufinden. Laut Angaben der eigenen Webseite (<http://www.studivz.net//about us/1/>) verzeichnet die Webseite acht Millionen registrierte Nutzer, mit einem wöchentlichen Zuwachs an Neuanmeldungen im sechsstelligen Bereich. Einwände bezüglich der genauen Zusammensetzung der Nutzer an dieser Stelle sind berechtigt. StudiVZ selbst liefert keine genauen Angaben bezüglich der demografischen Struktur der Nutzergruppe. Allgemeine Aussagen über diese Gruppe werden umso schwieriger, zieht man in Betracht, dass im Wintersemester 2007 / 2008 an deutschen Hochschulen 1.932.355 Personen immatrikuliert waren (Statistisches Bundesamt Deutschland, 2008), was lediglich knapp 25% aller studiVZ Nutzer ausmacht. Ein ausführlicher demografischer Fragebogenteil soll daher über die Studien hinweg die Auswahl vergleichbarer Stichproben sicherstellen.

---

<sup>25</sup> LCB = eine HTML Kodierung, die den technisch niedrigsten Browser berücksichtigt, um die Kompatibilität und Darstellungsform über verschiedene Versionen von Browser zu gewährleisten.

<sup>26</sup> = ursprünglich „Personal Home Page Tools“; steht für „Hypertext Preprocessor“. PHP ist eine Skriptsprache, die zur Erstellung von dynamischen Webanwendungen verwendet wird.

<sup>27</sup> "Social Software sind solche internetbasierten Anwendungen, die Informations-, Identitäts- und Beziehungsmanagement in den (Teil-) Öffentlichkeiten hypertextueller und sozialer Netzwerke unterstützen." (Schmidt 2006, S. 2)

Ein letztes populäres Argument im Bezug auf Onlineforschung betrifft die Ernsthaftigkeit der gegebenen Antworten. Die technisch gegebene Möglichkeit von Mehrfachantworten veranlasst viele Kritiker die Datenqualität von Onlinebefragungen anzuzweifeln, da sie bewusstes wiederholtes Falschantworten einiger Personen vermuten. In diesem Zusammenhang lässt sich Reips (2000) Argumentation anführen, laut derer es unwahrscheinlich ist, dass sich eine Versuchsperson per se über das eigentlich geteilte Verständnis nur einmal teilzunehmen hinwegsetzt, denn:

*„Generally, multiple participation by the same person is not very likely, as most Web experiments are not that thrilling. It seems unlikely that people would repeat long or boring experiments.“ (S.107)*

Zusammenfassend gilt festzustellen, dass Onlineforschung den Vergleich mit offline erhobenen Daten sehr wohl standhält. Mögliche Nachteile stehen dabei weniger statisch unveränderbar im Raum, sondern fungieren als Grundlage für einzuhaltende Richtlinien. Darüber hinaus bietet die Form der online basierten Datenerhebung diverse Vorteile, in denen sie der Offlineforschung sehr wohl überlegen ist. Da die Onlineforschung als äquivalente Variante zur Datengenerierung sowohl methodisch als auch theoretisch als abgesichert gilt und für die in dieser Arbeit durchzuführenden Studien eine Vielzahl von Probanden erforderlich war, kam diese Forschungsmethode zum Einsatz.

Die Darstellung und kritische Diskussion der einzelnen Stichprobensammlungen erfolgt jeweils im Rahmen der nacheinander durchgeführten Studien.



### 3 EMPIRISCHE STUDIEN

Der folgende Teil stellt die empirischen Untersuchungen dar. Er ist in zwei große Bestandteile untergliedert. Im ersten Abschnitt werden der BSRI und seine Validität einer ausführlichen Überprüfung unterzogen; weniger um seine Bedeutung als vielmehr seine Einsatzfähigkeit zur Messung des Geschlechtsrollenselbstkonzepts zu diskutieren. Zur Entwicklung des BSRI wurde nur die soziale Erwünschtheit als Kriterium herangezogen. Daher soll in erster Linie die Überprüfung des BSRI im Hinblick auf die aktuelle soziale Erwünschtheit stattfinden. Es ist jedoch ebenfalls belegt, dass die Messung sozialer Erwünschtheit auch immer Aspekte der Typizität mit erfasst. Geschlechterstereotype enthalten mit diesen präskriptiven und deskriptiven Eigenschaften zwei nicht identische Aspekte (Swazina et al., 2004), die jedoch auch nicht gänzlich voneinander zu lösen sind (siehe Seite 36 ff.). Die theoretische Grundlage für die Entwicklung von Messinstrumenten muss beide Aspekte berücksichtigen. Daher soll neben der Erwünschtheit auch die Typizität der BSRI-Eigenschaften überprüft werden. Zusätzlich werden aktuell erhobene Selbstbeschreibungsmaße diskutiert (Studie 1 und 2).

Diese Ergebnisse in Kombination mit der Vernachlässigung der negativen Items sollen schließlich die Grundlage für eine Neukonstruktion bilden. Im zweiten Teil der Arbeit wird das detaillierte Verfahren zur Konstruktion eines neuen Maßes dokumentiert. Dieses ist in mehrere Studien (3 - 7) untergliedert. Das Ergebnis soll ein ökonomisches, valides und reliables Maß zur Messung des GSK sein, welches zusätzlich zu den populären positiven Eigenschaften zur Erfassung von Maskulinität und Femininität in bewährten Instrumenten auch negative Eigenschaften beinhaltet. Dabei sollen alle beteiligten Attribute (1) den zugrunde gelegten diagnostischen Maßstäben entsprechen und (2) ein Abbild des deutschen Sprachgebrauchs und -verständnisses in Aktualität und Auftretenshäufigkeit der Adjektive darstellen. Tabelle 9 zeigt im Überblick die durchgeführten Studien und ihre jeweiligen Teilziele.

*Tabelle 9. Übersicht aller durchgeführten Studien*

<b>Studien</b>	<b>Inhalte</b>
<b>Studie 1</b>	Einschätzung von Typizität und sozialer Erwünschtheit der BSRI-Items
<b>Studie 2</b>	Selbstbeschreibung auf den BSRI-Items
<b>Studie 3</b>	Konstruktionsstudie - Generierung neuer GSK-Items
<b>Studie 4</b>	Einschätzung von Phase 1: Typizität (in 3 Teilgruppen) Phase 2: sozialer Erwünschtheit der neuen Items Phase 3: Selbstbeschreibung (in 3 Teilgruppen)
<b>Studie 5</b>	Validierung der neuen Teilskalen
<b>Studie 6</b>	Replikation der neuen Teilskalen und ihrer Validierung
<b>Studie 7</b>	Erweiterung der neuen Skalen um eine Erwartungs-mal-Wert-Komponente

Alle Studien sind nach dem klassischen Format Fragestellung, Methode, Ergebnisse und Diskussion strukturiert. In einem abschließenden Kapitel erfolgt dann die zusammenfassende Diskussion.

### 3.1 Kritische Überprüfung des BSRI

Aufgrund der im Theorieteil diskutierten Mängel des Bem Sex Role Inventory (Bem, 1974) zur Erfassung des Geschlechtsrollenselbstkonzepts (siehe S. 49) sollen die Eigenschaften einer differenzierten Prüfung unterzogen werden. Anlehnend an Bems Vorgehen, vor allem die soziale Erwünschtheit der Items für den deutschsprachigen Raum im 21. Jahrhundert zu überprüfen, sollen zusätzlich die Bewertungen der Typizität der einzelnen Items hinsichtlich der Skalen, denen sie angehören, untersucht werden. Schließlich wird in einem dritten Schritt überprüft, ob und wie die Selbstbeschreibung anhand des BSRI Unterschiede zwischen Geschlechtern und Gruppen aufzeigen kann, was den BSRI somit als valides und reliables Instrument zur Abbildung des geschlechtsbezogenen Selbstkonzepts auch heute noch ausweisen würde. Im Fazit besteht das Ziel darin die Validität des BSRI zu überprüfen.

Entlang der bisherigen kritischen Stimmen hinsichtlich der Messung des Geschlechtsrollenselbstkonzepts, insbesondere des BSRI, wird erwartet, dass einzelne Items der femininen und maskulinen Teilskala nicht nur nicht mehr zwischen den Geschlechtern zu differenzieren vermögen, sondern dass sie, aufgrund der Einschätzung ihrer Typizität und sozialen Erwünschtheit, nicht mehr ihrer ursprünglichen Skala und damit auch nicht dem Instrument als solches zugeordnet werden dürfen.

Dazu wurden zwei Studien durchgeführt, in denen diese Merkmale jeweils untersucht wurden. Zum Einsatz kam keine eigene Übersetzungen der Items von Bem (1974), sondern die deutsche Version des Bem Sex Role Inventory von Schneider-Düker und Kohler (1988). Obwohl dieses Instrument knapp 15 Jahre jünger ist als das englischsprachige Original, gilt die gleiche Kritik der unzureichenden Gütekriterien und der mangelnden Aktualität dieses Instruments in der heutigen Zeit, nicht zuletzt vor dem Hintergrund, dass Schneider-Düker und Kohler Bems Vorgehensweise ohne Änderungen wiederholten und damit auch lediglich das Kriterium der sozialen Erwünschtheit als Grundlage für die Itemauswahl heranzogen<sup>28</sup>.

#### 3.1.1 Studie 1: Einschätzung von Typizität und sozialer Erwünschtheit der BSRI-Items

##### 3.1.1.1 Fragestellung

Die deutschen Eigenschaften des BSRI wurden in zwei Schritten einer Überprüfung unterzogen. Zunächst wurden Probanden gebeten die Attribute bezüglich ihrer sozialen Erwünschtheit für die heutige Gesellschaft einzuschätzen, ähnlich wie es auch Bem (1974) und Schneider-Düker und Kohler (1988) in der Entwicklung der Instrumente vorgenommen hatten. Im zweiten Schritt wurden zusätzlich Beurteilungen der Typizität der Attribute erfragt.

##### 3.1.1.2 Methode

###### 3.1.1.2.1 Stichprobe

Insgesamt beurteilten 1.464 Personen (24% Männer, 76% Frauen) die BSRI-Items gemäß ihrer sozialen Erwünschtheit und Typizität. Im Schnitt waren die Probanden (Vpn) 22.92 Jahre alt ( $SD = 3.91$  Jahre, Min: 18, Max: 64 Jahre). 77% hatten die Schule mit dem Abitur verlassen,

---

<sup>28</sup> Eine Eigenschaft galt dann als geeignet, „wenn sie unabhängig vom Geschlecht der Beurteiler als signifikant erwünschter für das eine als für das andere Geschlecht beurteilt wurden“ (S. 258)

17% hatten später ein (Fach-)Hochschulstudium abgeschlossen. Die Gesamtstichprobe verteilt sich ungleich auf eine von vier Versuchsbedingungen (Einschätzung der Typizität vs. soziale Erwünschtheit, jeweils bezogen auf Männer und Frauen), was auf die Rekrutierungsmaßnahmen zurückzuführen ist (vgl. 3.1.1.2.3 Durchführung).

An der Studie zur Einschätzung der Typizität nahmen 464 Probanden teil. Die Studie zur der Einschätzung der sozialen Erwünschtheit verzeichnete mit 1.000 Versuchspersonen einen wesentlich größeren Zulauf als die Studie zur Typizität, was auf die Nachwirkungen der sehr erfolgreichen Rekrutierungsmaßnahmen zurückzuführen ist. In Tabelle 10 sind die demografischen Parameter der Teilstichproben aufgeführt.

Tabelle 10. Demografie der Probanden aus Studie 1 zur Überprüfung des BSRI

Fragebogenversion	Einschätzung der Typizität der Items für eine/n...		Einschätzung der sozialen Erwünschtheit der Items für eine/n...	
	...Mann 1	...Frau 2	...Mann 3	...Frau 4
<b>Geschlecht der Vpn</b>				
Männer	64	84	101	106
Frauen	160	156	395	398
Total	224	240	496	504
<b>Alter der Vpn</b>	$M = 23.19$ Jahre $SD = 4.48$ Min: 18, Max: 64 Jahre		$M = 22.80$ Jahre $SD = 3.61$ Min: 18, Max: 53 Jahre	
<b>Schulbildung</b>	77.8% Abitur		76.3% Abitur	
	8.5% Hochschulabschluss		10.1% Hochschulabschluss	
<b>sexuelle Orientierung</b>	92.4% heterosexuell		93.9% heterosexuell	
	3.6% homosexuell		2.2% homosexuell	
	3.6% bisexuell		3.9% bisexuell	

Die Versuchspersonen wurden innerhalb der Beurteilungen für Typizität oder sozialer Erwünschtheit zufällig auf die Bedingungen verteilt, so dass sich die Teilnehmerzahlen für die beiden Einschätzungsziele annähernd gleich gestalten. Beide Stichproben ähneln sich in den demografischen Merkmalen. Es konnten keine signifikanten Unterschiede im Hinblick auf das Alter, die Form des Schulabschlusses oder die sexuelle Orientierung festgestellt werden. Diese Variable wurde erfasst, da sowohl das biologische Geschlecht als auch die Geschlechtsorientierung Einfluss auf das soziale Geschlecht und damit auf die Bewertung von Typizität und sozialer Erwünschtheit der entsprechenden Items haben.

Bezüglich des Geschlechts gab es einen signifikanten Unterschied zwischen den Stichproben,  $\chi^2(1) = 21.63$ ,  $p < .001$ . Demnach schwankte zwar der Anteil der teilnehmenden Frauen zwischen 2:1 und 4:1 im Verhältnis zu den Männern; innerhalb einer Merkmalskombination (Typizität vs. sozialer Erwünschtheit) blieben die Anteile der einzelnen Geschlechter jedoch gleich verteilt. Beispielsweise schätzten 101 Männer die soziale Erwünschtheit der BSRI-Items in Hinblick auf einen Mann ein, während 106 weitere Männer diese Einschätzung mit Blick auf eine Frau der heutigen Gesellschaft vornahmen.

### 3.1.1.2.2 Instrumente

Den Probanden wurden die 60 Items des BSRI in der deutschen Übersetzung von Schneider-Düker und Kohler (1988) vorgelegt. Die Items sind Tabelle 11 und Tabelle 12 zu entnehmen. Es wurden vier Instruktionsbedingungen als between-subjects-Manipulation realisiert:

- (1) Typizität für Männer: Wir bitten Sie, die Typizität der folgenden Eigenschaften für Männer in unserer Gesellschaft im Allgemeinen einzuschätzen. Zum Beispiel könnte eine Aussage lauten *„Es ist typisch für Männer in unserer Gesellschaft gesetzestreu zu sein“*.
- (2) Typizität für Frauen: Wir bitten Sie, die Typizität der folgenden Eigenschaften für Frauen in unserer Gesellschaft im Allgemeinen einzuschätzen. Zum Beispiel könnte eine Aussage lauten *„Es ist typisch für Frauen in unserer Gesellschaft gesetzestreu zu sein“*. Die Eigenschaften wurden jeweils einzeln dargeboten. Für die Bewertung der Typizität wurde jeweils eine siebenstufige Skala mit den Abstufungen (1) *trifft gar nicht zu* bis (7) *trifft genau zu* vorgegeben.
- (3) Erwünschtheit für Männer: Wir bitten Sie, die folgenden Eigenschaften für Männer in unserer Gesellschaft im Allgemeinen einzuschätzen und zwar dahingehend, wie positiv bzw. negativ diese Eigenschaften für Männer sind. Zum Beispiel könnte eine Aussage lauten *„Für Männer in unserer Gesellschaft gesetzestreu zu sein - ist positiv (bzw. negativ)“*.
- (4) Erwünschtheit für Frauen: Wir bitten Sie, die folgenden Eigenschaften für Frauen in unserer Gesellschaft im Allgemeinen einzuschätzen und zwar dahingehend, wie positiv bzw. negativ diese Eigenschaften für Frauen sind. Zum Beispiel könnte eine Aussage lauten *„Für Frauen in unserer Gesellschaft gesetzestreu zu sein - ist positiv (bzw. negativ)“*. Die Erwünschtheitsratings wurden jeweils auf einer siebenstufigen Antwortskala mit den Abstufungen (1) *negativ* bis (7) *positiv* erfasst.

### 3.1.1.2.3 Durchführung

Diese Studien wurden online durchgeführt. Die Probanden erhielten über verschiedenste Rekrutierungsmaßnahmen einen Link, der sie zur ersten Seite des jeweiligen Fragebogens mit allgemeinen Informationen über den Forschungshintergrund leitete. Die vergleichsweise jungen und homogenen Stichproben begründen sich in der Bewerbung des Forschungsvorhabens in Studentennetzwerken. Der Vorteil dieser Rekrutierung bestand in einer großen Stichprobe. Hinzu kommt, dass der Einsatz klassischer Geschlechtsrollenmaße an Studierenden ohnehin eine große Popularität besitzt, weshalb bewusst diese Zielgruppe ausgewählt wurde.

Die Versuchspersonen wurden per Zufall einer der vier Instruktionsbedingungen zugeteilt. Nach einer kurzen Instruktion und Erklärung der Studie erfolgte zunächst die Erfassung demografischer Aspekte. Es schloss sich die Instruktion zur Einschätzung der Eigenschaftsbegriffe an. Jeweils 20 Items pro Seite wurden in einer Tabelle dargeboten. Ein Zähler der Versuchspersonen in Abhängigkeit vom Geschlecht, der im Hintergrund mitlief, sollte eine Gleichverteilung der Versuchspersonen über die Fragebogenversionen sicherstellen. Nach der Beantwortung aller Fragen wurde den Probanden für ihre Teilnahme gedankt. Zudem wurde ihnen die Möglichkeit gegeben, sich für eine Ergebnisrückmeldung in eine Interessentenliste einzutragen. Eine unabhängige Speicherung von E-Mail-Adressen und persönlichen Daten wurde gewährleistet.

### 3.1.1.3 Ergebnisse

Im Folgenden werden die Ergebnisse für die getrennte Betrachtung der maskulinen und femininen Items des BSRI unter den Gesichtspunkten der Typizität und Erwünschtheit

dargestellt. Die Variable *Version* bezeichnet das in der Instruktionsbedingung vorgegebene zu beurteilende Geschlecht (typisch/erwünscht für Männer vs. typisch/erwünscht für Frauen), das als between-subjects-Faktor realisiert war.

### 3.1.1.3.1 Analyse der maskulinen Eigenschaften: soz. Erwünschtheit

Die MANOVA für die maskulinen Eigenschaften ergab signifikante multivariate Effekte für die Variablen Geschlecht,  $F(20,949) = 3.32$ ,  $p < .001$ ,  $\eta^2 = .07$ , *Version*,  $F(20,949) = 18.47$ ,  $p < .001$ ,  $\eta^2 = .28$ , sowie für die Interaktion von Geschlecht und *Version*,  $F(20,949) = 3.87$ ,  $p < .001$ ,  $\eta^2 = .08$ . Für die Überprüfung der sozialen Erwünschtheit ist der Haupteffekt des Faktors *Version* von zentraler Bedeutung, der die a-priori-Klassifikation der Eigenschaften bestätigt. Er zeigt an, dass die jeweilige Eigenschaft von beiden Geschlechtern signifikant erwünschter für ein Geschlecht als für das andere angesehen wird. Probanden beiderlei Geschlechts hielten die maskulinen Eigenschaften insgesamt signifikant stärker erwünscht für Männer als für Frauen. In Tabelle 11 sind die mittleren Einschätzungen der sozialen Erwünschtheit der maskulinen Eigenschaften für Männer und Frauen durch beide Geschlechter dargestellt.

Tabelle 11. Statistische Kennwerte der Einschätzung soz. Erwünschtheit mas. Items (BSRI)

Eigenschaft	Trennschärfe	sozial erwünscht für Männer		sozial erwünscht für Frauen		Diff.	p Version	p Sex	sign. IA
		M	SD	M	SD				
hat Führungseigenschaften	.69	5.83	1.00	5.47	1.29	.36	<.001	<.001	<.001
tritt bestimmt auf	.72	5.72	1.07	5.25	1.29	.47	<.001	<.001	<.001
ehrgeizig	.64	5.93	1.04	5.70	1.17	.23	<.001	<.01	n.s.
Respekt einflößend	.66	5.20	1.28	4.38	1.43	.82	<.001	<.001	<.001
kann andere kritisieren ohne sich unbehaglich zu fühlen	.49	5.25	1.30	5.06	1.38	.18	<.05	n.s.	<.001
verteidige meine Meinung	.65	5.87	.92	5.77	1.18	.11	n.s.	<.001	n.s.
entschlossen	.72	6.08	.81	5.72	1.15	.36	<.001	<.001	<.01
sachlich	.58	5.44	1.14	4.99	1.36	.45	<.001	n.s.	<.01
nicht leicht beeinflussbar	.51	5.32	1.34	5.24	1.45	.07	n.s.	n.s.	<.05
unerschrocken	.66	5.76	1.06	5.05	1.29	.72	<.001	<.05	n.s.
intelligent	.44	6.27	.96	6.28	.89	.00	n.s.	n.s.	n.s.
hartnäckig	.62	5.33	1.10	5.13	1.25	.20	<.01	<.05	<.001
ist bereit, etwas zu riskieren	.54	5.54	1.00	4.97	1.34	.57	<.001	n.s.	n.s.
kraftvoll	.59	5.95	.95	5.01	1.31	.94	<.001	<.001	<.01
furchtlos	.64	5.67	1.12	4.81	1.39	.86	<.001	<.05	<.01
scharfsinnig	.57	5.92	.95	5.74	1.08	.18	<.01	n.s.	<.01
wetteifernd	.42	5.04	1.34	4.52	1.29	.52	<.001	n.s.	<.01
sicher	.59	6.00	.99	5.84	1.09	.16	<.05	<.01	n.s.
zeigt geschäftsmäßiges Verhalten	.59	5.39	1.25	4.97	1.30	.43	<.001	<.05	n.s.
konsequent	.63	5.77	1.01	5.67	1.10	.10	n.s.	<.05	<.001

Insgesamt gelten 16 der 20 maskulinen Eigenschaften als signifikant sozial erwünschter für Männer als für Frauen. Vier Items entsprachen nicht der vorgegebenen höheren Erwünschtheit für Männer: *intelligent* sowie *konsequent*, *verteidigt (seine/ihre) eigene Meinung* und *bin nicht leicht beeinflussbar*.

Der Haupteffekt Geschlecht, der besagt, dass Männer und Frauen sich darin unterschieden, wie sozial erwünscht eine einzelne Eigenschaft überhaupt für Menschen (über die Kategorien Männer / Frauen hinweg) eingeschätzt wurde, ging auf signifikante Unterschiede

zwischen Männern und Frauen in den Erwünschtheitsurteilen (über beide Geschlechter) für 13 Eigenschaften zurück<sup>29</sup>.

Der Interaktionseffekt von Geschlecht und Version war für 13 von 20 Eigenschaften signifikant und verlief nahezu gleichmäßig dergestalt, dass beide Geschlechter eine Eigenschaft für das eigene Geschlecht sozial erwünschter bewerteten, als es die Angehörigen des jeweils anderen Geschlechts taten. So hielten etwa die Männer die Eigenschaft *wetteifernd* für sozial erwünschter für Männer als Frauen dies taten. Umgekehrt hielten Frauen dieselbe Eigenschaft als sozial erwünschter für Frauen als Männer dies taten.

### 3.1.1.3.2 Analyse der maskulinen Eigenschaften: Typizität

Die MANOVA für die 20 maskulinen Eigenschaften ergab signifikante multivariate Effekte für die Variable, Version (Einschätzung für Männer vs. Frauen),  $F(20,441) = 31.94$ ,  $p < .001$ ,  $\eta^2 = .59$ , sowie für die Interaktion von Geschlecht und Version,  $F(20,441) = 6.14$ ,  $p < .001$ ,  $\eta^2 = .22$ . Auch für die vorliegende Fragestellung bezüglich der Typizität sind die Haupteffekte des Faktors *Version* von zentraler Bedeutung. Sie zeigen an, ob die jeweilige Eigenschaft von beiden Geschlechtern signifikant typischer für Männer als für Frauen angesehen wird.

Tabelle 12. Statistische Kennwerte der Einschätzung der Typizität mas. Items (BSRI)

Eigenschaft	Trennschärfe	typisch für Männer		typisch für Frauen		Diff.	p Version	p Sex	sign. IA
		M	SD	M	SD				
hat Führungseigenschaften	.56	4.95	1.05	4.82	1.30	.13	n.s.	<.001	<.001
tritt bestimmt auf	.53	4.78	1.04	4.19	.98	.59	<.001	n.s.	<.01
ehrgeizig	.35	4.84	1.13	5.09	1.11	-.25	<.05	n.s.	<.05
Respekt einflößend	.47	4.09	1.42	3.48	1.20	.61	<.001	n.s.	<.001
kann andere kritisieren ohne sich unbehaglich zu fühlen	.31	5.12	1.24	3.99	1.45	1.13	<.001	<.001	<.05
verteidige meine Meinung	.51	5.40	1.10	4.71	1.19	.69	<.001	n.s.	n.s.
entschlossen	.65	4.98	1.05	4.54	1.05	.44	<.001	n.s.	<.001
sachlich	.60	5.19	1.18	3.66	1.25	1.53	<.001	<.001	<.001
nicht leicht beeinflussbar	.39	3.90	1.38	3.57	1.18	.33	<.01	n.s.	<.05
unerschrocken	.64	4.58	1.11	3.62	1.19	.96	<.001	<.05	<.001
intelligent	.22	4.48	.81	5.20	1.11	-.72	<.001	n.s.	<.001
hartnäckig	.55	4.96	1.02	4.81	1.11	.15	n.s.	n.s.	<.01
ist bereit, etwas zu riskieren	.63	5.09	1.15	3.99	1.23	1.10	<.001	<.001	<.001
kraftvoll	.65	5.15	1.11	3.95	1.35	1.20	<.001	<.001	<.001
furchtlos	.59	4.64	1.15	3.43	1.12	1.21	<.001	n.s.	<.001
scharfsinnig	.26	4.18	.90	4.70	1.03	-.52	<.001	<.05	<.001
wetteifernd	.40	5.62	1.11	4.51	1.46	1.11	<.001	n.s.	<.05
sicher	.57	4.66	.94	4.25	1.08	.41	<.001	n.s.	<.001
zeigt geschäftsmäßiges Verhalten	.54	4.90	1.02	4.26	1.22	.64	<.001	<.01	<.001
konsequent	.57	4.69	1.12	4.14	1.25	.55	<.001	<.05	<.001

In Tabelle 12 sind die mittleren Einschätzungen der Typizität der maskulinen Eigenschaften für Männer und Frauen durch beide Geschlechter dargestellt. Der Haupteffekt *Version* ging auf signifikante univariate Effekte hinsichtlich der Typizität für Männer vs. Frauen für immerhin 18 der ursprünglich 20 Items zurück. 15 der univariaten Effekte fielen erwartungsgemäß im Sinne höherer Typizitätsratings für Männer aus. Zwei Items *hartnäckig* und *hat Führungseigenschaften* wiesen keine signifikant unterschiedlichen Bewertungen der Eigenschaften als typischer für das eine als für das andere Geschlecht auf. Schließlich zeigen

<sup>29</sup> Da der HE Geschlecht nichts über die differenzielle Bedeutung der Eigenschaften für Männer und Frauen aussagt, ist er im Rahmen dieser Analysen theoretisch irrelevant. Aus diesem Grund wird auf die Haupteffekte des Geschlechts in diesen und allen folgenden Analysen nicht weiter eingegangen.

die Ergebnisse für die drei Items *ehrgeizig*, *intelligent* und *scharfsinnig* – entgegen der Annahme des BSRI – höhere Typizitätsratings für Frauen als für Männer. Fünf Items (20%) erfüllen demnach nicht die Voraussetzungen der klar zuzuweisenden Typizität maskuliner Items für Männer.

Der Interaktionseffekt von Geschlecht und Version war für 19 von 20 Eigenschaften signifikant und verlief nahezu durchgängig (bis auf *kann andere kritisieren ohne sich unbehaglich zu fühlen*) in die gleiche Richtung: Erwartungsgemäß schätzten beide Geschlechter eine Eigenschaft für das eigene Geschlecht als typischer ein, als es die Angehörigen des jeweils anderen Geschlechts taten. So bewerteten etwa die Männer die Eigenschaft *sicher* als typischer für Männer, als Frauen dies taten. Umgekehrt erachteten Frauen dieselbe Eigenschaft als typischer für Frauen, als Männer dies taten.

Bei der Reliabilitätsprüfung der Skala fielen zudem die schlechten Trennschärfen einiger Items auf: *scharfsinnig*, *ehrgeizig* und *intelligent*. Deren Einschätzungen verdeutlichten bereits, dass diese Attribute eher den Frauen als den Männern zugeschrieben werden müssen. Darüber hinaus waren weitere Items ebenfalls wenig homogen im Vergleich zur Gesamtskala ( $r < .40$ , z. B. *kann andere kritisieren ohne sich unbehaglich zu fühlen* sowie *nicht leicht beeinflussbar*).

### 3.1.1.3.3 Analyse der femininen Eigenschaften: soz. Erwünschtheit

Die MANOVA für die femininen Eigenschaften ergab ebenfalls signifikante multivariate Effekte für die Variable Version,  $F(20,949) = 16.92$ ,  $p < .001$ ,  $\eta^2 = .26$ , sowie für die Interaktion von Geschlecht und Version,  $F(20,949) = 2.77$ ,  $p < .001$ ,  $\eta^2 = .06$ .

Der Haupteffekt Version zeigt, dass die jeweilige Eigenschaft von beiden Geschlechtern signifikant erwünschter für ein Geschlecht als für das andere angesehen wird. In Tabelle 13 sind die mittleren Einschätzungen der sozialen Erwünschtheit der femininen Eigenschaften für Männer und Frauen durch beide Geschlechter aufgeführt. Sowohl männliche als auch weibliche Probanden bewerteten 17 der 20 femininen Eigenschaften signifikant erwünschter für Frauen als für Männer. Die drei Items *selbstaufopfernd*, *abhängig* und *bescheiden* wurden als gleichermaßen erwünscht für beide Geschlechter eingeschätzt.

Der Interaktionseffekt von Geschlecht und Version war für zwölf von 20 Eigenschaften signifikant. Im Gegensatz zu den maskulinen Items zeigten sich die Interaktionen weniger gleichförmig, da sich verschiedene Zuschreibungsarten wiederfinden ließen. Im Vergleich zu den maskulinen Items zeigte sich bei keinem der femininen Items die Tendenz, diese Eigenschaft hinsichtlich der sozialen Erwünschtheit jeweils für das eigene Geschlecht vergleichsweise stärker zu beanspruchen. Eher das Gegenteil war der Fall: neun Items (*empfindsam*, *abhängig*) wurden sozial erwünschter für das jeweils andere Geschlecht gesehen. *Feinfühlig* schätzten Frauen sowohl für Männer als auch für Frauen sozial erwünschter ein als Männer dies taten, während das Item *empfänglich für Schmeicheleien* von beiden Geschlechtern als sozial erwünschter für Frauen eingeschätzt wurde.

Auffällig ist, dass die femininen Items im Wesentlichen geringe Maße als sozial erwünscht dem jeweils eigenen Geschlechtsgruppen zugeschrieben werden. Hinzu kommt, dass die Mittelwerte der Einschätzungen für die jeweiligen Geschlechter nur geringfügig über

den Skalenmittelwerten lagen. Das Item *abhängig* wies sich diesbezüglich durch besonders schlechte Kennwerte aus. Nicht allein ist die Trennschärfe als unzureichend zu bewerten, sondern auch die Mittelwerte, die sogar unterhalb des Skalenmittelwertes lagen. Das deutet an, dass diese Eigenschaft für beide Geschlechter unerwünscht ist.

Tabelle 13. Statistische Kennwerte der Einschätzung soz. Erwünschtheit fem. Items (BSRI)

Eigenschaft	Trennschärfe	sozial erwünscht für Männer		sozial erwünscht für Frauen		Diff.	p Version	p Sex	sign. IA
		M	SD	M	SD				
romantisch	.66	4.95	1.48	5.54	1.20	.58	<.001	n.s.	<.01
abhängig	.20	2.55	1.35	2.70	1.57	.15	n.s.	<.05	<.01
weichherzig	.64	3.97	1.47	4.87	1.43	.90	<.001	n.s.	<.01
bemüht sich, verletzte Gefühle zu besänftigen	.65	4.94	1.49	5.65	1.09	.72	<.001	<.05	n.s.
glücklich	.42	6.14	1.08	6.37	.89	.23	<.001	<.001	n.s.
feinfühlig	.65	4.84	1.55	5.78	1.05	.93	<.001	<.01	<.05
sinnlich	.69	4.86	1.51	5.83	1.06	.97	<.001	<.05	<.001
fröhlich	.49	5.85	1.08	6.10	.91	.25	<.001	<.001	n.s.
nachgiebig	.49	3.73	1.30	4.01	1.51	.28	<.01	n.s.	<.01
bescheiden	.44	4.63	1.47	4.77	1.39	.14	n.s.	n.s.	n.s.
empänglich für Schmeicheleien	.47	3.84	1.50	4.63	1.64	.79	<.001	<.05	<.001
empfindsam	.61	4.55	1.43	5.47	1.18	.92	<.001	n.s.	<.01
selbstaufopfernd	.44	4.08	1.57	4.18	1.68	.10	n.s.	<.001	n.s.
vermeidet barsche Worte	.47	4.63	1.65	4.99	1.48	.36	<.001	n.s.	n.s.
verspielt	.52	4.01	1.64	4.80	1.42	.79	<.001	n.s.	<.001
verführerisch	.58	5.18	1.39	5.79	1.10	.60	<.001	n.s.	<.01
achtet auf äußere Erscheinung	.39	5.45	1.19	5.96	.99	.51	<.001	<.05	<.05
leidenschaftlich	.61	5.66	1.23	5.88	.96	.23	<.001	<.01	<.001
herzlich	.67	5.27	1.38	6.02	.89	.75	<.001	<.001	n.s.
liebt Sicherheit	.42	4.69	1.39	5.20	1.21	.52	<.001	<.05	n.s.

### 3.1.1.3.4 Analyse der femininen Eigenschaften: Typizität

Die MANOVA für die femininen Eigenschaften ergab signifikante multivariate Effekte für die Variable Version,  $F(20,442) = 39.34$ ,  $p < .001$ ,  $\eta^2 = .64$ , sowie für die Interaktion von Geschlecht und Version,  $F(20,442) = 4.18$ ,  $p < .001$ ,  $\eta^2 = .16$ . Tabelle 14 listet die mittleren Typizitätswerte der femininen Eigenschaften für Männer und Frauen durch beide Geschlechter auf.

Tabelle 14. Statistische Kennwerte der Einschätzung der Typizität fem. Items (BSRI)

Eigenschaft	Trennschärfe	typisch für Männer		typisch für Frauen		Diff.	p Version	p Sex	sign. IA
		M	SD	M	SD				
romantisch	.72	3.70	.97	5.41	1.12	-1.71	<.001	n.s.	n.s.
abhängig	-.04	4.06	1.47	3.69	1.30	.37	<.01	<.05	n.s.
weichherzig	.62	3.57	1.12	4.83	1.24	-1.26	<.001	n.s.	<.05
bemüht sich, verletzte Gefühle zu besänftigen	.68	3.32	1.30	5.10	1.33	-1.78	<.001	<.01	<.001
glücklich	.09	4.59	1.01	4.41	.96	.18	n.s.	n.s.	<.05
feinfühlig	.73	3.44	.99	5.25	1.09	-1.81	<.001	<.001	n.s.
sinnlich	.71	3.62	1.10	5.44	1.09	-1.82	<.001	n.s.	n.s.
fröhlich	.37	4.67	.92	4.90	.94	-.23	<.01	n.s.	<.05
nachgiebig	.41	3.65	1.17	4.39	1.20	-.74	<.001	<.001	n.s.
bescheiden	.37	3.15	1.07	3.86	1.26	-.71	<.001	<.001	<.001
empänglich für Schmeicheleien	.38	5.11	1.40	5.70	1.08	-.59	<.001	n.s.	n.s.
empfindsam	.73	3.95	1.17	5.51	1.02	-1.56	<.001	<.05	n.s.
selbstaufopfernd	.63	3.39	1.18	4.83	1.38	-1.44	<.001	n.s.	<.001
vermeidet barsche Worte	.55	2.68	1.12	4.40	1.40	-1.72	<.001	n.s.	n.s.
verspielt	.14	4.58	1.50	4.60	1.14	-.02	n.s.	n.s.	n.s.
verführerisch	.60	4.06	1.32	5.51	1.16	-1.45	<.001	n.s.	<.01
achtet auf äußere Erscheinung	.63	4.77	1.10	6.00	.91	-1.23	<.001	n.s.	<.05
leidenschaftlich	.69	4.13	1.09	5.30	1.03	-1.17	<.001	n.s.	n.s.
herzlich	.73	3.97	.99	5.24	.98	-1.27	<.001	<.001	n.s.
liebt Sicherheit	.48	4.68	1.31	5.63	1.08	-.95	<.001	<.01	n.s.



Der Haupteffekt *Version* begründet sich in signifikanten Unterschieden bei 18 der 20 femininen Eigenschaften. Probanden beiderlei Geschlechts hielten die femininen Eigenschaften im Schnitt signifikant stärker für frauentypisch als für männertypisch. Ausnahmen bildeten die Items *glücklich* und *verspielt*, deren geschlechtsspezifischen Typizitätseinschätzungen am Signifikanzniveau scheiterten. Zudem wurde das Attribut *abhängig* als typischer für Männer eingeschätzt. Damit ist festzuhalten, dass drei der 20 Items der Vorgabe, typischer für Frauen als für Männer zu sein, nicht entsprechen.

Der Interaktionseffekt von Geschlecht und Version war für acht der 20 Eigenschaften signifikant. Wie es bereits für die maskulinen Eigenschaften der Fall war, lagen zumeist die Mittelwerte der Männer in der Einschätzung der positiven Eigenschaften als männertypisch über denen der Frauen, während sich für die Einschätzung der Eigenschaften als frauentypisch ein umgekehrtes Bild ergab (z. B. *weichherzig*). Bei zwei Eigenschaften jedoch schrieb ein Geschlecht die Eigenschaft jeweils dem anderen Geschlecht als typischer zu: *verführerisch* und *achtet auf äußere Erscheinung*.

Die Reliabilitätsprüfung der Skala offenbarte – neben den schon bekannten schwachen Items *glücklich*, *verspielt* und *abhängig* – auch die Eigenschaften *fröhlich*, *bescheiden* sowie *empänglich für Schmeicheleien* als fragwürdig, da deren geringe Trennschärfen die Homogenität der Gesamtskala beeinträchtigen.

### 3.1.1.3.5 Analyse der sozialen Erwünschtheit der Teilskalen

An dieser Stelle wird die soziale Erwünschtheit der Teilskalen betrachtet, da dies der Aspekt ist, auf dessen Grundlage die Items für den BSRI ausgewählt wurden. Dazu wurden die Einschätzungen der sozialen Erwünschtheit für die maskulinen bzw. femininen Items in Mittelwerten über die jeweils 20 Items des Original-BSRI zusammengefasst. Die deskriptiven Statistiken sind in Tabelle 15 dargestellt. Beide Teilskalen korrelierten signifikant negativ miteinander ( $r = -.14, p < .001$ )

Tabelle 15. Mittelwerte der Einschätzung der soz. Erwünschtheit der BSRI-Teilskalen FEM und MAS

	$\alpha$	<i>M</i>	<i>SD</i>
Maskulinität (Erwünschtheit)	.92	5.47	.77
Femininität (Erwünschtheit)	.90	4.96	.80

Die MANOVA für die Skalen Maskulinität (Mas) und Femininität (Fem) ergab einen erwarteten signifikanten Effekt für das Geschlecht,  $F(2,995) = 7.51, p < .001, \eta^2 = .02$ , die Version  $F(2,995) = 103.91, p < .001, \eta^2 = .17$  sowie die Interaktion des Geschlechts und der Version  $F(2,995) = 18.92, p < .001, \eta^2 = .04$ . In Tabelle 16 sind die Mittelwerte und Standardabweichungen für die Teilskalen getrennt nach den Geschlechtern aufgeführt. Im unteren Teil der Tabelle sind zum Vergleich die Werte aus den Studien von Bem (1974) sowie Schneider-Düker und Kohler (1988) aufgeführt.

Der univariate Haupteffekt Geschlecht für die AV Maskulinität besagt, dass Frauen diese insgesamt ( $M = 5.52, SD = .76$ ) als sozial erwünschter bewerteten als Männer ( $M = 5.32,$

$SD = .79$ ),  $F(1,996) = 12.94$ ,  $p < .001$ ,  $\eta^2 = .01$ . Der univariate Haupteffekt Geschlecht bei der Femininität war nicht signifikant ( $M_{\text{Männer}} = 4.89$ ,  $SD = .85$  vs.  $M_{\text{Frauen}} = 4.98$ ,  $SD = .78$ ).

Tabelle 16. Mittelwerte der Einschätzung der soz. Erwünschtheit der BSRI-Skalen FEM und MAS

	Maskulinität				Femininität			
	sozial erwünscht für							
	Mann		Frau		Mann		Frau	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Männer	5.71	.60	4.91	.76	4.44	.84	5.37	.54
Frauen	5.65	.63	5.38	.86	4.77	.83	5.19	.68
Männer	(Bem, 1974)		2.90		3.63		5.61	
Frauen	5.83		3.46		3.74		5.55	
Männer	(Schneider-Düker & Kohler, 1988)		3.22		3.23		5.29	
Frauen	5.35		2.84		3.17		5.59	

Die univariaten Haupteffekte der Version für Maskulinität,  $F(1,996) = 85.51$ ,  $p < .001$ ,  $\eta^2 = .08$ , und Femininität,  $F(1,996) = 133.29$ ,  $p < .001$ ,  $\eta^2 = .12$ , besagen jeweils, dass insgesamt die maskulinen Items erwünschter für Männer und die femininen Items erwünschter für Frauen eingeschätzt wurden.

Sowohl für Maskulinität,  $F(1,996) = 21.23$ ,  $p < .001$ ,  $\eta^2 = .02$ , als auch für die Femininität,  $F(1,996) = 18.65$ ,  $p < .001$ ,  $\eta^2 = .02$ , ergaben sich signifikante univariate Interaktionen von Geschlecht und Version. Für die Maskulinität bedeutet dies, dass Männer und Frauen die maskulinen Eigenschaften als signifikant weniger erwünscht für eine Frau einschätzten als für einen Mann. Für die Femininität bedeutet dies, dass Männer die femininen Attribute für Frauen deutlich erwünschter hielten als für Männer ( $M = 5.37$  vs.  $M = 4.44$ ), während Frauen den Unterschied, wenn auch signifikant, weniger stark ausgeprägt bewerteten ( $M = 5.19$  vs.  $M = 4.77$ ).

Insgesamt zeigte sich, dass Frauen in der Bewertung der Maskulinität für einen Mann in der Gesellschaft den ebenfalls einschätzenden Männern sehr ähnlich waren. In Bezug auf die Maskulinitätseinschätzung für Frauen schätzten die Frauen diese Eigenschaften zwar immer noch als sozial erwünschter für einen Mann ein, bewerteten die Eigenschaften aber auch sozial erwünschter für eine Frau als die Männer dies taten. Im Falle der Femininität ist das Bild etwas eindeutiger. Sowohl Männer als auch Frauen sahen die femininen Eigenschaften insgesamt sozial erwünschter für eine Frau als für einen Mann, wobei männliche Versuchspersonen diese Attribute für einen Mann in der Gesellschaft im stärkeren Maße als weniger erwünscht erachteten als Frauen. Zusätzlich kann festgestellt werden, dass alle berichteten Mittelwerte der Skalen, berücksichtigt man die Skalenbreite von eins bis sieben, signifikant über dem Mittelpunkt lagen und damit als insgesamt sozial erwünscht für beide Geschlechter angesehen werden müssen.

Bemerkenswert sind zudem die Ausmaße der Erwünschtheitszuschreibungen von Maskulinität und Femininität: während sowohl in den Stichproben von Bem (1974) als auch Schneider-Düker und Kohler (1988) Mittelwertsunterschiede von mindestens zwei Skalenpunkten in den Einschätzungen der Erwünschtheit für Männer und Frauen festzustellen waren, hat sich diese Differenz in der neuen Stichprobe auf Unterschiede von kleiner als eins reduziert. Stellt man die in Tabelle 16 aufgeführten Mittelwerte der drei Studien als für die

Datenform eher unübliches Liniendiagramm grafisch dar (Abbildung 10), werden diese Unterschiede deutlich.

Sowohl Bem (1974) als auch Schneider-Düker und Kohler (1988) mehr als zehn Jahre später konnten bei ihren Items klare Unterschiede bei den Erwünschtheitseinschätzungen der Teilskalen jeweils getrennt für Männer und Frauen nachweisen, während sich die deutschen Attribute von Schneider-Düker und Kohler (1988) weitere 20 Jahre später in ihrer Erwünschtheit für beide Geschlechter deutlich angenähert haben.

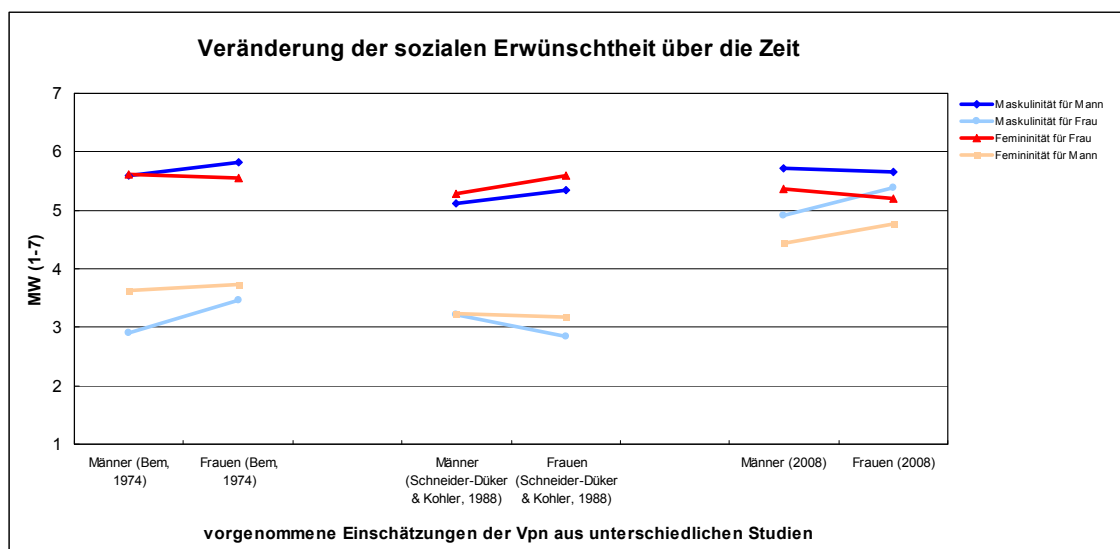


Abbildung 10. Darstellung der Veränderung der sozialen Erwünschtheit über die Zeit

### 3.1.1.3.6 Analyse der Typizität der Teilskalen

Obwohl Bem (1974) die Typizität der Teilskalen nicht berücksichtigt hatte, soll hier deren Betrachtung auf Skalenebene erfolgen, um so dem Kriterium der Typizität für die Neukonstruktion erste Anhaltspunkte zu liefern.

Dazu wurden die Einschätzungen der Typizität für die maskulinen bzw. femininen Items in Mittelwerten über die jeweils 20 Items des Original-BSRI zusammengefasst. Die deskriptiven Statistiken sind in Tabelle 17 dargestellt.

Tabelle 17. Mittelwerte der Einschätzung der Typizität der BSRI-Teilskalen FEM und MAS

	$\alpha$	$M$	$SD$
Maskulinität (Typizität)	.89	4.31	.51
Femininität (Typizität)	.89	4.42	.42

Zunächst wurde der Zusammenhang der beiden Teilskalen mittels einer Korrelationsanalyse überprüft. Maskulinität korrelierte signifikant positiv mit Femininität ( $r = .58, p < .01$ ). Da das Alter nicht mit Maskulinität und nur geringfügig signifikant mit der Femininitätsskala korreliert, wurde auch in den Analysen zur Typizität der BSRI-Attribute das Alter nicht berücksichtigt.

Die MANOVA für die Skalen Maskulinität und Femininität offenbarte den erwarteten signifikanten Effekt für das Geschlecht,  $F(2,467) = 6.37, p < .01, \eta^2 = .03$ , die Version  $F(2,467) = 82.04, p < .001, \eta^2 = .26$  sowie die Interaktion des Geschlechts und der Version

$F(2,467) = 24.69$ ,  $p < .001$ ,  $\eta^2 = .10$ . In Tabelle 18 sind die Mittelwerte und Standardabweichungen für die Teilskalen getrennt nach den Geschlechtern aufgeführt.

Der univariate Haupteffekt Geschlecht bei der Maskulinität war nicht signifikant ( $M_{\text{Männer}} = 4.28$ ,  $SD = .45$  vs.  $M_{\text{Frauen}} = 4.32$ ,  $SD = .53$ ). Der univariate Haupteffekt Geschlecht für die AV Femininität besagt, dass Frauen diese insgesamt ( $M = 4.47$ ,  $SD = .42$ ) als typischer bewerteten als Männer ( $M = 4.33$ ,  $SD = .40$ ),  $F(1,468) = 12.61$ ,  $p < .001$ ,  $\eta^2 = .03$ .

Tabelle 18. Mittelwerte der Einschätzung der Typizität der BSRI-Skalen FEM und MAS

	Maskulinität				Femininität			
	Mann		Frau		Mann		Frau	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Männer	4.26	.43	4.30	.46	4.53	.38	4.19	.36
Frauen	4.03	.39	4.63	.49	4.40	.39	4.55	.45

Der univariate Effekt der Version für Maskulinität,  $F(1,468) = 105.92$ ,  $p < .001$ ,  $\eta^2 = .19$  war zwar signifikant, überraschte jedoch mit dem Ergebnis, dass wider der Erwartung Maskulinität in den Einschätzungen im Allgemeinen typischer für Frauen ( $M = 4.51$ ,  $SD = .50$ ) als für Männer ( $M = 4.10$ ,  $SD = .41$ ) ist. Das Bild bezüglich der Femininität war umso bemerkenswerter, da der Effekt der Version keine Signifikanz erreichte,  $F(1,468) = 0.00$ ,  $p = \text{n.s.}$ ,  $\eta^2 = .00$ ), sich somit der Mittelwert in der Einschätzung hinsichtlich der Typizität der Items für Männer ( $M = 4.43$ ,  $SD = .39$ ) nicht vom Mittelwert der Typizität für Frauen ( $M = 4.42$ ,  $SD = .45$ ) unterschied.

Beide Interaktionseffekte waren ebenfalls signifikant. Es zeigte sich ein klares Muster dahingehend, dass sowohl hinsichtlich der Femininität,  $F(1,468) = 38.14$ ,  $p < .001$ ,  $\eta^2 = .08$  als auch der Maskulinität,  $F(1,468) = 40.81$ ,  $p < .001$ ,  $\eta^2 = .08$  Frauen und Männer die jeweiligen Items einer Skala grundsätzlich als typischer für das eigene Geschlecht beanspruchten. Lediglich in der Gruppe der Männer zeigte sich eine Auffälligkeit dahingehend, dass sie in der Typizität der maskulinen Eigenschaften keinen Unterschied für Frauen und Männer in der Einschätzung machten.

Im Gegensatz zur sozialen Erwünschtheit kann zusammengefasst werden, dass auch hinsichtlich der Typizität zwar alle berichteten Mittelwerte über dem Skalenmittelwert liegen und damit insgesamt als typisch angesehen werden, die Einschätzungen bezüglich der beiden Geschlechter jedoch etwas weniger eindeutig ausfallen.

### 3.1.1.4 Diskussion

Ziel dieser Studie war die Überprüfung der deutschen Eigenschaften des BSRI in Hinblick auf ihre soziale Erwünschtheit und ihre Typizität für die heutige Gesellschaft. Es wurde erwartet, dass einzelne Items der femininen und maskulinen Teilskalen nicht mehr in dem Maße zwischen den Geschlechtern differenzieren, wie Bem (1974) und Schneider-Düker und Kohler (1988) es als Voraussetzung für die Skalenbildung definiert hatten.

Die aktuelle Überprüfung der BSRI-Items wirft Zweifel an der Eignung des Instruments zur Erfassung unterschiedlicher sozial erwünschter Aspekte des GSK in der heutigen Zeit auf. So bietet sich beispielsweise in diesem Zusammenhang ein Vergleich mit den Daten aus Bems Originalstudie (Bem, 1974) sowie den Ergebnissen von Schneider-Düker und Kohler (1988) an,

die im unteren Teil von Tabelle 16 aufgeführt sind. Die Einschätzungen der geschlechtskongruenten Teilskalen (z. B. Femininität für Frauen) wurden von Männern und Frauen übereinstimmend vorgenommen. Auffällig dagegen sind jedoch die Zuschreibungen der Eigenschaften für das jeweils andere Geschlecht. Während Bem hier noch eindeutige Unterschiede dergestalt aufzeigen konnte, dass Femininität übereinstimmend von den Probanden als deutlich weniger sozial erwünscht für Männer als für Frauen angesehen wurde, hat sich diese Differenz in der Stichprobe aus der hier vorliegenden aktuellen Stichprobe stark verringert. So werden feminine Eigenschaften zwar immer noch als sozial erwünschter für Frauen bewertet, jedoch weitaus weniger nicht erwünscht auch für Männer.

Einer der auffälligsten Effekte war in diesem Zusammenhang zudem, dass die maskulinen Eigenschaften von den Frauen sozial erwünscht für die eigene Gruppe angesehen wurden. So beurteilten sie zwar die maskulinen Eigenschaften als sozial erwünschter für Männer als für Frauen, der Unterschied fiel jedoch im Vergleich zu früheren Studien deutlich geringer aus.

Insgesamt sind die Differenzen zwischen den Geschlechtern innerhalb einer Kategorie wesentlich kleiner geworden sind, wodurch das Instrument seine Prägnanz eingebüßt hat. Die seit nunmehr fast 40 Jahren währende Popularität des BSRI im Bereich der Geschlechtsrollenforschung hält einer empirischen Aktualitätsprüfung nicht stand.

Die Analysen zeigen auf, dass die femininen Items eine insgesamt niedrigere soziale Erwünschtheit aufweisen als die maskulinen Items. Einige skalenspezifische Attribute müssen nach diagnostischen Gesichtspunkten gar aus den Skalen ausgeschlossen werden. Bestimmte Items evozierten nicht den erwarteten Geschlechtsunterschied in der Typizitätsbeurteilung (z. B. maskuline Items *trete bestimmt auf*, *ehrgeizig*, *sicher*), wiesen unzureichende Faktorladungen auf ihren jeweiligen Teilskalen auf (feminine Items: *nachgiebig*, *bescheiden* und *liebe Sicherheit*) oder Faktorladungen auf jeweils anderen als der zugehörigen Teilskalen (*glücklich*). Das Item *abhängig* ist sogar über mehrere Kriterien als fragwürdig einzustufen. Eine mangelhafte Trennschärfe sowie Mittelwerte beider Geschlechter unterhalb des Skalenmittelwertes müssen unweigerlich – folgt man Bems Auswahlprozess – zum Ausschluss dieses Items führen. Eine solch detaillierte Itemanalyse haben beispielsweise auch Sethi und Allen (1984) schlussfolgern lassen, dass geltend für ihre Stichprobe einige Items (z. B. *adaptable*, *secretive*) keine Berechtigung zum Einschluss in den finalen Skalen haben. In einer Studie zu Wortnormierungen, die die semantische Bedeutung von Wörtern in den Vordergrund stellt, haben Leibetseder und Kollegen (Leibetseder et al., 1996) ebenfalls Auffälligkeiten bezüglich des Wortes *abhängig* gefunden und festgestellt, dass es insgesamt den geringsten Valenzwert besitzt und Männer wie auch Frauen dies übereinstimmend einschätzten.

In Bezug auf die Typizität konnten alle berichteten Mittelwerte als über dem Skalenmittelwert liegend nachgewiesen werden, gleichzeitig jedoch schätzten die Männer die maskulinen Items als nicht typischer für die eigene Gruppe ein. Auch hinsichtlich der Femininität blieben Fragen offen, da sich der Gesamtmittelwert in der Einschätzung der Items für Männer nicht vom Mittelwert der Typizität für Frauen unterschied.

Zusammenfassend lässt sich festhalten, dass die BSRI-Items eindeutige Mängel aufweisen und die Typizität nicht als gegeben vorausgesetzt werden kann. Während Bem (1974) und Schneider-Düker und Kohler (1988) diese nicht einer Überprüfung unterzogen hatten, zeigen vorliegende Ergebnisse, dass die berechnete Berücksichtigung der Typizität zu einer anderen Zuordnung der Items zu den Teilskalen führen würde.

Hinzu kommt, dass, obwohl Bem die Auswahl der Items auf ausschließlich positive Eigenschaften beschränkt hatte, nicht alle Beschreibungen in dieser eindeutigen Art und Weise mit einer Valenz für die entsprechenden Geschlechter belegt werden können. Des Weiteren differenziert der BSRI nicht mehr klar in der sozialen Erwünschtheit maskuliner und femininer Eigenschaften für die beiden Geschlechter.

### **3.1.2 Studie 2: Selbstbeschreibung mittels BSRI**

#### **3.1.2.1 Fragestellung**

Bereits Studie 1, die die BSRI-Items einer Aktualitätsüberprüfung entlang ausgewählter Kriterien des Konstruktionsprozesses von Bem (1974) sowie Schneider-Düker und Kohler (1988) unterzog, offenbarte deutliche Schwächen einiger Attribute, deren Verbleib in der Skala als mindestens fragwürdig, wenn nicht sogar unhaltbar diskutiert werden muss. In einer zweiten Studie sollte die Selbstbeschreibung auf den BSRI-Items weitere wichtige Informationen über Item- und Skalenkennwerte liefern. Während bei der Einschätzung der sozialen Erwünschtheit und der Typizität vor allem gesellschaftlich geteiltes, stereotypes Wissen aktiviert wird, sind es nun vor allem Überzeugungen von der eigenen Person, die bei der Selbstbeschreibung zum Tragen kommen. Die so vorgenommenen Einschätzungen könnten unter Umständen die in der Überprüfungsphase festgestellten Mängel einiger Items weiter verstärken oder dadurch relativieren, dass diese Items durch den individuellen Bezug im Rahmen der Selbstbeschreibung eine andere Bedeutung offenbaren.

Um als aktuelles Maß des GSK zu bestehen, sollten sich im BSRI-Einsatz folgende Ergebnisse herausstellen:

- (1) Interindividueller Vergleich: Frauen sollten sich auf der Femininitätsskala höher als Männer und Männer auf der Maskulinitätsskala höher als Frauen einschätzen.
- (2) Intraindividueller Vergleich: Frauen sollten sich auf der Femininitätsskala höher als auf der Maskulinitätsskala und Männer auf der Maskulinitätsskala höher als auf der Femininitätsskala einschätzen.

#### **3.1.2.2 Methode**

##### **3.1.2.2.1 Stichprobe**

Insgesamt nahmen 645 Personen an der Untersuchung teil: 207 von ihnen waren Männer, 438 Frauen. Die Versuchspersonen waren durchschnittlich 22.58 Jahre alt ( $SD = 2.93$  Jahre; Min: 18, Max: 43 Jahre), 78% gaben das Abitur als ihren höchsten Schulabschluss an, 8.4% konnten einen Hochschulabschluss vorweisen. Vergleichbar zur Studie 1 beschrieb sich die Mehrheit der Stichprobe (92%) bezüglich der sexuellen Orientierung als heterosexuell.

### 3.1.2.2.2 Instrumente

Es kam wiederum der BSRI in der deutschen Version (Schneider-Düker & Kohler, 1988) zum Einsatz. Die Instruktion für die Selbsteinschätzung lautete: „In diesem Teil der Befragung besteht Ihre Aufgabe darin, sich hinsichtlich jeder Eigenschaft selbst einzuschätzen, inwieweit diese ein Teil Ihrer Persönlichkeit ist. So könnte die Eigenschaft *fröhlich* zum Beispiel auch als folgende Frage formuliert sein: Halten Sie sich für einen *fröhlichen* Menschen?“ Erneut wurden alle 60 Items in Blöcken à 20 Items einzeln in tabellarischer Form präsentiert. Für die Beantwortung wurde eine siebenstufige Häufigkeitsskala (an die Originalstudien angelehnt) eingesetzt; (1) *fast nie*, (2) *gewöhnlich nicht*, (3) *manchmal, aber selten*, (4) *gelegentlich*, (5) *oft*, (6) *meistens*, (7) *immer*.

### 3.1.2.2.3 Durchführung

Die Probanden erhielten über verschiedene Onlinerekrutierungsmaßnahmen den Link zum Fragebogen, der nach den demografischen Informationen die 60 BSRI-Items zur Selbstbeschreibung enthielt. Nach Abschluss der Befragung gab es für alle teilnehmenden Personen die Möglichkeit sich für die Ergebnismeldung in eine Interessentenliste einzutragen. E-Mail-Adressen und Befragungsdaten wurden getrennt voneinander abgespeichert.

## 3.1.2.3 Ergebnisse

### 3.1.2.3.1 Skalenkennwerte von Maskulinität und Femininität

Zunächst sollen Analysen auf Itemebene zur Aufklärung der Struktur von Maskulinität und Femininität beitragen.

Tabelle 19. Statistische Kennwerte der Selbsteinschätzung anhand mas. BSRI-Items

Eigenschaft	Trennschärfe	Männer		Frauen		Differenz	p Sex
		M	SD	M	SD		
habe Führungseigenschaften	.60	4.83	1.16	4.65	1.20	.18	n.s.
trete bestimmt auf	.64	4.62	1.20	4.59	1.27	.03	n.s.
ehrgeizig	.44	4.97	1.31	5.14	1.17	-.17	n.s.
Respekt einflößend	.55	4.04	1.19	3.75	1.32	.29	<.05
kann andere kritisieren ohne mich unbehaglich zu fühlen	.46	4.63	1.52	4.05	1.54	.58	<.001
verteidige meine Meinung	.52	5.74	1.04	5.67	.97	.07	n.s.
entschlossen	.64	4.89	1.23	5.14	1.05	-.24	<.05
sachlich	.36	5.29	1.02	4.92	.94	.37	<.001
nicht leicht beeinflussbar	.34	4.67	1.42	4.55	1.36	.13	n.s.
unerschrocken	.50	4.98	1.11	4.39	1.25	.59	<.001
intelligent	.41	5.70	.92	5.66	.82	.04	n.s.
hartnäckig	.58	4.75	1.22	4.98	1.12	-.23	<.05
bin bereit, etwas zu riskieren	.49	4.70	1.34	4.46	1.26	.23	<.05
kraftvoll	.56	4.80	1.13	4.67	1.13	.14	n.s.
furchtlos	.47	4.48	1.17	3.93	1.25	.55	<.001
scharfsinnig	.54	5.34	1.08	5.16	.97	.18	n.s.
wetteifernd	.29	4.33	1.53	3.74	1.45	.58	<.001
sicher	.60	4.89	1.11	4.97	1.09	-.07	n.s.
zeige geschäftsmäßiges Verhalten	.37	4.36	1.38	4.09	1.37	.26	<.01
konsequent	.57	4.86	1.35	5.05	1.21	-.19	n.s.

Tabelle 19 gibt einen Überblick der Mittelwerte in der Selbstzuschreibung maskuliner Eigenschaften von Frauen und Männern. Wie bei den Typizitäts- und Erwünschtheitseinschätzungen in Studie 1 zeigten sich auch in der Selbsteinschätzung verschiedene Items als weniger für die Maskulinitätsskala geeignet als andere. So verfehlten zehn Items die Signifikanz

eines Geschlechtsunterschiedes (z. B: *trete bestimmt auf, ehrgeizig, sicher* oder das schon wiederholt als problematisch identifizierte Item *intelligent*). D. h. bezüglich dieser Items unterschieden sich Männer und Frauen in ihren Selbsteinschätzungen nicht. Zwei weitere Items (*hartnäckig, entschlossen*) wurden von den Frauen signifikant stärker als von den Männern zur Selbstbeschreibung herangezogen. Schließlich deutet die Trennschärfe des Items *wetteifern* an, dass es wenig zur Homogenität der maskulinen Skala beiträgt.

Tabelle 20 listet die statistischen Kennwerte für die femininen Eigenschaften auf. Bei insgesamt acht der 20 femininen Items unterschieden sich Männer und Frauen nicht in der Selbstzuschreibung (z. B. *abhängig* und *selbstaufopfernd*).

Tabelle 20. Statistische Kennwerte der Selbsteinschätzung anhand fem. BSRI-Items

Eigenschaft	Trennschärfe	Männer		Frauen		Differenz	p Sex
		M	SD	M	SD		
romantisch	.55	4.58	1.23	4.76	1.29	.18	n.s.
abhängig	.06	3.12	1.45	3.14	1.38	.02	n.s.
weichherzig	.46	4.83	1.28	5.09	1.25	.26	<.05
bemühe mich, verletzte Gefühle zu besänftigen	.45	5.37	1.19	5.50	1.11	.13	n.s.
glücklich	.22	5.03	1.32	5.25	1.04	.22	<.05
feinfühlig	.57	4.83	1.27	5.54	1.02	.70	<.001
sinnlich	.61	4.58	1.38	4.96	1.22	.38	<.01
fröhlich	.33	5.29	1.08	5.48	.88	.19	<.05
nachgiebig	.18	4.09	1.19	4.01	1.15	-.07	n.s.
bescheiden	.19	4.67	1.38	4.85	1.20	.18	n.s.
empänglich für Schmeicheleien	.26	4.99	1.46	5.26	1.30	.27	<.05
empfindsam	.53	4.98	1.26	5.60	1.06	.61	<.001
selbstaufopfernd	.34	4.38	1.30	4.38	1.32	.00	n.s.
vermeide barsche Worte	.25	4.44	1.47	4.59	1.42	.15	n.s.
verspielt	.32	5.03	1.35	4.89	1.28	-.14	n.s.
verführerisch	.39	4.01	1.41	4.42	1.25	.41	<.001
achte auf äußere Erscheinung	.26	5.56	1.10	5.88	.93	.33	<.001
leidenschaftlich	.55	4.93	1.31	5.23	1.17	.30	<.01
herzlich	.61	4.92	1.15	5.63	1.02	.71	<.001
liebe Sicherheit	.20	5.07	1.44	5.65	1.28	.58	<.001

Das Item *abhängig* zeichnete sich darüber hinaus erneut nicht nur durch eine mangelhafte Trennschärfe, sondern auch durch Mittelwerte beider Geschlechter unterhalb des Skalenmittelwertes aus. Das stellt die ursprünglich vorausgesetzte soziale Erwünschtheit dieses Items für die Skala wiederholt in Frage, dadurch dass – insgesamt betrachtet – die femininen Eigenschaften stärker zur Selbstbeschreibung herangezogen wurden als die maskulinen. Es fällt dann jedoch auf, dass die feminine Skala in dieser Stichprobe wesentlich weniger reliable ist als im Vergleich dazu die maskuline Skala. Neben *abhängig* sind auch die Trennschärfen der Eigenschaften *glücklich*, *nachgiebig* oder *liebe Sicherheit* wenig zufriedenstellend.

In einem nächsten Schritt wurden die nach dem BSRI definierten Skalen Maskulinität und Femininität einer Reliabilitätsprüfung unterzogen. Über die gesamte Stichprobe wies die Maskulinitätsskala ein zufriedenstellendes Cronbachs Alpha auf ( $r_{it} = .29 - .64$ ), während die feminine Skala immer noch gut, aber weniger homogen ausfiel ( $r_{it} = .06 - .61$ ). Diese Ergebnisse stimmen mit Bems Daten (1974) überein, deren feminine Teilskala ebenfalls eine geringere Reliabilität als die maskuline Teilskala aufwies. Das zeigte sich auch in den geschlechtsspezifischen Analysen (Tabelle 21).

Die Korrelationen der gebildeten Teilskalen offenbarten einen signifikanten negativen Zusammenhang zwischen Femininität und Alter ( $r = -.19$ ), d. h. mit höherem Alter fiel die



Zuschreibung femininer Attribute geringer aus. Maskulinität wies dagegen keinen Zusammenhang mit dem Alter auf. Des Weiteren zeigte sich, dass Femininität und Maskulinität in dieser Stichprobe nicht miteinander korreliert waren, was ebenfalls auch schon in der Konstruktionsstudie von Bem (1974) der Fall war.

Tabelle 21. Reliabilitäten und Mittelwerte von MAS und FEM (BSRI)

	Gesamt			Männer			Frauen			Vergleich		
	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$	M	SD	Diff.	p	$\eta^2$
Maskulinität	.88	4.76	.68	.89	4.90	.71	.88	4.69	.66	.21	<.001	.02
Femininität	.80	4.93	.58	.78	4.73	.62	.79	5.02	.54	-.29	<.001	.04
Mas	Bem <sup>30</sup>			4.97			4.57			.40		
Fem				4.44			5.01			-.57		
Mas	Schneider-Düker & Kohler <sup>31</sup>			4.30			4.14			.16		
Fem				4.17			4.46			-.29		

### 3.1.2.3.2 Geschlechtsunterschiede bei Maskulinität und Femininität

Eine multivariate Varianzanalyse für die beiden BSRI-Teilskalen mit dem Geschlecht als unabhängige Variable ergab einen signifikanten multivariaten Effekt,  $F(2,642) = 24.67$ ,  $p < .001$ ,  $\eta^2 = .07$ . Die beiden univariaten Effekte waren ebenfalls signifikant (Abbildung 11).

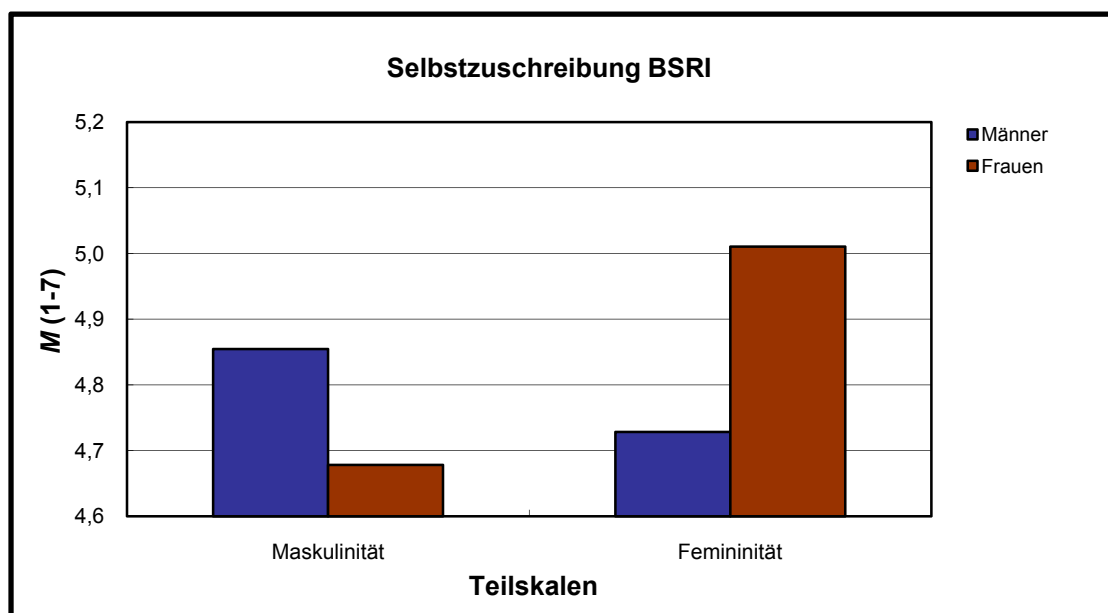


Abbildung 11. Selbstzuschreibung der BSRI-Eigenschaften

Männer schrieben sich die maskulinen Eigenschaften im stärkeren Maße zu als Frauen dies taten. Frauen wiesen hingegen höhere Werte auf der Femininitätsskala als die Männer auf. Der interindividuelle Vergleich bestätigt damit den BSRI als aktuelles Maß des GSK. Bemerkenswert ist, dass obwohl sich die Mittelwerte im Vergleich zu den Daten einer vergleichbaren Stichprobe aus der Entwicklung durch Schneider-Düker und Kohler (1988) leicht erhöhten, die Mittelwertdifferenzen zwischen den Geschlechtern nahezu konstant blieben. Frauen und Männer schrieben sich maskuline und feminine Eigenschaften zwar insgesamt stärker zu als noch vor 20 Jahren, jedoch ähnlich differenzierter Art wie es 1988 der Fall war.

<sup>30</sup> Werte aus der Stichprobe Stanford University (Bem, 1974; S. 161)

<sup>31</sup> Werte aus der Studenten-Stichprobe (Schneider-Düker und Kohler, 1988; S. 261)

Eine Varianzanalyse mit Messwiederholung offenbarte zusätzlich einen signifikanten Haupteffekt des Geschlechtsrollenselbstkonzepts derart, dass sich die Zuschreibungen von Maskulinität und Femininität voneinander unterscheiden,  $F(1,643) = 8.51$ ,  $p < .01$ ,  $\eta^2 = .01$ . Feminine Items wurden von allen Probanden stärker zur Selbstbeschreibung herangezogen als maskuline Items. Dieses Ergebnis widerspricht der Konstruktionsidee des BSRI, wonach die Messung des GSK mit zwei Teilskalen erfolgen soll, die jeweils gleichermaßen sozial erwünscht sind, demnach auch gleichstark vom jeweiligen Geschlecht für das eigene GSK beansprucht werden sollten. Des Weiteren zeigte sich eine signifikante Interaktion des GSK mit dem Geschlecht,  $F(1,643) = 42.01$ ,  $p < .01$ ,  $\eta^2 = .06$ . Frauen schrieben sich feminine Eigenschaften im stärkeren Maße zu als die maskulinen, während Männer sich auf der Maskulinitätsskala höher einschätzten als auf der Femininitätsskala.

### 3.1.2.3.3 Konfirmatorische Faktorenanalyse

In einem zweiten Schritt wurde eine konfirmatorische Faktorenanalyse mit AMOS berechnet (Abbildung 12), um die Voraussetzungen für die Skalenbildung zu überprüfen und die hypothetisierte Zweifaktorenstruktur zu bestätigen. Das Modell beinhaltete die Items der maskulinen und femininen Teilskalen, entsprechend Bems (1974) Theorie der unabhängigen Teilskalen.

Für die Analysen wurde zunächst die Gesamtstichprobe berücksichtigt, da Ergebnisse aus anderen Studien zeigten, dass sich das männliche und weibliche Geschlechtsrollenselbstkonzept durch starke strukturelle Ähnlichkeiten auszeichnen (Athenstaedt, 2003). Eine anschließend durchgeführte geschlechtsspezifische Überprüfung des Modells bestätigte diese Annahme. Abbildung 12 zeigt die detaillierte Faktorenstrukturen von Maskulinität und Femininität für die gesamte Stichprobe.

Insgesamt offenbarten die Analysen keinen guten Fit der Daten auf das Modell. Die Statistiken lagen alle über den theoretisch definierten Grenzwerten (Schermelleh-Engel, Moosbrugger, & Müller, 2003): Das Verhältnis von  $\chi^2/df$  ergab 4.99, der *CFI* betrug .64 und *RMSEA* erreichte einen Wert von .08, so dass dieses Modell für die von Bem vorgegebene Zuordnung der BSRI-Items nicht angenommen werden konnte.

Während die beiden Teilskalen nur einen geringen Zusammenhang aufwiesen, offenbarten die Faktorladungen einzelner Items Auffälligkeiten. So schwankten die Faktorladungen der maskulinen Eigenschaften stark von .29 bis .79. Bei der femininen Teilskala erwiesen sich die Faktorladungen als insgesamt betrachtet schwächer. Erneut fiel als erstes das Adjektiv *abhängig* auf, dessen Faktorladung mit .04 inakzeptabel ist. Darüber hinaus wiesen die Charakteristiken *nachgiebig*, *bescheiden* und *liebe Sicherheit* ebenfalls unzureichende Faktorladungen auf. Weder der Ausschluss der aus bisherigen Analysen als strittig geltenden Items, noch eine erhebliche Reduzierung der Itemanzahl, führte zu einer Verbesserung der Modellpassung ohne dass schlussendlich nur drei Items pro Teilskala verblieben.

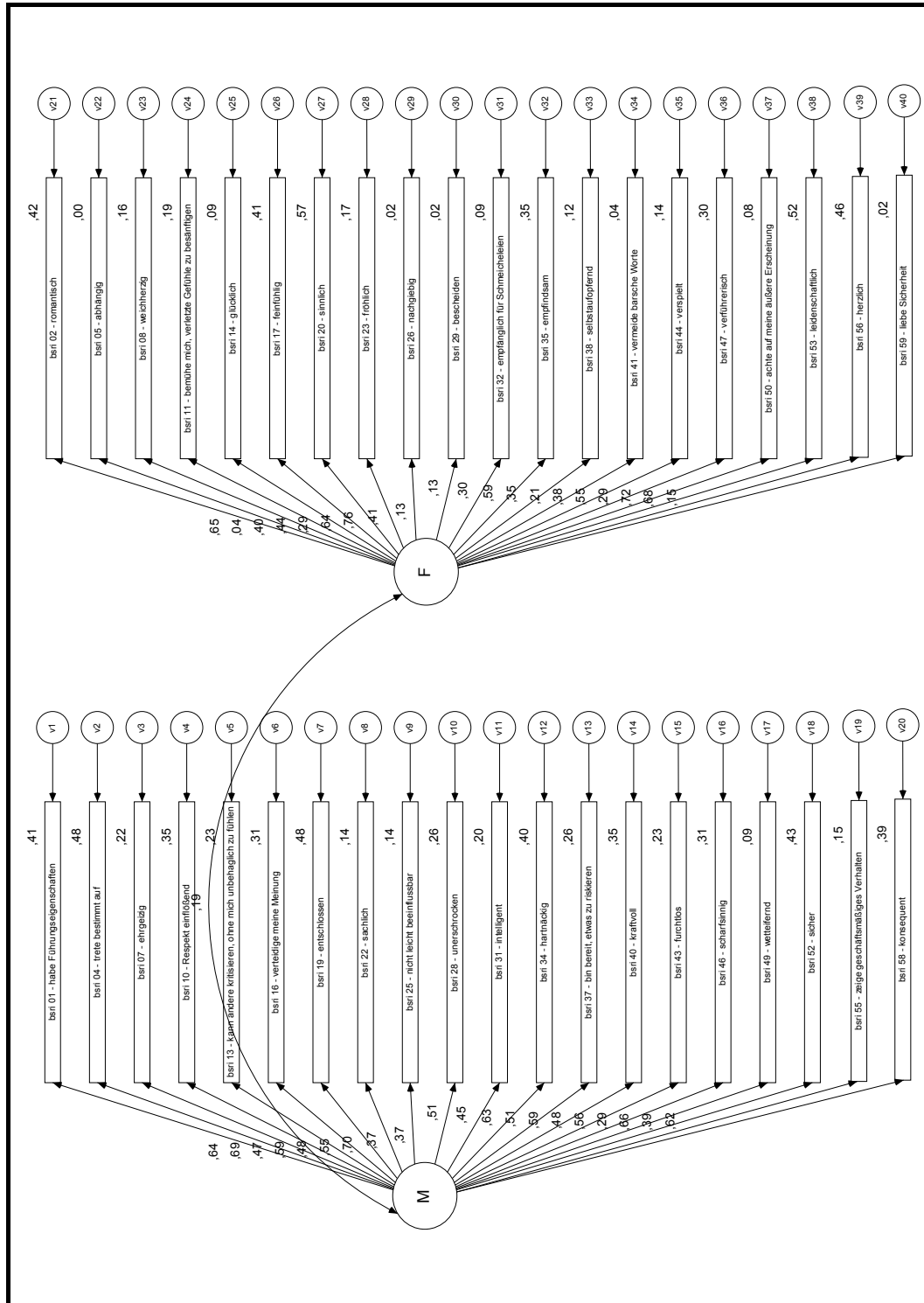


Abbildung 12. Konfirmatorische Faktorenanalyse des BSRi aus der Selbstbeschreibung

### 3.1.2.4 Diskussion

Während die kulturelle Abhängigkeit theoretisch bereits ausführlich diskutiert wurde, muss aufgrund der vorliegenden Ergebnisse der Schluss gezogen werden, dass die deutsche Übersetzung des BSRI dem heutigen Zeitgeist und dem, was gesellschaftlich als maskuline und feminine Werte definiert ist, nicht mehr entspricht. Auffällig ist in diesem Zusammenhang, dass die Erfassung der positiven Aspekte des Geschlechtsrollenelbstkonzepts mithilfe des BSRI nicht zu zufriedenstellenden Ergebnissen führt. So erweisen sich einige Items unter verschiedenen Gesichtspunkten wiederholt als mangelhaft, wie beispielsweise *abhängig*. Weder hält dieses Item den Ansprüchen der ursprünglichen Auswahlkriterien stand, noch kann es in der Selbstbeschreibung tatsächlich als ausschließlich positives Attribut bestätigt werden. Darüber hinaus sind auch auf der Skalenebene Einschränkungen festzustellen, zeigte sich doch die feminine Skala so viel weniger homogen im Vergleich zu den maskulinen Items.

Darüber hinaus reflektieren die Resultate der Selbstbeschreibung nur bedingt, was aus vergangenen Studien bezüglich der Verschiebung der Zuschreibungen bekannt ist. So hatte vor allem Twenges (1997) Metaanalyse die Diskussion um einen Anstieg der Maskulinitätswerte von Frauen genährt, der sich hier so nicht finden ließ. Vielmehr stellt sich, ausgehend von den vorliegenden Ergebnissen, die Frage, inwieweit die deutsche Version des BSRI dem Anspruch gerecht wird, zwei unabhängige Dimensionen des GSK zu erfassen, die gleichberechtigt nebeneinander existieren können. Dies ist vor allem bedeutsam vor dem Hintergrund, dass die femininen Eigenschaften in der vorliegenden Studie erneut am stärksten zur Selbstbeschreibung herangezogenen wurden.

Während die Ergebnisse der Selbstbeschreibung nicht gänzlich gegen einen Einsatz dieses Instruments sprechen, so legen sie doch eine klare Überarbeitung nahe; vor allem aufgrund der Nicht-Berücksichtigung negativer Attribute, deren Anteil am Selbstkonzept jedoch ausgiebig nachgewiesen wurde (siehe Theorie).

Für zukünftige Einsätze eines solchen Fragebogens kann das einfache Eliminieren einiger Items nicht die angestrebte Lösung sein, da mit diesem Vorgehen die Nachlässigkeiten aus der Konstruktionsphase – beispielsweise die Typizität nicht getrennt überprüft zu haben – nicht beseitigt werden und die Gefahr besteht, aus einem fehlerhaften Instrument lediglich die am wenigsten belasteten Items auszuwählen. Eine Neuentwicklung ist somit nicht nur wünschenswert, sondern notwendig, da nur durch sie eine adäquate Testkonstruktion gewährleistet werden kann. Diese Neukonstruktion wird umso bedeutsamer, berücksichtigt man, dass selbst die Erfassung positiver Aspekte nicht zufriedenstellend gelingt, müssen Items zusammengetragen werden, die die aktuellen Vorstellungen von Maskulinität und Femininität widerspiegeln. Mit Bezug auf die aktuellen Forschungsergebnisse und unter Berücksichtigung der Konstruktion des geschlechtlichen Selbst müssen im Rahmen dieser Neukonstruktion Items gefunden werden, die sowohl die positiven als auch negativen Aspekte der Geschlechtsrolle abzubilden vermögen. So soll die Vorhersageleistung dieses Teils der Selbstbeschreibung in Bezug auf andere psychologische Konzepte, die ebenfalls positiv und negativ in der Valenz sind, maximiert werden.

## 3.2 Neukonstruktion eines Instruments zur Erfassung positiver und negativer Aspekte des Geschlechtsrollenelbstkonzepts

### 3.2.1 Studie 3: Konstruktionsstudie – Generierung neuer GSK-Items

#### 3.2.1.1 Fragestellung

Im Bereich der Geschlechtsrollenstereotypenforschung haben sich vor allem zwei Arten von Kategorien herausgestellt: *Kompetenz* und *Aufgabenorientierung* als Beispiele der Kategorie positiv bewerteter männlicher Eigenschaften sowie *Personenorientierung* und *Emotionalität* als positiv bewertete weibliche Eigenschaften (Mecklenbräuker et al., 1994). In dem neuen Messinstrument sollen nun jedoch auch negativ bewertete Eigenschaften zur Differenzierung des Geschlechtsrollenelbstkonzepts herangezogen werden. Generell gelten überraschend wenige Adjektive als tatsächlich „wert-neutral“, da die meisten tatsächlich sehr eindeutig als positiv oder negativ bewertet werden können (Mecklenbräuker et al., 1994). In der freien Produktion von beschreibenden Attributen zeigt sich bisweilen, dass negative Charakteristiken ebenfalls durchaus verfügbar sind, jedoch unaufgefordert weniger häufig produziert werden als positive Eigenschaften (Helgeson, 1994a). Darüber hinaus sind eine Vielzahl von Studien bekannt (Antill et al., 1981; Athenstaedt, Altstötter-Gleich, & Hillberger, 2007; Buss, 1990; Dohi & Hirokawa, 2004; Fritz & Helgeson, 1998; Helgeson, 2003; Helgeson & Fritz, 1996; Helgeson & Fritz, 1998; Helgeson & Fritz, 1999; Holahan & Spence, 1980; Nagurney, 2007; Spence et al., 1979), die mit negativen Eigenschaften bereits arbeiteten und zeigen konnten, dass z. B. Unmitigated Agency und Unmitigated Communion nicht auf ihre Gegenstücke *Agency* und *Communion* reduziert werden dürfen (Helgeson & Fritz, 1999) und somit die individuelle Daseinsberechtigung negativer Teilskalen unterstrichen.

Angleitner und Ostendorf (1994) führten bezüglich der Vollständigkeit von Begriffssammlungen bereits existierende Adjektivlisten, wie beispielsweise die Liste persönlichkeitsbeschreibender Begriffe von Baumgarten (1933), eine deutsche Übersetzung der *Adjective Check List* (Gough & Heilbrun, 1965), die Eigenschaftswörterliste von Janke und Debus (1978), die Adjektivliste zur Selbstbeschreibung (Mummendey, Mielke, Maus & Hesener, 1977), die *German Adjective List* (John, 1982) sowie die Adjektivliste von Wippich und Bredenkamp auf (1977, 1979; alle zitiert in Angleitner & Ostendorf, 1994).

Keine der hier genannten Listen kam jedoch in der Konstruktionsphase zum Einsatz, um dem Umstand Rechnung zu tragen, dass zum einen Adjektive gefunden werden sollten, die aktuell und im Sprachgebrauch der Versuchspersonen sind. Zum anderen sollte gezielt nach Eigenschaftswörtern gesucht werden, denen per se die Möglichkeit inne liegt zwischen den Geschlechtern zu trennen. Um der weniger leichten Generierung auch negativer Beschreibungen zu begegnen, sollten Probanden explizit aufgefordert werden, diese selbst bei Schwierigkeiten zu produzieren.

Am Ende dieses Konstruktionsprozesses soll ein Instrument entwickelt worden sein, welches das Vokabular und die Aktualität von sowohl positiven als auch negativen Geschlechtsrollenstereotypen dieser Gesellschaft in der heutigen Zeit abzubilden vermag und

darüber hinaus in Bezug zu einer Vielzahl von psychologischen Konzepten gebracht werden kann.

Wie im vorherigen Abschnitt hergeleitet, wird aus zwei Gründen die Neukonstruktion einer deutschen Geschlechtsrollenskala angestrebt: (1) die aufgezeigten Schwierigkeiten mit dem BSRI als Maß der positiven Aspekte des GSK und (2) der theoretisch begründeten Notwendigkeit der Berücksichtigung negativer Aspekte des Geschlechtsrollenselbstkonzepts. Die Skalenkonstruktion erfolgt mittels vier Schritten in mehreren einzelnen Studien:

- (1) Erstellung eines aktuellen Pools an Eigenschaften (Studie 3)
- (2) Itemauswahl auf der Grundlage der Einschätzung von Typizität und Erwünschtheit sowie der differenziellen Selbstzuschreibung von Männern und Frauen (Studie 4)
- (3) Einsatz der Endskala in der Selbsteinschätzung und Validierung (Studie 5)
- (4) Replikation (Studie 6)

Alle Studien wurden mit unterschiedlichen Stichproben durchgeführt. In einem ersten – in diesem Abschnitt vorgestellten – Schritt (Studie 3) wurden zunächst Items zusammengetragen, die den erforderlichen Voraussetzungen der unterschiedlichen Typizität und Erwünschtheit für Männer und Frauen sowie der Aktualität im aktiven Sprachgebrauch gerecht werden sollten.

### **3.2.1.2 Methode**

#### **3.2.1.2.1 Stichprobe**

Um eine möglichst große Stichprobe zu erreichen und die Zahl zu generierender verschiedener Eigenschaften zu maximieren, wurden zahlreiche Rekrutierungsstrategien bemüht. Dabei lag aufgrund der besseren Erreichbarkeit von potentiellen Versuchspersonen der Schwerpunkt auf der Online-Studie, die über zentrale Studienseiten verlinkt wurde und vom steten Zugriff auf die Domain <http://www.w-lab.de><sup>32</sup> profitierte. Zusätzlich wurden Probanden auch offline befragt, um eine größere Altersstruktur abbilden zu können. Durch das sog. Schneeballprinzip<sup>33</sup> konnten viele Personen aus dem Bekannten- und Arbeitsumkreis rekrutiert werden.

Insgesamt gingen die Daten von 197 (82 Männer, 115 Frauen) Personen in die Analyse ein (15 davon füllten den Fragebogen offline aus). Der Mittelwert des Alters lag bei  $M = 29.15$  Jahren ( $SD = 11.55$  Jahre; Min: 18, Max: 71 Jahre). Insbesondere die älteren Probanden berichteten von erheblichen Schwierigkeiten bei der freien Generierung von Eigenschaften. Oftmals beschränkten sie sich darauf Beschreibungen vorzunehmen.

Das überwiegend junge Durchschnittsalter der Stichprobe spiegelt sich auch in den Schulabschlüssen wieder: mit 61.4% waren Personen mit Abitur überrepräsentiert. Weitere 16.2% verfügten bereits über einen Studienabschluss; 14.2% machten einen Real- sowie Gesamtschulabschluss und 6.6% waren Hauptschulabgänger. Studierende bildeten den größten Teil der Stichprobe (36.5%). Die berufliche Struktur reichte von Lagerarbeiter, Hebamme, Krankenschwester über Werkzeugmacher bis zu Chemielaboranten, Ärzten und Professoren. Alle Probanden waren deutscher Nationalität, zwei von ihnen gaben eine andere Muttersprache als Deutsch an. Da es in dieser Studie um die bloße Nennung von stereotypen

---

<sup>32</sup> eine für laufende Onlinestudien und umfangreiche Informationen zur Onlineforschung bekannte Webseite

<sup>33</sup> Jede Person informiert weitere mögliche Probanden.

Eigenschaften ging und eine maximale Vielfalt an Eigenschaften für die Testungen in den folgenden Studien erreicht werden sollte, wurden alle Personen in den Auswertungen berücksichtigt.

### 3.2.1.2.2 Instrumente

Alle Befragten wurden zunächst gebeten in einer freien Nennung typisch positive und negative Eigenschaften von Männern und Frauen zu generieren. Dazu erhielten sie folgende Instruktion: „Wir möchten von Ihnen wissen, welche Eigenschaften typisch für Männer und welche typisch für Frauen sind. Mit *„typisch für Männer“* meinen wir, dass eine Eigenschaft bei sehr vielen Männern und nur bei sehr wenigen Frauen vorkommt. Mit *„typisch für Frauen“* meinen wir, dass eine Eigenschaft bei sehr vielen Frauen und eben nur bei sehr wenigen Männern vorkommt“. Die Instruktion wurde bewusst kurz gehalten. Auf Beispiele für eine Eigenschaft wurde verzichtet, um ein mögliches Priming auf bestimmte Eigenschaften zu verhindern und eine größtmögliche kognitive Kreativität zu ermöglichen.

Zur besseren Strukturierung für die Versuchspersonen, wurden maximal sieben Textfelder pro Geschlecht und Valenz für die Antworten vorgegeben (Abbildung 13). Dieses Vorgehen berücksichtigte zudem die nachgewiesene Längenwirkung vorgegebener Antwortfelder auf Antwortlänge und -qualität (Spörrle, Gerber-Braun, & Försterling, 2007).

The image shows a screenshot of a web-based form for generating characteristics. The form is organized into four quadrants, each with a title and seven input fields:

- Top-Left:** "Typisch für Männer sind folgende positive Eigenschaften" with 7 input fields.
- Top-Right:** "Typisch für Frauen sind folgende positive Eigenschaften" with 7 input fields.
- Bottom-Left:** "Typisch für Männer sind folgende negative Eigenschaften" with 7 input fields.
- Bottom-Right:** "Typisch für Frauen sind folgende negative Eigenschaften" with 7 input fields.

At the bottom center of the form is a grey button labeled "Weiter".

Abbildung 13. Online-Seite zur Generierung von Eigenschaften

Im Idealfall sollten so mindestens 20 unterschiedliche Items pro Teilskala durch die befragten Personen generiert werden. Das sowohl für Online- als auch als schriftliche Befragung aufbereitete Material enthielt eine allgemeine Instruktion sowie Fragen zur Demografie (Schul- und berufliche Bildung, Alter, Geschlecht, etc.) und schließlich den Teil zur freien Generierung typischer Merkmale für beide Geschlechter.

### 3.2.1.3 Ergebnisse

Insgesamt 237 unterschiedliche freie Beschreibungen bildeten die Auswertungsgrundlage. Obwohl die Möglichkeit bestand, bis zu sieben Eigenschaften pro Geschlecht und Valenz anzugeben, machten lediglich neun Personen vollständig davon Gebrauch. Generell gilt es klar zu unterscheiden zwischen (1) tatsächlichen Eigenschaften im Sinne von Adjektiven (*mutig*, *leichtsinnig*) und (2) allgemeinen Charakteristiken, die von Komplimenten für das eigene oder andere Geschlecht bis hin zu inhaltlich ausführlichen Beschreibungen reichten. Diese allgemeinen Beschreibungen konnten jedoch nicht verwendet werden.

In einem umfangreichen Kategorisierungsprozess wurden die Nennungen der Eigenschaften zunächst inhaltsanalytisch sortiert und ausgezählt: (a) jeweils die Nennungen einer Eigenschaft insgesamt über alle Gruppen und Kategorien und darauf folgend (b) die Nennungen aufgeschlüsselt nach eingeschätzter Kombination und des Geschlechts der Probanden.

Zur Auswahl geschlechtstypischer Merkmale mit positiver und negativer Valenz wurden an die gegebenen Eigenschaften verschiedene Kriterien angelegt: Die ausgewählten Eigenschaften mussten nicht nur eine hohe Nennungshäufigkeit für ein Geschlecht aufweisen (in Anlehnung an Klapprott (1994) wurden nur Eigenschaften mit einer Nennungshäufigkeit von mindestens 2% ausgewählt), sondern gleichzeitig kaum oder gar nicht für das jeweils andere Geschlecht als typisch benannt worden sein. Darüber hinaus galt es Eigenschaften zu identifizieren, die, wenn nicht gleich häufig, dann zumindest per se sowohl von Männern als auch Frauen als typisch für die jeweilige Merkmalskombination (z. B. positiv für Männer) aufgeführt wurden. Indem ein eigenschaftsbasiertes Instrument angestrebt wurde, das eine hohe Praktikabilität bezüglich der Beschreibung und Antwortskalen innerhalb des zukünftigen Instruments besitzen sollte, wurden zudem Wortgruppen oder Sätze ausgeschlossen<sup>34</sup>. Die nach diesem Vorgehen ausgewählten Begriffe pro Kategorie sind in Tabelle 22 aufgelistet.

Das Ziel der Studie war die größtmögliche Unabhängigkeit der Teilskalen. Insbesondere die negativen sollen keine bloße Umkehr der positiven Teilskalen darstellen, da somit ihre Erfassung redundant und ihr Beitrag zur Varianzaufklärung minimiert wäre. Daher soll eine Verwendung von Antonymen und eine gleichzeitige Verwendung von ein- und derselben Eigenschaften in zwei Kategorien generell vermieden werden. Dieser Anspruch könnte unter Umständen einen Ausschluss betroffener Eigenschaften nach sich ziehen, insbesondere dann, wenn sie innerhalb einer Geschlechtergruppe eine Rolle spielten (z. B. *emotional*: sowohl positiv als auch negativ für Frauen).

---

<sup>34</sup> Mit Ausnahme der Beschreibung „gute Zuhörer“, die aufgrund der sehr häufigen Nennung für den Moment in der Grundauswahl verblieb.



Tabelle 22. Ausgewählte Eigenschaften nach Generierung pro Teilskala

Maskuline Skala		Feminine Skala	
positiv 20	negativ 20	positiv 20	negativ 20
mutig	verschlossen	multitaskfähig	besitzergreifend
ruhig	faul	sensibel	irrational
rational	gewalttätig	gerecht	ängstlich
selbstbewusst	rücksichtslos	kreativ	nachgiebig
beschützend	überheblich	einfühlsam	naiv
zielorientiert	schweigsam	verständnisvoll	launisch
ehrgeizig	gefühllos	kommunikativ	selbstzweifelnd
pragmatisch	leichtsinnig	mitfühlend	hinterlistig
geradlinig	unzuverlässig	ordentlich	nachtragend
zielstrebig	unordentlich	liebepoll	eitel
durchsetzungsfähig	unsensibel	empathisch	kompliziert
stark	ungeduldig	organisiert	unsicher
praktisch	starrsinnig	warmherzig	empfindlich
direkt	egoistisch	fleißig	hektisch
gelassen	rechthaberisch	gefühlvoll	intrigant
konsequent	dominant	fürsorglich	zickig
sachlich	aggressiv	sanftmütig	lästern
handwerklich	machohaft	vielseitig	leichtgläubig
logisch	arrogant	intuitiv	
souverän	ignorant	gute(r) Zuhörer(in)	
unkompliziert	stur		
<b>21</b>	<b>21</b>	<b>20</b>	<b>18</b>

In diesem Zusammenhang erfordern nicht nur gegensätzliche Wörter eine besondere Handhabung, sondern auch einfache Suffixe und Präfixe, die in der deutschen Sprache zu einigen Schwierigkeiten führen könnten. Es existierten Kombinationen von Wörtern, die die Teilskalen in unterschiedlicher Art in Beziehung setzten. Die Eigenschaft *emotional* wird demnach durchaus als beispielsweise typischer für Frauen angesehen als für Männer, dabei jedoch sowohl positiv als auch negativ eingeschätzt. Es verbleibt damit aber zumindest innerhalb der geschlechtstypischen Kategorie.

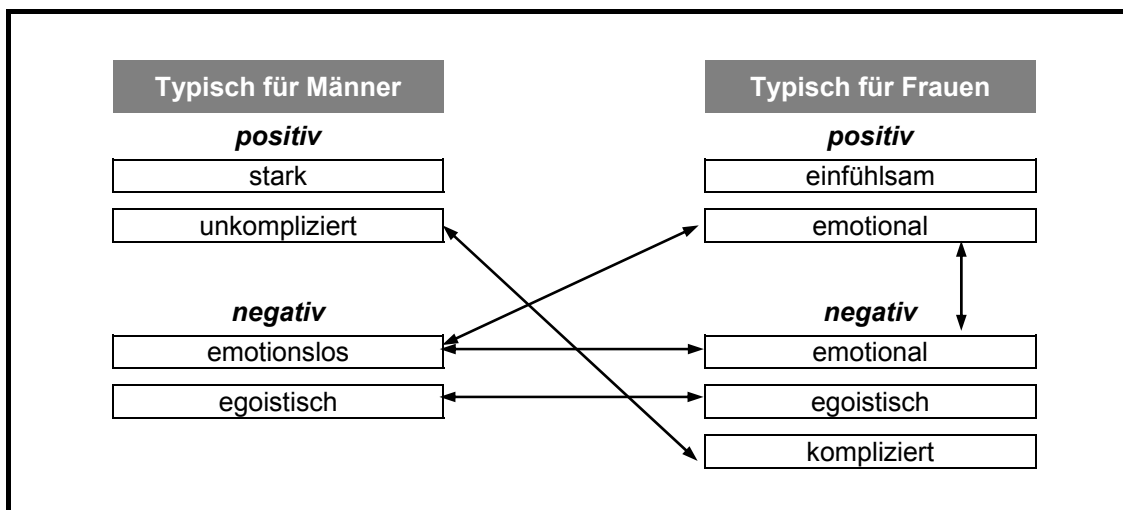


Abbildung 14. Beziehungen ausgewählter Eigenschaften zueinander

Demgegenüber würde *emotionslos*, durch das Suffix „los“ als fehlende Emotionalität (= fehlende positive Eigenschaft) charakterisiert, den negativen männlichen Eigenschaften zugeteilt werden und dadurch eine negative Korrelation zu den positiven weiblichen Eigenschaften verstärken (Beziehung zwischen F+ und M-). Ähnlich verhält es sich mit *kompliziert* und *unkompliziert*.

Während *kompliziert* als negative Eigenschaft für Frauen gilt, wird das durch das Präfix „un“ charakterisierte Antonym *unkompliziert* in diesem Fall als positive Eigenschaft für Männer gewertet. In diesem Fall wird eine Beziehung zwischen F- und M+ hergestellt (Abbildung 14). Obwohl diese Wortverwandtschaften einer möglichen Unabhängigkeit der Skalen entgegenstehen, wurden sie nicht aus dem Konstruktionsprozess ausgeschlossen. Da diese Eigenschaften frei generiert wurden, dem aktuellen Sprachgebrauch entsprechen und tatsächlich zur Selbstbeschreibung herangezogen wurden, verblieben sie zunächst in der Studie. Darüber hinaus wurde erwartet, dass die nachfolgenden Analysen zeigen würden, welche der Eigenschaften den weiteren statistischen Kriterien gegebenenfalls nicht standhalten und daher nicht in die endgültigen Skalen eingehen dürften. Die Einschätzungen der Typizität und sozialen Erwünschtheit würden dann zeigen, welche dieser Eigenschaften insgesamt als positiv oder negativ anzusehen sind und inwieweit sie ausreichend typischer für ein Geschlecht sind. In der finalen Auswahlphase müssten dann jedoch, sofern die hier aufgelisteten Eigenschaften nicht vorher ausgeschlossen werden, diese Wortverwandtschaften als weiteres Auswahlkriterium berücksichtigt werden.

### **3.2.1.4 Diskussion**

In einem ersten Schritt wurde aufgrund freier Nennungen von Männern und Frauen eine Auswahl von positiven und negativen sowohl maskulinen als auch femininen Eigenschaften produziert, die das Ausgangsmaterial für die Neukonstruktion der Skala zur Erfassung der positiven und negativen Konstrukte des GSK darstellen. In den nachfolgend präsentierten Studien sollen sie dann für ihre Eignung zur Messung des Geschlechtsrollenselbstkonzepts hinsichtlich ihrer Typizität und sozialen Erwünschtheit eingeschätzt werden.

Die Schwierigkeit bei der Auswahl bestand dabei vor allem darin, (1) zum einen ähnliche Eigenschaften in einer Kategorie zusammenzufassen und (2) zum anderen den durchaus unterschiedlichen Nuancen von ähnlich klingenden und bisweilen fälschlicherweise synonym verwendete Eigenschaften gerecht zu werden und diese erst im weiteren Itemauswahlprozess durch statistische Kennwerte gegebenenfalls auszusortieren.

Nach den zugrunde gelegten Kriterien waren ca. 20 Items pro Teilskala an diesem Konstruktionspunkt angestrebt worden. Andere Studien offenbarten bereits mögliche Schwierigkeiten, ad hoc negative Eigenschaften zu produzieren. So hatte beispielsweise Helgeson (1994a) in ihrer Arbeit zur Prototypikalität zeigen können, dass bei der Beschreibung von sechs fiktiven Personen (z. B. maskuliner Mann, maskuline Frau, femininer Mann, etc.) auch negative Eigenschaften herangezogen wurden, jedoch weniger leicht und häufig als positive Charakteristiken. Betrachtet man die Nennungshäufigkeiten etwas genauer, fiel es den Probanden auch in dieser Studie insgesamt einfacher positive Eigenschaften zu generieren. Dies traf für die eigene Geschlechtergruppe im stärkeren Maße zu, während über alle Gruppen hinweg weniger negative Eigenschaften frei genannt wurden. Umso eingeschränkter war die Generierung negativer Eigenschaften für das eigene Geschlecht – zumindest für die Männer. Diese Tendenz begründet sich in der bereits dargestellten sozialen Identitätstheorie SIT (Tajfel & Turner, 1979; 1986), gemäß derer die eigene Gruppe der Motivation eines stabilen und

hohen Selbstwertes folgend eine positive Darstellung anstrebt, während die Außengruppe Abwertung erfährt. Demgegenüber agierten die Frauen weniger im Einklang mit der SIT. Im Vergleich zu den Männern generierten sie vergleichsweise mehr negative Eigenschaften für ihre eigene Gruppe.

Ähnlich wie auch in der Studie von Wolff und Taylor (1979), die diesen Umstand auf die größere verbale Sprachkompetenz bei Frauen zurückführten, produzierten die Frauen insgesamt mehr freie Nennungen als die Männer. Insbesondere, wenn man neben der größeren Anzahl von Frauen in der Stichprobe zusätzlich den selbstwertdienlichen Fokus auf die Fremdgruppe berücksichtigt, lässt sich die größere Menge generierter maskuliner Items erklären. Eine der bereits in dieser Konstruktionsstudie festgestellten Herausforderungen bestand in der zu erreichenden größtmöglichen Unabhängigkeit der Teilskalen, die eine einfache Umkehr der negativen Teilskalen von den positiven Teilskalen im Sinne von redundanter Information ausschließt.

Auch Helgeson (1994a) hatte in ihrer Studie von einer offensichtlichen bipolaren Konzeption frei generierter Eigenschaften berichtet. So stellte sie fest, dass eine Vielzahl der Charakteristiken einer einzuschätzenden männlichen Zielperson den Antonymen einer weiblichen Person entsprachen (z. B. *not caring* vs. *caring*, *self-confident* vs. *insecure*, *not emotional* vs. *emotional*; S. 669) – obwohl jede Versuchsperson jeweils nur eine Person fokussieren und somit keinen direkten Vergleichsmaßstab zwischen den Geschlechtern herstellen musste. Die Frage, die jedoch offen bleibt, ist, inwieweit ein Individuum bei der Generierung von Charakteristiken eines Geschlechts nicht auch unmittelbar den Vergleich zum anderen Geschlecht anstellt, um die passenden Attribute auszuwählen. Dies könnte womöglich bei der Generierung der Eigenschaften die Auftretenshäufigkeit von Antonymen über Geschlechter und Bewertungskategorien hinweg erklären.

Aus einer anderen Perspektive heraus diskutierten Deaux und LaFrance (1998) anhand des von Biernat und Manis (1994) berichteten Effekts des *shifting standards* am Beispiel der Einschätzung von Aggressivität, welche entscheidende Rolle ein nicht explizit induzierter Vergleichsmaßstab spielt. Demnach wäre es möglich, dass unterschiedliche Maßstäbe an die Aggressivität von Frauen und Männern angelegt werden. Bezüglich der stereotypen Erwartung, dass Männer insgesamt als aggressiver gelten, würde dies bedeuten, dass eine viel stärkere Demonstration dieses Verhaltens seitens der einzuschätzenden Person bedarf, um dies in der Bewertung auch widerzuspiegeln. Auch Prentice und Carranza (2002) diskutierten in ihrer Studie die Möglichkeit, dass Männer und Frauen mit unterschiedlichen Standards, wie beispielsweise höhere soziale und gleichzeitig geringere Leistungsstandards für Frauen, eingeschätzt und gemessen werden.

Während es in diesen Beispielen weniger um die freie Produktion von Eigenschaften ging, deutet sich die Schwierigkeit von (un-)bewussten Vergleichsmaßstäben im Zusammenhang mit der Einschätzung von Geschlechtern und der daraus resultierenden Einschränkung an, derart, dass womöglich auch ad hoc generierte Charakteristiken einer Gruppe nicht so frei vom Eindruck der jeweils anderen Gruppe sind, wie es bisweilen wünschenswert wäre.

Gerade die Beziehungen und die semantische Ähnlichkeit von Attributen über die Geschlechter als auch über die Valenzkategorien hinweg leisten einen erheblichen Beitrag zur Definition der Beziehung der Teilskalen zueinander. Es bleibt daher abzuwarten, inwieweit Eigenschaften dieser Art im Rahmen der nächsten Schritte der Konstruktionsphase bestätigt werden oder im Verlauf möglicherweise doch ausgeschlossen werden (müssen).

Die Paarung linguistischer Antonyme unterschiedlicher Valenz (z. B. *faul* und *fleißig*) stellt sowohl theoretisch als auch statistisch eine Herausforderung dar. Zum einen ist gerade in Fällen verschiedenartiger Wortstämme nicht eindeutig, inwieweit inhaltlich wirklich Pole des gleichen Attributs abgebildet sind, was die Nutzung semantischer Differentiale wie beim PAQ (Spence et al., 1974) problematisch macht. Zum anderen gilt es als schwierig nachzuweisen, inwieweit diese positiven und negativen Eigenschaften äquidistant vom Mittelpunkt, d. h. im gleichen Maße jeweils positiv und negativ sind (Pahl & Eiser, 2005).

In den nächsten Schritten der Validierung sind genau diese Aspekte möglicher umgekehrter Wortpaarungen über Teilskalen hinweg mit besonderer Sorgfalt zu betrachten. Während es zum einen darum gehen soll, über statistische Kennwerte aus der Selbstbeschreibung die skalenstärksten und aussagenkräftigsten Items auszuwählen, muss zum anderen die Familiarität von Items aus unterschiedlichen Teilskalen als zusätzliches Kriterium und womöglich sogar in seiner Priorität vor den statistischen Kennwerten, herangezogen werden.

### **3.2.2 Studie 4: Itemauswahl auf der Basis von Typizität, sozialer Erwünschtheit und differentieller Selbstzuschreibung von Männern und Frauen**

#### **3.2.2.1 Fragestellung**

In der folgenden Studie 4 sollten die Items aus der Konstruktionsphase in der Selbstbeschreibung eingesetzt werden, um durch die Überprüfung statistischer Kennwerte die Items für die Zielskalen auszuwählen. Dabei wurden die Items sowohl hinsichtlich der Typizität, jeweils getrennt für Männer und Frauen, als auch der sozialen Erwünschtheit eingeschätzt.

In diesem Zusammenhang ist vor allem die Unterscheidung zwischen (a) Erwünschtheit hinsichtlich eines allgemeinen, angenehmen Gefühls, welches in der Person selbst liegt und (b) der Angenehmheit von Eigenschaften bei anderen Personen – die eigentliche soziale Erwünschtheit (Mecklenbräuker et al., 1994) – wichtig. Beim BSRI wurden beispielsweise Items aufgrund ihrer sozialen Erwünschtheit ausgewählt: Beide Geschlechter mussten unabhängig voneinander eine Eigenschaft als erwünschter für das eine Geschlecht als für das andere klassifizieren. Inwieweit tatsächlich die soziale Erwünschtheit der Einschätzung und nicht etwa ein persönliches Wertesystem zugrunde gelegt wurde, bleibt unklar und hinsichtlich der uneindeutigen Instruktion sogar fragwürdig. Mecklenbräuker et al. (1994) untersuchten daher in ihrer Studie die Zusammenhänge der unterschiedlichen Formen der vier möglichen Erwünschtheitsvariablen: zwei Versuchspersonengeschlechter jeweils in Bezug auf zwei mögliche Zielpersonengeschlechter (z. B. Erwünschtheit bei Männern, eingeschätzt von Frauen). Die Autoren stellten dabei fest, dass es kaum Unterschiede zwischen den Einschätzungen gab, und

dennoch fanden die Autoren einen bedeutsamen Effekt. So unterschieden sich die maskulinen Items kaum bezüglich der verschiedenen Erwünschtheitsratings, einige feminine Eigenschaften jedoch schon. Ein Beispiel dafür wäre *feinfühlig*, was Frauen für erwünschter bei Männern halten, während Männer diese Eigenschaft für Frauen weniger wünschenswert einschätzen.

Insgesamt wurde diese Studie in drei große Phasen unterteilt: (1) In einem ersten Teil wurden die Items bezüglich der Typizität, jeweils getrennt für die Geschlechter, eingeschätzt. (2) In zweiten Teil fand die allgemeine Beurteilung der Erwünschtheit der Attribute für Personen in der heutigen Gesellschaft statt, und (3) in der dritten Phase schließlich die Selbstbeschreibung einiger Probanden mit den neuen Items.

Im Gegensatz zur Typizität, die geschlechtsspezifisch – also getrennt in Bezug auf Männer und Frauen – bewertet werden sollte, wurde in der Befragung zur sozialen Erwünschtheit das Kriterium verändert. Dazu wurden alle relevanten Items in nur einer Studie dargeboten und sollten nun, nicht mehr wie beispielsweise von Bem (1974) umgesetzt, einzeln für Männer und Frauen, sondern hinsichtlich der generellen sozialen Erwünschtheit dieser Charakteristiken in der heutigen Zeit eingeschätzt werden. Mit dieser Erfassung sollte überprüft werden, inwieweit die Valenzen der Eigenschaften, so wie sie in der Pilotstudie von den Versuchspersonen generiert wurden, sich in der Meinung einer größeren Gruppe wiederfinden lassen. Die ursprünglich getrennte Einschätzung der sozialen Erwünschtheit in Bezug auf Männer und Frauen wurde verworfen, da sich empirisch und theoretisch kaum argumentieren lässt, warum sowohl Typizität als auch soziale Erwünschtheit einzeln für Männer und Frauen eingeschätzt werden sollten. Bezüglich der Typizität können diese Bewertungen einfach vorgenommen werden, während in Hinblick auf die Valenz diese Beurteilungen schwerer fallen, da sie – wie aus der Gruppenforschung bekannt – nicht frei sind vom individuellen oder gruppenspezifischen Selbstwert, inklusive der daraus resultierenden Ablehnung der Fremdgruppe oder Aufwertung der Eigengruppe. Studien zum Geschlechtsrollenselbstkonzept zeigten zudem, dass die Eigenschaften heutzutage durchaus entgegen ihrer ursprünglichen Kategorisierung unterschiedlich sozial definiert sind. So ist es für Frauen in der heutigen Zeit angemessener, auch *aggressiv* und *durchsetzungsfähig* (ursprünglich maskuline Items) zu sein als noch vor 30 Jahren. Darüber hinaus ist es sozial vertretbar, dass Männer auch feminine Eigenschaften wie *empathisch* zur Selbstbeschreibung heranziehen. Eine detaillierte Erfassung der Bewertung der sozialen Erwünschtheit wäre im größeren Maße – als dies beispielsweise bei der Typizität der Fall ist – von der eigenen Rolle und deren Bewertung abhängig.

Zudem haben eigene Analysen in der vorliegenden Arbeit gezeigt, dass die Erfassung von sozialer Erwünschtheit getrennt hinsichtlich beider Geschlechter zum Ausschluss nahezu aller zu überprüfenden Items führen würde. Aus diesem Grunde wurde die soziale Erwünschtheit der Items nicht mehr für Männer und Frauen getrennt beurteilt, sondern nur noch als allgemeines Valenzmaß erhoben. Dieses Vorgehen sollte gewährleisten, dass die unterschiedlich positiven und negativen Items auch berechtigterweise den Skalen entsprechend ihrer allgemeinen sozialen Erwünschtheit zugeordnet wurden.

Grundsätzlich werden bei der Entwicklung der Skalen mehrere Kriterien angelegt:

- (1) Wortgruppen sollen vermieden werden.
- (2) Die Teilskalen sollten semantisch differenziert (ohne große inhaltliche Überschneidungen zu den jeweils anderen Teilskalen), unabhängig sowie kurz und ökonomisch sein.
- (3) Die Items sollten hinsichtlich der Typizität signifikant, mit größtmöglicher Differenz zwischen den Geschlechtern differenzieren.
- (4) Die Itemmittelwerte sollten über dem Skalenmittelpunkt liegen.
- (5) Im Hinblick auf die soziale Erwünschtheit sollte die Valenz der positiven Attribute signifikant über, die der negativen signifikant unter dem Skalenmittelpunkt liegen.

Das sehr strenge Auswahlverfahren führte zum Ausschluss einer großen Anzahl von Eigenschaften, indem aus dem ursprünglich großen Pool der Items aus der Konstruktionsphase viele Eigenschaften den Kriterien nicht standhielten. Vor dem Hintergrund, dass sich die ursprünglich 80 ausgewählten Items aus der Generierung aufgrund ihrer unzureichenden statistischen Kennwerte sehr schnell reduzierten, wurden nachträglich weitere Items in die Bewertungsstudie mit aufgenommen.

So sollte sichergestellt werden, dass ein hinreichend große Ausgangsstichprobe an Charakteristiken zur Verfügung stand. Dies erforderte wiederum ein adaptives Überprüfen der Eigenschaften dahingehend, dass verschiedene, aufeinanderfolgende Studien unterschiedliche Items überprüfen. Dazu wurden, zusätzlich zu den Items aus der Konstruktionsphase, Eigenschaften ausgewählt, die aufgrund umfangreicher Recherchen (weitere Nennungen aus der Pilotstudie, die weniger häufig waren, jedoch aus anderen Instrumenten bereits bekannt waren; aufgrund aktueller Kennwerte von Items anderer bereits bestehender Instrumenten, Dudenanalysen, etc.) für ein neues Instrument als geeignet schienen. Die Art der Erfassung für diese nachträglich getesteten Eigenschaften blieb gleich, während die Items selbst angepasst wurden.

### **3.2.2.2 Methode**

#### **3.2.2.2.1 Stichprobe**

##### *Typizitätseinschätzung (Phase 1)*

Unter Berücksichtigung der notwendigen Erweiterung um neue Items, ergaben sich in Bezug auf die Typizitätsratings drei Teilgruppen, die hier getrennt aufgeführt werden, damit jeweils gruppenspezifisch die multivariaten Effekte berichtet werden können. So beurteilte Gruppe 1 beispielsweise *logisch* und *sachlich*, während Gruppe 2 die Bewertungen für *wetteifernd* und *furchtlos* vornahm (die zu beurteilenden Items sind über die Bezeichnung *ID* in Tabelle 25 bis Tabelle 36 entsprechend ihrer Gruppen identifiziert). Die Aufteilung der zu beurteilenden Items ist zudem auch der Tatsache geschuldet, dass die Versuchspersonen jeweils nur einen Ausschnitt aller Eigenschaften bearbeiteten, um Ermüdungserscheinungen der repetitiven und wenig abwechslungsreichen Befragung vorzubeugen. An der Einschätzung zur Typizität für Männer und Frauen beteiligten sich insgesamt 1.212 Personen (Tabelle 23).

Tabelle 23. Übersicht der Demografie zur Überprüfung der Typizität (3 Teilstichproben)

Einschätzung der Typizität der Items für ein / e						
	Gruppe 1		Gruppe 2		Gruppe 3	
	Mann	Frau	Mann	Frau	Mann	Frau
Männer	144	144	50	53	102	93
Frauen	145	150	64	61	95	111
Total	289	294	114	114	197	204
<b>Alter der Vpn</b>	M = 23.15 Jahre SD = 4.09		M = 23.08 Jahre SD = 3.94		M = 23.15 Jahre SD = 3.45	
	Min: 18, Max: 64 Jahre		Min: 18, Max: 43 Jahre		Min: 18, Max: 40 Jahre	
<b>Schulbildung</b>	79.0%	Abitur	70.5%	Abitur	67.7%	Abitur
	6.9%	Hochschulabschluss	12.8%	Hochschulabschluss	6.0%	Hochschulabschluss
<b>sexuelle Orientierung</b>	93.4%	heterosexuell	92.2%	heterosexuell	96.6%	heterosexuell
	2.2%	homosexuell	3.9%	homosexuell	1.2%	homosexuell
	4.4%	bisexuell	3.9%	bisexuell	2.2%	bisexuell

Die Zahl der Personen für die Einschätzungen der Typizität der Eigenschaften für Männer vs. Frauen wichen kaum voneinander ab, da die Zufallszuordnung im Zuge der Onlinebefragung sicherstellte, dass sich die Versuchspersonen nach dem Zufallsprinzip einer von zwei Fragebogenversionen zugeteilt wurden und sich das *N* in allen Zellen gleichmäßig erhöht, unabhängig von der Abbrecherquote in den unterschiedlichen Bedingungen. Die Befragung folgte innerhalb der Teilstudien der gleichen Prozedur. In allen drei Gruppen waren die demografischen Merkmale der Stichproben vergleichbar, was auf die Rekrutierungsmaßnahmen zurückzuführen ist. In der ersten Teilstichprobe zur Einschätzung der Typizität waren dies 289 Probanden (144 Männer, 145 Frauen), die die ausgewählten Eigenschaften aus der Generierungsphase danach einzuschätzen hatten, wie typisch diese für Männer in der heutigen Gesellschaft sind. In der zweiten Teilstichprobe beurteilten 294 Personen (144 Männer, 150 Frauen) die Charakteristiken danach, inwieweit diese für Frauen typisch sind, usw.

#### Beurteilung der sozialen Erwünschtheit (Phase 2)

Weitere 277 Personen nahmen an der Studie zur Einschätzung der sozialen Erwünschtheit der Items teil. Obwohl der Altersdurchschnitt ebenfalls bei ca. 23 Jahren (*SD* = 3.77 Jahre, Min: 18, Max: 43 Jahre) lag, fiel die Abiturquote mit 63% etwas geringer aus. Auch in dieser Gruppe gab die Mehrheit an heterosexuell zu sein (94%). Wie bereits angemerkt, wurden in dieser Phase alle Items nicht nach Versionen für Männer und Frauen geteilt eingeschätzt, weshalb nur eine einzelne Stichprobe nötig war.

#### Selbstbeschreibung (Phase 3)

Im dritten Schritt wurden schließlich insgesamt 1.687 Personen gebeten, die ihnen dargebotenen Adjektive im Rahmen einer Selbstbeschreibung zu bewerten. Ziel war es, diejenigen Items zu identifizieren, die von Männern und Frauen in unterschiedlichem Maße zur Selbstbeschreibung verwendet wurden. Wie auch schon bei der Typizitätseinschätzung erfolgte das Testen durch der sich reduzierenden Itemstichprobe iterativ dahingehend, dass zusätzliche Items in weiteren Befragungen nachträglich aufgenommen wurden, wodurch sich drei Teilstichproben ergaben (Tabelle 24). Auf eine Zusammenlegung der Stichproben wurde im Interesse des Berichts multivariater Effekte verzichtet.

Tabelle 24. Übersicht der Demografie zur Selbstbeschreibung (3 Teilstichproben)

Selbsteinschätzung						
	Gruppe 1		Gruppe 2		Gruppe 3	
Männer	274		135		168	
Frauen	421		143		546	
Total	695		278		714	
<hr/>						
<b>Alter der Vpn</b>	M = 22.74 Jahre		M = 23.01 Jahre		M = 22.52 Jahre	
	SD = 3.35		SD = 3.71		SD = 3.09	
<b>Schulbildung</b>	Min: 18, Max: 53 Jahre		Min: 18, Max: 41 Jahre		Min: 18, Max: 39 Jahre	
	78.0%	Abitur	75.0%	Abitur	75.0%	Abitur
<b>sexuelle Orientierung</b>	6.8%	Hochschulabschluss	7.6%	Hochschulabschluss	7.6%	Hochschulabschluss
	95.8%	heterosexuell	92.4%	heterosexuell	94.4%	heterosexuell
	1.7%	homosexuell	2.2%	homosexuell	2.5%	homosexuell
	2.5%	bisexuell	5.4%	bisexuell	3.1%	bisexuell

### 3.2.2.2 Instrumente

Die Instruktionen für die Einzelstudien ähnelten denen in Studie 1 zur Bewertung des BSRI.

#### *Instruktion Typizität*

Die Anleitung zur Einschätzung bezüglich der Typizität der neuen Items lautete: „Wir bitten Sie, die Typizität der folgenden Eigenschaften für Frauen / Männer in unserer Gesellschaft im Allgemeinen einzuschätzen. Zum Beispiel könnte eine Aussage lauten ‚Es ist typisch für Frauen / Männer in unserer Gesellschaft *gesetzestreu* zu sein“. Für die Beantwortung wurde eine siebenstufige Skala von (1) *trifft gar nicht zu* bis (7) *trifft genau zu* vorgegeben.

#### *Instruktion soziale Erwünschtheit*

Die Anweisungen für die Frage nach der sozialen Erwünschtheit der Items war ähnlich und lautete: „Wir bitten Sie, die folgenden Eigenschaften für eine Person dahingehend einzuschätzen, wie positiv bzw. negativ diese Eigenschaft ist. Zum Beispiel könnte eine Aussage lauten ‚Für eine Person in unserer Gesellschaft *gesetzestreu* zu sein – ist positiv / negativ“. Erneut wurde eine siebenstufige Antwortskala mit den Abstufungen (1) *negativ* bis (7) *positiv* vorgegeben.

#### *Instruktion Selbstbeschreibung*

Die Instruktion für die Selbstbeschreibung lautete: „Ihre Aufgabe besteht darin, sich hinsichtlich jeder Eigenschaft selbst einzuschätzen, d. h. inwieweit sie ein Teil Ihrer Persönlichkeit ist. So könnte die Eigenschaft *fröhlich* zum Beispiel auch als folgende Frage formuliert sein: Halten Sie sich für einen *fröhlichen* Menschen?“ Die Probanden hatten anschließend die Möglichkeit, sich auf einer siebenstufigen Skala von (1) *trifft gar nicht zu* bis (7) *trifft genau zu* einzuschätzen.

### 3.2.2.2.3 Durchführung

#### *Typizität*

Bezüglich der Typizität wurden die Versuchspersonen per Zufall einer von vier Fragebogenversionen zugeteilt. Erneut wurde die mögliche gleichzeitige Bewertung von sowohl Frauen als auch Männern mittels Permutation der Teile verhindert, um so einen Bias der einen auf die jeweilige andere Einschätzung zu verhindern.



Nach der Instruktion zur Studiendurchführung erfolgte zunächst die Erfassung demografischer Angaben, ehe in einer kurzen Einleitung die Versuchspersonen dann gebeten wurden die dargebotenen Adjektive dahingehend zu bewerten, wie typisch diese für einen Mann / eine Frau unserer Gesellschaft sind. Nach einer Erklärung der Skala und den Antwortmöglichkeiten wurden die Items tabellarisch, jeweils 20 Items pro Onlineseite, dargestellt. Nach der Beantwortung aller Fragen wurde den Probanden für ihre Teilnahme gedankt. Zudem hatten sie die Möglichkeit sich in eine Interessentenliste einzutragen, um eine Ergebnismeldung zu erhalten.

#### *Erwünschtheit und Selbstbeschreibung*

In der zweiten und dritten Teilstudie wurden die Einschätzungen der sozialen Erwünschtheit (wie auch der Selbstbeschreibung) erbeten. Das Prozedere ähnelte dem der Typizitätseinschätzung. So beantworteten alle Probanden zunächst einige Fragen zur Demografie. Im Anschluss erhielten sie die Instruktion, wie die Beurteilung der sozialen Erwünschtheit bzw. Selbstbeschreibung vorzunehmen ist. Es folgte dann wiederum die Darbietung der Eigenschaften in Blöcken von jeweils 20 Items pro Onlineseite.

#### **3.2.2.3 Ergebnisse**

Aufgrund des Umfangs der insgesamt drei separaten Befragungen zur Erfassung der Typizität, sozialen Erwünschtheit und Selbstbeschreibung werden die Resultate nacheinander präsentiert. Den Ergebnissen zur Einschätzung der Typizität folgen die der Bewertung zur sozialen Erwünschtheit und der Selbstbeschreibung. Die statistischen Kennwerte aus diesen Analysen wurden für die Items der jeweiligen Teilskalen in Tabellen zusammengefasst. Dazu sind in den Tabellen die MANOVA-Effekte der Items aufgeführt. Da in einigen Teilstudien das Alter der Probanden die Bewertung der Items beeinflusste, wurde es als Kovariate in den Analysen berücksichtigt.

##### *3.2.2.3.1 Ergebnisse der Einschätzungen zur Typizität der neuen Items*

###### *Typizitätsbeurteilungen der positiven maskulinen Eigenschaften*

In der Studie zur Feststellung der Typizität wurden die getesteten Haupteffekte in allen Gruppen (1 - 3) signifikant. Die MANOVA für die positiven maskulinen Eigenschaften ergab signifikante multivariate Effekte für die Variable Version,  $F(21,535) = 40.45$ ,  $p < .001$ ,  $\eta^2 = .61$  sowie für die Interaktion von Geschlecht und Version,  $F(21,535) = 5.88$ ,  $p < .001$ ,  $\eta^2 = .19$ . Für die vorliegende Aufgabenstellung sind vor allem die Haupteffekte des Faktors Version von zentraler Bedeutung, entlang dieser augenscheinlich wird, dass die jeweilige Eigenschaft von beiden Geschlechtern als signifikant typischer für Männer als für Frauen angesehen wird. In Tabelle 25 sind die mittleren Einschätzungen der Typizität der positiven maskulinen Eigenschaften für Männer und Frauen durch beide Geschlechter unter Berücksichtigung des Alters als Kovariate dargestellt. Die Eigenschaften sind nach der Größe der Unterschiede in der Geschlechtstypizität sortiert. Der Haupteffekt Version ging in der ersten Gruppe auf signifikante Unterschiede in 18 der 21 Eigenschaften dahingehend zurück, dass diese Items als typischer für Männer als für

Frauen eingeschätzt wurden (z. B. *beschützend*, *direkt*). Den deutlichsten Unterschied bezüglich der Geschlechtscharakteristik erreichte dabei das Item *handwerklich*. Das ursprünglich als typisch positive Eigenschaft für Männer frei generierte Item *ehrgeizig* wurde keinem Geschlecht als typischer zugeschrieben. Der Interaktionseffekt von Geschlecht und Version war für 19 der 21 Eigenschaften signifikant und zeigte nahezu durchgängig das gleiche Muster: die Eigenschaft wurde sowohl von den Männern als auch den Frauen als typischer für Männer eingeschätzt als für Frauen.

Tabelle 25. Itemstatistiken bzgl. der Typizität der neuen positiven maskulinen Teilskala

Maskulin positiv	ID	Einschätzung der Typizität der Eigenschaften jeweils getrennt für Männer und Frauen								Abweichung vom Skalenmittelwert der Typizitätseinschätzung		
		Trennschärfe	M für Männer	SD	M für Frauen	SD	Diff <sub>M</sub>	HE version	IA sex* version	Diff <sub>M</sub>	t	p
handwerklich	1	.63	5.32	1.14	3.11	1.25	2.21	<.001	<.001	.22	3.04	<.01
logisch	1	.76	5.20	1.31	3.38	1.41	1.82	<.001	<.001	.29	4.35	<.001
sachlich	1	.72	4.84	1.12	3.44	1.17	1.40	<.001	<.001	.14	2.53	<.05
analytisch	3	.53	4.96	1.33	3.61	1.38	1.35	<.001	<.001	.27	3.63	<.001
wetteifernd	2	.35	5.62	.98	4.28	1.28	1.34	<.001	n.s.	.94	10.80	<.001
kameradschaftlich	3	.67	5.83	.99	4.54	1.23	1.29	<.001	<.001	1.14	17.16	<.001
beschützend	1	.59	5.28	1.23	4.01	1.57	1.27	<.001	<.001	.65	10.23	<.001
unkompliziert	1	.49	4.13	1.60	2.87	1.18	1.26	<.001	<.001	-.50	-7.74	<.001
direkt	1	.63	5.06	1.23	3.83	1.31	1.23	<.001	<.001	.45	7.85	<.001
rational	1	.60	4.91	1.23	3.73	1.35	1.18	<.001	<.001	.32	5.79	<.001
praktisch	1	.63	5.21	1.24	4.09	1.29	1.12	<.001	<.001	.65	11.42	<.001
entscheidungsfreudig	3	.52	4.51	1.37	3.40	1.41	1.11	<.001	<.001	-.08	-1.04	n.s.
realistisch	3	.66	5.00	1.22	3.99	1.18	1.01	<.001	<.001	.48	7.37	<.001
furchtlos	2	.54	4.13	1.06	3.15	1.13	.98	<.001	<.001	-.38	-4.71	<.001
kraftvoll	2	.67	4.86	1.00	3.90	1.40	.96	<.001	<.01	.37	4.24	<.001
lösungsorientiert	3	.61	5.43	1.10	4.53	1.30	.90	<.001	<.001	.98	15.04	<.001
geradlinig	1	.64	4.64	1.25	3.79	1.20	.85	<.001	<.001	.22	4.61	<.001
selbstsicher	3	.54	5.18	1.08	4.34	1.14	.84	<.001	n.s.	.75	12.48	<.001
stark	1	.61	5.06	1.11	4.26	1.23	.80	<.001	<.001	.66	12.70	<.001
pragmatisch	1	.50	4.72	1.09	4.00	1.06	.72	<.001	<.001	.36	8.45	<.001
tatkräftig	3	.71	5.50	1.06	4.80	1.09	.70	<.001	<.001	1.16	20.15	<.001
gelassen	1	.44	4.56	1.08	3.96	1.02	.60	<.001	<.05	.26	5.79	<.001
durchsetzungsfähig	1	.53	5.02	.96	4.46	1.06	.56	<.001	<.01	.74	17.98	<.001
konsequent	1	.62	4.58	1.18	4.02	1.21	.56	<.001	<.001	.30	6.21	<.001
mutig	1	.48	4.56	1.08	4.16	1.23	.40	<.001	<.01	.36	7.63	<.001
ruhig	1	.26	4.16	1.19	3.77	1.06	.39	<.001	<.05	-.04	-1.33	n.s.
souverän	1	.57	4.74	.94	4.37	1.02	.37	<.001	<.001	.56	13.43	<.001
hartnäckig	2	.27	4.86	1.09	5.18	1.10	-.32	<.05	n.s.	.99	13.20	<.001
selbstbewusst	1	.43	4.99	1.13	4.76	1.02	.23	<.01	n.s.	.88	19.91	<.001
zielstrebig	1	.55	4.95	1.09	4.83	1.13	.12	n.s.	<.001	.89	19.53	<.001
zielorientiert	1	.43	4.92	1.17	4.86	1.15	.06	n.s.	n.s.	.89	18.93	<.001
ehrgeizig	1	.39	5.03	1.06	5.06	.95	-.03	n.s.	<.05	1.05	25.05	<.001

Ausnahme stellten die Items *zielstrebig*, *zielorientiert* und *ehrgeizig* dar, die von beiden Geschlechtern für das eigene Geschlecht als typischer beurteilt wurden als es die Angehörigen des jeweils anderen Geschlechts taten. So hielten etwa die Männer die Eigenschaft *zielstrebig* typischer für Männer, als Frauen dies taten. Umgekehrt schätzten Frauen dieselbe Eigenschaft als typischer für Frauen ein, als Männer dies taten. Beim Item *zielorientiert* wurde dieser Unterschied jedoch nicht signifikant. Zusätzlich ist anzumerken, dass diese Items grundsätzlich als Eigenschaften für die Endversion der positiven maskulinen Skala ungeeignet sind, da die Unterschiede der Typizitätseinschätzungen für Männer und Frauen zu gering ausfielen und der Haupteffekt der Version somit die Signifikanz verfehlte.

In ähnlicher Weise ließen sich diese Ergebnisse auch für die zusätzlichen Eigenschaften aus den Stichprobengruppen 2 und 3 zeigen (vgl. Tabelle 25) - so dass insgesamt 32 Items einer Überprüfung unterzogen wurden. Durch diese nachträglich durchgeführten Befragungen und das Hinzunehmen von Eigenschaften wie *analytisch*, *wetteifernd* oder *gelassen* wurde sichergestellt, dass genügend Items für die Endversion der positiven maskulinen Skala zur Verfügung standen.

Für einen abschließenden Auswertungsschritt sei erneut auf Tabelle 25 verwiesen, in der die Abweichungen der Typizitätseinschätzungen vom Skalenmittelwert aufgeführt sind. Als

eines der Kriterien zur Auswahl der Items sollten die Mittelwerte in den Bewertungen der Typizität idealerweise oberhalb des Skalenmittelwertes liegen. Wie der Tabelle zu entnehmen ist, trifft dies auf alle positiven maskulinen Eigenschaften mit Ausnahme der Items *entscheidungsfreudig* und *ruhig* zu. Deren Gesamtmittelwert der Einschätzung war nicht signifikant vom Skalenmittelwert verschieden. Dies ist jedoch nur bedingt problematisch, da beide Items insgesamt zumindest als typischer für Männer als für Frauen eingeschätzt wurden und in den Einschätzungen für die Männer auch über dem Skalenmittelwert lagen. Die Items *unkompliziert* und *furchtlos* dagegen lagen in ihren Typizitätseinschätzungen signifikant unterhalb des Skalenmittelwertes, was ihre Eignung für die Skala in Frage stellt.

#### Typizitätsbeurteilungen der negativen maskulinen Eigenschaften

Die MANOVA der Gruppe 1 (Tabelle 26) für die negativen maskulinen Eigenschaften ergab signifikante multivariate Effekte für die Variable Version,  $F(21,538) = 48.78$ ,  $p < .001$ ,  $\eta^2 = .66$ , sowie für die Interaktion von Geschlecht und Version,  $F(21,538) = 4.26$ ,  $p < .001$ ,  $\eta^2 = .14$ . Der Haupteffekt für die Variable Version besagt, dass sowohl Männer als auch Frauen die negativen maskulinen Eigenschaften als typischer für Männer einstufen. In Tabelle 26 sind die mittleren Einschätzungen der Typizität der 21 negativen maskulinen Eigenschaften für Männer und Frauen (von Gruppe 1 durch ID = 1 gekennzeichnet) durch beide Geschlechter unter Berücksichtigung des Alters als Kovariate dargestellt.

Tabelle 26. Itemstatistiken bzgl. der Typizität der neuen negativen maskulinen Teilskala

Maskulin negativ	ID	Trennschärfe	Einschätzung der Typizität der Eigenschaften jeweils getrennt für Männer und Frauen						HE version	IA sex* version	Abweichung vom Skalenmittelwert der Typizitätseinschätzung		
			M für Männer	SD	M für Frauen	SD	Diff <sub>M</sub>	Diff <sub>F</sub>			Diff <sub>M</sub>	t	p
wortkarg	3	.38	4.74	1.32	2.15	1.24	2.59	<.001	n.s.	-.55	-6.02	<.001	
prahlerisch	2	.66	5.42	1.14	3.04	1.28	2.38	<.001	n.s.	.24	2.11	<.05	
machohaft	1	.61	4.39	1.22	2.30	1.20	2.09	<.001	n.s.	.66	-9.88	<.001	
rauh	2	.74	4.43	1.09	2.36	1.05	2.07	<.001	<.05	-.61	-6.19	<.001	
vulgär	3	.60	4.60	1.34	2.69	1.29	1.91	<.001	n.s.	-.35	-4.29	<.001	
gewalttätig	1	.61	3.90	1.35	2.28	1.07	1.62	<.001	<.001	.91	-15.81	<.001	
angeberisch	3	.52	5.24	1.06	3.67	1.32	1.57	<.001	<.05	.44	6.04	<.001	
grob	3	.68	3.91	1.32	2.41	1.03	1.50	<.001	n.s.	-.85	-12.26	<.001	
machtbesessen	3	.61	4.86	1.31	3.39	1.37	1.47	<.001	n.s.	.11	1.42	n.s.	
unaufmerksam	3	.50	4.57	1.28	3.12	1.42	1.45	<.001	n.s.	-.19	-2.42	<.05	
selbstüberschätzend	3	.65	5.05	1.25	3.63	1.37	1.43	<.001	<.05	.31	4.15	<.001	
unordentlich	1	.45	4.97	1.19	3.57	1.14	1.40	<.001	n.s.	-.26	4.76	<.001	
faul	1	.47	4.50	1.26	3.11	1.27	1.39	<.001	n.s.	.18	-3.07	<.01	
aggressiv	1	.68	4.09	1.34	2.75	1.13	1.34	<.001	<.05	.58	-9.98	<.001	
unsensibel	1	.63	3.78	1.25	2.56	1.02	1.22	<.001	n.s.	.85	-15.94	<.001	
leichtsinnig	1	.46	4.83	1.31	3.69	1.14	1.14	<.001	n.s.	-.26	4.73	<.001	
schroff	3	.70	3.79	1.37	2.67	1.13	1.12	<.001	n.s.	-.78	-11.32	<.001	
emotionslos	3	.60	3.00	1.45	1.97	1.07	1.03	<.001	n.s.	-1.53	-22.27	<.001	
kühl	2	.61	4.08	1.19	3.07	1.16	1.01	<.001	<.001	-.43	-4.97	<.001	
schweigsam	1	.28	4.08	1.34	3.13	1.22	.95	<.001	n.s.	.42	-7.70	<.001	
gefühllos	1	.47	3.09	1.43	2.25	1.12	.84	<.001	n.s.	1.30	-23.92	<.001	
verschlossen	1	.34	4.27	1.35	3.44	1.12	.83	<.001	n.s.	.14	-2.69	<.01	
rücksichtslos	1	.61	3.96	1.20	3.20	1.31	.76	<.001	n.s.	.41	-7.91	<.001	
unzuverlässig	1	.50	3.90	1.27	3.15	1.10	.75	<.001	<.001	.48	-9.66	<.001	
überheblich	1	.54	4.43	1.27	3.75	1.16	.68	<.001	n.s.	-.07	1.32	n.s.	
geltungsbedürftig	2	.42	5.22	1.18	4.65	1.38	.57	<.01	n.s.	.96	10.98	<.001	
ignorant	1	.51	3.99	1.24	3.44	1.13	.55	<.001	n.s.	.29	-5.85	<.001	
laut	2	.48	5.03	1.39	4.50	1.22	.53	<.01	n.s.	.75	8.59	<.001	
intolerant	2	.47	3.97	1.44	3.48	1.15	.49	<.01	n.s.	-.25	-2.87	<.01	
dominant	1	.43	4.75	1.19	4.27	1.05	.48	<.001	n.s.	-.51	10.67	<.001	
egoistisch	1	.58	4.57	1.32	4.12	1.03	.45	<.001	n.s.	-.33	6.71	<.001	
stur	1	.39	4.92	1.10	4.71	1.07	.21	<.05	<.05	-.81	17.94	<.001	
rechthaberisch	1	.40	4.88	1.27	4.70	1.11	.18	n.s.	<.05	-.78	15.73	<.001	
starrsinnig	1	.40	4.69	1.08	4.57	1.06	.12	n.s.	<.01	-.62	13.76	<.001	
arrogant	1	.40	4.07	1.27	4.15	1.09	-.08	n.s.	n.s.	-.12	2.44	<.05	
ungeduldig	1	.27	4.47	1.27	4.42	1.26	.05	n.s.	n.s.	-.40	7.72	<.001	

Der Haupteffekt Version ging in der ersten befragten Gruppe auf signifikante Unterschiede in 17 Eigenschaften zurück: Diese Items wurden als typischer für Männer als für Frauen eingeschätzt (*ignorant*, *rücksichtslos*). Den eindeutigsten Unterschied bezüglich der Geschlechts-

charakteristik erreichte dabei das Item *aggressiv*. Die ursprünglich als typisch negative Eigenschaften für Männer frei generierten Items *ungeduldig*, *arrogant*, *rechthaberisch* und *starrsinnig* wurden dagegen nicht dem einen oder anderen Geschlecht als typischer zugeschrieben. Damit erfüllen sie nicht die Voraussetzungen für die Aufnahme als negative maskuline Attribute in die Endversion der Skala.

Der Interaktionseffekt von Geschlecht und Version war für sechs Eigenschaften (*gewalttätig*, *unzuverlässig*, *starrsinnig*, *rechthaberisch*, *aggressiv*, *stur*) signifikant. Vier dieser Eigenschaften (*gewalttätig*, *unzuverlässig*, *aggressiv*, *stur*) wurden von beiden Geschlechtern als typischer für Männer eingeschätzt, während die Items *starrsinnig* und *rechthaberisch* bezüglich der Typizität jeweils stärker dem jeweils anderen Geschlecht zugeschrieben wurden. Dies deutet bereits die Validierung der negativen Valenz der Begriffe an. In ähnlicher Weise ließen sich diese Ergebnisse auch für die zusätzlichen Eigenschaften aus den Befragungen (wie z. B. *intolerant*, *laut*, *grob*, *emotionslos*) der beiden zusätzlichen Stichprobengruppen zeigen (vgl. Tabelle 26).

An dieser Stelle sei erneut darauf hingewiesen, dass Items wie *unsensibel* besondere Aufmerksamkeit bei der Auswahl erfordern, da diese als Antonym potentiell ein Item der positiven femininen Teilskala sein könnten, was dem Ziel der Skalenkonstruktion, voneinander unabhängige Teilskalen zu entwickeln, die nicht lediglich ein negatives Spiegelbild bereits vorhandener Items darstellen, widersprechen würde. Darüber hinaus haben Items wie *wortkarg* sehr hohe Typizitätsratings bezogen auf die Männer, was deren Eignung für die Selbstbeschreibung dahingehend einschränkt, dass diese Eindeutigkeit die Zuschreibung dieser Eigenschaft durch die Frauen weniger wahrscheinlich macht und damit zu stark zwischen den Geschlechtern differenziert.

Im Gegensatz zu den positiven maskulinen Items muss bezüglich der Ergebnisse nicht nur festgestellt werden, dass einige Items nicht signifikant vom Skalenmittelpunkt verschieden eingeschätzt wurden, sondern dass die Einschätzung jener Items, die Signifikanz erreichten, bisweilen sogar unterhalb des Skalenmittelpunktes lagen (z. B. *emotionslos*, *gefühllos*).

#### *Typizitätsbeurteilungen der positiven femininen Eigenschaften*

Die MANOVA der eingeschätzten positiven femininen Eigenschaften aus der ersten Stichprobengruppe ergab, ebenso wie in den vorherigen Analysen, signifikante multivariate Effekte für die Variable Version,  $F(20,538) = 78.91$ ,  $p < .001$ ,  $\eta^2 = .75$  sowie für die Interaktion von Geschlecht und Version,  $F(20,538) = 7.97$ ,  $p < .001$ ,  $\eta^2 = .23$ . In Tabelle 27 finden sich die mittleren Einschätzungen der Typizität der aus der ersten Teilstichprobe (ID = 1) getesteten 20 positiven femininen Eigenschaften für Männer und Frauen durch beide Geschlechter unter Berücksichtigung des Alters als Kovariate. Der Haupteffekt Version spiegelte sich in dieser ersten Gruppe in signifikanten Unterschieden für die 20 Variablen wider. Probanden beiderlei Geschlechts hielten die femininen Eigenschaften signifikant typischer für Frauen als für Männer (z. B. *intuitiv*, *fleißig*). Den eindeutigsten Unterschied bezüglich der Geschlechtscharakteristik erreichte dabei das Item *multitaskfähig*.

Wie auch schon in der Darstellung der Analysen für die positiven maskulinen Eigenschaften sei auf das Verhältnis der Itemeinschätzungen zum Skalenmittelpunkt hingewiesen. Die Typizitätseinschätzungen aller positiven femininen Eigenschaften liegen deutlich über dem Skalenmittelwert, wobei diese Differenz für zwei der Items (*ordentlich*, *mütterlich*) nicht signifikant wurde.

Tabelle 27. Itemstatistiken bzgl. der Typizität der neuen positiven femininen Teilskala

Feminin positiv	ID	Trennschärfe	Einschätzung der Typizität der Eigenschaften jeweils getrennt für Männer und Frauen						Abweichung vom Skalenmittelwert der Typizitätseinschätzung			
			M für Männer	SD	M für Frauen	SD	Diff <sub>M</sub>	HE version	IA sex* version	Diff <sub>M</sub>	t	p
emotional	3	.71	3.59	1.36	6.11	1.01	2.52	<.001	n.s.	.86	9.90	<.001
multitaskfähig	1	.62	3.17	1.29	5.49	1.09	2.32	<.001	<.001	.33	4.95	<.001
mütterlich	3	.58	2.90	1.46	5.15	1.27	2.25	<.001	<.001	.03	0.36	n.s.
intuitiv	1	.67	3.39	1.33	5.51	.99	2.12	<.001	<.001	.47	7.13	<.001
sinnlich	2	.82	3.53	1.16	5.53	.95	2.00	<.001	n.s.	.53	5.50	<.001
romantisch	2	.81	3.68	1.17	5.60	.97	1.92	<.001	n.s.	.66	6.95	<.001
kommunikativ	1	.66	4.13	1.24	5.88	.95	1.75	<.001	<.05	.98	17.44	<.001
verführerisch	2	.69	4.19	1.28	5.84	.96	1.65	<.001	<.05	1.05	11.39	<.001
mitfühlend	1	.80	3.85	1.09	5.45	.85	1.60	<.001	<.05	.63	12.46	<.001
gute Zuhörer	1	.71	3.93	1.48	5.51	1.06	1.58	<.001	<.001	.71	11.39	<.001
ordentlich	1	.56	3.26	1.07	4.82	1.10	1.56	<.001	n.s.	.02	0.44	n.s.
einfühlsam	1	.79	3.91	1.07	5.46	.95	1.55	<.001	<.01	.66	12.99	<.001
sensibel	1	.62	4.24	1.27	5.69	.96	1.45	<.001	n.s.	.95	17.85	<.001
kreativ	1	.57	3.92	1.18	5.36	.95	1.44	<.001	<.01	.62	12.22	<.001
empfindsam	2	.70	3.91	1.18	5.35	.95	1.44	<.001	n.s.	.63	8.09	<.001
gefühlvoll	1	.80	4.17	1.18	5.58	.85	1.41	<.001	<.05	.87	16.84	<.001
zärtlich	3	.78	4.53	1.30	5.74	.92	1.21	<.001	n.s.	1.16	18.41	<.001
gepflegt	3	.63	4.73	1.17	5.92	.86	1.19	<.001	<.05	1.34	22.59	<.001
aufmerksam	3	.65	3.97	1.16	5.08	1.06	1.11	<.001	<.01	.54	8.66	<.001
leidenschaftlich	2	.64	4.45	1.15	5.55	.99	1.10	<.001	n.s.	1.02	12.81	<.001
warmherzig	1	.72	4.40	1.05	5.44	.90	1.04	<.001	n.s.	.91	19.82	<.001
verständnisvoll	1	.71	4.10	1.08	5.13	.99	1.03	<.001	<.001	.60	12.90	<.001
sozial	3	.73	4.54	1.17	5.57	.95	1.03	<.001	n.s.	1.06	18.18	<.001
sanftmütig	1	.64	3.91	1.01	4.93	.89	1.02	<.001	<.01	.42	9.32	<.001
empathisch	1	.55	3.79	.90	4.80	1.04	1.01	<.001	<.05	.29	6.40	<.001
fürsorglich	1	.67	4.52	1.15	5.47	.82	0.95	<.001	n.s.	1.00	21.63	<.001
vielseitig	1	.59	4.42	1.17	5.36	1.01	0.94	<.001	<.001	.90	18.23	<.001
fleißig	1	.53	4.29	1.09	5.15	.92	0.86	<.001	n.s.	.72	15.48	<.001
liebvoll	1	.65	4.75	.96	5.46	.84	0.71	<.001	n.s.	1.09	26.88	<.001
organisiert	1	.36	4.28	1.28	4.81	1.06	0.53	<.001	<.001	.54	10.74	<.001
gerecht	1	.38	4.21	1.01	4.43	1.20	0.22	<.05	<.001	.28	6.36	<.001

#### Typizitätsbeurteilungen der negativen femininen Eigenschaften

Die MANOVA der ersten befragten Gruppe (Tabelle 28) für die negativen femininen Eigenschaften ergab signifikante multivariate Effekte für die Variable Version,  $F(18,544) = 47.91$ ,  $p < .001$ ,  $\eta^2 = .61$  sowie für die Interaktion von Geschlecht und Version,  $F(18,544) = 6.37$ ,  $p < .001$ ,  $\eta^2 = .17$ . In Tabelle 28 sind die mittleren Einschätzungen der Typizität der negativen femininen Eigenschaften für Männer und Frauen durch beide Geschlechter unter Berücksichtigung des Alters als Kovariate angeführt.

Der Haupteffekt Version (z. B. negative feminine Eigenschaften werden den Frauen stärker zugeschrieben, durch beide Geschlechter) wurde für 17 der insgesamt in dieser ersten Gruppe überprüften 18 Items gefunden (z. B. *ängstlich*, *irrational*, *hektisch*). Lediglich das Item *besitzergreifend* wurde als gleichermaßen typisch sowohl für Frauen als auch für Männer betrachtet und erfüllt damit nicht die Voraussetzungen, um als negatives feminines Attribut in die Endversion der Skala aufgenommen zu werden. Der eindeutigste Unterschied bezüglich der Geschlechtscharakteristik zeigte sich beim Item *lästernd*.

Der Interaktionseffekt von Geschlecht und Version war für acht der 18 Eigenschaften signifikant. Sieben dieser Eigenschaften (z. B. *irrational*, *selbstzweifelnd*) wurden von beiden Geschlechtern als typischer für Frauen eingeschätzt, während *nachgiebig* im Hinblick auf die Typizität jeweils stärker dem eigenen Geschlecht zugeschrieben wurde.

Ähnlich wie bei den maskulinen Items haben einige Eigenschaften (z. B. *geschwätzig*) sehr hohe Typizitätsratings für Frauen, was deren Eignung für die Selbstbeschreibung möglicherweise dahingehend einschränkt, dass diese Eindeutigkeit das Zuschreiben dieser Charakteristik seitens der Männer weniger wahrscheinlich macht und damit zu stark zwischen den Geschlechtern trennt.

Ähnlich wie bei den negativen maskulinen Items, gab es Adjektive, die teilweise nicht signifikant vom Skalenmittelpunkt verschieden eingeschätzt wurden (*ängstlich*) oder in ihren Einschätzungen signifikant unterhalb des Mittelpunktes der Typizitätsskala lagen (*abergläubisch*, *leichtgläubig*). In Bezug auf das Item *ängstlich* ist dies weniger von Bedeutung, da es im Hinblick auf die Typizität für Frauen durchaus vom Mittelpunkt der Skala verschieden war, während beispielsweise *abergläubisch* als insgesamt weniger typisch bewertet wurde.

Diese Ergebnisse zeigten sich auch für die zusätzlichen Eigenschaften (wie *weinerlich*, *zänkisch* oder *gehässig*) aus den Stichprobengruppen 2 und 3 (vgl. Tabelle 28).

Tabelle 28. Itemstatistiken bzgl. der Typizität der neuen negativen femininen Teilskala

Feminin negativ	ID	Trennschärfe	Einschätzung der Typizität der Eigenschaften jeweils getrennt für Männer und Frauen						Abweichung vom Skalenmittelwert der Typizitätseinschätzung			
			M für Männer	SD	M für Frauen	SD	Diff <sub>M</sub>	HE version	IA sex* version	Diff <sub>M</sub>	t	p
geschwätzig	3	.65	3.16	1.56	5.74	1.44	2.58	<.001	n.s.	.50	5.81	<.001
lästernd	1	.71	3.18	1.46	5.48	1.12	2.30	<.001	n.s.	.09	4.80	<.001
zickig	1	.75	3.09	1.69	5.32	1.13	2.23	<.001	<.01	.34	2.85	<.001
intrigant	1	.69	2.56	1.31	4.74	1.19	2.18	<.001	n.s.	.22	-4.89	<.001
überechtig	3	.64	2.81	1.14	4.77	1.09	1.97	<.001	<.05	-.19	-2.58	<.05
launisch	1	.66	3.96	1.37	5.69	1.10	1.73	<.001	n.s.	.25	13.01	<.001
abergläubisch	2	.58	2.25	1.06	3.94	1.44	1.69	<.001	n.s.	-.91	-9.06	<.001
hinterlistig	1	.67	2.88	1.37	4.55	1.29	1.67	<.001	n.s.	.22	-4.76	<.001
kompliziert	1	.56	3.97	1.66	5.59	1.20	1.62	<.001	<.001	-.11	11.46	<.001
orientierungslos	3	.58	2.93	1.50	4.38	1.56	1.45	<.001	<.001	-.33	-3.92	<.001
zänkisch	3	.67	3.13	1.31	4.49	1.18	1.36	<.001	n.s.	-.17	-2.37	<.05
nachtragend	1	.64	3.54	1.49	4.86	1.13	1.32	<.001	<.05	1.17	3.70	<.001
überempfindlich	3	.70	3.89	1.48	5.20	1.24	1.31	<.001	<.01	.55	7.27	<.001
hektisch	1	.61	3.46	1.18	4.69	.96	1.23	<.001	n.s.	-.34	1.52	<.001
neidisch	3	.57	3.93	1.60	5.17	1.32	1.23	<.001	<.001	.57	7.13	<.001
selbstzweifelnd	1	.51	3.65	1.33	4.83	1.22	1.18	<.001	<.05	-.29	4.52	<.001
gehässig	3	.67	3.31	1.56	4.47	1.44	1.16	<.001	n.s.	-.12	-1.51	n.s.
selbstkritisch	3	.24	3.75	1.56	4.90	1.52	1.16	<.001	<.001	.34	4.14	<.001
empfindlich	1	.52	4.35	1.35	5.34	.85	.99	<.001	<.05	.08	16.54	<.001
ängstlich	1	.54	3.52	1.16	4.48	1.13	.96	<.001	n.s.	-.15	-0.13	n.s.
irrational	1	.51	3.22	1.33	4.14	1.51	.92	<.001	<.001	-.01	-5.69	<.001
leichtgläubig	1	.46	3.66	1.39	4.53	1.18	.87	<.001	n.s.	-.34	-1.65	<.001
weinerlich	2	.46	3.32	1.27	4.09	1.15	.77	<.001	<.01	-.28	-3.29	<.01
naiv	1	.44	3.57	1.42	4.18	1.30	.61	<.001	n.s.	.79	-2.33	<.05
eitel	1	.38	4.89	1.23	5.46	.97	.57	<.001	<.001	.79	24.71	<.001
unsicher	1	.39	3.64	1.17	4.12	1.04	.48	<.001	n.s.	.85	-2.35	<.05
nörglerisch	2	.39	4.18	1.25	4.55	1.27	.37	<.05	n.s.	.35	4.07	<.001
nachgiebig	1	.03	3.73	1.30	4.03	1.26	.30	<.01	<.001	-.13	-2.99	<.01
besitzergreifend	1	.18	4.73	1.24	4.73	1.15	.00	n.s.	n.s.	-.34	15.00	<.001

### Diskussion

Ein Aspekt ist bei der Typizitätseinschätzung auffällig und soll daher an dieser Stelle kurz diskutiert werden: die Abweichungen der Typizitätseinschätzungen der negativen maskulinen und femininen Eigenschaften vom Skalenmittelwert (Tabelle 26 und Tabelle 28). Im Gegensatz zu den positiven maskulinen und femininen Items muss bezüglich der Ergebnisse nicht nur festgestellt werden, dass einige Items nicht signifikant vom Skalenmittelpunkt verschieden bewertet wurden, sondern dass diejenigen Items, die Signifikanz erreichten, bisweilen sogar unterhalb des Skalenmittelpunktes lagen. Obwohl diese Items grundsätzlich als typischer für Männer als für Frauen (und umgekehrt) eingeschätzt werden, könnte die niedrige Ausprägung ihrer Typizitätseinschätzung bereits den Einfluss der sozialen Erwünschtheit der Items andeuten. So wurden diese Eigenschaften zwar dem entsprechenden Geschlecht eindeutig als

typischer zugeschrieben, dies jedoch in geringerem Maße im Vergleich zu sowohl anderen positiven als auch negativen Eigenschaften.

Im Zuge der Itemauswahl werden Attribute gesucht, die für die beiden Geschlechter unterschiedlich typisch und entweder sozial erwünscht oder sozial unerwünscht sind. Der Nachweis der Typizität ist mit den vorliegenden Analysen vollbracht, es fehlt die Betrachtung der Valenz im Sinne der sozialen Erwünschtheit, um klar zeigen zu können, inwieweit die Attribute positiv oder negativ sind. Wie bereits in der theoretischen Herleitung aufgeführt, wurde in diesem Zusammenhang auf eine getrennte Befragung der sozialen Erwünschtheit für Männer und Frauen verzichtet, da die Geschlechtsspezifität bereits in den Typizitätseinschätzungen abgebildet ist und eine erneute frauen- und männerbezogene Einschätzung redundant wäre.

### 3.2.2.3.2 Ergebnisse der Einschätzungen zur sozialen Erwünschtheit der neuen Items

#### Soziale Erwünschtheit der positiven maskulinen Eigenschaften

Bezüglich der Überprüfung der sozialen Erwünschtheit sei auf Tabelle 29 verwiesen, in der die Mittelwerte beider Geschlechter aufgeführt sind. Zur Berechnung der Abweichungen dieses Mittelwerts vom Skalenmittelwert wurden einfache t-Tests durchgeführt (mit entsprechender Korrektur des Alpha-Niveaus). In der Tabelle sind jeweils Trennschärfen, Mittelwerte beider Geschlechter, die t-Werte und die Signifikanzniveaus der statistischen Überprüfung angegeben.

Tabelle 29. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen positiven maskulinen Teilskala

Maskulin positiv	Einschätzung der sozialen Erwünschtheit der Eigenschaften					
	ID	Trennschärfe	$M_{\text{Gesamt}}$	SD	t	p
handwerklich	1	.25	4.90	1.17	13.19	<.001
logisch	1	.43	5.29	1.05	19.89	<.001
sachlich	1	.49	5.15	1.12	17.76	<.001
analytisch	3	.47	4.75	1.27	9.67	<.001
wetteifernd	2	.29	4.15	1.31	1.29	n.s.
kameradschaftlich	3	.35	5.96	.91	36.26	<.001
beschützend	1	.38	5.45	1.14	21.49	<.001
unkompliziert	1	.23	5.48	1.19	20.69	<.001
direkt	1	.17	5.29	1.03	21.20	<.001
rational	1	.46	4.99	1.12	14.66	<.001
praktisch	1	.54	5.44	1.00	24.48	<.001
entscheidungsfreudig	3	.47	5.61	.88	16.26	<.001
realistisch	3	.56	5.66	.90	31.12	<.001
furchtlos	2	.35	4.64	1.10	9.50	<.001
kräftvoll	2	.52	5.39	1.03	22.77	<.001
lösungsorientiert	3	.63	5.50	1.03	23.84	<.001
geradlinig	1	.50	5.20	1.25	16.26	<.001
selbstsicher	3	.37	5.82	.98	30.31	<.001
stark	1	.53	5.48	1.04	23.80	<.001
pragmatisch	1	.35	4.69	1.29	8.94	<.001
tatkräftig	3	.61	5.91	.88	36.18	<.001
gelassen	1	.34	5.75	1.01	29.83	<.001
durchsetzungsfähig	1	.55	5.71	.88	31.62	<.001
konsequent	1	.49	5.62	1.04	26.43	<.001
mutig	1	.54	5.68	.90	31.19	<.001
ruhig	1	.22	4.42	1.25	6.04	<.001
souverän	1	.60	5.74	.89	32.19	<.001
hartnäckig	2	.35	5.06	1.12	15.89	<.001
selbstbewusst	1	.40	6.38	.76	52.39	<.001

*Selbstbewusst* und *kameradschaftlich* erwiesen sich in dieser Gruppe der Eigenschaften im stärksten Maße als sozial erwünscht. Wie der Tabelle 29 zu entnehmen ist, trifft dies auf alle positiven maskulinen Eigenschaften mit Ausnahme des Items *wetteifernd* zu, welches zwar ebenfalls über dem Skalenmittelwert liegt, dabei jedoch die Signifikanz verfehlt. Es ist erneut festzustellen, dass die Beurteilungswerte signifikant über dem Skalenmittelwert lagen und damit

grundsätzlich als sozial erwünscht gelten. Wichtig sind in diesem Zusammenhang die Trennschärfen der Items vor dem Hintergrund der Zusammenfassung dieser Charakteristiken zu einer möglichen Erwünschtheitsskala. Einige Items wie *direkt*, *ruhig* oder *unkompliziert* erwiesen sich als wenig homogen für eine mögliche Gesamtskala.

#### Soziale Erwünschtheit der negativen maskulinen Eigenschaften

In den Analysen zur sozialen Erwünschtheit der negativen maskulinen Attribute zeigte sich, dass das Item *dominant* das am ehesten sozial erwünschte Item ist (wenn auch unterhalb des Skalenmittels), während Items wie *gewalttätig* und *unzuverlässig* am stärksten sozial unerwünscht sind.

Tabelle 30. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen negativen maskulinen Teilskala

Einschätzung der sozialen Erwünschtheit der Eigenschaften						
Maskulin negativ	ID	Trennschärfe	$M_{\text{Gesamt}}$	SD	t	p
wortkarg	3	.22	2.65	1.06	-21.86	<.001
prahlerisch	2	.58	2.20	1.20	-25.06	<.001
machohaft	1	.46	2.54	1.43	-17.95	<.001
rauh	2	.50	3.19	1.11	-12.98	<.001
vulgär	3	.35	2.66	1.40	-16.44	<.001
gewalttätig	1	.45	1.22	.65	-72.81	<.001
angeberisch	3	.56	2.12	1.11	-28.19	<.001
grob	3	.65	2.26	1.05	-28.40	<.001
machtbesessen	3	.53	1.95	1.17	-30.02	<.001
unaufmerksam	3	.40	1.96	.86	-40.13	<.001
selbstüberschätzend	3	.57	2.31	1.01	-28.25	<.001
unordentlich	1	.21	3.02	1.05	-15.24	<.001
faul	1	.12	2.41	1.17	-21.96	<.001
aggressiv	1	.57	2.47	1.45	-18.63	<.001
unsensibel	1	.50	2.14	1.10	-29.03	<.001
leichtsinig	1	.43	2.76	1.24	-17.29	<.001
schroff	3	.53	2.32	1.00	-28.79	<.001
emotionslos	3	.50	2.05	1.08	-31.39	<.001
kühl	2	.50	2.77	1.19	-17.47	<.001
schweigsam	1	.11	3.44	1.22	-8.15	<.001
gefühllos	1	.32	1.96	1.38	-23.64	<.001
verschlossen	1	.21	2.49	1.09	-24.15	<.001
rücksichtslos	1	.49	1.85	1.28	-29.65	<.001
unzuverlässig	1	.27	1.63	.77	-48.86	<.001
überheblich	1	.57	2.00	1.13	-30.13	<.001
geltungsbedürftig	2	.41	3.06	1.37	-11.78	<.001
ignorant	1	.43	1.87	1.09	-32.91	<.001
laut	2	.18	3.40	1.23	-8.45	<.001
intolerant	2	.46	1.82	1.06	-35.28	<.001
dominant	1	.45	3.77	1.43	-3.02	<.01
egoistisch	1	.48	2.93	1.47	-12.77	<.001
stur	1	.37	3.07	1.32	-12.26	<.001

Wie aus Tabelle 30 ersichtlich, zeigten die Analysen mit t-Tests, dass alle negativen maskulinen Eigenschaften signifikant unter dem Skalenmittel liegen, d. h. dass alle Items als sozial unerwünscht eingeschätzt wurden. Die geringsten Trennschärfen der Eigenschaften wiesen die Items *schweigsam*, *faul* und *laut* auf.

#### Soziale Erwünschtheit der positiven femininen Eigenschaften

*Gerecht* und *gepflegt* stellten sich anhand der Mittelwertvergleiche (Tabelle 31) als die am sozial erwünschtesten femininen Eigenschaften heraus. Insgesamt lagen alle Eigenschaften oberhalb des Skalenmittelwertes. Eine Ausnahme bildete lediglich das Item *mütterlich*, welches die Signifikanz verfehlte. Die positiven femininen Eigenschaften gelten folglich als sozial erwünscht. *Multitaskfähig* und *organisiert* fallen aufgrund ihrer sehr geringen Trennschärfen im Hinblick auf die allgemeine soziale Erwünschtheit auf.



Tabelle 31. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen positiven femininen Teilskala

Einschätzung der sozialen Erwünschtheit der Eigenschaften						
Feminin positiv	ID	Trennschärfe	$M_{\text{Gesamt}}$	SD	t	p
emotional	3	.54	5.12	1.13	17.19	<.001
multitaskfähig	1	.05	5.51	1.13	22.55	<.001
mütterlich	3	.26	4.04	1.40	.60	n.s.
intuitiv	1	.41	5.30	1.03	21.02	<.001
sinnlich	2	.51	5.60	1.00	27.31	<.001
romantisch	2	.55	5.39	1.14	20.55	<.001
kommunikativ	1	.35	6.01	.88	37.97	<.001
verführerisch	2	.37	5.55	1.10	24.19	<.001
mitfühlend	1	.62	5.81	.96	31.86	<.001
gute Zuhörer	1	.51	5.96	.89	38.03	<.001
ordentlich	1	.13	5.15	1.07	19.04	<.001
einfühlsam	1	.72	5.80	.99	31.59	<.001
sensibel	1	.53	5.00	1.19	14.05	<.001
kreativ	1	.29	5.66	1.08	26.42	<.001
empfindsam	2	.50	4.95	1.20	13.71	<.001
gefühlvoll	1	.60	5.84	1.09	29.48	<.001
zärtlich	3	.65	5.84	.96	32.68	<.001
gepflegt	3	.38	6.24	.87	43.17	<.001
aufmerksam	3	.43	6.19	.71	51.02	<.001
leidenschaftlich	2	.57	6.04	.95	36.62	<.001
warmherzig	1	.60	5.95	.96	34.93	<.001
verständnisvoll	1	.43	6.07	.83	42.92	<.001
sozial	3	.54	6.02	.97	36.23	<.001
sanftmütig	1	.57	5.09	1.19	15.56	<.001
empathisch	1	.24	4.60	1.36	7.52	<.001
fürsorglich	1	.65	5.75	.90	33.01	<.001
vielseitig	1	.26	6.17	.83	44.81	<.001
fleißig	1	.12	5.70	.91	32.17	<.001
liebvoll	1	.71	6.13	.89	40.63	<.001
organisiert	1	.07	5.47	.93	25.91	<.001
gerecht	1	.52	6.24	.86	45.07	<.001

*Soziale Erwünschtheit der negativen femininen Eigenschaften*

Für die negativen femininen Eigenschaften (Tabelle 32) wurde ebenfalls überprüft, inwieweit sich die Mittelwerte beider Geschlechter signifikant vom Skalenmittelwert unterschieden.

Tabelle 32. Itemstatistiken bzgl. der soz. Erwünschtheit der neuen negativen femininen Teilskala

Einschätzung der sozialen Erwünschtheit der Eigenschaften						
Feminin negativ	ID	Trennschärfe	$M_{\text{Gesamt}}$	SD	t	p
geschwätzig	3	.24	3.56	1.42	-5.18	<.001
lästernd	1	.43	2.18	1.17	-26.34	<.001
zickig	1	.40	2.22	1.12	-27.13	<.001
intrigant	1	.23	1.84	1.19	-30.28	<.001
überechtigt	3	.33	2.56	1.00	-24.43	<.001
launisch	1	.40	2.45	1.00	-26.45	<.001
abergläubisch	2	.27	2.77	1.26	-17.25	<.001
hinterlistig	1	.22	1.62	.99	-41.68	<.001
kompliziert	1	.25	3.06	1.36	-11.28	<.001
orientierungslos	3	.44	2.50	1.10	-22.65	<.001
zänkisch	3	.45	2.12	1.06	-29.67	<.001
nachtragend	1	.42	2.19	1.06	-29.53	<.001
überempfindlich	3	.49	2.12	.89	-36.00	<.001
hektisch	1	.50	2.74	.94	-22.40	<.001
neidisch	3	.40	2.35	1.03	-27.18	<.001
selbstzweifelnd	1	.38	2.52	1.36	-17.65	<.001
gehässig	3	.46	1.79	1.11	-33.60	<.001
selbstkritisch	3	-.12	5.45	1.11	21.98	<.001
empfindlich	1	.26	3.28	1.17	-10.43	<.001
ängstlich	1	.37	2.98	1.08	-16.30	<.001
irrational	1	.26	2.99	1.23	-13.78	<.001
leichtgläubig	1	.37	2.29	1.00	-29.25	<.001
weinerlich	2	.45	2.22	1.01	-29.78	<.001
naiv	1	.33	2.40	1.09	-25.26	<.001
eitel	1	.16	3.23	1.45	-9.29	<.001
unsicher	1	.45	2.52	1.08	-23.47	<.001
nörglerisch	2	.49	1.91	.91	-39.14	<.001
nachgiebig	1	.04	4.02	1.06	.11	n.s.

Insgesamt betrachtet erwiesen sich die negativen femininen Eigenschaften erwartungsgemäß weniger sozial erwünscht als die positiven. Zunächst ist festzustellen, dass die Items *hinterlistig* oder *gehässig* als die am wenigsten sozial erwünschten Items beurteilt wurden. Lediglich die Items *nachgiebig* und *selbstkritisch* wiesen Mittelwerte auf, die insgesamt oberhalb des Skalenmittelwertes lagen, grundsätzlich als sozial erwünscht gelten müssen und somit nicht die Voraussetzungen für die negative feminine Teilskala erfüllen.

Berücksichtigt man zusätzlich die Trennschärfen der Items, ist für die Eigenschaft *selbstkritisch* zudem ein negativer Koeffizient festzustellen, was – zusammen mit dem hohen Beurteilungsmittelwert – nahelegt, dass die Probanden dieses Item möglicherweise eher in der Kategorie positiv maskulin verankern würden.

### *Diskussion*

Erneut wird an dieser Stelle die Problematik unabhängiger Teilskalen sowie der Vergleichbarkeit der sozialen Erwünschtheit negativer maskuliner und femininer Eigenschaften deutlich. Bei den negativen maskulinen Eigenschaften haben Männer diese als insgesamt sozial erwünschter eingeschätzt als es die Frauen taten.

Dieses Antwortverhalten lässt sich vor dem Hintergrund des *self-serving bias* diskutieren. Demnach wissen Männer wohl um die Typizität dieser gesellschaftlich weniger sozial erwünschten Items für das eigene Geschlecht und werten diese daher indirekt auf, sofern man – berücksichtigt man die Höhe der Einschätzungen – von einer Aufwertung sprechen kann.

Das Bild bei den negativen femininen Eigenschaften ist ähnlich. Auch bei diesen Items fällt die Beurteilung der Männer (wenn auch insgesamt unterhalb des Skalenmittelwertes) leicht positiver aus als bei den Frauen. Muss man zunächst feststellen, dass die femininen negativen Eigenschaften insgesamt als offensichtlich sozial erwünschter gelten als die maskulinen negativen Eigenschaften, bleibt die Frage nach der strengeren Einschätzung der Frauen, was sich unter anderem auf die Theorien der Geschlechtsrollen und -einschätzungen zurückführen lässt. So ist zwar eine Rollenveränderung der Frauen über die Jahre (Diekman & Eagly, 2000; Diekman, Eagly, Mladinic, & Ferreira, 2005; Eagly & Steffen, 2000; Twenge, 1997; 2001) mit beispielsweise verbesserten Berufs- und Karrieremöglichkeiten festzustellen, doch sind diese Möglichkeiten der Karriere und des Einkommens beschränkt. Demnach mag eine mögliche Erklärung lauten, dass Frauen zwar sehr wohl über eine Repräsentation der typischen negativen femininen Eigenschaften (gerade auch für die eigene Gruppe) verfügen, diese aber – aufgrund der vermuteten geringeren gesellschaftlichen Bewertung – selbst als noch extremer sozial unerwünscht verurteilen, als der andere Teil der Gesellschaft dies ohnehin schon tut. Sie üben eine strengere Kritik an der eigenen Gruppe vor.

#### *3.2.2.3.3 Ergebnisse der Selbsteinschätzungen der neuen Items*

##### *Selbstbeschreibung auf den positiven maskulinen Eigenschaften*

Insgesamt wurden - ähnlich wie bei den Einschätzungen zur Typizität - aufgrund des umfangreichen Itemausschlusses im Verlauf der Studienkonstruktion drei aufeinanderfolgende Stich-

proben befragt (Gruppe 1 - 3). Die Haupteffekte wurden für alle drei Gruppen signifikant. Erneut sollen lediglich beispielhaft die Ergebnisse der ersten Teilstichprobe etwas ausführlicher dargestellt werden. Die MANOVA für die positiven maskulinen Eigenschaften mit dem Alter als Kovariate ergab den erwarteten signifikanten multivariaten Effekt für das Geschlecht,  $F(21,645) = 9.89$ ,  $p < .001$ ,  $\eta^2 = .24$ . In Tabelle 33 sind die mittleren Selbstzuschreibungen von Männern und Frauen auf den positiven maskulinen Eigenschaften für Männer und Frauen unter Berücksichtigung des Alters als Kovariate dargestellt.

Tabelle 33. Itemstatistiken bzgl. der Selbstbeschreibung der neuen positiven maskulinen Teilskala

Selbstbeschreibung von Männern und Frauen								
Maskulin positiv	ID	Trennschärfe	$M_{\text{Männer}}$	$SD$	$M_{\text{Frauen}}$	$SD$	$\text{Diff}_M$	$p$
handwerklich	1	.27	4.09	1.15	4.02	1.24	.07	n.s.
logisch	1	.47	5.57	1.16	5.07	1.05	.50	<.001
sachlich	1	.52	5.25	.92	4.71	1.05	.54	<.001
analytisch	3	.23	5.25	1.39	4.65	1.54	.60	<.001
wetteifernd	2	.26	4.55	1.52	3.85	1.54	.71	<.001
kameradschaftlich	3	.43	5.86	.98	5.88	.94	-.02	n.s.
beschützend	1	.33	5.16	1.16	4.87	1.13	.29	<.01
unkompliziert	1	.30	4.60	1.56	4.40	1.48	.20	n.s.
direkt	1	.38	4.90	1.31	4.82	1.35	.08	n.s.
rational	1	.39	5.23	1.15	4.82	1.24	.41	<.001
praktisch	1	.55	5.04	1.17	4.94	1.13	.10	n.s.
entscheidungsfreudig	3	.47	4.27	1.46	3.87	1.54	.40	<.05
realistisch	3	.39	5.48	1.12	5.40	1.17	.08	n.s.
furchtlos	2	.47	4.14	1.51	3.28	1.36	.86	<.001
kraftvoll	2	.56	4.55	1.43	3.91	1.51	.64	<.001
lösungsorientiert	3	.44	5.72	1.03	5.61	1.15	.11	n.s.
geradlinig	1	.52	4.60	1.34	4.80	1.23	-.20	<.05
selbstsicher	3	.53	4.73	1.37	4.48	1.42	.25	<.05
stark	1	.51	4.61	1.20	4.63	1.12	-.02	n.s.
pragmatisch	1	.29	4.52	1.34	4.39	1.17	.13	n.s.
tatkräftig	3	.59	5.37	1.17	5.48	1.11	-.11	n.s.
gelassen	1	.35	5.34	1.08	4.67	1.20	.67	<.001
durchsetzungsfähig	1	.54	4.83	1.10	4.87	1.06	-.04	n.s.
konsequent	1	.57	4.90	1.13	4.80	1.15	.10	n.s.
mutig	1	.47	4.40	1.17	4.18	1.10	.22	<.05
ruhig	1	.04	4.98	1.35	4.22	1.43	.76	<.001
souverän	1	.60	4.76	1.11	4.72	1.05	.04	n.s.
hartnäckig	2	.44	5.11	1.25	5.13	1.13	-.03	n.s.
selbstbewusst	1	.54	4.88	1.22	4.86	1.18	.02	n.s.
zielstrebig	1	.56	4.91	1.21	5.37	1.18	-.46	<.001
zielorientiert	1	.58	5.07	1.19	5.43	1.07	-.36	<.001
ehrgeizig	1	.44	4.80	1.26	5.24	1.21	-.44	<.001

Insgesamt wurde das Item *logisch* im Mittel von beiden Geschlechtern am stärksten zur Selbstbeschreibung herangezogen, *handwerklich* dagegen am wenigsten, jedoch immer noch knapp über dem Skalenmittelwert. Der Haupteffekt Geschlecht ging dabei auf signifikante Unterschiede in elf der Items in dieser ersten Teilstichprobe zurück. Sieben dieser Eigenschaften offenbarten den erwarteten Geschlechterunterschied. So beanspruchten die Männer beispielsweise erwartungsgemäß das Item *ruhig* für die eigene Selbstbeschreibung stärker als dies die Frauen für sich taten. Während einige weitere Items zumindest die gleiche Richtung der Differenz zwischen Männern und Frauen zeigten, wurden andere Items wiederum sogar stärker von Frauen als von den Männern reklamiert – entgegen der Hypothesen. So sahen Frauen sich als signifikant *zielstrebig*, *ehrgeiziger* und *zielorientierter*. Berücksichtigt man die Ergebnisse der Typizitätseinschätzungen der Gruppe 1 überrascht dieses Ergebnis jedoch weniger, da sich dort bereits gezeigt hatte, dass besonders diese drei Items von beiden Geschlechtern für das eigene Geschlecht als typischer eingeschätzt wurden. Daher scheinen diese Charakteristiken für die Endversion der positiven maskulinen Skala grundsätzlich ungeeignet.

Das multimodale Vorhaben sicherte demnach durch das Anlegen verschiedenster Kriterien die ideale Auswahl der Items. Dies führte aber, wie bereits dargelegt, zum Ausschluss vieler Variablen und zur Reduktion der Auswahl derer, die für die Endversionen der Skalen in Betracht kommen. So beschrieben sich beispielsweise zwar beide Geschlechter mit dem Item *ruhig*, was der hohen sozialen Erwünschtheit geschuldet ist, gleichzeitig zeigte sich die erwartete Differenz einer stärkeren Beanspruchung seitens der Männer (zurückzuführen auf die höhere Typizität für diese Gruppe), was für ein gutes Item spricht. Jedoch offenbart ein Blick auf die Trennschärfen im Rahmen der Reliabilitätsanalysen für eine mögliche positive maskuline Skala, dass dieses Item als eines der ersten aus der Skala aufgrund mangelnder Homogenität ausgeschlossen werden sollte. Aus diesem Grund wurden, ähnlich wie in den Befragungen zur Bestimmung der Typizitätseinschätzung, zusätzliche Items (wie *analytisch*, *furchtlos* oder *tatkräftig*) aus der Befragung weiterer Teilstichproben der gleichen Parameterbestimmung unterzogen.

#### Selbstbeschreibung auf den negativen maskulinen Eigenschaften

Die MANOVA für die negativen maskulinen Eigenschaften mit dem Alter als Kovariate ergab für die erste Teilstichprobe ebenfalls den erwarteten signifikanten multivariaten Effekt für das Geschlecht,  $F(21,655) = 11.57$ ,  $p < .001$ ,  $\eta^2 = .27$ . Die mittleren Selbstzuschreibungen von Männern und Frauen auf den negativen maskulinen Eigenschaften (unter Berücksichtigung des Alters als Kovariate) finden sich in Tabelle 34.

Tabelle 34. Itemstatistiken bzgl. der Selbstbeschreibung der neuen negativen maskulinen Teilskala

Selbstbeschreibung von Männern und Frauen								
Maskulin negativ	ID	Trennschärfe	$M_{\text{Männer}}$	SD	$M_{\text{Frauen}}$	SD	Diff <sub>M</sub>	p
wortkarg	3	.07	3.37	1.68	2.80	1.70	.57	<.001
prahlerisch	2	.60	3.09	1.61	2.51	1.36	.58	<.01
machohaft	1	.43	2.38	1.30	1.53	.97	.85	<.001
rauh	2	.60	3.30	1.45	2.46	1.15	.84	<.001
vulgär	3	.44	3.08	1.61	2.18	1.36	.90	<.001
gewalttätig	1	.50	1.55	.95	1.36	.68	.19	<.01
angeberisch	3	.49	3.36	1.39	2.71	1.24	.64	<.001
grob	3	.62	2.61	1.21	2.04	1.08	.57	<.001
machtbesessen	3	.49	2.92	1.69	2.26	1.37	.66	<.001
unaufmerksam	3	.34	3.29	1.47	2.89	1.41	.40	<.01
selbstüberschätzend	3	.48	3.09	1.47	2.38	1.20	.71	<.001
unordentlich	1	.27	4.34	1.71	4.28	1.68	.06	n.s.
faul	1	.38	4.66	1.33	4.18	1.32	.48	<.001
aggressiv	1	.51	2.28	1.21	2.13	1.18	.15	n.s.
unsensibel	1	.48	2.46	1.17	2.03	1.00	.43	<.001
leichtsinnig	1	.34	3.37	1.37	3.13	1.22	.24	<.05
schruff	3	.62	2.60	1.27	2.17	1.19	.43	<.001
emotionslos	3	.44	2.43	1.61	1.53	1.03	.90	<.001
kühl	2	.41	3.68	1.53	2.77	1.47	.91	<.001
schweigsam	1	.06	3.73	1.53	3.38	1.42	.35	<.01
gefühllos	1	.40	2.18	1.18	1.59	.95	.59	<.001
verschlossen	1	.17	3.60	1.57	3.40	1.34	.20	n.s.
rücksichtslos	1	.60	2.19	1.15	1.91	.93	.28	<.001
unzuverlässig	1	.40	2.31	1.24	2.05	1.05	.26	<.01
überheblich	1	.57	3.22	1.37	2.65	1.12	.57	<.001
geltungsbedürftig	2	.47	4.01	1.62	3.75	1.55	.25	n.s.
ignorant	1	.54	2.54	1.20	2.18	1.01	.36	<.001
laut	2	.35	3.86	1.63	4.15	1.64	-.28	n.s.
intolerant	2	.36	2.47	1.40	2.25	1.47	.22	n.s.
dominant	1	.43	3.67	1.34	3.70	1.43	-.03	n.s.
egoistisch	1	.59	3.24	1.41	3.23	1.26	.01	n.s.
stur	1	.45	3.63	1.45	4.06	1.39	-.43	<.001
rechthaberisch	1	.59	3.89	1.45	3.82	1.41	.07	n.s.
starrsinnig	1	.48	3.57	1.39	3.69	1.28	-.12	n.s.
arrogant	1	.62	2.66	1.45	2.37	1.23	.29	<.01
ungeduldig	1	.36	3.66	1.39	4.11	1.41	-.45	<.001

Insgesamt wurde das Item *gewalttätig* im geringsten Maße von Frauen und Männern zur Selbstbeschreibung herangezogen, während dagegen die Eigenschaft *faul* von beiden Geschlechtern verhältnismäßig stark für die eigene Gruppe beansprucht wurde, sogar stärker als vergleichsweise positive maskuline Items wie *handwerklich* oder *mutig*. Der Haupteffekt Geschlecht ging in dieser ersten Teilstichprobe auf signifikante Unterschiede bei 14 der Items zurück. Mit Ausnahme von zwei Items (*ungeduldig*, *stur*) zeigte sich die Verteilung erwartungsgemäß. Männer schrieben sich die Items signifikant stärker zu als Frauen dies taten, z. B. *unsensibel* oder *überheblich*. Diese Ergebnisse zeigten sich in ähnlicher Form auch für die beiden übrigen Teilstichproben.

An den Mittelwerten der Items wird sehr schnell deutlich, dass – obwohl sie negativ in ihrer Valenz sein sollten – sie bisweilen überraschend wenig von den Probanden im klassischen Sinne abgelehnt wurden. So fallen die Mittelwerte der Items zwar vergleichsweise geringer aus als bei den positiven Items, jedoch nicht derart niedrig, wie man es aus dem theoretischen Vorgehen heraus hätte erwarten können. So müssen einige Items aufgrund zu hoher Mittelwerte für die Endversion der Skala als ursprünglich negative erachtete Eigenschaften ausgeschlossen werden.

Erneut offenbart ein Blick auf die Trennschärfen der Items ein weiteres Problem. So führte das Item *schweigsam* zwar zu dem erhofften Geschlechtsunterschied in der Selbstbeschreibung dahingehend, dass Männer höhere Werte aufwiesen als Frauen, und die Selbstzuschreibung liegt – wenn auch eigentlich zu knapp – unterhalb des Skalenmittels, doch mit einer Trennschärfe von lediglich .06 würde dieses Items zugunsten einer zufriedenstellenden Reliabilität der Endskala nicht berücksichtigt werden. Daher wurden in den Folgebefragungen zusätzliche Items (wie *kühl*, *laut* oder *wortkarg*) hinsichtlich ihrer Eignung für die Skala überprüft, für die sich ähnliche Ergebnisse ergaben und deren Kennwerte ebenfalls Tabelle 34 entnommen werden können.

#### *Selbstbeschreibung auf den positiven femininen Eigenschaften*

Die MANOVA für die positiven femininen Eigenschaften mit dem Alter als Kovariate ergab, zunächst erneut nur für die ersten Teilstichprobe berechnet, einen signifikanten multivariaten Effekt für das Geschlecht,  $F(20,627) = 13.45$ ,  $p < .001$ ,  $\eta^2 = .30$ . In Tabelle 35 sind die mittleren Selbstzuschreibungen von Männern und Frauen auf den positiven femininen Eigenschaften getrennt für Männer und Frauen zusammengefasst. Insgesamt wurde das Item *gute Zuhörer* am stärksten sowohl von Frauen als auch Männern zur Selbstbeschreibung herangezogen. Demgegenüber wurde die Eigenschaft *ordentlich* von beiden Geschlechtern in deutlich geringerem Maße zur Selbstbeschreibung genutzt.

Der Haupteffekt Geschlecht ging dabei auf signifikante Unterschiede bei 16 der Items zurück, allesamt mit dem erwarteten Verlauf der Selbstzuschreibung. Frauen beurteilten sich auf allen Items signifikant stärker als Männer dies taten, z. B. *intuitiv*, *fleißig* oder *gefühlvoll*. Im Trend zeigte sich das auch für die verbliebenen vier Items (z. B. *ordentlich*), deren Zuschreibung von Männern und Frauen jedoch die Signifikanz verfehlte.

Tabelle 35. Itemstatistiken bzgl. der Selbstbeschreibung der neuen positiven femininen Teilskala

Feminin positiv	ID	Trennschärfe	Selbstbeschreibung von Männern und Frauen				Diff <sub>M</sub>	p
			M <sub>Männer</sub>	SD	M <sub>Frauen</sub>	SD		
emotional	3	.51	4.39	1.52	5.66	1.29	1.27	<.001
multitaskfähig	1	.32	4.39	1.39	5.39	1.03	1.00	<.001
mütterlich	3	.47	2.78	1.48	4.26	1.71	1.48	<.001
intuitiv	1	.39	4.68	1.19	5.14	1.09	.46	<.001
sinnlich	2	.67	4.50	1.38	4.82	1.29	.32	n.s.
romantisch	2	.55	4.76	1.48	5.13	1.54	.36	n.s.
kommunikativ	1	.38	5.05	1.31	5.39	1.13	.34	<.001
verführerisch	2	.50	4.02	1.63	4.70	1.39	.67	<.001
mitfühlend	1	.60	5.09	1.00	5.74	.88	.65	<.001
gute Zuhörer	1	.49	5.66	1.11	6.03	.88	1.00	<.001
ordentlich	1	.22	3.87	1.72	4.10	1.66	.23	n.s.
einfühlsam	1	.68	4.99	1.11	5.63	.92	.64	<.001
sensibel	1	.32	4.72	1.39	5.33	1.14	.61	<.001
kreativ	1	.37	4.51	1.32	4.86	1.30	.35	<.001
empfindsam	2	.50	4.58	1.21	5.16	1.17	.58	<.001
gefühlvoll	1	.70	5.06	1.13	5.73	.93	.67	<.001
zärtlich	3	.62	5.22	1.31	5.72	1.12	.50	<.001
gepflegt	3	.44	5.75	.98	6.20	.87	.45	<.001
aufmerksam	3	.36	5.08	1.18	5.43	1.04	.35	<.001
leidenschaftlich	2	.58	5.12	1.32	5.66	1.11	.54	<.001
warmherzig	1	.66	5.26	1.09	5.75	.88	.49	<.001
verständnisvoll	1	.58	5.46	.93	5.70	.83	.24	<.001
sozial	3	.64	5.12	1.30	5.72	1.13	.60	<.001
sanftmütig	1	.42	4.70	1.17	4.87	1.20	.17	n.s.
empathisch	1	.29	4.38	1.22	4.77	1.37	.39	<.001
fürsorglich	1	.61	5.17	1.08	5.48	.93	.31	<.001
vielseitig	1	.46	5.31	1.12	5.40	1.02	.09	n.s.
fleißig	1	.41	4.27	1.29	4.76	1.18	.49	<.001
liebvoll	1	.64	5.21	.99	5.68	.87	.47	<.001
organisiert	1	.32	4.72	1.32	5.09	1.21	.37	<.001
gerecht	1	.37	5.64	.89	5.68	.80	.04	n.s.

Hinsichtlich der Trennschärfen bleibt festzustellen, dass sich bei den positiven femininen Eigenschaften bisher das homogenste Bild zeigte.

Zusätzlich überprüfte Items aus den weiteren Teilstichproben 2 und 3 sind beispielsweise *sinnlich*, *zärtlich* oder *aufmerksam*, deren Skalenkennwerte der Tabelle 35 entnommen werden können.

#### Selbstbeschreibung auf den negativen femininen Eigenschaften

Die MANOVA für die negativen femininen Eigenschaften mit dem Alter als Kovariate ergab ebenfalls den erwarteten signifikanten multivariaten Effekt für das Geschlecht,  $F(18,660) = 11.29$ ,  $p < .001$ ,  $\eta^2 = .24$  (berechnet für die erste Teilstichprobe).

Die mittleren Selbstzuschreibungen von Männern und Frauen auf den negativen maskulinen Eigenschaften für Männer und Frauen finden sich in Tabelle 36. Insgesamt wurde das Item *intrigant* am wenigsten von Frauen und Männern zur Selbstbeschreibung herangezogen. Demgegenüber wurden die Eigenschaften *empfindlich* und *nachgiebig* von beiden Geschlechtern vergleichsweise stark für sich beansprucht. Ihre durchschnittlichen Einschätzungen lagen im Durchschnitt nicht nur oberhalb des Skalenmittelwertes, sondern sogar in der Einschätzung höher als vergleichsweise positive feminine Items, wie beispielsweise *ordentlich*.

Der Haupteffekt Geschlecht ging dabei in dieser ersten Teilstichprobe auf signifikante Unterschiede bei 13 der Items zurück. Mit Ausnahme von zwei Items (*intrigant*, *hinterlistig*) waren die Unterschiede gemäß der Erwartung. Frauen beschrieben sich auf den Items signifikant stärker als Männer dies taten (*orientierungslos* oder *zickig*).

Tabelle 36. Itemstatistiken bzgl. der Selbstbeschreibung der neuen negativen femininen Teilskala

Feminin negativ	ID	Trennschärfe	Selbstbeschreibung von Männern und Frauen				Diff <sub>M</sub>	p
			M <sub>Männer</sub>	SD	M <sub>Frauen</sub>	SD		
geschwätzig	3	.21	3.54	1.45	4.26	1.70	.72	<.001
lästernd	1	.33	3.06	1.32	3.53	1.33	.47	<.001
zickig	1	.57	2.46	1.33	3.42	1.38	.96	<.001
intrigant	1	.27	2.01	1.16	1.78	.99	-.23	<.01
übersichtlich	3	.32	3.41	1.37	3.65	1.54	.23	n.s.
launisch	1	.52	3.42	1.38	4.14	1.28	.72	<.001
abergläubisch	2	.17	1.90	1.20	2.55	1.66	.66	<.001
hinterlistig	1	.31	2.12	1.24	1.80	1.00	-.32	<.001
kompliziert	1	.51	3.76	1.62	4.19	1.46	.43	<.001
orientierungslos	3	.36	2.32	1.25	2.82	1.67	.50	<.001
zänkisch	3	.53	2.38	1.29	2.64	1.31	.26	<.05
nachtragend	1	.40	3.19	1.56	3.28	1.38	.09	n.s.
überempfindlich	3	.65	3.13	1.64	3.95	1.64	.83	<.001
hektisch	1	.46	3.06	1.34	3.58	1.22	.52	<.001
neidisch	3	.47	2.99	1.52	3.24	1.35	.25	n.s.
selbstzweifelnd	1	.50	3.85	1.48	4.11	1.45	.26	<.05
gehässig	3	.38	2.82	1.58	2.48	1.42	-.33	<.05
selbstkritisch	3	.14	5.14	1.28	5.56	1.14	.42	<.001
empfindlich	1	.56	3.88	1.33	4.46	1.27	.58	<.001
ängstlich	1	.55	2.99	1.28	3.74	1.29	.75	<.001
irrational	1	.43	2.86	1.33	3.08	1.16	.22	<.05
leichtgläubig	1	.42	2.94	1.44	3.30	1.46	.36	<.001
weinerlich	2	.50	2.15	1.06	3.41	1.50	1.26	<.001
naiv	1	.38	2.76	1.27	2.97	1.37	.21	<.05
eitel	1	.27	3.81	1.62	3.99	1.47	.18	n.s.
unsicher	1	.52	3.70	1.42	3.88	1.34	.18	n.s.
nörglerisch	2	.39	3.39	1.49	3.75	1.45	.36	<.05
nachgiebig	1	.07	4.24	1.12	4.13	1.24	-.11	n.s.
besitzergreifend	1	.40	3.58	1.39	3.73	1.36	.15	n.s.

Wie auch schon bei den Mittelwerten der negativen maskulinen Items wird erneut deutlich, dass – obwohl sie negativ in ihrer Valenz sein sollten – die negativen femininen Items bisweilen doch überraschend stark von den Probanden zur Selbstbeschreibung herangezogen wurden. So fiel zwar der Mittelwert der negativen femininen Eigenschaften deutlich geringer aus als bei den positiven Items, jedoch nicht so gering, wie man aus dem theoretischen Vorgehen heraus hätte erwarten können. So müssen einige Items aufgrund zu hoher Werte für ursprünglich als negativ erachtete Eigenschaften für die Endversion der Skala ausgeschlossen werden.

Die Betrachtung der Itemtrennschärfen zeichnete – ähnlich wie bei den positiven femininen Eigenschaften – insgesamt ein homogenes Bild. Durch die Befragung der beiden zusätzlichen Teilstichproben wurden weitere Items (wie *weinerlich*, *zänkisch* oder *gehässig*) hinsichtlich ihrer Eignung für die Skala überprüft. Ihre Skalenskennwerte, die sich in die bisherigen Analysen in ähnlicher Form eingliedern lassen, finden sich in Tabelle 36.

#### 3.2.2.3.4 Auswahl der Eigenschaften für die Endversion der Skala

Basierend auf den Analysen wurden aus den jeweils vier Kategorien (feminin vs. maskulin, positiv vs. negativ) sechs Eigenschaften für die finale Skala ausgewählt. Die Auswahl erfolgte, wie bereits an anderer Stelle dargelegt, nach klar definierten Kriterien wie beispielsweise Items, die signifikant typischer für das eine Geschlecht als für das andere sein sollten, dabei über dem Skalenmittelwert liegen müssten, und gleichzeitig für eine größtmögliche Differenz zwischen der Typizität für Männer und Frauen sorgen sollten.

In Tabelle 37 ist der Entscheidungsprozess für die Endauswahl der Items als Ausschnitt dargestellt. Für alle vier Teilskalen lassen sich so jeweils in den drei Beurteilungsphasen proble-

matische Items auflisten, die einzelne oder sogar mehrere Voraussetzungen für eine Aufnahme in die Endversion des Messinstruments verletzen.

Tabelle 37. Auszuschließende Items

Mas+	Mas-	Fem+	Fem-
<b>Typizität</b>			
<i>ehrig</i> (nicht typischer) <i>zielstrebig, zielorientiert</i> und <i>ehrig</i> (jeweils typischer fürs eigene Geschlecht beansprucht) <i>entscheidungsfreudig, ruhig</i> (nicht sign. vom Skalenmittelwert als typischer eingeschätzt) <i>unkompliziert, furchtlos</i> (unterhalb des Skalen-M für Typizität)	<i>ungeduldig, arrogant, rechthaberisch, starsinnig</i> (nicht dem einen oder anderen Geschlecht als typischer zugeschrieben) <i>unsensibel</i> (schwierig wegen Vorsilbe) <i>wortkarg</i> (zu hohe Typizitätswerte für Männer) <i>emotionslos, gefühllos</i> (unterhalb des Skalen-M für Typizität)	<i>ordentlich, mütterlich</i> (nicht signifikant oberhalb des Skalen-M für Typizität)	<i>besitzergreifend</i> (gleichermaßen typisch für beide Geschlechter) <i>nachgiebig</i> (jeweils dem eigenen Geschlecht als typischer zugeschrieben) <i>geschwätzig</i> (zu hohe Typizitätswerte für Frauen) <i>ängstlich</i> (nicht oberhalb des Skalen-M für Typizität) <i>leichtgläubig</i> (unterhalb des Skalen-M für Typizität)
<b>soziale Erwünschtheit</b>			
<i>wetteifernd</i> (nicht signifikant oberhalb des Skalen-M für soziale Erwünschtheit) <i>direkt, ruhig, unkompliziert</i> (wenig homogen für Erwünschtheits-skala)	<i>schweigsam, faul, laut</i> (wenig homogen für Erwünschtheits-skala)	<i>mütterlich</i> (nicht signifikant oberhalb des Skalen-M für soziale Erwünschtheit) <i>multitaskfähig, organisiert</i> (wenig homogen für Erwünschtheits-skala)	<i>nachgiebig, selbstkritisch</i> (signifikant oberhalb des Skalen-M für soziale Erwünschtheit) <i>selbstkritisch</i> (wenig homogen für Erwünschtheitskala)
<b>Selbstbeschreibung</b>			
<i>handwerklich</i> (am wenigsten zur Selbstbeschreibung herangezogen) <i>zielstrebig, zielorientiert</i> und <i>ehrig</i> (Frauen beanspruchten diese Items stark für sich)	<i>gewalttätig</i> (am wenigsten zur Selbstbeschreibung herangezogen) <i>faul</i> (von beiden Geschlechtern zur Selbstbeschreibung beansprucht) <i>ungeduldig, stur</i> (Frauen beanspruchten diese Items stark für sich)	<i>ordentlich</i> (am wenigsten zur Selbstbeschreibung herangezogen) <i>sinnlich, romantisch</i> (von beiden Geschlechtern zur Selbstbeschreibung beansprucht)	<i>intrigant</i> (am wenigsten zur Selbstbeschreibung herangezogen) <i>empfindlich, nachgiebig</i> (von beiden Geschlechtern zur Selbstbeschreibung beansprucht) <i>intrigant, hinterlistig</i> (von Frauen nicht stärker zur Selbstbeschreibung beansprucht)

Durch diesen Auswahlprozess sowie das bereits beschriebene Testen weiterer Items im Zuge der Skalenkonstruktion ergab sich nach diesen Kriterien für die vier Teilskalen schließlich folgende Auswahl:

Tabelle 38. Auswahl der Items für die vier Teilskalen

Mas+	Mas-	Fem+	Fem-
logisch rational sachlich lösungsorientiert analytisch praktisch	überheblich angeberisch rücksichtslos machtbesessen prahlerisch schroff	einfühlsam gefühlvoll zärtlich emotional leidenschaftlich liebepoll	orientierungslos naiv überevorsichtig überempfindlich ängstlich selbstzweifelnd

### 3.2.2.4 Zusammenfassende Diskussion

Insbesondere die getrennt vorgenommene Erfassung der sozialen Erwünschtheit der Items für Männer und Frauen erwies sich als schwierig. Folgt man der Argumentation von Hosoda und Stone (2000), die unter anderem den Bezug zu Spence (1985) herstellten, steht fest, dass die Geschlechterschematheorie ein multifaktorielles und multidimensionales Konzept darstellt. Insofern lässt sich nicht notwendigerweise erwarten, dass geschlechtsbezogene Überzeugungen und Verhaltensweisen über verschiedene Kontexte und Themen hinweg zwischen Männern und Frauen konsistent sind. Auch andere Studien griffen die Problematik der sozialen



Erwünschtheit auf. So stellten beispielsweise Prentice und Carranza (2002) fest, dass eine Vielzahl von Charakteristiken in jeder Kategorie existierten, die nicht in der komplementären Kategorie des anderen Geschlechts auftauchten. Bedeutsam sind in diesem Zusammenhang vor allem Items, bei denen es keine Überschneidung für Geschlechter gibt, wie beispielsweise *intelligent* oder *humorvoll*, die als weniger sozial erwünscht für Frauen aber nicht automatisch erwünschter für Männer gelten, ebenso wie *optimistisch* weniger erwünscht für Männer scheint, gleichwohl aber auch nicht automatisch erwünschter für Frauen – ein Muster, was sich bisweilen auch in den vorliegenden Daten wiederfinden ließ. Diese Ergebnisse zeigen vor allem auf, dass soziale Präskriptionen für Frauen und Männer nicht einfach spiegelbildlich sind, d.h. die Standards des einen Geschlechts sind mitunter weniger streng als für das andere. Darüber hinaus gab es in der Studie von Prentice und Carranza (2002) für eine Vielzahl von Items eine hohe Korrespondenz von sozialer Erwünschtheit und Typizität. Dennoch wurden auch Items identifiziert, die sich nicht unterschieden in Hinblick auf soziale Erwünschtheit, wohl aber in den Typizitätsratings. Dies traf speziell bezüglich typisch positiver Eigenschaften für Frauen zu, was als Indiz dafür gewertet werden kann, dass Frauen – verglichen zu sozialen Standards – womöglich insgesamt als positiver erachtet werden. Auch Krahe et al. (2007) kamen bei ihrem Vorgehen der Skalenkonstruktion zu dem Schluss, dass die getrennte Einschätzung der Eigenschaften für Mädchen und Jungen eine wesentlich höhere Schwelle für die Identifizierung geschlechtstypischer Eigenschaften darstellte als Studien, die zur Bewertung der Geschlechtstypizität beispielsweise eine bipolare Skala mit Endpunkten *männlich* vs. *weiblich* zugrunde legten.

Ehrenberg, Cataldegirmen und Klauer (2001) stellten in ihrer Studie fest, dass die Urteilsdimensionen Valenz und Geschlechtstypikalität nicht unabhängig voneinander sind; und die Mittelwerte zu  $r = .60$  ( $p < .01$ ) korrelierten. Obwohl in der Geschlechterstereotypenforschung maskuline häufig positiver bewertet werden als die femininen Eigenschaften (z. B. Spence, Helmreich & Stapp, 1975), fand sich in der Studie von Ehrenberg et al. über alle Stimuli und Personen hinweg ein starker Zusammenhang zwischen der Positivität und Femininität der Einschätzungen. Die durch die positiven Verhaltensweisen implizierten Eigenschaften scheinen überwiegend eher mit der weiblichen Geschlechtsrolle assoziiert zu sein.

Zusammenfassend lässt sich festhalten, dass die Auswahlkriterien für die Items der Zielskala als sehr anspruchsvoll zu bewerten sind, weshalb sich schnell Probleme in dem Versuch offenbarten, diesen strengen Kriterien Rechnung zu tragen. Das Merkmal der sozialen Erwünschtheit wurde zur scheinbar unüberwindbaren Hürde, vor allem bei der Bewertung der negativen Eigenschaften, wie aus den Kommentaren der Probanden zu entnehmen war. Aufgrund des bereits angeführten strengen Auswahlverfahrens kam es zum Ausschluss einer großen Anzahl von Eigenschaften, da viele Items aus der Konstruktionsphase den Kriterien nicht standhielten. Dies erforderte wiederum ein adaptives Überprüfen zusätzlicher Items durch weitere Teilstichproben. Wenngleich das Nachtesten bestimmter Items die Befragung mehrerer Teilstichproben notwendig machte, sicherte dieses Vorgehen schlussendlich die Auswahl der

sechs angestrebten Items pro Teilskala bei gleichzeitiger Berücksichtigung der zuvor festgelegten Qualitätsrichtlinien.

### 3.2.3 Studie 5: Validierung der neuen Teilskalen

#### 3.2.3.1 Fragestellung

Ziel dieser Studie war es, die Reliabilität und Validität der neuen Teilskalen zu überprüfen. Die für die Endversion der Skala ausgewählten Items sollten von den Probanden dahingehend eingeschätzt werden, in welchem Maße diese einen Teil ihres Selbst darstellen. Zur Bestimmung der Reliabilität wurde neben der internen Konsistenz der Teilskalen für einen Teil der Stichprobe eine Wiederholungsmessung nach acht Wochen durchgeführt, um auch die Re-Test-Reliabilität des Instruments prüfen zu können.

Für die Validitätsprüfung des neuen Instruments kam eine theoriegeleitete Auswahl von Konstrukten zum Einsatz, durch die die unterschiedliche Vorhersageleistung der einzelnen Skalen zu Konstrukten außerhalb des Geschlechtsrollenselbst nachgewiesen werden sollte. Bezüglich der Validierung sind in Tabelle 39 die einzelnen Konstrukte und deren spezifische Vorhersagen aufgeführt, die im folgenden Abschnitt ausführlich dargestellt werden.

Tabelle 39. Zuordnung und Vorhersagen der Validierungskonstrukte (Studie 5)

Konstrukte	Vorhersagen
<b>Prädiktoren für MAS+</b> berufl. Ehrgeiz (AVEM) Hoffen auf Erfolg (AMS)	Die positive Maskulinitätsskala ist spezifischer Prädiktor und klärt mehr Varianz bezüglich der Konstrukte beruflicher Ehrgeiz sowie Hoffen auf Erfolg auf, indem sie eine bessere Vorhersageleistung erzielt als die drei übrigen Konstrukte.
<b>Prädiktoren für MAS-</b> Aggressivität (AQ) Werteskala Macht	Die negative Maskulinitätsskala sagt das Konstrukt Aggressivität sowie das Konzept der machtbezogenen Werte besser vorher als Mas+ und die beiden Femininitätsskalen.
<b>Prädiktoren für FEM+</b> soziale Werte Empathie (Sorge)	Die positive feminine Teilskala hat spezifischen Vorhersagewert für die Konstrukte der sozialen Werte sowie Empathie und ist für beide Konzepte ein besserer Prädiktor als die drei übrigen Konstrukte.
<b>Prädiktoren für FEM-</b> ESWS Ängstl. Depressivität (BSF)	Die negative Femininitätsskala sagt die Konstrukte emotionale Selbstwertschätzung und Ängstliche Depressivität besser vorher und klärt mehr Varianz auf als die maskulinen Teilskalen und Fem+.

#### 3.2.3.2 Methode

##### 3.2.3.2.1 Stichprobe

Die Gesamtstichprobe setzte sich aus fünf Teilstichproben zusammen, an denen insgesamt 800 Personen (272 Männer und 528 Frauen) teilnahmen. Über die Teilgruppen hinweg und damit erneut vergleichbar zu allen in dieser Arbeit präsentierten Studien, waren die Versuchspersonen im Schnitt 23.08 Jahre ( $SD = 3.86$ , Min: 18, Max: 47 Jahre) alt. Bezüglich des Schulabschlusses berichteten 61%, die Schule mit dem Abitur abgeschlossen zu haben; weitere 8% konnten einen Hochschulabschluss vorweisen. 91% gaben Heterosexualität als ihre sexuelle Orientierung an.

### 3.2.3.2.2 Instrumente

#### Neues Instrument zur Erfassung positiver und negativer Aspekte des GSK

Zum Einsatz kamen die aus der Konstruktionsphase abgeleiteten und als typisch eingeschätzten sechs Items der jeweils vier Teilskalen. Diese 24 Attribute wurden in tabellarischer Form einzeln dargeboten (Tabelle 40). Zu Beginn erhielten die Probanden eine ausführliche Instruktion: „Sie finden nun eine Reihe von Eigenschaften, mit denen man sich selber beschreiben kann. So könnte die Eigenschaft *fröhlich* zum Beispiel auch als folgende Frage formuliert sein: ‚Halten Sie sich für einen *fröhlichen* Menschen?‘ Lesen Sie sich bitte jede Eigenschaft durch und schätzen Sie anschließend ein, inwieweit diese auf Sie zutrifft oder nicht.“. Im Anschluss hatten die Versuchspersonen die Möglichkeit, auf einer fünfstufigen Antwortskala von (1) *trifft gar nicht zu* bis (5) *trifft genau zu* ihre Selbstbeschreibung anzugeben.

Tabelle 40. Darstellung der Erfassung der Geschlechtsrollenitems

Ich bin ...	Wie sehr trifft diese Eigenschaft auf Sie zu					Wie wichtig ist es für Sie diese Eigenschaft zu besitzen				
	trifft gar nicht zu				trifft genau zu	gar nicht wichtig				sehr wichtig
... liebevoll	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... naiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... ängstlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zusätzlich wurde die Erfassung einer Wichtigkeitskomponente integriert, indem die Probanden gebeten wurden, zu jedem Item auch Angaben bezüglich der Wichtigkeit dieser Eigenschaft für das eigene Selbst zu machen (Instruktion: „*Wie wichtig ist es für Sie diese Eigenschaft zu besitzen*“). Dazu wurde direkt neben der fünfstufigen Antwortskala zur Selbsteinschätzung eine weitere Skala mit ebenfalls fünf Antwortmöglichkeiten von (1) *gar nicht wichtig* bis (5) *sehr wichtig* dargeboten, auf der die Probanden ihre Einschätzungen vornehmen konnten. Da die Daten der Wichtigkeitserfassung für den Validierungsprozess nicht relevant sind, werden sie an dieser Stelle der Ergebnisdarstellung nicht weiter berücksichtigt, sondern erst in Studie 7 (siehe ab S. 191) aufgegriffen.

#### BSRI

Um die konvergente Validität überprüfen zu können, wurde in der vorliegenden Studie ebenfalls der BSRI in seiner deutschen Version eingesetzt (Bem, 1974; Schneider-Düker, 1978; Schneider-Düker & Kohler, 1988). Die Instruktion, bezogen auf die BSRI-Items lautete: „In diesem Teil der Befragung besteht Ihre Aufgabe darin, bei jeder Eigenschaft einzuschätzen, inwieweit sie ein Teil Ihrer Persönlichkeit ist. So könnte die Eigenschaft *fröhlich* zum Beispiel auch als folgende Frage formuliert sein: ‚Halten Sie sich für einen *fröhlichen* Menschen?‘“. Für die Einschätzungen wurden die Eigenschaften jeweils einzeln dargeboten. Für die Beantwortung wurde in dem Fall eine fünfstufige Skala mit den Abstufungen (1) *fast nie* bis (5) *immer* vorgegeben. Damit bleibt die Art der Antwortskala als Häufigkeitserfassung der ursprünglichen Formulierung Bems ähnlich, gleichzeitig wurde das siebenstufige Antwortformat um zwei Stufen reduziert, um die Vergleichbarkeit zu den Antworten der neuen Geschlechtsrollenitems zu gewährleisten (ebenfalls fünfstufig). Bei der Erfassung wurde

aufgrund der Ökonomie und der ohnehin ausreichend langen Befragung nur auf die ursprünglich als feminine und maskuline entwickelten Attribute zurückgegriffen.

#### *Validierung MAS+*

##### *AMS (Achievement Motive Scale – Hoffen auf Erfolg)*

Zur Messung der Hoffnung auf Erfolg wurde die Achievement Motive Scale (AMS) von Gjesme und Nygard (1970) in der deutschen Übersetzung von Götttert und Kuhl (1980) eingesetzt. In der vorliegenden Studie wurde in Anlehnung an Engeser (2005) nur die Skala *Hoffnung auf Erfolg* mit je fünf Aussagen (z. B. „*Ich möchte gern vor eine etwas schwierige Arbeit gestellt werden.*“) über fünffach gestufte Ratingskalen von (1) *trifft überhaupt nicht zu* bis (5) *trifft vollkommen zu* erhoben, die mit Items anderer Validierungskonstrukte gemischt wurden.

##### *AVEM (beruflicher Ehrgeiz)*

Bei dem AVEM handelt es sich um ein Instrument zur interventionsbezogenen Diagnostik beruflichen Bewältigungsverhaltens (Schaarschmidt & Fischer, 1996). In der vorliegenden Studie kamen lediglich die sechs Items der Teilskala des *beruflichen Ehrgeizes* zum Einsatz. Diese Dimension beinhaltet Komponenten des Strebens nach Zielen und des Weiterkommens im Beruf (z. B. „*Ich möchte beruflich weiter kommen, als es die meisten meiner Bekannten geschafft haben.*“) und wurde hypothetisch in einen Zusammenhang mit der positiven maskulinen Teilskala gebracht. Wie auch Items anderer zur Validierung herangezogenen Konstrukte in dieser Studie wurden die Aussagen in einem umfassenden Befragungsteil integriert und zusammen mit den Aussagen anderer Konstrukte dargeboten. Die übereinstimmende Antwortskala wies ein fünfstufiges Format auf von (1) *trifft überhaupt nicht zu* bis (5) *trifft vollkommen zu*.

#### *Validierung MAS-*

##### *Aggressivität (AQ)*

Bezogen auf die Trait-Aggressivität und der dispositionalen Aggressionsneigung kam in der vorliegenden Studie eine Teilskala aus der von Möller (2006) entwickelten Skala zum Einsatz. Insgesamt hatte Möller (2006) einen Fragebogen mit 14 Items konzipiert, der auf dem *Aggression Questionnaire* (Buss & Warren, 2000) aufbaute, und aus vier Komponenten bestand: *Ärger*, *Verbale Aggression*, *Relationale Aggression* und *Physische Aggression*. Erneut wurde aus Gründen der Ökonomie der Befragung und der zielgerichteten Verwendung der Validierungskonstrukte lediglich eine der vier Teilkomponenten eingesetzt: *Physische Aggression* mit vier Items (z. B. „*Wenn mich jemand richtig provoziert, kann ich auch zuschlagen.*“). Da die Items mit denen anderer Validierungskonstrukte gemischt wurden, hatte die Antwortskala ein fünfstufiges Format von (1) *trifft überhaupt nicht zu* bis (5) *trifft vollkommen zu*. Die Instruktion zu diesem Teil der Befragung wurde aufgrund der allgemeinen Gültigkeit neutral wie folgt formuliert: „Im Folgenden bitten wir Sie, folgende Aussagen dahingehend zu beantworten, inwieweit Sie Ihnen zustimmen. Bitte markieren Sie die Antworten, die Ihrer Meinung am besten entsprechen.“

### *Werteskala Macht*

Die Wichtigkeit bestimmter machtbezogener Werte (und zur Validierung von Mas-) für die Probanden wurde mittels eines an den Schwartz Value Survey (Schwartz, 1992) angelehnten Fragebogens von Schröder (2005) erfasst. Das hier eingesetzte Konzept der *machtbezogenen Werte* umfasste fünf Aussagen zu sozial weniger ausgerichteten Tugenden, Gefühlen und Meinungen von Personen mit Schwerpunkt auf Durchsetzungsfähigkeit, Reichtum und Führung. Dazu wird jeweils ein spezifischer Wert aufgeführt und anschließend mit wenigen Worten kurz erklärt (z. B. *Soziale Macht – Kontrolle über andere, Dominanz*). Für diesen Befragungsteil wurde eine fünfstufige Antwortskala vorgegebenen, mit den Optionen (1) *unwichtig*, (2) *eher unwichtig*, (3) *unentschieden*, (4) *eher wichtig*, (5) *wichtig*.

Die Probanden erhielten zu diesem Teil folgende Instruktion: „Dieser Teil beschäftigt sich mit Ihren Werten, Gefühlen und Meinungen. Zunächst bitten wir Sie, sich folgende Frage zu stellen: ‚*Welche Werte sind für mich wichtig als Leitprinzipien meines Lebens, welche sind für mich weniger wichtig?*‘ Dieser Teil enthält eine Liste mit Werten. Nach Nennung eines jeden Wertes ist in Klammern eine kurze Erläuterung gegeben zum Verständnis der Wertbegriffe. Die Liste ist keineswegs vollständig, d. h. es ist gut möglich, dass Sie persönlich wichtige Werte nicht wiederfinden. Es geht jedoch nur um Ihre Einschätzung der hier aufgelisteten Werte.“

### *Validierung FEM+*

#### *Soziale Werte*

Die Wichtigkeit bestimmter sozialer Werte (und zur Validierung von Fem+) für die Probanden wurde ebenfalls mittels eines an den Schwartz Value Survey (Schwartz, 1992) angelehnten Fragebogens von Schröder (2005) erfasst. Das Konzept der *sozialen Werte* umfasste sieben Aussagen zu sozial ausgerichteten Tugenden, Gefühlen und Meinungen von Personen. Erneut wurde jeweils ein spezifischer Wert aufgeführt und anschließend mit wenigen Worten kurz erklärt (z. B. *Verantwortungsbewusstsein: zuverlässig, verlässlich sein*). Auch für diesen Befragungsteil wurde eine fünfstufige Antwortskala gewählt, mit den Optionen (1) *unwichtig*, (2) *eher unwichtig*, (3) *unentschieden*, (4) *eher wichtig*, (5) *wichtig*. Die Instruktion war die gleiche wie für die machtbezogene Werteskala.

#### *Empathie (Sorge)*

Für die Validierung vom FEM+ kam ein Fragebogen von Maes, Schmitt und Schmal (1995) zum Einsatz, dessen Komponente *Sorge um die Gefühle anderer* (z. B.: „*Ich mache mir Sorgen um Menschen, denen es schlechter geht als mir.*“) mit neun Items relevant war. Der Abschnitt des Fragebogens war mit dieser Instruktion überschrieben: „Im Folgenden bitten wir Sie, die Aussagen dahingehend zu beantworten, inwieweit diese zutreffen. Bitte markieren Sie die Antworten, die am besten Ihrer Meinung entsprechen.“ Es kam die bereits in anderen Teilen eingesetzte fünfstufige Antwortskala von (1) *trifft gar nicht zu* bis (5) *trifft genau zu* zum Einsatz.

### *Validierung FEM-*

#### *ESWS (Emotionale Selbstwirksamkeit)*

Zur Validierung der negativen femininen Teilskala wurde das Konzept der emotionalen Selbstwirksamkeit eingesetzt, welches der multidimensionalen Selbstwertkala (MSWS) entnommen wurde. Dieses von Schütz und Sellin (2006) für den deutschen Sprachraum entwickelte Verfahren basiert auf der Multidimensional Self-Concept Scale (Fleming & Courtney, 1984) und erfasst mittels eines Selbstbeurteilungsfragebogens verschiedene Facetten der subjektiven Selbstwertschätzung bei Jugendlichen und Erwachsenen. Die hier eingesetzte Teilskala Emotionale Selbstwirksamkeit (ESWS, Cronbachs Alpha = .84) besteht aus sieben Items, welche einzeln mit einer siebenstufigen Ratingskala von (1) *gar nicht* bis (7) *sehr* dargeboten werden. Die diesem Teil vorangestellte Instruktion lautete: „Im Folgenden finden Sie Fragen, die persönlichen Einstellungen und Eigenschaften betreffen. Bitte lesen Sie diese sorgfältig durch und entscheiden Sie, wie sehr diese Eigenschaften auf Sie zutreffen.“

#### *Stimmung (Ängstliche Depressivität)*

*Ängstliche Depressivität* wurde mit der entsprechenden Teilskala aus dem Berliner Stimmungsfragebogen (BSF) erfasst, bei dem es sich um eine Weiterentwicklung (Hörhold & Klapp, 1993) des Mehrdimensionalen Stimmungsfragebogens von Hetcheltjen und Mertesdorf (1973) handelt. Die Probanden wurden in der Instruktion nicht – wie in den Originalfragebögen – nach der augenblicklichen Stimmung gefragt, sondern allgemein um eine Einschätzung hinsichtlich ihrer Person gebeten. Die hier eingesetzte Skala besteht aus fünf Items (z. B. *deprimiert, betrübt*). Die Skala wurde direkt an die Erfassung der Geschlechtsrollenitems mit gleicher Instruktion („*Bitte geben Sie bei jeder der folgenden Eigenschaften an, inwieweit sie ein Teil Ihrer Persönlichkeit ist.*“) und Antwortskala (siebenstufig von (1) *trifft überhaupt nicht zu* bis (7) *trifft vollkommen zu*) präsentiert. Aus Gründen der großen Anzahl von Items und des daraus resultierenden Zeitaufwandes für die Probanden wurden immer nur einzelne Konstrukte präsentiert, die zusammen das Gesamtbild ergeben.

#### *3.2.3.2.3 Durchführung*

Da für eine Online-Untersuchung ohne Belohnung der Fragebogen mit allen Validierungskonstrukten zu lang sein würde, mit der daraus resultierenden Gefahr eines erhöhten, unerwünschten Drop-Outs, wurden die Probanden per Zufall einer von mehreren Varianten zugewiesen, die randomisiert maximal drei bis sechs Konstrukte enthielten. Somit erfolgte die Befragung in insgesamt fünf Versionen, in denen unterschiedliche Kombinationen der Validierungskonstrukte dargeboten wurden.

Die Grundstruktur innerhalb der fünf Varianten ist identisch: zunächst wurden die Probanden über das Ziel der Untersuchung und die zugesicherte Anonymität informiert. Dann erhielten sie den Hinweis, dass sie die Untersuchung jederzeit beenden könnten. Nach Angabe ihrer demografischen Informationen bearbeiteten alle Personen zuerst die Geschlechtsrollenitems hinsichtlich ihrer Selbstzuschreibung (Tabelle 40). Die Zahl der Probanden, die die einzelnen Maße bearbeitet hatten, unterschied sich aufgrund der Verteilung

der Konstrukte auf unterschiedliche Versionen. Es ist jedoch abgesichert, dass für jedes Konstrukt eine hinreichend große Stichprobe vorliegt (vgl. Tabelle 41).

Dann erst unterschieden sich die fünf Gruppen (1-5) dahingehend, dass ein jeder Proband eine Reihe von Validierungskonstrukten bearbeitete. Da diese bisweilen als bekannt galten und ihre Beantwortung die anderen Konstrukte beeinflussen könnte, wurde nicht nur die Reihenfolge der Konstrukte permutiert, sondern auch ihre Zusammenstellung mit Auswirkung auf die Anzahl der zu bearbeitenden Konstrukte. Zudem waren auch die Argumente der Ökonomie und Länge des Fragebogens ausschlaggebend für diese Art der Konstruktzusammenstellung. So hat beispielsweise eine Gruppe von Probanden sowohl die Fragebögen zur Aggressivität, Hoffen auf Erfolg als auch zur emotionalen Selbstwertschätzung ausgefüllt. Dieses Vorhaben führte schließlich zu unterschiedlichen Versuchspersonenzahlen für die einzelnen Konstrukte, die zusammen mit den Antwortkategorien und den Versuchspersonenzahlen in Tabelle 41 aufgeführt sind.

Tabelle 41. Übersicht der eingesetzten Validierungskonstrukte

Konstrukt	Gesamt	N		Skalenniveau
		Männer	Frauen	
<b>Prädiktoren für MAS+</b>				
berufl. Ehrgeiz (AVEM)	425	148	277	1-5
Hoffen auf Erfolg (AMS)	561	191	370	1-5
<b>Prädiktoren für MAS-</b>				
Aggressivität (AQ)	562	192	370	1-5
Werteskala Macht	227	84	143	1-5
<b>Prädiktoren für FEM+</b>				
Empathie (Sorge)	425	148	277	1-5
soziale Werte	227	84	143	1-5
<b>Prädiktoren für FEM-</b>				
ESWS	565	194	371	1-7
Ängstl. Depressivität (BSF)	141	47	95	1-7

### 3.2.3.3 Ergebnisse

#### 3.2.3.3.1 Explorative Faktorenanalyse der neuen Teilskalen

Zunächst wurde eine Faktorenanalyse zur Überprüfung der theoretisch hergeleiteten Teilskalen mit Maximum-Likelihood-Methode und schiefwinkliger Rotation durchgeführt. Diese Umsetzung geht von einer Korrelation der Faktoren und ist in Fällen der Datenreduktion bei theoretischen Anhaltspunkten mit anzunehmender Korrelation zwischen den Faktoren anzuwenden (Moosbrugger & Schermelleh-Engel, 2007). Die Analyse ergab eine eindeutige Struktur von vier Faktoren mit Eigenwerten  $> 1$ , die insgesamt 54.50% der Varianz aufklärten.

Die positiven Ladungen der Eigenschaften, alle größer als .40, ließen sich in der Strukturmatrix (Tabelle 42) sehr gut den vier Teilskalen zuordnen. Faktor 1 bildet die Items der positiven maskulinen Teilskala ab, Faktor 2 die Items der Skala Mas-, Faktor 3 die Items der Skala Fem+ sowie Faktor 4 die Skala negative Femininität. Mit dieser Rotationslösung wurden substantielle positive Sekundärladungen auf anderen Faktoren vermieden. Dies ist besonders bedeutsam vor dem Hintergrund, dass gerade Sekundärladungen auf den gegenpoligen Skalen, wie beispielsweise Fem+ und Mas-, die Unabhängigkeit der Skalen beeinflussen würden. Auffällig ist, dass die Faktorladungen über alle Items der femininen negativen Teilskala insgesamt die geringsten Faktorladungen aufweisen.

Tabelle 42. Matrix der Faktorladungen der Eigenschaften (Explorative Faktorenanalyse, PC)

	Faktoren			
	Mas+	Mas-	Fem+	Fem-
logisch	.74			
sachlich	.73			
rational	.70			
lösungsorientiert	.70			
analytisch	.68			
praktisch	.64			
angeberisch		.81		
prahlerisch		.78		
überheblich		.75		
machtbesessen		.68		
rücksichtslos		.68		
schruff		.54		
gefühlvoll			.87	
zärtlich			.82	
liebepoll			.82	
einfühlsam			.76	
leidenschaftlich			.74	
emotional			.65	
ängstlich				.77
übereorsichtig				.71
zweifelnd				.70
überempfindlich				.69
orientierungslos				.53
naiv				.49

Moosbrugger und Schermelleh-Engel (2007) argumentierten, dass die klassische Hauptkomponentenanalyse (PC) vor allem dem Ziel nachgeht, möglichst viel Varianz der beobachteten Variablen aufzuklären und dabei von einer messfehlerfreien Erhebung dieser ausgeht. Die Hauptachsenanalyse (PAF) dagegen bemüht sich um das eigentliche Aufdecken latenter Faktoren mit ihrem Beziehungsmuster zueinander und berücksichtigt dabei die Messfehlervarianz der Variablen.

Tabelle 43. Matrix der Faktorladungen der Eigenschaften (Explorative Faktorenanalyse, PAF)

	Faktoren			
	Mas+	Mas-	Fem+	Fem-
logisch	.68			
sachlich	.68			
rational	.65			
lösungsorientiert	.62			
analytisch	.61			
praktisch	.56			
angeberisch		.78		
prahlerisch		.73		
überheblich		.68		
machtbesessen		.61		
rücksichtslos		.60		
schruff		.45		
gefühlvoll			.87	
zärtlich			.77	
liebepoll			.77	
einfühlsam			.72	
leidenschaftlich			.66	
emotional			.59	
ängstlich				.73
übereorsichtig				.64
zweifelnd				.61
überempfindlich				.59
orientierungslos				.43
naiv				.40

Daher wurde zur Überprüfung der Struktur eine weitere Faktorenanalyse mit PAF berechnet, die insgesamt 45.87% der Varianz aufklärte (Tabelle 43) und zu einer vergleichbaren Strukturmatrix führte. Eine explorative Faktorenanalyse gilt als struktursuchendes, ohne auf einer Theorie



basierendes Verfahren, welches zur Aufdeckung nicht bekannter Zusammenhänge zwischen Variablen und Faktoren anzuwenden ist, jedoch keinen Modelltest ermöglicht.

Im Falle der hier vorliegenden Datenstruktur und theoretischen Herleitung gilt es jedoch auch die hypothetisch definierte Struktur zu überprüfen. Aus diesem Grund soll eine konfirmatorische Faktorenanalyse durchgeführt werden, die das Testen der faktoriellen Struktur, wie diese theoriegeleitet bereits vor der Analyse feststeht, vornimmt und bereits existierende Hypothesen über das Zusammenwirken der Faktoren zu bestätigen versucht.

3.2.3.3.2 Konfirmatorische Faktorenanalyse der neuen Teilskalen

Hinsichtlich der neuen Geschlechtsrollenitems soll ihre Zuordnung zu den respektiven Teilskalen durch konfirmatorische Faktorenanalysen bestätigt werden. Im Gegensatz zu explorativen Analysen erlaubt die CFA eine gezielte Hypothesentestung und erfordert methodisch die Postulierung eines Modells sowie die Festlegung der Faktorenanzahl und der Zugehörigkeit der Itemladungen zu den Faktoren.

Mittels konfirmatorischer Faktorenanalysen können somit theoretisch postulierte Modelle vor allem durch eine genaue Definition der Zusammenhänge von beobachteten Items und latenten Faktoren gegeneinander geprüft werden (Abbildung 15).

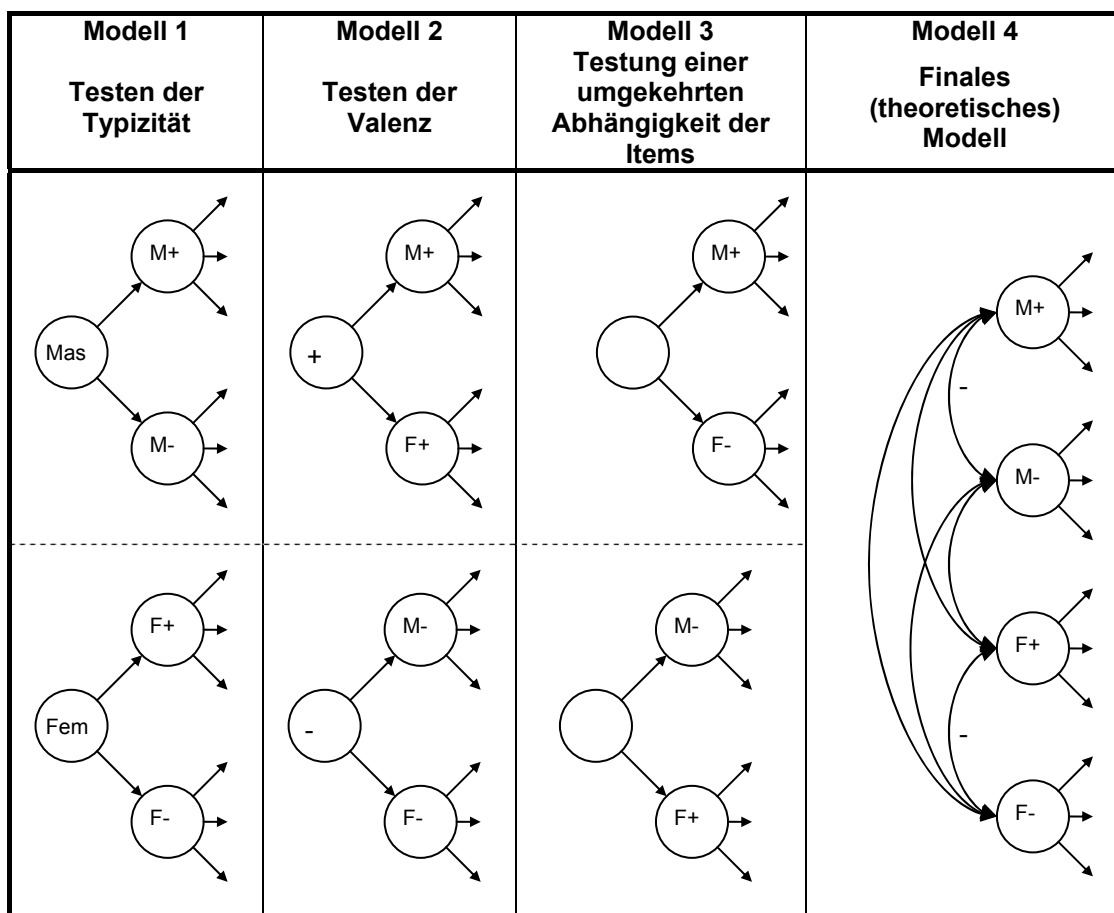


Abbildung 15. Modelle zur Überprüfung der Beziehungen der Teilskalen zueinander

Das einfachste Modell geht davon aus, dass alle Items – ungeachtet der Subskalenstruktur – auf einem einzigen Faktor laden. In Modellen mit zwei Faktoren sind verschiedene

Kombinationen möglich: in einem Modell ist die Facette der Typizität das entscheidende Merkmal derart, dass die jeweils positiven und negativen Teilskalen einer Dimension (Maskulinität oder Femininität) in engerem Zusammenhang zueinander stehen (als die Valenz), während in einem weiteren Modell jeweils beide positiven sowie negativen Teilskalen die größere Beziehung zueinander aufweisen (und nicht etwa die Typizität).

In einem dritten Zweifaktorenmodell kann die umgekehrte Beziehungsstruktur berücksichtigt werden, wonach jeweils negative Maskulinität und positive Femininität die größten gemeinsamen Varianzanteile aufweisen würden. Dieses Modell spiegelt die Sorge wider, dass die negativen Teilskalen jeweils nur die umgedrehten Gegenstücke der positiven Teilskalen der jeweils anderen Dimension sein könnten. Im letzten Modell schließlich wird das Beziehungskonstrukt aller vier Teilskalen als maximal unabhängig voneinander betrachtet, gleichbedeutend mit dem Bestreben, durch die Hinzunahme negativer Teilskalen zusätzlichen Aufklärungsgewinn zu erhalten. Lässt man Korrelationen zwischen den latenten Variablen zu, erhält man folgende Parameter für die unterschiedlichen Modelle, jeweils auch getrennt berechnet für Männer und Frauen (Tabelle 44). Die Kennwerte sind zur besseren Darstellung der feinen Unterschiede mit drei Dezimalstellen angegeben.

Tabelle 44. Modellprüfung mit den Items der neuen Skala

	n.s. p <sub>χ<sup>2</sup></sub>	< 3.00 χ <sup>2</sup> /df	> .950 CFI	< .050 RMSEA
<b>1. Ein-Faktoren-Modell</b>	<b>.000</b>	<b>4.11</b>	<b>.909</b>	<b>.062</b>
Männer	.000	2.12	.902	.064
Frauen	.000	2.87	.905	.060
<b>2. Zwei-Faktoren-Modell (Typizität)</b>	<b>.000</b>	<b>3.11</b>	<b>.939</b>	<b>.051</b>
Männer	.000	1.67	.942	.050
Frauen	.000	2.26	.937	.049
<b>3. Zwei-Faktoren-Modell (Valenz)</b>	<b>.000</b>	<b>3.40</b>	<b>.930</b>	<b>.055</b>
Männer	.000	1.64	.945	.048
Frauen	.000	2.55	.922	.054
<b>4. Zwei-Faktoren-Modell (umgekehrt)</b>	<b>.000</b>	<b>2.95</b>	<b>.943</b>	<b>.049</b>
Männer	.000	1.63	.945	.048
Frauen	.000	2.19	.940	.048
<b>5. Vier-Faktoren-Modell</b>	<b>.000</b>	<b>2.25</b>	<b>.964</b>	<b>.040</b>
Männer	.000	1.57	.951	.046
Frauen	.000	1.87	.957	.041

$$N_{\text{Männer}} = 272, N_{\text{Frauen}} = 528$$

Der Modell-Fit als Maß überprüft, inwieweit die implizierte Kovarianz-Matrix der Bevölkerung der empirisch beobachteten Matrix entspricht (Schermelleh-Engel et al., 2003; Stapleton, 1997; Weston & Gore, 2006). Dabei zielt die Modellprüfung auf ein nicht-signifikantes Chi-Quadrat als Ergebnis ab, da dieser Wert nachweist, dass das hypothetisierte Modell durch die Daten tatsächlich abgebildet wird. Vor dem Hintergrund der Tatsache, dass der Chi-Quadrat-Wert in solchen Modellsimulationen statistisch komplex ist und u. a. von der Anzahl der Parameter, den Freiheitsgraden oder der Stichprobengröße abhängt, werden weitere Kennwerte zur Modelltestung herangezogen, um so mehr, wenn die Chi-Quadrat-Testung zu einem unerwünschten signifikanten Ergebnis führt (Schermelleh-Engel et al., 2003). Das Chi-Quadrat war in den Analysen tatsächlich signifikant für alle Modelle, weshalb – entsprechend den methodischen Empfehlungen – auf Kennwerte, wie das Verhältnis von  $\chi^2/df$  (welches so klein wie möglich sein sollte), CFI (Bentler, 1990) und RMSEA zurückgegriffen wurde. Dabei wurden strenge Kriterien des Verhältnisses von  $\chi^2/df < 3.00$ , einem CFI als Fit-Index  $> .95$  sowie in

Kombination mit einem *RMSEA*-Wert  $< .50$  als Voraussetzungen für einen zufriedenen Fit festgelegt (Hu & Bentler, 1999; MacKenzie, Podsakoff, & Jarvis, 2005; Moosbrugger & Schermelleh-Engel, 2007; Schermelleh-Engel et al., 2003), die höher lagen als beispielsweise der vorgeschlagene Wertebereiche für den Fit-Index von  $.80$  bis  $.90$  (Bentler & Bonett, 1980).

Obwohl der Fit für alle Modelle – bis auf eines – akzeptabel war, und dies ohne den Einsatz von Itemparcel, die den Modell-Fit durch Reduktion der zu integrierenden Variablen zu verbessern vermögen (Choi et al., 2007; Meade & Kroustalis, 2006), erwies sich das postulierte Vier-Faktoren-Modell als das am besten repräsentierte (Tabelle 44). Das Einfaktorenmodell zeigte erwartungsgemäß den schwächsten Modell-Fit ( $\chi^2/df = 4.11$ ;  $p \leq .001$ ;  $CFI = .91$ ;  $RMSEA = .062$ ). Das hypothetisch konstruierte Vier-Faktoren-Modell erwies sich im Vergleich zu allen möglichen Modellen als das stärkste, mit den besten Kennwerten für den Modell-Fit ( $\chi^2/df = 2.25$ ;  $p \leq .001$ ;  $CFI = .97$ ;  $RMSEA = .040$ ), was das Ergebnismuster der explorativen Faktorenanalyse widerspiegelt. Dies galt auch für die getrennt berechneten Analysen für Männer und Frauen, für die sich die Parameter nur gering unterschieden.

Die Analysen zeigten (Abbildung 16), dass alle Items den ursprünglich zugeschriebenen Faktoren auch tatsächlich zugeordnet wurden. Dabei variierten die Faktorladungen der einzelnen Items von  $.39$  bis  $.91$ . Die Items der positiven maskulinen Teilskala waren in ihren Ladungen etwas homogener als beispielsweise die Ladungen der negativen femininen Items bezüglich ihres respektablen Faktors.

Die Beziehungen der Teilskalen zueinander fielen hypothesenkonform gering aus. Betrachtet man die inhaltlich typischen Verknüpfungen, so zeigte sich, dass die positive und die negative maskuline Teilskala einen Korrelationskoeffizienten von  $r = .17$  aufwiesen. Die beiden femininen Teilskalen offenbarten mit  $r = .21$  einen nur gering größeren Zusammenhang. Auf Basis der gemeinsamen Valenz ließ sich eine negative Beziehung der beiden positiven Teilskalen ( $r = -.19$ ) sowie eine sehr geringe, positive Beziehung der beiden negativen Teilskalen ( $r = .06$ ) ablesen. Die stärksten Beziehungen im Modell fanden sich schließlich jeweils zwischen den Skalen über Typizität und Valenz hinweg: so wiesen sowohl die positive maskuline und die negative feminine Teilskala ( $r = -.24$ ) als auch die negative maskuline und die positive feminine Teilskala ( $r = -.23$ ) erwartungsgemäß annähernd gleich starke negative Beziehungen zueinander auf. Damit zeigten auch die Ergebnisse der CFA, dass die neu entwickelten Teilskalen die zugrunde liegenden theoretischen Differenzierungen nach Typizität und Valenz überzeugend abbilden.

Diese Ergebnisse gewinnen um so mehr an Bedeutung, betrachtet man gleichzeitig die Resultate der Modellprüfung mit den klassischen BSRI-Items (Abbildung 12, S. 129), deren Modell-Fit deutlich schwächere Kennwerte aufwies. Damit ist in Bezug auf die postulierte vierfaktorielle Struktur das neu entwickelte Instrument dem BSRI eindeutig überlegen.

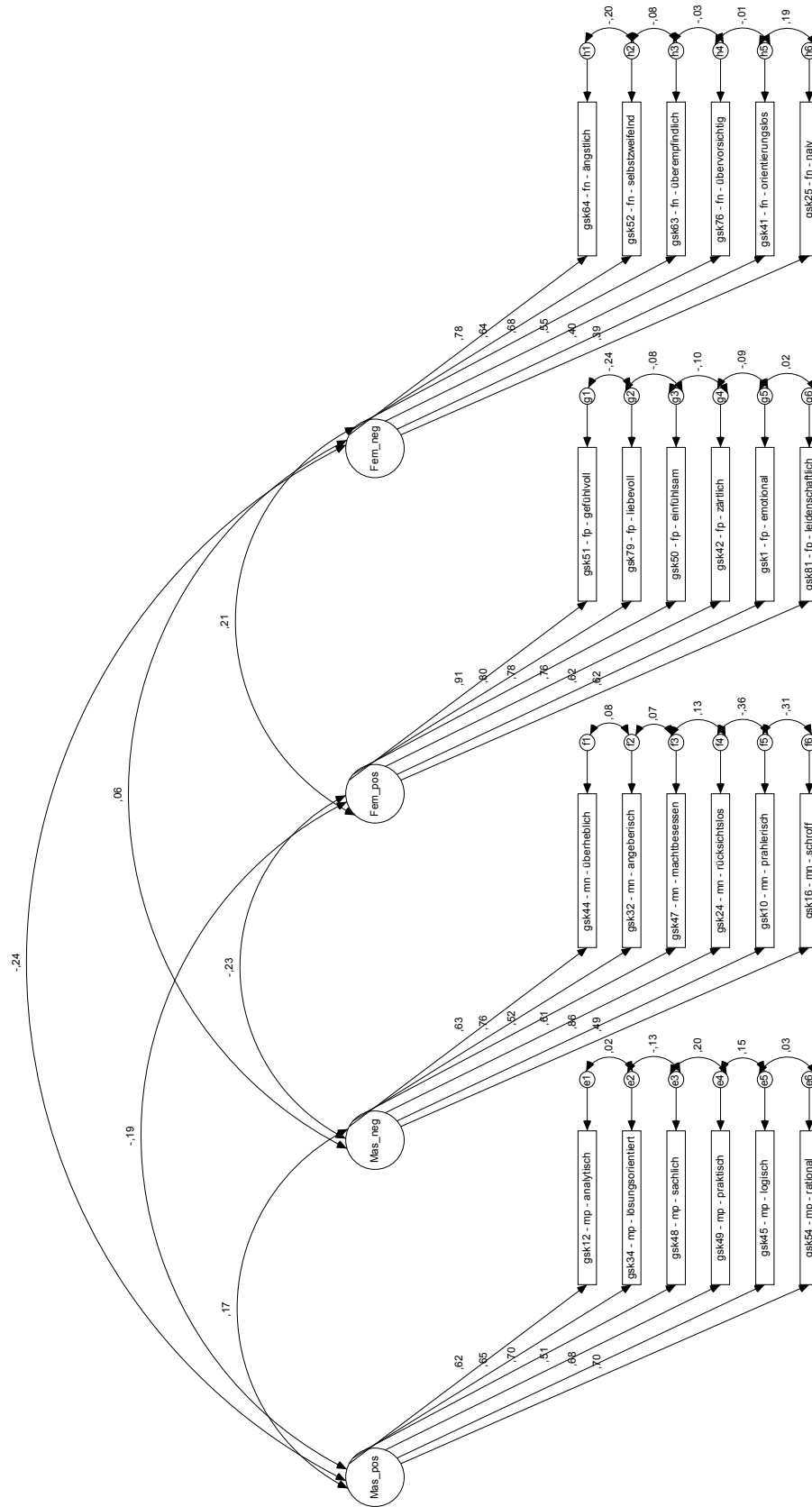


Abbildung 16. Faktorladungen aller Items der vier Teilskalen im Vier-Faktoren-Modell

### 3.2.3.3.3 Statistische Kennwerte der Teilskalen

Im nächsten Schritt der Datenanalyse wurden die Reliabilitäten der neuen Teilskalen über die gesamte Stichprobe für die Selbstbeschreibung bestimmt. Zusätzlich wurde geprüft, ob sich Frauen und Männer hypothesenkonform auf den Teilskalen unterschieden. Die Ergebnisse dieser Analysen sind in Tabelle 45 zusammengefasst.

Tabelle 45. Skalenskennwerte aller vier neuen Teilskalen in der Selbstbeschreibung

Skala	Gesamtstichprobe			Männer			Frauen			Vergleich		
	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$	M	SD	Diff.	p	$\eta^2$
Mas+	.81	3.58	.68	.79	3.84	.62	.80	3.43	.67	.41	<.001	.08
Mas-	.80	2.01	.70	.82	2.23	.77	.77	1.90	.63	.33	<.001	.06
Fem+	.88	4.03	.70	.87	3.72	.67	.86	4.19	.65	-.47	<.001	.10
Fem-	.74	2.66	.76	.74	2.43	.72	.73	2.78	.74	-.35	<.001	.05

Die vier Teilskalen mit jeweils sechs Items wiesen mit  $\alpha = .74 - .88$  Reliabilitätswerte auf, die angesichts der – aus Gründen der Testökonomie als Ziel vorgegebenen – Kürze der Skala als gut anzusehen sind. Die Trennschärfekoeffizienten aller Items lagen bei über .30. Alle Skalenanalysen wurden sowohl für die Gesamtstichprobe als auch getrennt für Männer und Frauen vorgenommen, um gerade Reliabilitätswerte, die möglicherweise aufgrund des höheren Anteils von Frauen in der Stichprobe zustande kamen, von vornherein ausschließen zu können. Während die Ergebnisse insgesamt über die Gruppen hinweg konsistent sind, fällt bereits an dieser Stelle auf, dass in allen Gruppen die Reliabilität der negativen femininen Teilskala am geringsten ausfällt.

Eine multivariate Varianzanalyse für die Selbstbeschreibung auf den vier Teilskalen mit dem Geschlecht als unabhängige Variable und Alter als Kovariate ergab einen signifikanten multivariaten Effekt für das Geschlecht,  $F(4, 794) = 47.83$ ,  $p < .001$ ,  $\eta^2 = .19$ . Alle vier univariaten Effekte waren signifikant, wie in Tabelle 45 dargestellt. Auf den beiden maskulinen Teilskalen wiesen die Männer höhere Werte auf als die Frauen, während erwartungsgemäß Frauen auf den beiden femininen Teilskalen höhere Werte aufwiesen als die Männer. Der multivariate Effekt des Alters wurde ebenfalls signifikant,  $F(4, 794) = 5.54$ ,  $p < .001$ ,  $\eta^2 = .03$  und beruhte auf signifikanten Unterschieden bei den beiden maskulinen Skalen. In Bezug auf Mas+ bedeutet dies, dass mit zunehmendem Alter die Zuschreibungen geringfügig höher ausfielen, während Mas- bei dieser Stichprobe im höheren Alter geringer ausfiel.

Eine weitere Varianzanalyse mit der Valenz (positiv-negativ) und der Geschlechtstypizität der Teilskala (männlich-weiblich) als Messwiederholungsfaktoren sowie dem Geschlecht der Probanden ergab zunächst einen Haupteffekt der Valenz: Männer und Frauen schrieben sich insgesamt negative Attribute in geringerem Ausmaß zu als positive,  $F(1, 798) = 3030.24$ ,  $p < .001$ ,  $\eta^2 = .79$ . Darüber hinaus zeigte sich die bereits erwähnte signifikante Interaktion von Geschlechtstypizität und Geschlecht,  $F(1, 798) = 188.73$ ,  $p < .05$ ,  $\eta^2 = .19$ .

Auch die einfache Interaktion von Typizität und Valenz wurde signifikant,  $F(1, 798) = 27.43$ ,  $p < .001$ ,  $\eta^2 = .03$ . Die Probanden beanspruchten somit nicht nur positive, sondern auch die negativen geschlechtskongruenten Eigenschaften stärker zur Selbstbeschreibung als negative geschlechtsinkongruente Eigenschaften. Dieses Ergebnis wurde durch die signifikante Dreifachinteraktion von Valenz, Geschlechtstypizität und

Geschlecht eingeschränkt,  $F(1, 798) = 4.37, p < .05, \eta^2 = .01$ . Demnach ist der Effekt, geschlechtstypische Eigenschaften sowohl im positiven wie auch im negativen Bereich stärker zur Selbstbeschreibung heranzuziehen als geschlechtsuntypische, nur für Frauen gültig, womit die aufgestellte Hypothese vier nur bedingt bestätigt werden kann.

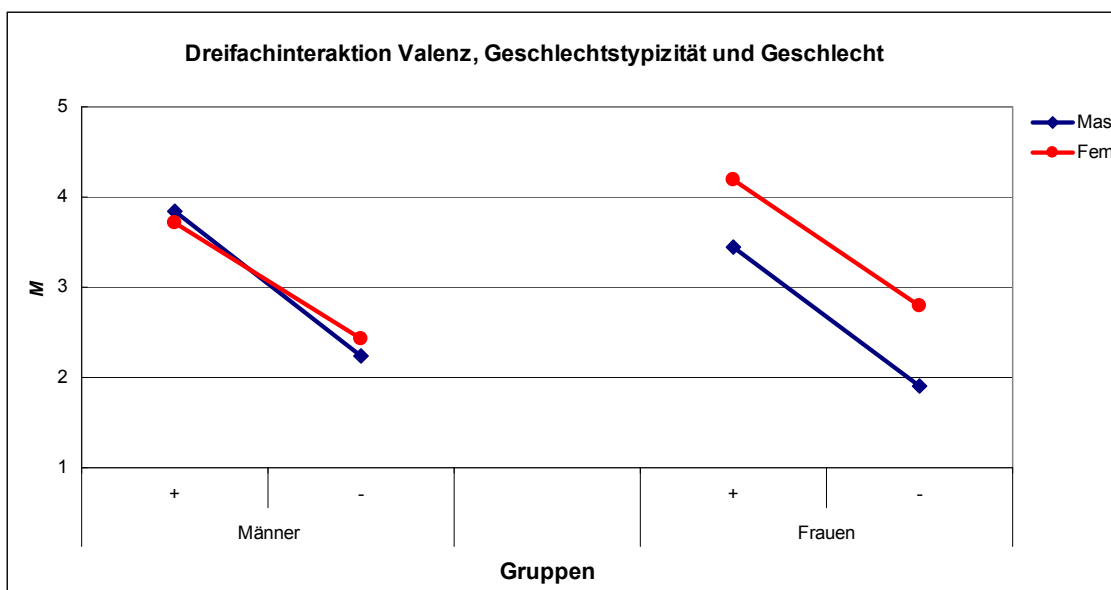


Abbildung 17. Dreifachinteraktion von Valenz, Geschlechtstypizität und Geschlecht

#### 3.2.3.3.4 Re-Test-Reliabilität

Von der ursprünglichen Stichprobe hatten 355 Personen (44%) eine E-Mail-Adresse hinterlassen, um an einer bereits angekündigten zweiten Erhebung teilzunehmen. Der nach ca. acht Wochen verschickten Einladung per E-Mail folgten 36% der Probanden. Insgesamt lagen Daten von 128 Personen zur Auswertung vor. Das Geschlechterverhältnis der Re-Test-Stichprobe entspricht dem der Ausgangsstichprobe mit 34% Männern und 66% Frauen, ebenso ist der Altersdurchschnitt mit 23.82 Jahren ( $SD = 4.44$  Jahre) sehr ähnlich. In Tabelle 46 sind die Kennwerte der vier Teilskalen zum zweiten Messzeitpunkt dargestellt.

Tabelle 46. Skalenkennwerte der vier Teilskalen in der Selbstbeschreibung zu T2

Selbst	Gesamt			Männer			Frauen			Vergleich		
	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$	M	SD	Diff.	p	$\eta^2$
Mas+ T2	.85	3.60	.63	.83	3.85	.65	.83	3.47	.59	.38	<.01	.08
Mas- T2	.83	1.96	.69	.80	2.23	.62	.82	1.82	.68	.41	<.01	.08
Fem+ T2	.86	4.09	.58	.83	3.88	.57	.85	4.19	.56	-.31	<.01	.07
Fem- T2	.73	2.61	.71	.79	2.35	.69	.68	2.74	.69	-.39	<.01	.07

Die Re-Test-Reliabilitäten der Teilskalen ( $\alpha = .68 - .86$ ) erwiesen sich acht Wochen nach der ersten Messung als zufriedenstellend. Die negative feminine Teilskala zeigte sich dabei erneut am schwächsten in ihrer Homogenität. Die Unterschiede zwischen den Geschlechtern auf den vier Teilskalen blieben nach der ersten Messung konstant (z. B.  $Diff_{Mas+ T1} = .41$  vs.  $Diff_{Mas+ T2} = .38$ ). Frauen hatten sowohl höhere Werte auf der positiven femininen;  $F(1,125) = 8.94, p < .01, \eta^2 = .07$  als auch negativen femininen Teilskala,  $F(1,125) = 8.75, p < .01, \eta^2 = .07$  im Vergleich zu den Männern, die wiederum höhere Werte auf den maskulinen Teilskalen erzielten,  $F(1,125) = 10.09, p < .01, \eta^2 = .08$  (für Mas+);  $F(1,125) = 10.88, p < .01, \eta^2 = .08$  (für Mas-) als

die weiblichen Probanden. In der Abbildung 18 sind die Mittelwertvergleiche zu beiden Messzeitpunkten, jeweils getrennt für Frauen und Männer, dargestellt.

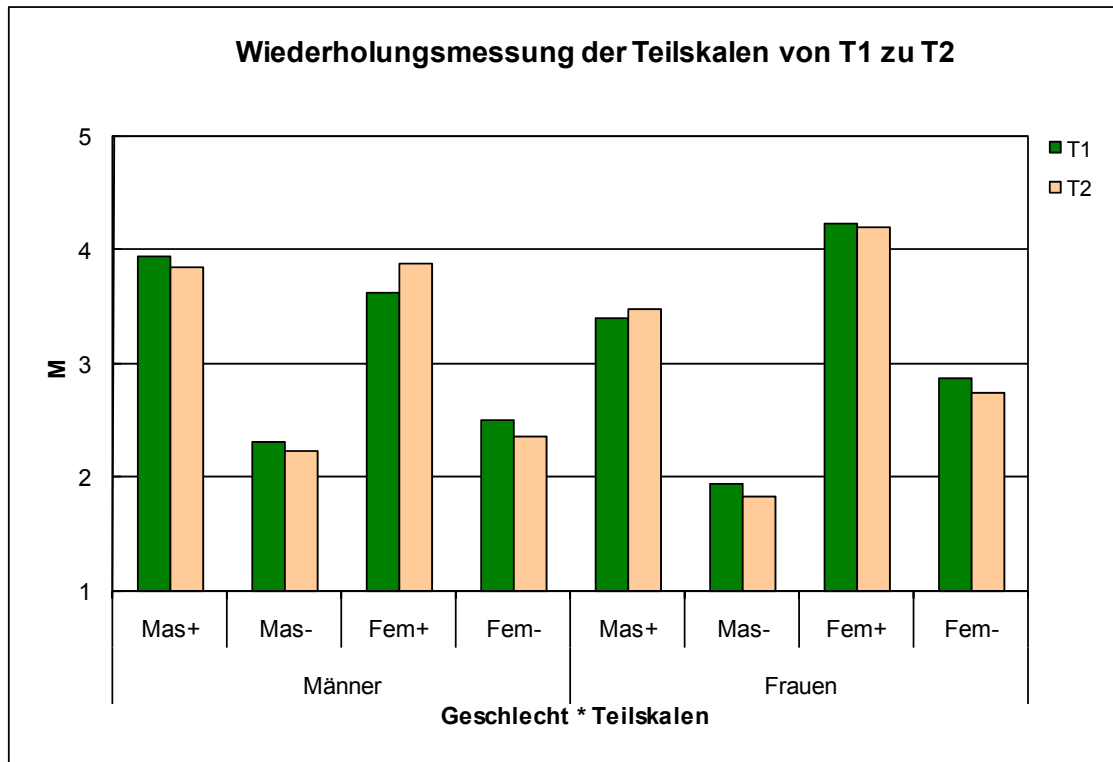


Abbildung 18. Mittelwerte der Teilskalen zu T1 und T2 im Vergleich

Die geringen Unterschiede der Ergebnisse von T1 zu T2 spiegeln sich auch in den Korrelationsbeziehungen der Teilskalen wider (Tabelle 47): Alle Teilskalen korrelierten jeweils hoch positiv mit ihren Wiederholungsmessungen.

Tabelle 47. Korrelationen der Teilskalen zu T1 und T2

	Mas+ T2	Mas- T2	Fem+ T2	Fem- T2
Mas+	.70**	.26**	-.16	-.35**
Mas-	.15	.63**	-.22*	.01
Fem+	-.17	-.35**	.61**	.24**
Fem-	-.37**	-.21*	.24**	.71**
Mas+ T2		.26**	.01	-.39**
Mas- T2			-.39**	-.05
Fem+ T2				.23**

Anmerkung: \*\* $p < .001$ , \* $p < .01$

In den Regressionsanalysen mit der Selbstbeschreibungen zu T2 (z. B. Mas+<sub>T2</sub>) als Kriterium und drei schrittweise hinzugefügten Prädiktoren: Geschlecht, entsprechende Teilskala zu T1 (z. B. Mas+<sub>T1</sub>) sowie die restlichen Teilskalen zu T1 (z. B. Mas-<sub>T1</sub>, Fem+<sub>T1</sub>, Fem-<sub>T1</sub>) erwiesen sich erwartungsgemäß jeweils die entsprechenden Teilskalen aus T1 als die stärksten Prädiktoren:

$$\text{Mas+}_{T1} (\beta = .69^{**}, B = .62, SE B = .06, \Delta R^2 = .41, p < .001)$$

$$\text{Mas-}_{T1} (\beta = .59^{**}, B = .54, SE B = .07, \Delta R^2 = .33, p < .001)$$

$$\text{Fem+}_{T1} (\beta = .61^{**}, B = .50, SE B = .06, \Delta R^2 = .30, p < .001)$$

$$\text{Fem-}_{T1} (\beta = .69^{**}, B = .67, SE B = .06, \Delta R^2 = .44, p < .001).$$

Insgesamt kann zusammengefasst werden, dass die Befunde über den achtwöchigen Re-Test-Zeitraum dem Instrument eine gute Stabilität nachwiesen. In Kombination mit den Ergebnissen

zur internen Konsistenz ist damit festzuhalten, dass mit dem neuen Instrument nach den üblichen Standards der Reliabilitätsprüfung ein zuverlässiges Messinstrument vorliegt.

### 3.2.3.3.5 Korrelative Zusammenhänge der Teilskalen

Die Korrelationsmatrix (Tabelle 48) zeigt zunächst, dass es zumindest die beiden maskulinen Teilskalen geringfügig aber signifikant mit dem Alter korrelierte ( $r_s < .11$ ). Darüber hinaus unterstützte die Matrix Ergebnisse, die bereits aus der konfirmatorischen Faktorenanalyse bekannt sind: zwischen nahezu allen Teilskalen gibt es signifikante, jedoch geringe Korrelationen. Die engsten Beziehungen bestanden demnach jeweils zwischen den beiden entgegengesetzten Teilskalen Fem+ und Mas- sowie Mas+ und Fem-, wobei die Höhe der Korrelation aufzeigte, dass die Teilskalen durchaus distinkte Aspekte des geschlechtsbezogenen Selbstkonzepts erfassen. Die Teilskalen mit gleicher Typizität (Mas+ und Mas-, Fem+ und Fem-) korrelierten jeweils schwach positiv miteinander. Bezüglich der Teilskalen mit gleicher Valenz war das Bild weniger eindeutig: positive Maskulinität und positive Femininität korrelierten schwach negativ miteinander, während Mas- und Fem- unabhängig voneinander waren.

Tabelle 48. Korrelationen der Teilskalen

	Mas+	Mas-	Fem+	Fem-
Alter	.11**	-.10**	-.02	-.05
Mas+		.19**	-.15**	-.22**
Mas-			-.23**	.05
Fem+				.15**

Anmerkung: \*\* $p < .01$ , \* $p < .05$

### 3.2.3.3.6 Statistische Kennwerte der Validierungskonstrukte

In einem nächsten Schritt der Datenanalyse wurden die Reliabilitäten der eingesetzten Validierungsinstrumente sowie die mittleren Ausprägungen sowohl über die gesamte Stichprobe als auch getrennt für Männer und Frauen überprüft. Des Weiteren wurde untersucht, inwieweit sich Männer und Frauen auf den Konstrukten unterschieden. Die Ergebnisse dieser Analysen sind in Tabelle 49 zusammengefasst.

Tabelle 49. Statistische Kennwerte der Validierungskonstrukte für die vier Teilskalen

	Gesamt			Männer			Frauen			Unterschied		
	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	$\eta^2$
<b>Prädiktoren für MAS+</b>												
berufl. Ehrgeiz (AVEM)	.87	3.43	.86	.88	3.58	.89	.86	3.35	.86	7.08	<.01	.02
Hoffen auf Erfolg (AMS)	.92	3.31	.93	.91	3.48	.89	.92	3.22	.93	1.08	<.01	.02
<b>Prädiktoren für MAS-</b>												
Aggressivität (AQ)	.74	2.34	1.03	.73	2.69	1.05	.73	2.16	.96	35.94	<.001	.06
Werteskala Macht	.72	3.18	.68	.77	3.29	.75	.68	3.12	.62	3.28	n.s.	.01
<b>Prädiktoren für FEM+</b>												
Empathie (Sorge)	.88	3.60	.72	.88	3.35	.74	.87	3.73	.69	28.05	<.001	.06
soziale Werte	.60	4.45	.37	.70	4.37	.42	.48	4.51	.32	7.68	<.01	.03
<b>Prädiktoren für FEM-</b>												
ESWS	.92	3.08	1.41	.92	2.86	1.35	.92	3.20 <sup>35</sup>	1.43	7.38	<.01	.01
Ängstl. Depressivität (BSF)	.87	3.64	1.30	.87	3.44	1.34	.86	3.74	1.28	1.67	n.s.	.01

<sup>35</sup> Hohe Mittelwerte bedeuten an dieser Stelle geringe emotionale Selbstwertschätzung im Sinne von „an sich selbst zweifeln“ oder „überzeugt zu sein ein nutzloser Mensch zu sein“.



Für alle zum Einsatz gekommenen Validierungskonstrukte wurden gute bis sehr gute Reliabilitätswerte ermittelt, sowohl für die Gesamtstichprobe als auch getrennt für die beiden Geschlechter. Darüber hinaus gab es keine systematischen Abfälle der Cronbachs Alphas in Abhängigkeit von der Position der Skalen, weshalb Ermüdungserscheinungen während des Bearbeitens des Fragebogens zum Ende hin ausgeschlossen werden können. In Bezug auf die Validierungskonstrukte zeigten sich bei einigen Skalen Zusammenhänge mit dem Geschlecht, derart, dass Männer signifikant höhere Werte auf Skalen wie beruflichen Ehrgeiz, Hoffen auf Erfolg oder Aggressivität berichteten, während Frauen empathischer waren, über mehr soziale Werte verfügten und sich weniger emotional wertschätzten.

Bei den anderen Konstrukten (Ängstliche Depressivität, Werteskala Macht) zeigten sich diese Geschlechtseffekte nicht, was jedoch deren Validierungswert nicht einschränkt, da in erster Linie der Einfluss des Geschlechtsrollenselbstkonzepts überprüft werden sollte.

### 3.2.3.3.7 Korrelationen der Validierungskonstrukte mit den Teilskalen

Auf der Ebene der bivariaten Korrelationen ergaben sich durchgängig die postulierten Zusammenhänge der Teilskalen mit den zum Einsatz gekommenen Validierungskonstrukten (Tabelle 50, die relevanten Korrelationen sind grau unterlegt). So korrelierte MAS+ am stärksten positiv mit der Hoffnung auf Erfolg und beruflichem Ehrgeiz. Die negative maskuline Teilskala dagegen korrelierte positiv mit der Aggressivität und der machtbezogenen Werteskala.

Tabelle 50. Korrelationen der Teilskalen mit den Validierungskonstrukten

	Mas+	Mas-	Fem+	Fem-
berufl. Ehrgeiz (AVEM)	.29**	.22**	-.05	-.20**
Hoffen auf Erfolg (AMS)	.49**	.13**	-.07	-.26**
Aggressivität (AQ)	-.04	.31**	-.10*	.02
Werteskala Macht	.15*	.30**	.07	-.15*
Empathie (Sorge)	-.14**	-.32**	.46**	.29**
soziale Werte	-.09	-.26**	.50**	.01
ESWS	-.15**	.07	.03	.56**
Ängstl. Depressivität (BSF)	-.10	.20*	.16	.52**

Anmerkung: \*\* $p < .01$ , \* $p < .05$

Die positive feminine Teilskala wies erwartungsgemäß eine hohe positive Korrelation zu Maßen wie Empathie und sozialen Werten auf. Mit anderen Konstrukten stand Fem+ dagegen in keiner oder nur sehr geringer Beziehung. Die negative feminine Teilskala korrelierte hypothesenkonform mit emotionaler Selbstwertschätzung derart, dass die hohe Ausprägung auf Fem- mit wenig emotionaler Selbstwertschätzung<sup>36</sup> einhergeht. Ebenso zeigte Fem- starke positive Zusammenhänge zu dem Konstrukt Ängstliche Depressivität. Je höher die negative Femininität, desto größer waren die berichteten Beschwerden. Darüber hinaus deuteten sich zudem negative Zusammenhänge von Fem- zu Konstrukten an, die entsprechend der theoriegeleiteten Hypothesen ursprünglich der positiven maskulinen Teilskala zugeschrieben wurden, wie dem Leistungsmotiv oder der Hoffnung auf Erfolg.

<sup>36</sup> Ein hoher Korrelationskoeffizient bei hohen Werten auf Fem- ist - aufgrund der Kodierung der Antworten - mit einer geringen Ausprägung aus ESWS zu interpretieren.

### 3.2.3.3.8 Konstruktvaliditätsanalysen mittels hierarchischer Regressionen

Zur Überprüfung der Konstruktvalidität der geschlechtsrollenspezifischen Teilskalen wurde im Rahmen der Regressionsanalysen der Ansatz des *destruktiven Testens* (Anderson & Anderson, 1996) verwendet. Dieses Vorgehen prüft, inwieweit der spezifisch vorhergesagte Zusammenhang eines Validierungskonstrukts mit der entsprechenden Teilskala des GSK auch dann noch statistisch signifikanten Bestand hat, wenn den Regressionsmodellen sukzessiv weitere, konkurrierende Prädiktoren hinzugefügt werden. Bei dieser Methode geht es weniger darum, den angenommenen Zusammenhang zwischen zwei Konzepten aufzubrechen, sondern darum zu überprüfen, wie zuverlässig dieser Zusammenhang auch unter Berücksichtigung weiterer theoretischer und methodischer Einflussvariablen ist. Die Beziehung von zwei Zielvariablen wird durch das schrittweise Hinzufügen theoretisch relevanter Kovariate herausgefordert. Kann der Zusammenhang dieser beiden Zielvariablen leicht aufgelöst werden, ist die Validität dieser Beziehung in Frage zu stellen. In der vorliegenden Arbeit sah das Vorgehen folgende Schritte vor: als Erstes wurde die jeweils assoziierte Teilskala des Validierungskonstrukts in die Regressionsgleichung aufgenommen, um in den folgenden Schritten nachzuweisen, dass die Hinzunahme der weiteren Teilskalen (im einem zweiten Schritt die beiden Teilskalen entsprechender Typizität und Valenz, im dritten Schritt die Teilskala entgegengesetzt in Typizität und Valenz sowie im vierten und letzten Schritt die des biologischen Geschlechts) zu keinem aussagekräftigen Zugewinn der Aufklärungsleistung führt.

Tabelle 51. Vorhersage von AVEM-BE (berufl. Ehrgeiz) durch Teilskalen (Mas+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas+	.36	.06	.29**
Schritt 2			
Mas+	.32	.06	.26**
Fem+	.21	.06	.17**
Mas-	.04	.06	.03
Schritt 3			
Mas+	.27	.06	.22**
Fem+	.23	.06	.19**
Mas-	.07	.06	.06
Fem-	-.19	.05	-.16*
Schritt 4			
Mas+	.27	.06	.22**
Fem+	.23	.06	.19**
Mas-	.07	.06	.06
Fem-	-.19	.06	-.16*
Geschlecht	-.02	.09	-.01

Anmerkung:  $R^2 = .08^{**}$ ; Schritt 2  $\Delta R^2 = .03^*$ ; Schritt 3  $\Delta R^2 = .03^*$ ; Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$

In Tabelle 51 ist dieses Vorgehen am Beispiel des durch den AVEM erhobenen beruflichen Ehrgeizes aufgeführt. Erwartungskonform leistete die positive maskuline Teilskala den größten Beitrag zur Varianzaufklärung des beruflichen Ehrgeizes ( $\beta = .29^{**}$ ,  $B = .36$ ;  $SE B = .06$ ,  $R^2 = .08$ ,  $p < .001$ ). Je höher die Ausprägung der positiven Maskulinität, desto höher der berichtete berufliche Ehrgeiz. Das Hinzunehmen weiterer Teilskalen offenbarte, dass darüber hinaus sowohl die positive feminine Teilskala ( $\beta = .17^{**}$ ,  $B = .21$ ;  $SE B = .06$ ,  $\Delta R^2 = .03$ ,  $p < .001$ ) als auch die negative feminine Teilskala ( $\beta = -.16^{**}$ ,  $B = -.19$ ;  $SE B = .05$ ,  $\Delta R^2 = .03$ ,  $p < .01$ ) etwas zur Varianzaufklärung beitragen konnten. Je höher die Ausprägungen positiver Femininität sowie je geringer die berichteten Werte der negativen Femininität, desto größer der

berufliche Ehrgeiz. Entsprechend der Hypothesen war der Modellzugewinn durch das Aufnehmen von Fem+ und Fem- insgesamt kleiner als die Varianzaufklärung, die die positive Maskulinität allein bereits leisten konnte. Darüber hinaus konnte festgestellt werden, dass das biologische Geschlecht im letzten Schritt keine weitere Varianz in der Ausprägung des beruflichen Ehrgeizes zu erklären vermochte.

In ganz ähnlicher Weise präsentierte sich das Bild in Bezug auf das Konstrukt Hoffen auf Erfolg (Tabelle 52). Die positive maskuline Teilskala hatte erneut den größten Anteil zur Varianzaufklärung des Konstrukts ( $\beta = .49^{**}$ ,  $B = .66$ ;  $SE B = .05$ ,  $R^2 = .24$ ,  $p < .001$ ). Je höher die Ausprägung der positiven Maskulinität, desto größer war die Hoffnung auf Erfolg. Das Hinzunehmen weiterer Teilskalen offenbarte, dass darüber hinaus lediglich die negative feminine Teilskala ( $\beta = -.15^{**}$ ,  $B = -.19$ ;  $SE B = .05$ ,  $\Delta R^2 = .02$ ,  $p < .01$ ) einen weiteren, wenn auch sehr geringen Beitrag zur Varianzaufklärung leisten konnte. Je niedriger die Ausprägung auf der negativen femininen Teilskala, desto geringer fiel die Hoffnung auf Erfolg aus. Das biologische Geschlecht trug nichts zur Varianzaufklärung bei.

Tabelle 52. Vorhersage von AMS - Hoffen auf Erfolg durch Teilskalen (Mas+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas+	.66	.05	.49**
Schritt 2			
Mas+	.67	.05	.49**
Fem+	.04	.05	.03
Mas-	.05	.05	.04
Schritt 3			
Mas+	.61	.05	.45**
Fem+	.06	.05	.05
Mas-	.08	.05	.06
Fem-	-.19	.05	-.15**
Schritt 4			
Mas+	.62	.05	.45**
Fem+	.07	.05	.05
Mas-	.07	.05	.05
Fem-	-.19	.05	-.15*
Geschlecht	.04	.08	.02

Anmerkung:  $R^2 = .24^{**}$ ; Schritt 2  $\Delta R^2 = \text{n.s.}$ ; Schritt 3  $\Delta R^2 = .02^*$ ; Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$

Zur Validierung der negativen maskulinen Teilskala wurde die Reihenfolge der schrittweise eingehenden Faktoren in die Regressionsanalysen entsprechend angepasst.

Im Anschluss an die Mas- Skala im ersten Schritt wurden in Schritt zwei die Skalen Fem- und Mas+ aufgenommen und erst im dritten Schritt dann Fem+, ehe im letzten Schritt der Einfluss des Geschlechts überprüft wurde. Dabei zeigte sich unter anderem, dass je höher die negative Maskulinität ( $\beta = .31^{**}$ ,  $B = .46$ ;  $SE B = .06$ ,  $R^2 = .10$ ,  $p < .001$ ) und je geringer die negative Femininität ( $\beta = -.12^*$ ,  $B = -.17$ ;  $SE B = .06$ ,  $\Delta R^2 = .01$ ,  $p < .01$ ), desto höher die Ausprägungen der Aggressivität (Tabelle 53).

Im Gegensatz zur vorherigen Analyse leistete auch das biologische Geschlecht einen eigenen Beitrag ( $\beta = -.24^{**}$ ,  $B = -.52$ ;  $SE B = .09$ ,  $\Delta R^2 = .05$ ,  $p < .001$ ), blieb aber in der Aufklärungsstärke im Modell hinter dem Einfluss der negativen Maskulinität zurück ( $R^2 = .10$  vs.  $R^2 = .05$ ).

Tabelle 53. Vorhersage von Aggressivität (AQ) durch Teilskalen (Mas-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas-	.46	.06	.31**
Schritt 2			
Mas-	.50	.06	.33**
Fem-	-.17	.06	-.12*
Mas+	-.02	.06	-.01
Schritt 3			
Mas-	.49	.06	.33**
Fem-	-.18	.06	-.12*
Mas+	-.01	.06	-.01
Fem+	-.04	.06	-.03
Schritt 4			
Mas-	.45	.06	.30**
Fem-	-.24	.06	-.16**
Mas+	.03	.06	.02
Fem+	.04	.06	.03
Geschlecht	-.52	.09	-.24**

Anmerkung:  $R^2 = .10^{**}$ ; Schritt 2  $\Delta R^2 = .01^*$ , Schritt 3  $\Delta R^2 = \text{n.s.}$ , Schritt 4  $\Delta R^2 = .05^{**}$ ;  $**p < .001$ ;  $*p < .01$

Bei den Analysen hinsichtlich der Werteskala Macht (Tabelle 54) konnte das biologische Geschlecht keinen eigenen Einfluss geltend machen.

Hypothesenkonform leistete die negative maskuline Teilskala den größten Beitrag ( $\beta = .30^{**}$ ,  $B = .29$ ;  $SE B = .06$ ,  $R^2 = .10$ ,  $p < .001$ ). Personen mit höheren Werten im Machtbereich hatten hohe negative maskuline Werte, ebenso wie geringe positive maskuline Werte ( $\beta = -.14^*$ ,  $B = -.14$ ;  $SE B = .06$ ,  $\Delta R^2 = .03$ ,  $p < .01$ ) oder höhere positive feminine Werte ( $\beta = .19^*$ ,  $B = .17$ ;  $SE B = .06$ ,  $\Delta R^2 = .03$ ,  $p < .001$ ).

Tabelle 54. Vorhersage von Werteskala Macht durch Teilskalen (Mas-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas-	.29	.06	.30**
Schritt 2			
Mas-	.29	.06	.30**
Fem-	.05	.07	.05
Mas+	-.14	.06	-.14*
Schritt 3			
Mas-	.33	.06	.34**
Fem-	.06	.07	.06
Mas+	-.16	.06	-.16*
Fem+	.17	.06	.19*
Schritt 4			
Mas-	.32	.06	.34**
Fem-	.05	.07	.05
Mas+	-.14	.06	-.15*
Fem+	.18	.06	.20*
Geschlecht	-.09	.10	-.06

Anmerkung:  $R^2 = .10^{**}$ ; Schritt 2  $\Delta R^2 = .03^*$ , Schritt 3  $\Delta R^2 = .03^{**}$ , Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$

Die Validierung der positiven femininen Teilskala sah folgende Reihenfolge der schrittweise eingehenden Faktoren in die Regressionsanalysen vor: Fem+ in einem ersten Schritt, gefolgt von Mas+ und Fem-. Die negative maskuline Teilskala wurde im dritten Schritt eingefügt und das biologische Geschlecht folgte im letzten Schritt.

Zur Vorhersage von Empathie (Tabelle 55) stellte sich hypothesenkonform die positive feminine Teilskala als bester Prädiktor heraus ( $\beta = .46^{**}$ ,  $B = .47$ ;  $SE B = .04$ ,  $R^2 = .21$ ,  $p < .001$ ); je höher die berichtete positive Femininität, desto empathischer (in Bezug auf Sorge um ihre Mitmenschen) beschrieben sich die Probanden. Dies galt ebenso für die positive

Maskulinität ( $\beta = .22^{**}$ ,  $B = .21$ ;  $SE B = .04$ ,  $\Delta R^2 = .05$ ,  $p < .001$ ): Personen, die sich stark auf den positiven maskulinen Eigenschaften positionierten, schätzten sich auch empathisch ein, als Personen mit gering ausgeprägter positiver Maskulinität. Erneut sei angemerkt, dass der Aufklärungswert von Mas+ deutlich geringer ausfiel als der von Fem+. Zusätzlich offenbarte die Regressionsanalyse, dass ebenso Probanden mit gering ausgeprägten negativen maskulinen Eigenschaften stärkere besorgte Empathiewerte aufwiesen ( $\beta = -.25^{**}$ ,  $B = -.25$ ;  $SE B = .04$ ,  $\Delta R^2 = .06$ ,  $p < .001$ ).

Tabelle 55. Vorhersage von Empathie (Sorge) durch Teilskalen (Fem+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem+	.47	.04	.46**
Schritt 2			
Fem+	.43	.04	.42**
Mas+	.21	.04	.22**
Fem-	-.03	.05	-.03
Schritt 3			
Fem+	.37	.04	.37**
Mas+	.24	.04	.25**
Fem-	.02	.04	.02
Mas-	-.25	.04	-.25**
Schritt 4			
Fem+	.36	.04	.35**
Mas+	.23	.04	.24**
Fem-	.04	.05	.03
Mas-	-.25	.04	-.24**
Geschlecht	.08	.07	.05

Anmerkung:  $R^2 = .21^{**}$ ; Schritt 2  $\Delta R^2 = .05^{**}$ , Schritt 3  $\Delta R^2 = .06^{**}$ , Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$

Eine hohe Ausprägung von positiver Femininität ging zudem mit hohen sozialen Werten ( $\beta = .50^{**}$ ,  $B = .25$ ;  $SE B = .03$ ,  $R^2 = .25$ ,  $p < .001$ ; Tabelle 56) einher. Negative Maskulinität konnte im Vergleich nur einen sehr geringen Einfluss geltend machen, im Sinne, dass hohe Ausprägung auf Mas- mit gering ausgeprägten sozialen Werten assoziiert war.

Tabelle 56. Vorhersage von sozialen Werten durch Teilskalen (Fem+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem+	.25	.03	.50**
Schritt 2			
Fem+	.25	.03	.50**
Mas+	-.03	.03	-.06
Fem-	-.02	.03	-.03
Schritt 3			
Fem+	.23	.03	.47**
Mas+	-.02	.03	-.04
Fem-	.00	.03	.00
Mas-	-.07	.03	-.14*
Schritt 4			
Fem+	.23	.03	.46**
Mas+	-.03	.03	-.05
Fem-	.00	.03	.01
Mas-	-.07	.03	-.14*
Geschlecht	.03	.05	.04

Anmerkung:  $R^2 = .25^{**}$ ; Schritt 2  $\Delta R^2 = \text{n.s.}$ , Schritt 3  $\Delta R^2 = .02^*$ , Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$

Die Validierung der negativen femininen Teilskala sah schließlich folgende Reihenfolge der Faktoren in den Regressionsanalysen vor: Fem- in einem ersten Schritt sowie Mas- und Fem+ darauf folgend. Im dritten Schritt wurde Mas+ aufgenommen und in einem vierten das biologische Geschlecht.

Negative Femininität trat als stärkster Prädiktor von emotionaler Selbstwertschätzung zutage (Tabelle 57): je höher die eigene berichtete negative Femininität, desto geringer fiel die emotionale Selbstwertschätzung aus ( $\beta = .56^{**37}$ ,  $B = 1.06$ ;  $SE B = .07$ ,  $R^2 = .31$ ,  $p < .001$ ).

Tabelle 57. Vorhersage von ESWS durch Teilskalen (Fem-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem-	1.06	.07	.56**
Schritt 2			
Fem-	1.07	.07	.56**
Mas-	-.10	.07	-.05
Fem+	.06	.07	.03
Schritt 3			
Fem-	1.05	.07	.56**
Mas-	-.10	.07	-.05
Fem+	.07	.07	.04
Mas+	-.09	.08	-.04
Schritt 4			
Fem-	1.06	.07	.56**
Mas-	-.09	.08	-.05
Fem+	.07	.07	.03
Mas+	-.10	.08	-.05
Geschlecht	-.07	.12	-.02

Anmerkung:  $R^2 = .31^{**}$ ; Schritt 2  $\Delta R^2 = n.s.$ , Schritt 3  $\Delta R^2 = n.s.$ , Schritt 4  $\Delta R^2 = n.s.$ ;  $**p < .001$ ;  $*p < .01$

In einer letzten Regressionsanalyse wurde schließlich der Zusammenhang der vier Teilskalen zur ängstlichen Depressivität überprüft (Tabelle 58). Erneut waren es Personen mit hoch ausgeprägter negativer Femininität, die von stärkerer ängstlicher Depressivität berichteten ( $\beta = .52^{**}$ ,  $B = .88$ ;  $SE B = .12$ ,  $R^2 = .27$ ,  $p < .001$ ).

Tabelle 58. Vorhersage von Ängstlicher Depressivität durch Teilskalen (Fem-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem-	.88	.12	.52**
Schritt 2			
Fem-	.82	.13	.48**
Mas-	.22	.14	.11
Fem+	.25	.14	.13
Schritt 3			
Fem-	.82	.13	.48**
Mas-	.22	.14	.11
Fem+	.25	.14	.14
Mas+	-.01	.14	.00
Schritt 4			
Fem-	.84	.14	.49**
Mas-	.24	.15	.12
Fem+	.23	.15	.12
Mas+	-.02	.14	-.01
Geschlecht	-.11	.24	-.04

Anmerkung:  $R^2 = .27^{**}$ ; Schritt 2  $\Delta R^2 = n.s.$ , Schritt 3  $\Delta R^2 = n.s.$ , Schritt 4  $\Delta R^2 = n.s.$ ;  $**p < .001$ ;  $*p < .01$

### 3.2.3.3.9 Konvergente Validität

Zur Überprüfung der konvergenten Validität wurden die Ergebnisse der neuen Teilskalen mit denen der klassischen BSRI-Skalen kontrastiert (Tabelle 59). Beide Instrumente zielen darauf ab das geschlechtsbezogene Selbstkonzept zu erfassen. Die Skalen des BSRI (Bem, 1974) konnten zufriedenstellende Reliabilitäten vorweisen:  $BSRI_{MAS}$  erwies dabei insgesamt etwas

<sup>37</sup> Positive  $\beta$ -Koeffizienten bedeuten in diesem Fall eine geringe emotionale Selbstwertschätzung derart, dass die Probanden berichten, stark an sich zu zweifeln und keine positive Einstellung zu sich selbst zu haben.

homogener als  $BSRI_{FEM}$ . Erwartungsgemäß waren Männer in ihrer Selbsteinschätzung maskuliner,  $F(1,798) = 29.26$ ,  $p < .001$ ,  $\eta^2 = .04$  als Frauen. Bezüglich der Femininität,  $F(1,798) = 51.04$ ,  $p < .001$ ,  $\eta^2 = .06$  wiesen Frauen signifikant höhere Werte als Männer auf. Die überprüften Korrelationen der sechs Teilskalen zueinander folgten dem hypothetisierten Muster, was gerade für die jeweils äquivalenten positiven Teilskalen besonders eindeutig ausfiel und somit die Messung ähnlicher Konstrukte mittels der beiden Instrumente bestätigt.

Tabelle 59. Deskriptive Statistik und Korrelationen der BSRI-Skalen mit den GSK-Teilskalen

	BSRI - Mas	BSRI - Fem
$\alpha$	.89	.79
N	800	800
$M_{Männer}$	3.51	3.38
$M_{Frauen}$	3.29	3.61
<i>Korrelation mit</i>		
Mas+	.52**	-.09**
Mas-	.27**	-.27**
Fem+	.01	.71**
Fem-	-.51**	.20**

Anmerkung: \*\* $p < .01$ , \* $p < .05$

So hatte Mas+ einen stärkeren positiven inhaltlichen Bezug zu der maskulinen als zu der femininen Teilskala des BSRI, ebenso wie Fem+ eine starke positive Korrelation zur femininen BSRI-Teilskala aufwies, während zur maskulinen BSRI-Skala keine Beziehung bestand.

Mas- stand erwartungsgemäß in einer positiven Korrelation zu  $BSRI_{MAS}$ , zeigte gleichzeitig jedoch eine ebenso starke, wenn auch negative Korrelation zu  $BSRI_{FEM}$ . Personen, die sich maskuline BSRI-Items zuschrieben, taten dies auch ähnlich stark bezüglich der maskulinen, jedoch negativen Eigenschaften, während die auf dem BSRI eher feminin eingestuften Probanden die negativen maskulinen Eigenschaften im geringeren Maße zur Beschreibung des eigenen Selbst heranzogen. In ähnlicher Weise verhielt es sich mit der Beziehung der neuen negativen femininen Teilskala zu den BSRI-Skalen; Fem- korrelierte ebenfalls hypothesenkonform positiv mit  $BSRI_{FEM}$ , gleichzeitig jedoch auch negativ und wesentlich stärker mit  $BSRI_{MAS}$ . Das Akzeptieren femininer BSRI-Attribute für das eigene Selbst ging einher mit dem Zuschreiben auch negativer femininer Eigenschaften der neuen Teilskala. Personen wiederum, die sich auf der BSRI eher maskulin positionierten, lehnten die negativen femininen Beschreibungen stärker ab. Wichtig ist aber, dass die Korrelationen eher niedrig ausfallen, was dafür spricht, dass die negativen Teilskalen des neuen Instruments eigenständige Aspekte des GSK abbilden.

Die Beziehung der komplementären Teilskalen (Mas- und  $BSRI_{FEM}$ , Fem- und  $BSRI_{MAS}$ ) wurde jeweils bestimmt durch einen negativen Zusammenhang, der nachweislich für die unterschiedliche Valenz der Teilskalen steht, sowie der Höhe des Korrelationskoeffizienten, der die gemeinsame Varianz dokumentiert. Die hier erzielten Koeffizienten waren damit zum einen Beleg für die konvergente Validität, derart, dass die entsprechenden geschlechtsspezifischen Teilskalen mit ihrem jeweiligen Gegenpart ausreichend Varianz teilen und damit grundsätzlich ein ähnliches Konstrukt messen. Zum anderen deuteten die Komplementärbeziehungen an, dass auch diese Skalen (z. B. Mas- und  $BSRI_{FEM}$ ) zwar durchaus über gemeinsame Anteile

verfügen, die neuen negativen Teilskalen dennoch nicht zwangsläufig lediglich eine Umkehr der ursprünglichen ausschließlich positiven Skalen bedeuten. Die negativen Skalen stellen ihrerseits sowohl valenzbasiert als auch inhaltlich einen Zugewinn in der Erklärung des geschlechtsbezogenen Selbstkonzepts dar.

### **3.2.3.4 Diskussion**

In Studie 5 wurden zunächst weitere Daten zur Selbstbeschreibung erfasst, die allen 24 Items zufriedenstellende Faktorladungen und in konkurrierend geprüften konfirmatorischen Modellen das angenommene Vierfaktorenmodell (entsprechend der vier Teilskalen) als das am besten repräsentierende nachwies. Zudem zeichneten sich alle Teilskalen durch gute interne Konsistenzwerte und Re-Test-Reliabilitäten aus, zusätzlich auch in getrennten Analysen für Männer und Frauen. Entsprechend der Hypothesen schrieben sich Männer die maskulinen Eigenschaften im stärkeren Maße zu als Frauen, während Frauen sich die femininen Eigenschaften im stärkeren Maße zuschrieben als Männer dies taten. Ebenfalls hypothesenkonform konnte auch ein Effekt der Valenz bestätigt werden. So schrieben sich Männer und Frauen die positiven Merkmale des GSK stärker zu als negative und darüber hinaus beanspruchten Frauen eher die negativen Attribute für das eigene Selbst, die typischer für Frauen als für Männer sind. Lediglich für die Männer konnte dieser Interaktionseffekt so nicht bestätigt werden (Interaktion vom Valenz und Typizität).

Entscheidend für die Entwicklung der neuen Teilskalen waren zudem ihre Beziehungen zueinander. Demnach sollte es sich bei den negativen Teilskalen nicht einfach um eine Umkehr ihrer respektiven positiven Gegenstücke handeln, was ihre Erfassung redundant machen würde. Ein Nachweis der Unterschiedlichkeit der vier Teilskalen sowohl bezüglich der Typizität als auch ihrer Valenz war daher erforderlich. Entsprechend der Hypothesen konnte so unter anderem zunächst gezeigt werden, dass positive Maskulinität und negative Femininität zwar negativ, jedoch tatsächlich nur gering miteinander korreliert waren. Zwei der wichtigsten Beziehungen sind die zwischen Fem+ und Mas- sowie Mas+ und Fem-. In beiden Fällen fanden sich zwar negative Korrelationen, die in ihren Koeffizienten jedoch zufriedenstellend niedrig ausfielen. Der Anteil der gemeinsamen Varianz betrug jeweils nicht mehr als 5%, was für die eigenständige Relevanz der negativen Teilskalen spricht. Darüber hinaus konnte die inhaltliche Unabhängigkeit der Teilskalen zusätzlich im Validierungsprozess bestätigt werden. Auch eine bereits durchgeführte Wiederholungsmessung nach wenigen Wochen bestätigte die Qualitäten und Stabilität der Teilskalen über die Zeit hinweg.

Die Überprüfung der Konstruktvalidität wurde durch das Einbeziehen zweier theoriegeleiteter Validierungskonstrukte pro Teilskala umgesetzt. Jeder Teilskala wurden Konstrukte zugeordnet, denen bereits durch vorliegende Forschungsergebnisse relevante Beziehungen zum biologischen Geschlecht und dem Geschlechtsrollenselbstkonzept nachgewiesen wurden. Erwartungsgemäß konnten die postulierten Zusammenhänge von Teilskalen und Konstrukten belegt werden. So wiesen die neuen Maskulinitätsskalen höhere Zusammenhänge mit Konstrukten wie beruflichem Ehrgeiz, Hoffen auf Erfolg sowie auch



(physischer) Aggressivität und Macht auf als die Femininitätsskalen. Dementsprechend konnte den Femininitätsskalen ein stärkerer Bezug zu Konstrukten wie sozialen Werten und Empathie sowie emotionaler Selbstwertschätzung und ängstlicher Depressivität nachgewiesen werden als den Maskulinitätsskalen.

Das Entscheidende in der Validierungsphase waren jedoch die spezifischen Beziehungen der Konstrukte zu den Teilskalen, entsprechend Typizität und Valenz. Es war zu zeigen, dass jede Teilskala den höchsten Beitrag zur Vorhersage der jeweiligen Validierungskonstrukte leistete, der auch nach Einführung der übrigen Teilskalen erhalten blieb. Dies ließ sich für alle Teilskalen und Validierungskonstrukte zeigen. Zum Einsatz kam der Ansatz des destruktiven Testens (Anderson & Anderson, 1996), der die Stärke einer Verbindung zwischen zwei Konstrukten (z. B. Fem+ und Empathie) durch das sukzessive Hinzufügen möglicher konkurrierender Konzepte (z. B. Mas+) prüft. Trotz dieser erhöhten Schwierigkeit, dass auch die jeweils anderen Teilskalen Zusammenhänge zu den Validierungskonstrukten haben könnten, erwiesen sich die zu ihren Konstrukten eindeutig hypothetisch zugeordneten Teilskalen jeweils als die stärksten Prädiktoren. So standen Aggressivität und das Machtmotiv in einem spezifischen Zusammenhang mit der negativen maskulinen Teilskala. Ebenso wurde für die Konstrukte Hoffen auf Erfolg oder beruflicher Ehrgeiz ein Zusammenhang mit der positiven maskulinen Teilskala vorhergesagt. Im gleichen Maße schließlich war auch die Verknüpfung zwischen Empathie und der positiven Femininitätsskala nachweislich größer als die zur negativen Femininitätsskala, während negative Femininität in stärkerem Zusammenhang zu Problemen mit der emotionalen Selbstwertschätzung und depressiver Ängstlichkeit stand.

### **3.2.4 Studie 6: Replikation der neuen Teilskalen**

#### **3.2.4.1 Fragestellung**

In Studie 6 ging es darum, die Qualitäten und Reliabilitäten der neuen Teilskalen erneut zu überprüfen und die Validität einer weiteren Testung durch zusätzliche Konstrukte zu unterziehen. Dazu wurde eine zusätzliche Gruppe von Probanden gebeten, die neuen GSK-Items als Teil ihres Selbst einzuschätzen. Es wurden neue Validierungskonstrukte nach dem Kriterium ausgewählt, dass sie eine konzeptuelle Nähe zu den in Studie 5 verwendeten Konstrukten aufweisen sollten aber durch neue Instrumente erfasst wurden. Dementsprechend wurde zur Validierung der Teilskala Mas+ der Aspekt der Leistungsmotivation ausgewählt, für die Teilskala Mas- ein alternatives Maß der Aggressivität, für Fem+ ein Konzept der Fürsorglichkeit und für Fem- ein Maß der Trait-Ängstlichkeit.

#### **3.2.4.2 Methode**

##### **3.2.4.2.1 Stichprobe**

An der Gesamtstichprobe beteiligten sich 839 Personen (292 Männer und 547 Frauen). Erneut präsentierte sich mit einem Durchschnittsalter von 23.06 Jahren ( $SD = 4.17$  Jahre, Min: 18, Max: 45 Jahre) eine vorwiegend studentische Stichprobe, von der neben den 62% mit einem

Abiturabschluss weitere 8% über einen Hochschulabschluss verfügten. Übereinstimmend gaben 91% Heterosexualität als ihre sexuelle Orientierung an, 2,8% bezeichneten sich als homosexuell, 6,7% als bisexuell.

### 3.2.4.2.2 Instrumente

#### Neues Instrument zur Erfassung des GSK

Die Befragung erfolgte erneut in einer Onlinestudie, in der die Probanden zunächst die 24 neuen Geschlechtsrollenitems hinsichtlich ihrer Selbstzuschreibung einschätzten, deren Darstellung und Erfassung denen aus den vorherigen Studien entsprach (Tabelle 60).

Tabelle 60. Darstellung der Erfassung der Geschlechtsrollenitems (Studie 6)

Ich bin ...	Wie sehr trifft diese Eigenschaft auf Sie zu				
	trifft gar nicht zu		trifft genau zu		
... liebevoll	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... naiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... ängstlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Zum Einsatz kamen vier Validierungskonstrukte, von denen jeweils eines einer spezifischen Teilskala zugeordnet wurde. Die Darbietungsreihenfolge der Validierungskonstrukte wurde permutiert, alle Probanden beurteilten sämtliche Maße.

#### Validierung MAS+

##### Leistungsorientierung (FPI)

Zur Validierung der positiven Maskulinität wurde das Konzept der *Leistungsorientierung* aus der 7. Auflage des FPI (Fahrenberg et al., 2001) eingesetzt. Die entsprechende Skala besteht aus 12 Items (z. B. „*Ich habe gerne mit Aufgaben zu tun, die schnelles Handeln verlangen.*“), denen ein dichotomer Antwortmodus mit den Optionen (1) *stimmt* oder (2) *stimmt nicht* zugeordnet war.

#### Validierung MAS-

##### Aggressivität (FPI)

Aggressivität als Validierungskonstrukt in Bezug auf negative Maskulinität wurde aus der 7. Auflage des FPI (Fahrenberg et al., 2001) verwendet. Die Skala umfasste 12 Items, z. B. „*Wenn jemand meinem Freund etwas Böses tut, bin ich dabei, wenn es heimgezahlt wird.*“ Auch diesen Items ist eine dichotome Antwortskala zugeordnet mit den Optionen (1) *stimmt* oder (2) *stimmt nicht*.

#### Validierung FEM+

##### Fürsorglichkeit (Nurturance) (PRF)

Zur Validierung von FEM+ kam die Teilskala Nurturance (Fürsorglichkeit) der Personality Research Form (PRF) (Jackson, 1974; Stumpf et al., 1985) zum Einsatz. Das Konzept beinhaltet 16 Aspekte wie Mitgefühl, Güte oder Unterstützung anderer (Beispielitem: „*Ich finde Kinder lästig, weil sie so viel Pflege erfordern.*“). Den Probanden wurde zur Einschätzung eine zweistufige Antwortskala (1) *Richtig* oder (2) *Falsch* dargeboten.

### Validierung FEM-

#### Trait-Ängstlichkeit

Zur Erfassung des Trait-Ängstlichkeit wurde die Trait-Skala des State-Trait-Angstinventar (STAI) in der deutschen Fassung von Laux, Glanzmann, Schaffner und Spielberger (1981) verwendet, die aus 20 Items besteht. 13 dieser Items sind dabei in Richtung Angst formuliert und sieben in Richtung Angstfreiheit. Probanden wurden gebeten zu beschreiben, wie sie sich im Allgemeinen fühlen. Die Beantwortung der Aussagen (z. B. „*Ich glaube, mir geht es schlechter als anderen Leuten.*“) erfolgte auf einer vierstufigen Skala mit Häufigkeitsangaben von (1) *fast nie* bis (4) *fast immer*. Der STAI wurde schlicht mit „Wir bitten Sie, einige Ihrer üblichen Verhaltensweisen, Einstellungen und Gewohnheiten zu beschreiben.“ eingeleitet, dem eine Erklärung der Antwortskala folgte.

Dass die Wahl der Validierungskonstrukte im Sinne distinkter Konzepte für die vier Teilskalen als gelungen angesehen werden kann, zeigten die geringen Korrelationen der Konstrukte untereinander (Tabelle 61).

Tabelle 61. Korrelationstabelle der Validierungskonstrukte für Studie 6

	Aggr. (FPI)	Fürs. (PRF)	Angst (STAI)
Leistungsorientierung (FPI)	.20**	.07*	.25**
Aggressivität (FPI)		-.14**	-.17**
Fürsorglichkeit (PRF)			.02

Anmerkung: \*\*p < .01, \*p < .05

#### 3.2.4.2.3 Durchführung

Ähnlich wie in allen bisher durchgeführten Studien der vorliegenden Arbeit wurden zunächst alle Probanden über das Ziel der Studie informiert, der Anonymität versichert und darauf hingewiesen, dass sie die Untersuchung jederzeit beenden könnten. Nach Angabe ihrer demografischen Informationen beantworteten alle Personen zuerst die Geschlechtsrollenitems hinsichtlich ihrer Selbstzuschreibung, ehe sie schließlich Angaben zu allen Validierungskonstrukten machten, die zur Vermeidung von Reihenfolgeeffekten permutiert wurden.

### 3.2.4.3 Ergebnisse

#### 3.2.4.3.1 Statistische Kennwerte der Teilskalen

Im ersten Schritt der Datenanalyse wurden zunächst die internen Konsistenzen der neuen Teilskalen für die Selbstbeschreibung sowohl über die gesamte Stichprobe als auch getrennt für Männer und Frauen bestimmt. Die Ergebnisse dieser Analysen sind in Tabelle 62 zusammengefasst.

Tabelle 62. Skalenkennwerte der 4 Teilskalen in der Selbstbeschreibung

Skala	Gesamtstichprobe			Männer			Frauen			Vergleich		
	$\alpha$	M	SD	$\alpha$	M	SD	$\alpha$	M	SD	Diff.	p	$\eta^2$
Mas+	.82	3.69	.69	.81	3.91	.65	.81	3.57	.68	.34	<.001	.05
Mas-	.78	2.11	.70	.76	2.34	.71	.77	1.98	.67	.36	<.001	.06
Fem+	.88	4.03	.75	.86	3.73	.73	.88	4.18	.72	-.45	<.001	.08
Fem-	.73	2.84	.77	.78	2.57	.81	.67	2.98	.71	-.41	<.001	.07

Anmerkung:  $N_{\text{Männer}} = 547$ ,  $N_{\text{Frauen}} = 292$

Die vier Teilskalen wiesen nicht nur erneut zufriedenstellende Reliabilitätswerte auf, sondern unterstrichen damit eindrucksvoll die Replizierbarkeit der Teilskalen in einer wiederholten

Messung mit einer neuen Stichprobe. Die getrennten Analysen für Männer und Frauen zeichneten ein ähnliches Bild, wobei wiederum auf die insgesamt am geringsten ausfallende Reliabilität der negativen femininen Teilskala, insbesondere in der Gruppe der Frauen, hingewiesen werden sollte.

Eine multivariate Varianzanalyse für die Selbstbeschreibung auf den vier Teilskalen mit dem Geschlecht als unabhängige Variable und Alter als Kovariate ergab einen signifikanten multivariaten Effekt für das Geschlecht,  $F(4, 833) = 40.00$ ,  $p < .001$ ,  $\eta^2 = .16$ . Alle vier univariaten Effekte waren signifikant, wie in Tabelle 62 dargestellt. Auf den beiden maskulinen Teilskalen wiesen Männer höhere Werte auf als Frauen, während erwartungsgemäß Frauen auf den beiden femininen Teilskalen höhere Werte zeigten als die Männer. Der multivariate Effekt des Alters wurde ebenfalls signifikant,  $F(4, 833) = 4.36$ ,  $p < .01$ ,  $\eta^2 = .00$ ; und ging auf signifikante Unterschiede bei der positiven maskulinen Skala zurück.

Eine weitere Varianzanalyse mit der Valenz (positiv-negativ) und der Geschlechtstypizität der Teilskala (männlich-weiblich) als Messwiederholungsfaktoren sowie dem Geschlecht der Probanden ergab zunächst einen Haupteffekt der Valenz: Männer wie Frauen schrieben sich insgesamt negative Attribute in geringerem Ausmaß zu als positive,  $F(1, 837) = 2743.71$ ,  $p < .001$ ,  $\eta^2 = .77$ . Darüber hinaus zeigte sich erwartungsgemäß auch eine signifikante Interaktion von Geschlechtstypizität und Geschlecht,  $F(1, 837) = 160.63$ ,  $p < .001$ ,  $\eta^2 = .16$ . Die Probanden beanspruchten somit nicht nur positive, sondern auch die negativen geschlechtskongruenten Eigenschaften stärker zur Selbstbeschreibung als die geschlechtsinkongruenten Eigenschaften.

### 3.2.4.3.2 Konfirmatorische Faktorenanalyse der neuen Teilskalen

Wie bereits in Studie 5 wurde eine konfirmatorische Faktorenanalyse zur Absicherung der Dimensionen des GSK durchgeführt. Das Vier-Faktoren-Modell konnte als jenes mit dem besten Modell-Fit ( $\chi^2/df = 2.91$ ;  $p \leq .001$ ;  $CFI = .95$ ;  $RMSEA = .048$ ) sowohl für Frauen als auch Männer bestätigt werden (Tabelle 63).

Tabelle 63. Modellprüfung mit den Items der neuen Skala (Studie 6)

	<b>n.s.</b> $p_{\chi^2}$	<b>&lt; 3.00</b> $\chi^2/df$	<b>&gt; .950</b> CFI	<b>&lt; .050</b> RMSEA
<b>1. Ein-Faktoren-Modell</b>	<b>.000</b>	<b>4.84</b>	<b>.900</b>	<b>.068</b>
Männer	.000	2.39	.890	.069
Frauen	.000	3.97	.871	.074
<b>2. Zwei-Faktoren-Modell (Typizität)</b>	<b>.000</b>	<b>3.83</b>	<b>.927</b>	<b>.058</b>
Männer	.000	1.95	.926	.057
Frauen	.000	2.63	.930	.055
<b>3. Zwei-Faktoren-Modell (Valenz)</b>	<b>.000</b>	<b>4.68</b>	<b>.905</b>	<b>.066</b>
Männer	.000	2.18	.907	.064
Frauen	.000	3.10	.910	.062
<b>4. Zwei-Faktoren-Modell (umgekehrt)</b>	<b>.000</b>	<b>3.46</b>	<b>.937</b>	<b>.054</b>
Männer	.000	1.72	.944	.050
Frauen	.000	2.41	.939	.051
<b>5. Vier-Faktoren-Modell</b>	<b>.000</b>	<b>2.91</b>	<b>.952</b>	<b>.048</b>
Männer	.000	1.60	.954	.045
Frauen	.000	2.18	.950	.046

Anmerkung:  $N_{\text{Männer}} = 292$ ,  $N_{\text{Frauen}} = 547$ ; (Ergebnisse aus Studie 5 unterhalb der Diagonalen)

Während die Kennwerte dabei geringfügig schlechter ausfielen als in Studie 5 (z. B.  $\chi^2/df = 2.91$  vs. 2.25), erfüllen sie dennoch die festgelegten Minimalkriterien. Dabei erwiesen sich die die

Beziehungen der Teilskalen insgesamt betrachtet als ähnlich (z. B. Mas+ und Mas-:  $r = .20$ , im Vergleich zu  $r = .17$  aus der ersten Studie), lediglich die Beziehung der positiven maskulinen mit der negativen femininen Teilskala fiel etwas größer aus.

### 3.2.4.3.3 Korrelative Zusammenhänge der Teilskalen

Die Korrelationsmatrix (Tabelle 64) zeigte zunächst, dass in dieser Studie nur Mas+ eine geringfügige aber signifikante Korrelation mit dem Alter aufwies. Darüber hinaus unterstützte die Matrix Ergebnisse, die bereits aus der konfirmatorischen Faktorenanalyse bekannt sind: zwischen nahezu allen Teilskalen bestehen signifikante, jedoch geringe Korrelationen.

Tabelle 64. Korrelationen der Teilskalen

	Mas+	Mas-	Fem+	Fem-
Alter	.11**	-.10**	-.02	-.05
Mas+		.19**	-.15**	-.22**
Mas-			-.23**	.05
Fem+				.15**

	Mas+	Mas-	Fem+	Fem-
Alter	.12**	-.03	-.04	.00
Mas+		.16**	-.13**	-.37**
Mas-			-.27**	-.10**
Fem+				.25**

Anmerkung: \*\* $p < .01$ , \* $p < .05$ ; (Ergebnisse aus Studie 5 oberhalb der Diagonalen)

Zum Vergleich sind oberhalb der Diagonalen die Korrelationen aus Studie 5 aufgeführt, denen das Bild der vorliegenden Studie ähnelt. Die engsten Beziehungen bestanden demnach erneut jeweils zwischen den beiden entgegengesetzten Teilskalen Fem+ und Mas- sowie Mas+ und Fem-. Die Teilskalen mit gleicher Typizität (Mas+ und Mas-, Fem+ und Fem-) korrelierten jeweils schwach positiv miteinander. Positive Maskulinität und positive Femininität sowie Mas- und Fem- korrelierten jeweils schwach negativ miteinander. Insgesamt konnte die Korrelationsstruktur aus Studie 5 auch in Studie 6 repliziert werden.

### 3.2.4.3.4 Statistische Kennwerte und Korrelationen der Validierungskonstrukte

In einem nächsten Schritt wurden die Reliabilitäten der eingesetzten Instrumente sowie die mittleren Ausprägungen über die gesamte Stichprobe überprüft. Des Weiteren wurden die Geschlechtsunterschiede auf den Validierungskonstrukten betrachtet und getestet, inwieweit sich Männer und Frauen hinsichtlich der Konstrukte unterschieden. Die Ergebnisse dieser Analysen sind in Tabelle 65 zusammengefasst.

Für alle zum Einsatz gekommenen Validierungskonstrukte wurden gute bis sehr gute Reliabilitätswerte ermittelt, sowohl für die Gesamtstichprobe als auch getrennt für die beiden Geschlechter. Darüber hinaus konnten systematische Abfälle der Alphas zum Ende des Fragebogens hin ausgeschlossen werden. Erneut waren es die mittels des FPI erhobenen Skalen, die die schwächsten internen Konsistenzen aufwiesen. Geschlechtsunterschiede zeigten sich in der erwarteten Richtung auf allen Konstrukten mit Ausnahme von Leistungsorientierung.

Tabelle 65. Statistische Kennwerte der Validierungskonstrukte für Studie 6

	Gesamt			Männer			Frauen			Unterschied		
	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	<i>F</i>	<i>p</i>	$\eta^2$
<b>Prädiktoren für MAS+</b>												
Leistungsorientierung	.65	1.43	.22	.68	1.43	.22	.64	1.43	.21	.13	n.s.	<.001
<b>Prädiktoren für MAS-</b>												
Aggressivität	.70	1.58	.23	.68	1.53	.24	.67	1.46	.22	21.65	<.001	<.01
<b>Prädiktoren für FEM+</b>												
Fürsorglichkeit	.71	1.34	.19	.71	1.41 <sup>38</sup>	.19	.68	1.31	.18	61.51	<.001	n.s.
<b>Prädiktoren für FEM-</b>												
STAI	.93	2.26	.57	.94	2.15	.60	.92	2.31	.54	16.98	<.001	<.01

Anmerkung: *N* = 839

Auf der Ebene der bivariaten Korrelationen ergaben sich die postulierten Zusammenhänge der Teilskalen mit den zum Einsatz gekommenen Validierungskonstrukten (Tabelle 66). Die positive maskuline Teilskala korrelierte am stärksten mit der Leistungsorientierung. Die Teilskala Mas-dagegen korrelierte negativ mit der Aggressivität. Das negative Vorzeichen ist der Art der Kodierung dieser Skala geschuldet und muss dahingehend interpretiert werden, dass hohe Aggressivitäts-Werte mit hohen negativen maskulinen Werten assoziiert sind. FEM+ wies erwartungsgemäß eine hohe Korrelation zur Fürsorglichkeit auf, deren negatives Vorzeichen als ein positiver Zusammenhang zu interpretieren ist. Die negative feminine Teilskala zeigte eine starke positive Korrelation zur Trait-Angst.

Tabelle 66. Korrelationen der Teilskalen mit den Validierungskonstrukten

	Mas+	Mas-	Fem+	Fem-
Leistungsorientierung (FPI)	-.38**	-.23**	-.05	.33**
Aggressivität (FPI)	-.06	-.39**	.04	.06
Fürsorglichkeit (PRF)	.10**	.31**	-.39**	-.13**
Trait-Ängstlichkeit	-.25**	.02	.04	.63**

Anmerkung: \*\**p* < .01, \**p* < .05

### 3.2.4.3.5 Konstruktvaliditätsanalysen mittels hierarchischer Regressionen

Wie bereits in Studie 5 wurde der spezifische Vorhersagebeitrag der Teilskalen zu den Validierungskonstrukten nach der Methode des destruktiven Testens (Anderson & Anderson, 1996) über hierarchische Regressionsanalysen überprüft. In Schritt 1 wurde die jeweils assoziierte Teilskala des Validierungskonstrukts in die Regressionsgleichung aufgenommen, um in den folgenden Schritten nachzuweisen, dass die Hinzunahme der weiteren Teilskalen (im zweiten Schritt die beiden Teilskalen entsprechender Typizität und Valenz, im dritten Schritt die Teilskala entgegengesetzt in Typizität und Valenz sowie im vierten und letzten Schritt schließlich die des biologischen Geschlechts) zu keinem aussagekräftigen Zugewinn der Aufklärungsleistung führt.

Erwartungskonform trug die positive maskuline Teilskala den größten Beitrag zur Varianzaufklärung der Leistungsorientierung (Tabelle 67) bei ( $\beta = -.38^{**}$ ,  $B = -.12$ ;  $SE B = .04$ ,  $R^2 = .14$ ,  $p < .001$ ). Je höher die Ausprägung der positiven Maskulinität, desto höher die berichtete Leistungsorientierung (deren negatives Vorzeichen lediglich auf die Art der Antwortkodierungen zurückzuführen ist).

<sup>38</sup> Aufgrund der Kodierung der Antworten bedeuten hier geringe Zahlen hohe Ausprägung in Richtung des Konstrukts

Das Hinzunehmen weiterer Teilskalen offenbarte, dass darüber hinaus alle weiteren Teilskalen ebenso wie das biologische Geschlecht einen eigenen Beitrag verzeichnen konnten. Für die positive Femininität und die negative Maskulinität war deren Vorhersageleistung ebenfalls in einer Erhöhung der Leistungsorientierung festzumachen, während die negative feminine Teilskala mit einer Verringerung dieser in Zusammenhang steht. Dennoch ist auch im letzten Schritt der Analyse Mas+ der stärkste Prädiktor.

Tabelle 67. Vorhersage von Leistungsorientierung durch Teilskalen (Mas+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas+	-.12	.04	-.38**
Schritt 2			
Mas+	-.11	.01	-.36**
Fem+	-.05	.01	-.16**
Mas-	-.07	.01	-.21**
Schritt 3			
Mas+	-.09	.01	-.27**
Fem+	-.06	.01	-.21**
Mas-	-.07	.01	-.22**
Fem-	.07	.01	.27**
Schritt 4			
Mas+	-.09	.01	-.30**
Fem+	-.05	.01	-.18**
Mas-	-.08	.01	-.25**
Fem-	.08	.01	.29**
Geschlecht	-.08	.02	-.17**

Anmerkung:  $R^2 = .14^{**}$ ; Schritt 2  $\Delta R^2 = .05^{**}$ ; Schritt 3  $\Delta R^2 = .06^*$ ; Schritt 4  $\Delta R^2 = .02$ ;  $**p < .001$ ;  $N = 817$

Zur Überprüfung des Einflusses der negativen maskulinen Teilskala zeigte sich in Hinblick auf Aggressivität erwartungsgemäß, dass je höher die negative Maskulinität ( $\beta = -.39^{**}$ ,  $B = -.13$ ;  $SE B = .01$ ,  $R^2 = .15$ ,  $p < .001$ ), desto höher die Ausprägungen bei physischer Aggression (Tabelle 68). Erneut verzeichneten die positive feminine Teilskala und das biologische Geschlecht einen geringen, jedoch zu vernachlässigenden eigenen Beitrag.

Tabelle 68. Vorhersage von Aggressivität durch Teilskalen (Mas-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas-	-.13	.01	-.39**
Schritt 2			
Mas-	-.13	.01	-.39**
Fem-	.01	.01	.03
Mas+	.00	.01	.01
Schritt 3			
Mas-	-.13	.01	-.41**
Fem-	.01	.01	.04
Mas+	.00	.01	.01
Fem+	-.02	.01	-.08*
Schritt 4			
Mas-	-.13	.01	-.40**
Fem-	.01	.01	.03
Mas+	.01	.01	.02
Fem+	-.03	.01	-.09*
Geschlecht	.04	.02	.08*

Anmerkung:  $R^2 = .15^{**}$ ; Schritt 2  $\Delta R^2 = n.s.$ ; Schritt 3  $\Delta R^2 = .01^*$ ; Schritt 4  $\Delta R^2 = .02^*$ ;  $**p < .001$ ;  $*p < .05$ ;  $N = 812$

Zur Vorhersage von Fürsorglichkeit (Tabelle 69) stellte sich hypothesenkonform die positive feminine Teilskala als der eindeutig beste Prädiktor heraus ( $\beta = -.39^{**}$ ,  $B = -1.51$ ;  $SE B = .12$ ,  $R^2 = .15$ ,  $p < .001$ ); je höher die berichtete positive Femininität, als desto fürsorglicher beschrieben sich die Probanden (erneut ist das negative Vorzeichen auf die Art der Kodierung

zurückzuführen). Darüber hinaus konnte ebenso die negative Maskulinität in umgekehrter Weise einen Einfluss verzeichnen ( $\beta = .21^{**}$ ,  $B = .85$ ;  $SE B = .12$ ,  $\Delta R^2 = .04$ ,  $p < .001$ ): Personen, die sich stark auf den negativen maskulinen Eigenschaften positionierten, schätzten sich als weniger fürsorglich ein. Erneut sei angemerkt, dass der Aufklärungswert von Mas+ deutlich geringer ausfiel als der von Fem+.

Tabelle 69. Vorhersage von Fürsorglichkeit durch Teilskalen (Fem+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem+	-1.51	.12	-.39**
Schritt 2			
Fem+	-1.46	.13	-.38**
Mas+	.19	.14	.04
Fem-	-.11	.13	-.03
Schritt 3			
Fem+	-1.26	.13	-.33**
Mas+	.07	.14	.02
Fem-	-.12	.13	-.03
Mas-	.85	.13	.21**
Schritt 4			
Fem+	-1.17	.13	-.31**
Mas+	.01	.14	.00
Fem-	-.06	.13	-.02
Mas-	.77	.13	.19**
Geschlecht	-.69	.20	-.11*

Anmerkung:  $R^2 = .15^{**}$ ; Schritt 2  $\Delta R^2 = \text{n.s.}$ , Schritt 3  $\Delta R^2 = .04^{**}$ , Schritt 4  $\Delta R^2 = .01^*$ ;  $**p < .001$ ;  $*p < .01$ ,  $N = 817$

Noch klarer verhielt es sich mit der Vorhersageleistung von Angst (STAI) durch negative Femininität (Tabelle 70). Erwartungsgemäß berichteten Personen mit hoch ausgeprägter negativer Femininität am stärksten von Ängsten ( $\beta = .63^{**}$ ,  $B = .46$ ;  $SE B = .02$ ,  $R^2 = .40$ ,  $p < .001$ ). Erneut konnte die positive Femininität in umgedrehter Weise einen kleinen Einfluss geltend machen, der aber vor dem Hintergrund des Beitrags von Fem- deutlich geringer ausfiel.

Tabelle 70. Vorhersage von STAI durch Teilskalen (Fem-)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Fem-	.46	.02	.63**
Schritt 2			
Fem-	.48	.02	.66**
Mas-	.04	.02	.05
Fem+	-.09	.02	-.11**
Schritt 3			
Fem-	.48	.02	.65**
Mas-	.04	.02	.05
Fem+	-.09	.02	-.11**
Mas+	-.03	.02	-.03
Schritt 4			
Fem-	.48	.02	.65**
Mas-	.04	.02	.05
Fem+	-.09	.02	-.12**
Mas+	-.03	.02	-.03
Geschlecht	.01	.03	.01

Anmerkung:  $R^2 = .40^{**}$ ; Schritt 2  $\Delta R^2 = .02^{**}$ , Schritt 3  $\Delta R^2 = \text{n.s.}$ , Schritt 4  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .01$ ;  $N = 816$

### 3.2.4.4 Diskussion

Die Befunde aus Studie 6 können die Ergebnisse der ersten Validierungsstudie überzeugend bestätigen. Alle Teilskalen des neuen Inventars zur Messung des Geschlechtsrollenselbstkonzepts zeichneten sich erneut durch gute Reliabilitäten aus und



konnten als Bestandteil eines Vierfaktorenmodells durch konfirmatorische Faktorenanalysen identifiziert werden. Männer schrieben sich erwartungsgemäß maskuline Eigenschaften im stärkeren Maße zu als Frauen, während Frauen die femininen Eigenschaften im stärkeren Maße für sich beanspruchten als Männer dies taten. Die sicherlich als komplex geltenden Beziehungen der Teilskalen konnten in ähnlicher Weise nachgewiesen werden: es bestanden jeweils stärkere Korrelationen (negativer Art) zwischen Fem- und Mas+ (sowie Mas- und Fem+) als zwischen allen anderen möglichen Kombinationen. Die Zusammenhänge zwischen den jeweils entgegengesetzten Teilskalen fielen zwar geringfügig höher aus als in Studie 5, blieben aber immer noch deutlich unter der Marke, die die Relevanz eigenständiger negativer Teilskalen hätte in Frage stellen können. Zudem konnte erneut der Beweis geliefert werden, dass alle vier Teilskalen eigene Beiträge zur Vorhersage von theoretisch bekannten geschlechtsrollentypischen Validierungskonstrukten leisteten. Zur Überprüfung der Konstruktvalidität wurden in Studie 6 vier neue Konstrukte vorgestellt, die aber inhaltlich verwandt zu den in Studie 5 betrachteten sind. Dies waren für die Teilskala Mas+ das Konstrukt der Leistungsorientierung, für Mas- Aggressivität, gemessen mittels des FPIs, für Fem+ das Konzept der Fürsorglichkeit sowie für die negative feminine Teilskala die Trait-Angst. So konnte gezeigt werden, dass die negative maskuline Teilskala erwartungsgemäß einen höheren Zusammenhang zum Konstrukt Aggressivität aufwies als jede andere Teilskala. Dementsprechend waren hohe Ausprägungen negativer femininer Eigenschaften stärker verknüpft mit dem als feminin geltenden Validierungskonstrukt der Trait-Ängstlichkeit.

Bestätigt wurden die korrelativen Ergebnisse erneut durch das destruktive Testen (Anderson & Anderson, 1996), welches nicht nur den Zusammenhang von zwei Konstrukten im Rahmen einer Regressionsanalyse überprüft, sondern zudem erschwerte Bedingungen schafft durch das sukzessive Hinzufügen möglicher konkurrierender Konzepte (z. B. Mas+), die diese existierende Beziehung der Konstrukte aufbrechen könnten. Trotz dieser anspruchsvollen Analysen bewährten sich die hypothetisierten Konstrukte in den Beziehungen zu den jeweils zugeordneten Teilskalen.

Diese Replikationsstudie unterstreicht die Qualität der vier neuen Teilskalen als ein sparsames Inventar und die Relevanz für die psychologische Forschung im Bereich des Geschlechtsrollenselbstkonzepts. Es verdeutlicht zudem, dass durch das Hinzufügen der negativen Teilskalen ein deutlicher Zugewinn an erklärter Varianz in den ausgewählten Validierungskonstrukten erzielt werden kann und bestätigt damit die theoretische Erweiterung des GSK um den Aspekt negativer Merkmale.

### **3.2.5 Studie 7: Erweiterung der neuen Skalen um eine Erwartungs-mal-Wert-Komponente**

#### **3.2.5.1 Fragestellung**

Wie bereits in Abschnitt 2.3.2 *Die Bedeutung negativer Aspekte der Geschlechtsrollenidentität* ab Seite 60 ff. theoretisch dargelegt, wurde in bisherigen Konzeptualisierungen und Messungen des GSK auf die Erfassung der subjektiven Bedeutsamkeit eines Merkmals für die individuelle Selbstkonstruktion verzichtet. Es wurde unterstellt, dass die Valenz geschlechtstypischer

Attribute ausreichend ist, um individuelle Unterschiede im GSK abzubilden. Analog zu Erwartungs-mal-Wert-Modellen der Motivations- und Einstellungsforschung (Heckhausen, 1989) kann jedoch alternativ postuliert werden, dass die Kombination der Valenz eines Merkmals mit der subjektiven Bedeutsamkeit des Merkmals für die individuelle Selbstkonstruktion eine bessere Vorhersage von psychologischen Korrelaten des GSK erlaubt. Diese Hypothese sollte in einer abschließenden Studie empirisch geprüft werden.

Ausgangspunkt ist die Annahme, dass Personen die sozial geteilten Eigenschaftszuschreibungen mehr oder weniger stark für sich selbst als positiv oder negativ bewerten. Dies wird vor allem vor dem Hintergrund der nun auch negativ einzuschätzenden Eigenschaften relevant. Wenn Personen gebeten werden, sich auf diesen Eigenschaften zu positionieren und anzugeben, wie sehr diese Teil ihres Selbstkonzepts sind, könnte damit gleichermaßen eine gewisse Unzufriedenheit nach dem Prinzip einhergehen, dass man sich zwar negative Eigenschaften wie *schroff* und *ängstlich* zuschreiben muss, man aber mit der eigenen Ausprägung diesbezüglich unzufrieden ist oder diese als unwichtig für das eigene Selbst betrachtet. Dieses Problem soll nun durch die Einführung einer Wichtigkeitskomponente im Zuge der Selbstbeschreibung angegangen werden. Angelehnt an das bekannte Erwartungs-mal-Wert-Modell wird zusätzlich zur eigenen Einschätzung bezüglich der vorliegenden Eigenschaften auch noch die persönliche (und nicht mehr stereotype) Wichtigkeit erfragt. Deshalb wird erwartet, dass das Produkt aus Selbsteinschätzung und persönlicher Wichtigkeit einen größeren Aufklärungswert besitzt als die einfache Selbstzuschreibung.

### 3.2.5.2 Methode

An der Überprüfung der Erwartungs-mal-Wert-Erfassung des GSK nahmen die Versuchspersonen aus Studie 5 teil. Da die Erfassung der Wichtigkeitskomponente in den Ablauf der Validierungsstudie (siehe Abschnitt 3.2.3.2.2 Instrumente, S. 160) durch eine erweiterte Erfassung integriert wurde, sind die Angaben zu Stichprobenszusammensetzung, Instrumenten und Durchführung identisch. Lediglich die Art der kombinierten Erfassung der Geschlechtsrollenitems (Tabelle 71) muss kurz aufgeführt werden, da Probanden neben der Einschätzung „Wie sehr trifft diese Eigenschaft auf Sie zu“ zusätzlich zu jedem Item Angaben bezüglich der Wichtigkeit dieser Eigenschaft für das eigene Selbst machen sollten.

Tabelle 71. Darstellung der Erfassung der Geschlechtsrollenitems

Ich bin ...	Wie sehr trifft diese Eigenschaft auf Sie zu					Wie wichtig ist es für Sie diese Eigenschaft zu besitzen				
	trifft gar nicht zu			trifft genau zu		gar nicht wichtig			sehr wichtig	
... liebevoll	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... naiv	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
... ängstlich	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Dazu wurde direkt neben der fünfstufigen Antwortskala zur Selbsteinschätzung eine weitere Skala mit fünf Möglichkeiten von (1) *gar nicht wichtig* bis (5) *sehr wichtig* dargeboten, auf der die Probanden ihre Einschätzungen vornehmen konnten.

### 3.2.5.3 Ergebnisse

#### 3.2.5.3.1 Korrelative Zusammenhänge der Teilskalen und der Wertekomponenten

Die Korrelationsmatrix (Tabelle 72) zeigt keine nennenswerten Korrelationen zwischen den Wichtigkeitszuschreibungen auf Itemebene, da in Bezug auf die Wichtigkeitseinschätzung nur die Produktbildungen mit den itemspezifischen Selbstzuschreibungen relevant sind (weshalb an dieser Stelle auch keine Reliabilitäten der Wichtigkeitszuschreibung präsentiert werden).

Tabelle 72. Korrelationen der Teilskalen und der individuellen Wichtigkeitskomponente

	Mas+	Mas-	Fem+	Fem-	Mas+ W	Mas- W	Fem+ W	Fem- W
Alter	.11**	-.10**	-.02	-.05	.12**	.04	-.03	.06
Mas+		.19**	-.15**	-.22**	.59**	.14**	-.12**	.05
Mas-			-.23**	.05	.13**	.30**	-.19**	.10**
Fem+				.15**	-.04	-.01	.71**	.05
Fem-					-.07*	.00	.07*	.10**
Mas+ W						.05	.06	-.03
Mas- W							-.13**	.62**
Fem+ W								-.00

Anmerkung: \*\* $p < .01$ , \* $p < .05$

Auffällig ist jedoch, dass die Gesamtwichtigkeitszuschreibung von Fem- mit der von Mas- stark positiv miteinander korrelierte, was vermuten lässt, dass beide Skalen gleichermaßen akzeptiert oder abgelehnt werden.

#### 3.2.5.3.2 Validitätsanalysen der individuellen Wertkomponente - Wichtigkeit

In erweiterten Analysen sollte schließlich gezeigt werden, dass die Erfassung der individuellen Wertkomponente im Rahmen der Selbstzuschreibung einen zusätzlichen Beitrag zu der Varianzaufklärung von möglichen Validierungskonstrukten leisten kann; derart, dass nicht nur die eigene Positionierung auf geschlechtsbezogene Attribute, sondern zudem die Angabe der Wichtigkeit eben dieser Eigenschaften für das Selbstkonzept ebenfalls entscheidend sind. Dazu wurde pro Item die multiplikative Verknüpfung von Selbstzuschreibung und Wichtigkeitsurteil gebildet. In der Tabelle 73 sind die Korrelationen der gemittelten Produkte pro Teilskala mit den jeweiligen Validierungskonstrukten angegeben.

Tabelle 73. Korrelationen der gemittelten Produkte der Teilskalen mit den Validierungskonstrukten

	gemittelte Produkte aus Selbst und Wichtigkeit			
	Mas+ S*W	Mas- S*W	Fem+ S*W	Fem- S*W
berufl. Ehrgeiz (AVEM)	-.07	.09	.02	.06
Hoffen auf Erfolg (AMS)	.01	-.01	.08	.14**
Aggressivität (AQ)	.05	.07	.02	-.03
Werteskala Macht	.16*	.18**	.01	-.03
Empathie (Sorge)	.10*	-.12*	-.20**	-.04
Soziale Orientierung (FPI)	.20	.05	.19	-.12
ESWS	-.05	-.02	-.02	-.12**
Ängstl. Depressivität (BSF)	-.15	-.04	-.02	-.03

Anmerkung: \*\* $p < .01$ , \* $p < .05$

Die Ergebnisse stellen bereits die Position der zusätzlich erfassten Wichtigkeit in Interaktion mit den individuellen Bewertungen der Items in Frage, da sich geringe bis gar keine Korrelationen zu den Konstrukten finden ließen. So deuteten zwar ein durchschnittlich negativer Zusammenhang der gemittelten positiven femininen Interaktionsterme mit sozialer Empathie sowie eine positive Beziehung zwischen den negativen maskulinen Interaktionstermen zu

machtbezogenen Werten in Richtung Relevanz der Wichtigkeit zusätzlich zur Selbstbeschreibung, während sie gleichzeitig dann jedoch die nahezu einzigen erwähnenswerten Korrelationen darstellten.

Nachdem in Studie 5 bereits der Nachweis für die unterschiedliche Vorhersagefähigkeit der neuen Teilskalen hinsichtlich der Konstrukte eindrucksvoll erfolgte, wurden nun die Zusammenhänge der kombinierten Selbst- und Wichtigkeitsbewertungen der Geschlechtsrollenitems zu den Validierungskonstrukten überprüft. Dazu wurden in einzelnen Regressionsanalysen für jedes Konstrukt (z. B. beruflicher Ehrgeiz) zunächst in einem ersten Schritt die bereits bestätigten spezifischen Teilskalen aus der Selbstbeschreibung (z.B. Mas+ S) in das Modell aufgenommen. In einem zweiten Schritt folgten dann die jeweils entsprechenden, aus den Wichtigkeitseinschätzungen gemittelten Teilskalen (z. B. Mas+ W),; ehe im dritten und letzten Schritt schließlich die Interaktionsterme, ebenfalls pro Teilskala gemittelt über die Item-spezifischen Produkte aus Selbst- und Wichtigkeitsbewertung (Mas+ S\*W), hinzugefügt wurden.

In Tabelle 74 ist dieses Vorgehen am Beispiel des durch den AVEM erhobenen beruflichen Ehrgeizes aufgeführt. Erwartungskonform und aus Studie 5 bekannt leistete die positive maskuline Selbstbeschreibung einen entscheidenden Beitrag zur Varianzaufklärung des beruflichen Ehrgeizes. Je höher die Ausprägung der eigenen Zuschreibung positiver Maskulinität, desto höher der berichtete berufliche Ehrgeiz.

Tabelle 74. Vorhersage von berufl. Ehrgeiz durch individuelle Wertkomponente (Mas+)

Modellparameter Variable	B	SE B	$\beta$
Schritt 1			
Mas+ Selbst	.36	.06	.28**
Schritt 2			
Mas+ Selbst	.27	.07	.22**
Mas+ Wichtigkeit	.16	.08	.12*
Schritt 3			
Mas+ Selbst	.26	.07	.21**
Mas+ Wichtigkeit	.17	.08	.12*
IAT Mas+ S * W	-.06	.08	-.04

Anmerkung:  $R^2 = .08^{**}$ ; Schritt 2  $\Delta R^2 = .01^*$ , Schritt 3  $\Delta R^2 = \text{n.s.}$ ;  $**p < .001$ ;  $*p < .05$

Das Hinzunehmen der individuellen Wertkomponente ( $\beta = .12^*$ ,  $B = .16$ ;  $SE B = .08$ ,  $\Delta R^2 = .01$ ,  $p < .05$ ) konnte darüber hinaus weitere Varianz aufklären, wenngleich der Zugewinn des  $R^2$  gering ist. So stieg die Vorhersageleistung des beruflichen Ehrgeizes zwar minimal in Abhängigkeit der hohen zugeschriebenen Wichtigkeit von Mas+; es deutete sich aber bereits an, was der Interaktionsterm schlussendlich bestätigte: die Kombination aus Selbstzuschreibung und individuell bewerteter Wichtigkeit der einzelnen Items konnte die erwartete Vorhersageleistung auf das Validierungskonstrukt nicht nur nicht erhöhen, sondern verfehlte darüber hinaus deutlich die Signifikanz als Beleg des eigenen statistischen Beitrags in der Regressionsanalyse.

Die Betrachtung dieser Analysen für alle weiteren Validierungskonstrukte und deren respektiven Teilskalen (sowohl in Selbst- als auch Wichtigkeitsbewertung) ergab ähnliche Befunde. In Tabelle 75 sind exemplarisch alle Regressionsberechnungen in Kürze aufgeführt.

Den Teilskalen sind die entsprechenden Validierungskonstrukte sowie die schrittweise hinzugefügten Variablen in den hierarchischen Regressionsanalysen zugeordnet.

So ist in der ersten Zeile die bereits in Tabelle 74 abgebildete Analyse des beruflichen Ehrgeizes und Mas+ dargestellt (es werden nur signifikante Ergebnisse dargestellt). Aufgeführt werden die Veränderungen des  $R^2$ , sofern mindestens statistisch signifikant auf dem  $p < .05$  Niveau sowie die dazugehörigen signifikanten Beta-Gewichte des jeweils letzten statistisch signifikanten Modells. So lässt sich schnell überschauen, dass in den meisten Analysen (z. B. Fem+ und Empathie, Fem- in Bezug Ängstliche Depressivität) jeweils nur die entsprechende Selbstbeschreibung etwas zur Erklärung des Validierungskonstrukts beitragen konnte; Ergebnisse, die bereits aus dem destruktiven Testen der entsprechenden Teilskalen für die Konstrukte bekannt waren. Darüber hinaus gab es, wie am Beispiel des beruflichen Ehrgeizes und der positiven maskulinen Teilskala dargestellt, einige Konstrukte, bezüglich derer die Wichtigkeit einen eigenen Beitrag in der Vorhersageleistung aufweisen konnte. Auch in Bezug auf die emotionale Selbstwertschätzung (ESWS) konnte neben der Selbstbeschreibung ( $\beta = .56$ ,  $B = 1.06$ ,  $SE B = .07$ ,  $R^2 = .31$ ,  $p < .001$ ) die Wichtigkeit ( $\beta = -.07$ ,  $B = -.14$ ,  $SE B = .07$ ,  $\Delta R^2 = .04$ ,  $p < .05$ ) etwas von der Varianz aufklären.

Tabelle 75. Vorhersage der Validierungskonstrukte durch Selbst- und Wichtigkeitszuschreibung

	Zugewinn Selbst	Zugewinn Wichtigkeit	Zugewinn Produkt	$\beta_{\text{Selbst}}$	$\beta_{\text{Wichtigkeit}}$	$\beta_{\text{Produkt}}$
<b>Prädiktoren für MAS+</b>						
berufl. Ehrgeiz (AVEM)	$\Delta R^2 = .08$	$\Delta R^2 = .01$	-	.22	.12	-
Hoffen auf Erfolg (AMS)	$\Delta R^2 = .24$	-	-	.49	-	-
<b>Prädiktoren für MAS-</b>						
Aggressivität (AQ)	$\Delta R^2 = .10$	$\Delta R^2 = .01$	-	.27	.09	-
Werteskala Macht	$\Delta R^2 = .09$	$\Delta R^2 = .06$	$\Delta R^2 = .02^{39}$	.17	.28	.15
<b>Prädiktoren für FEM+</b>						
Empathie (Sorge)	$\Delta R^2 = .21$	-	-	.46	-	-
soziale Werte	$\Delta R^2 = .25$	$\Delta R^2 = .05$	-	.25	.33	-
<b>Prädiktoren für FEM-</b>						
ESWS	$\Delta R^2 = .31$	$\Delta R^2 = .01$	-	.57	-.07	-
Ängstl. Depressivität. (BSF)	$\Delta R^2 = .27$	-	-	.52	-	-

Das eigentliche Ziel, die Vorhersageleistung durch Hinzunahme des Produkts aus Selbstbeschreibung und Wichtigkeitsbeurteilung zu maximieren, konnte dagegen nicht erreicht werden. Lediglich für die machtbezogene Werteskala und Mas- ließ sich andeutungsweise die mögliche Relevanz einer individuell erfassten Wichtigkeitseinschätzung selbstbeschreibender Geschlechtsrollenattribute erahnen. So führte neben der negativen maskulinen Selbstbeschreibung ( $\beta = .30$ ,  $B = .29$ ,  $SE B = .06$ ,  $R^2 = .09$ ,  $p < .001$ ) und der entsprechenden individuellen Wichtigkeitsbewertung ( $\beta = .27$ ,  $B = .24$ ,  $SE B = .06$ ,  $\Delta R^2 = .06$ ,  $p < .001$ ) schließlich auch der Interaktionsterm aus beiden Komponenten ( $\beta = .15$ ,  $B = .17$ ,  $SE B = .07$ ,  $\Delta R^2 = .02$ ,  $p < .05$ ) zu einer, wenn auch geringen, signifikanten Erhöhung des  $R^2$ . Dieses Bild erfuhr jedoch Einschränkung durch die Tatsache, dass dieses Ergebnis lediglich auf die Einschätzungen seitens der Männer zurückging. Betrachtete man dagegen diese Analyse getrennt für Frauen, verschwand der Einfluss des Produkts.

<sup>39</sup> gilt nur für Männer

### 3.2.5.3.3 Gruppenvergleiche mittels Mediansplit zum Konzept Selbst \* Wichtigkeit

Während die Rolle der Wichtigkeit von Eigenschaften im Selbstkonzept aufgrund der Regressionsberechnungen gering scheint, blieb die Frage, inwieweit dieses Konstrukt lediglich einer Überarbeitung in Konzeption und Erfassung bedarf oder gänzlich aufgegeben werden sollte. In einem weiteren Ansatz sollte daher die Selbstzuschreibung in Abhängigkeit der Wichtigkeitszuschreibung untersucht werden.

Tabelle 76. Mediane aus Selbst- und Wichtigkeitszuschreibung (N = 800)

	MAS+	MAS-	FEM+	FEM-
Selbst	3.67	1.83	4.00	2.67
Wichtigkeit	3.83	1.50	4.33	1.83

Dazu wurden per Mediansplit-Methode Gruppen gebildet, die dadurch gekennzeichnet waren, dass ihre Einschätzungen in der Zustimmung der Items für das eigene Selbst und deren bewerteter Wichtigkeit für jede Teilskala ober- oder unterhalb des Medians lagen. Somit ergeben sich insgesamt 16 Gruppen: 4 pro Teilskala: eine kongruent hohe aus Selbst- und Wichtigkeitseinschätzung, eine kongruent niedrige, eine hohe Selbstzuschreibung / niedrige Wichtigkeit, eine niedrige Selbstzuschreibung / hohe Wichtigkeit.

In Tabelle 76 sind die Mediane pro Teilskala und in Abbildung 19 die prozentuale Verteilung der Probanden in Abhängigkeit ihrer Selbstzuschreibung und Wichtigkeitseinschätzung ober- vs. unterhalb des Medians für die vier Teiskalen dargestellt.

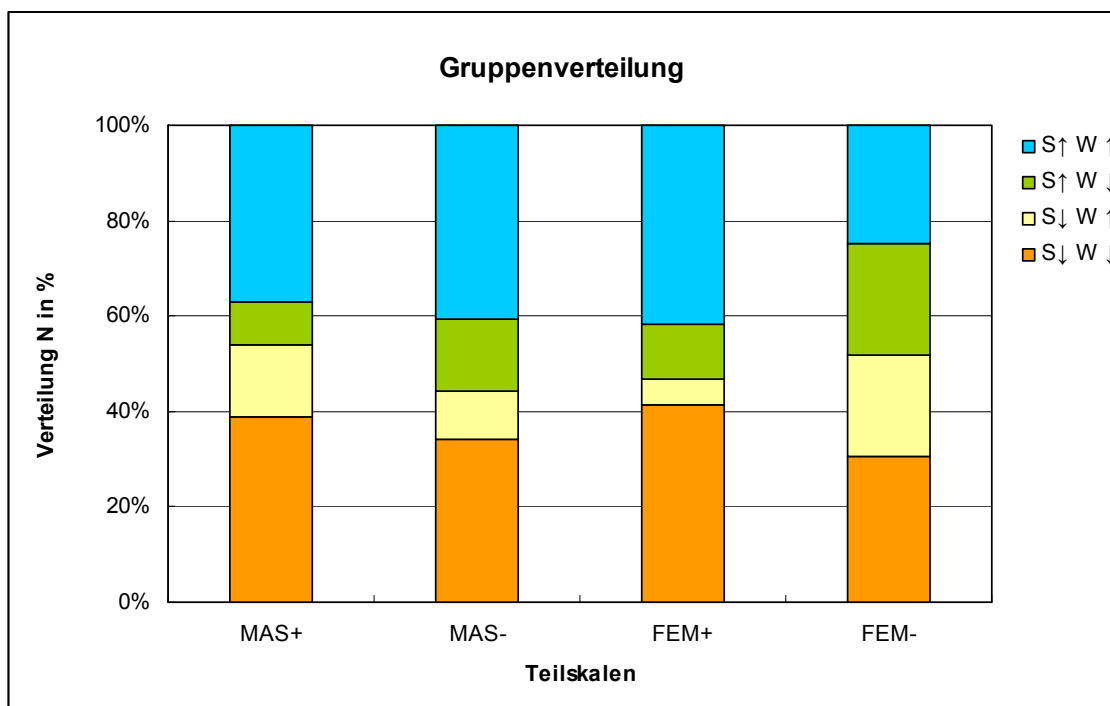


Abbildung 19. Verteilung der Probanden in % in Abhängigkeit der Selbst- und Wichtigkeitseinschätzung

So waren beispielsweise 272 Probanden (36%) dadurch charakterisiert, dass sie – bezogen auf positive Maskulinität – die Items sowohl stark für das eigene Selbst als zutreffend beanspruchten und gleichzeitig diesen Items auch eine hohe persönliche Wichtigkeit zuschrieben (blaue Markierung). Dagegen gab es 67 Probanden (9%), die sich zwar hoch auf den positiven

maskulinen Eigenschaften einschätzten, diesen Attributen jedoch eine vergleichsweise geringe subjektive Wichtigkeit für die eigene Selbstkonstruktion beimaßen (grüne Markierung). Auffällig in der Abbildung ist vor allem die Verteilung der Mittelkategorien, also solcher Personen, die eine inkongruente Selbsteinschätzung vornahmen, indem sie entweder bei gleichzeitig hoher Selbstzuschreibung der Attribute diesen eine geringe Wichtigkeit beimaßen (grün) oder im umgekehrten Fall, die Attribute für wenig selbstbeschreibend hielten, diese für das eigene Selbst jedoch sehr wichtig beurteilten (gelb). Während bei den Teilskalen Mas+, Mas- und Fem+ diese Gruppen kleiner (zwischen 17-25% gemessen an der jeweiligen Gesamtstichprobe der Teilskala) ausfielen als die mit kongruenten Zuschreibungen, stellte sich die Gruppenverteilung bezüglich der Einschätzungen für Fem- sehr viel homogener dar. Die Gruppenverteilung erfolgte mit 21 - 30% nahezu gleichmäßig.

Um den möglichen Einfluss der individuellen Wertkomponente näher zu beleuchten, wurden die Unterschiede der Mittelwerte auf den Validierungskonstrukten dieser Gruppen für die jeweiligen Teilskalen untersucht. Wichtig waren in diesem Zusammenhang vor allem die Unterschiede zwischen den Gruppen, deren Selbstbeurteilung in die gleiche Richtung geht, deren Wichtigkeitszuschreibung sich aber unterscheidet (z. B. Selbst ↑ und Wichtigkeit ↓ vs. Selbst ↑ und Wichtigkeit ↑). Es waren demnach vor allem solche Differenzen bedeutsam, in denen beide Gruppen hinsichtlich der Selbstbeschreibung gleichermaßen ober- vs. unterhalb des Medians lagen und ihre Verschiedenartigkeit bei der Beurteilung der Validierungskonstrukte ausschließlich auf gegensätzliche Wichtigkeitsbewertung der Eigenschaften zurückgeführt werden konnte. Leider ließ sich auch mithilfe dieses Ansatzes dieser Einfluss der Wichtigkeit nicht nachweisen. Dazu wurden ANOVAs mit der Gruppenzugehörigkeit (1 - 4) und den jeweiligen Validierungskonstrukten gerechnet. Es zeigte sich, dass ausschließlich die Selbstzuschreibung maßgeblich die Ausprägungen auf den Konstrukten beeinflusste, während die (womöglich davon unabhängig eingeschätzte) Wichtigkeit nicht zu den Unterschieden bei beiden Konstrukten beitragen konnte.

Eine naheliegende Vermutung bezieht sich auf den inhaltlichen Aspekt dieser als Wichtigkeit erfassten Wertkomponente und deren Nähe zu anderen, bereits in der Selbstbeschreibung enthaltenen Item-Parametern. Der Vorteil des vorliegenden Skalenkonstruktionsverfahrens liegt zum einen darin, dass für jedes Item verschiedene Parameter erfasst wurden (z. B. Typizität, Erwünschtheit, etc.) und zum anderen in der Möglichkeit, diese verschiedenen Parameter auf Ebene der Teilskalen miteinander zu vergleichen. Aufgrund der unterschiedlichen Stichproben und Antwortskalen (fünf- vs. siebenstufig) sind die Voraussetzungen für statistische Vergleiche nicht erfüllt, eine explorative Betrachtung der Mittelwerte offenbarte jedoch eine durchaus aufschlussreiche Erklärung.

Im Rahmen der Konstruktionsphase wurden für alle später ausgewählten Items sowohl Einschätzungen der Typizität als auch sozialen Erwünschtheit erfasst. Beide Aspekte waren Grundlage für die finale Auswahl der Items, so dass in den zur Selbstbeschreibung vorgelegten Items diese beiden Parameter integriert sein sollten. Der geringe bis sogar nicht vorhandene Einfluss der individuell eingeschätzten Wichtigkeit eben dieser Items lässt dann wiederum die

Spekulation zu, inwieweit diese Wichtigkeits-Einschätzung womöglich durch die eigene Positionierung auf diesen Items in der Selbstbeschreibung bereits repräsentiert sein könnte. In der Abbildung 20 sind die Parameter auf Skalenebene abgetragen. Jede Teilskala ist durch vier Parameter repräsentiert (Typizität, soziale Erwünschtheit, Selbstbeschreibung und Wichtigkeit), die über die jeweils sechs Items gemittelt wurden.

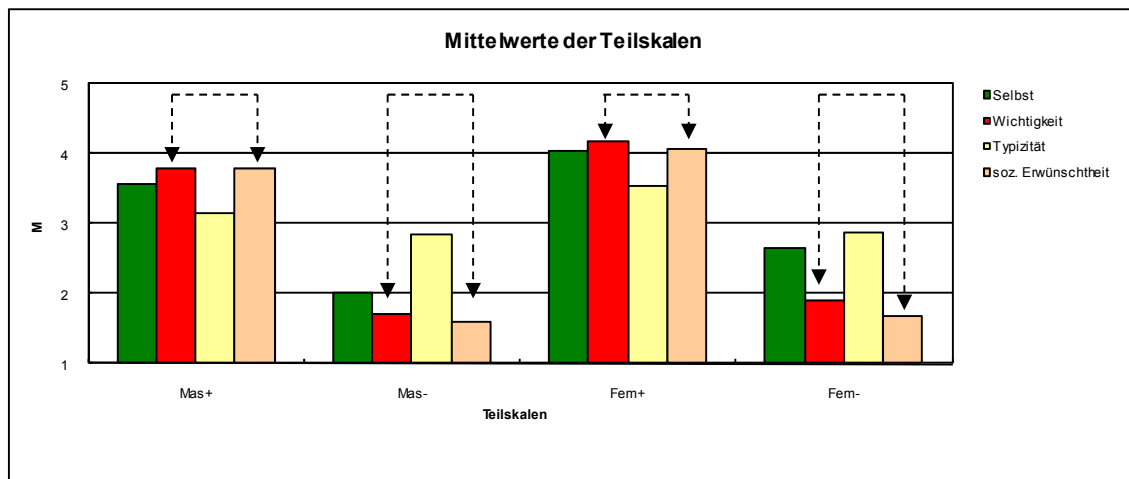


Abbildung 20. Vergleich der gemittelten 4-Item-Parameter aller vier Teilskalen

Die unterschiedlichen Antwortskalen wurden in der Darstellung angepasst. Es ist wichtig anzumerken, dass diese Parameter nicht alle von ein- und derselben Stichprobe stammen (der auftretende Bias durch diese Art der Erfassung wurde bereits diskutiert), sondern jeweils unterschiedliche Probanden diese Einschätzungen vornahmen.

Die Stichproben ähnelten sich zwar in ihren Geschlechts-, Alters- und Bildungsstrukturen, dennoch darf die Betrachtung über eine rein explorative, stark eingeschränkte Interpretation der hier dargestellten Werte nicht hinausgehen. Die vier Parameter der Teilskalen unterstützten die bereits formulierte Spekulation: während sich die Selbst- und Wichtigkeitseinschätzungen durchaus unterscheiden und damit doch eigene Aspekte des Selbstkonzepts erfassen, wird vor allem die große Ähnlichkeit der Wichtigkeitsbewertung der Items mit denen der sozialen Erwünschtheit deutlich.

Auf allen vier Teilskalen lagen die Unterschiede beider Mittelwerte zwischen .01 und .24. Somit scheint sich die Vermutung zu bestätigen, dass die Erfassung der Wichtigkeit oder zumindest die Art der individuellen Erfassung in der vorliegenden Studie, der der sozialen Erwünschtheit sehr ähnlich ist und dadurch durch die Teilskalen und deren ausgewählten Items per se repräsentiert ist.

Obwohl dieser Vergleich keiner statistischen Überprüfung unterzogen werden kann, fällt zudem die negative feminine Teilskala auf, deren Differenz zwischen Wichtigkeit und sozialer Erwünschtheit am größten ausfällt und somit zumindest für diese Teilskala Aspekte wie Wichtigkeit und Erwünschtheit vermutlich doch unterschiedliche Aspekte im Rahmen der Selbstzuschreibung bedeuten könnten.



### **3.2.5.4 Diskussion**

#### *Problematik der individuellen Wertkomponente*

Zusammenfassend muss die Schlussfolgerung gezogen werden, dass die Erfassung der individuellen Wertekomponente keinen entscheidenden eigenen Beitrag innerhalb des geschlechtsbezogenen Selbstkonzepts zu leisten scheint. Während offen bleibt, inwieweit dies tatsächlich der Art der Erfassung geschuldet sein mag, scheint die Frage, warum der Anteil dieser Wertkomponente so weit hinter den hypothetischen Erwartungen zurückbleibt, durch die abschließenden Analysen der Itemparameter beantwortet. Die Erfassung und Werte der Wichtigkeit sind denen der sozialen Erwünschtheit aus der Konstruktionsstudie sehr ähnlich und dadurch in den Teilskalen und deren ausgewählten Items per se schon repräsentiert. Die vorgenommene Selbst-Positionierung könnte diese Wichtigkeit bereits beinhalten, weshalb die zusätzliche, losgelöste Erfassung der Wichtigkeitskomponente keinerlei neue Information des Selbstkonzepts zu erfassen vermochte.

Darüber hinaus bemerkenswert ist auch, dass die Gruppenverteilung nach dem Mediansplit in Abhängigkeit von kombinierter Selbst- und Wichtigkeitszuschreibung gerade in den beiden gemischten Gruppen (Selbst hoch bei geringer Wichtigkeit, Wichtigkeit hoch bei geringer Selbstzuschreibung) geringer ausfiel als in den beiden Gruppen, in denen die Items sowohl nach Selbst als auch Wichtigkeit übereinstimmend beansprucht oder abgelehnt wurden. Während sich dieses Bild für die Teilskalen MAS+, MAS- und FEM+ ähnelte, fiel auf, dass bezüglich der negativen femininen Teilskala das Gruppenverhältnis über alle vier Ausprägungskombinationen nahezu gleich ausfiel. Es scheint hinsichtlich dieser Teilskala entweder einfacher zu sein, Items für sich zu beanspruchen und sie dennoch für weniger wichtig zu halten wie dementsprechend Items für das eigene Selbst als bestimmend abzulehnen und ihnen dennoch eine große Wichtigkeit zuzuschreiben. Auf der anderen Seite jedoch könnte die gleichmäßige Verteilung in die vier entstehenden Gruppen die Schwierigkeiten der negativen femininen Teilskala für das eigene Selbstkonzept verdeutlichen. Es ist denkbar, dass die gelöste Zuschreibung von Selbst und Wichtigkeit für Probanden problematisch ist, da diese Teilskala als gemeinhin am wenigsten geschätzt gilt und Männer und Frauen sich möglicherweise schwer tun, eine klare Position diesbezüglich zu beziehen, was die erhöhte Anzahl der Probanden in den beiden Mischgruppen erklären könnte.

## 4 DISKUSSION

In dem vorliegenden Diskussionskapitel sollen zunächst die Fragestellung sowie die Ergebnisse kurz aufgegriffen und zusammengefasst werden. In einem anschließenden, in mehrere Abschnitte strukturierten Teil werden dann einzelne Aspekte der Konstruktion und Validierung der neuen Teilskalen herausgegriffen und vor dem Hintergrund der bestehenden Literatur kritisch diskutiert.

Ziel dieser Studie war die Entwicklung und Validierung eines neuen Instruments mit sowohl positiven als auch negativen Eigenschaften zur Erfassung des Geschlechtsrollenselbstkonzepts bei Erwachsenen. Gegenüber bestehenden Instrumenten wie dem BSRI (Studie 1 und 2) oder ad-hoc zusammengestellten Skalen hat das neue Inventar eine Vielzahl von Vorteilen:

- (A) In der Arbeit wurde elaboriert der theoretische Hintergrund auch negativer Aspekte des geschlechtsbezogenen Selbstkonzepts diskutiert und die Notwendigkeit eines Inventars, welches diese negativen Komponenten berücksichtigt, herausgearbeitet.
- (B) Das vorliegende Instrument basiert auf geschlechtstypischen Eigenschaften, die durch den mehrstufigen Auswahlprozess sowohl die aktuellen Vorstellungen als auch das aktuelle Vokabular dessen abbilden, was im deutschsprachigen Raum als typisch für Männer und Frauen gilt (Studie 3 und 4).
- (C) Das Inventar erfasst dabei Selbstbeschreibungen bezüglich sowohl positiver als auch negativer Attribute, die theoretisch nachweislich Teil des Geschlechtsrollenselbstkonzepts sind.
- (D) Der Fragebogen verzichtet bewusst auf Wortgruppen und besteht nur aus einzelnen Adjektiven.
- (E) Zu allen gewählten Attributen liegen eine Vielzahl von Itemparametern aus mehreren Studien vor (u. a. Typizität, Erwünschtheit, Selbstbeschreibung, Wichtigkeit, Zufriedenheit), die die Grundlage der Itemauswahl bildeten.
- (F) Die vier neuen Teilskalen wurden mittels einer großen Zahl an sehr unterschiedlichen psychologischen Erlebens- und Verhaltenskonstrukten validiert, die die Unabhängigkeit der Skalen belegen und diese für einen Einsatz in der Forschung empfehlen.
- (G) Es liegt zusätzlich bereits eine Replikation der Validierung vor, die sowohl gefundene Ergebnisse (wie beispielsweise die Unabhängigkeit der Teilskalen) bestätigt, als auch die Relevanz der Messung des Geschlechtsrollenselbstkonzepts auch durch nun negative Eigenschaften in Bezug auf weitere Konstrukte unter Beweis stellt.

In den Studien 1 und 2 wurden in Analysen zur Typizität, sozialen Erwünschtheit und der Selbstbeschreibung zunächst die Grenzen des BSRI aufgezeigt, die eine der Ausgangsbedingungen für die Entwicklung des neuen Instruments bildeten.

Im Zuge dieses Entwicklungsprozesses kann zunächst festgestellt werden, dass die für das Inventar ausgewählten Eigenschaften in der Generierungs- und Überprüfungsphase mit hoher Übereinstimmung von beiden Geschlechtern als jeweils entsprechend typisch männlich bzw.

typisch weiblich bewertet wurden (Studie 3 und 4). Die Einschätzungen bezüglich der Typizität erfolgten zwecks Vermeidung einer künstlichen Polarisierung aufgrund des Vergleichsmaßstabes nicht gleichzeitig für beide Geschlechter, sondern für sich allein genommen, was strengere Maßstäbe an die Auswahl der Attribute anlegte.

Schließlich wurden in Studie 4 erste Daten zur Selbstbeschreibung erfasst, um die aufgrund der Auswahlkriterien festgelegten Items hinsichtlich ihrer Skalenzugehörigkeit und Faktorenstruktur überprüfen zu können. Alle 24 Items zeigten zufriedenstellende Faktorladungen auf und die entsprechenden Teilskalen konnten durch sehr gute interne Konsistenzen bestätigt werden. Konkurrierend geprüfte konfirmatorische Modelle wiesen das hypothetisierte Vierfaktorenmodell (entsprechend der vier Teilskalen) als das am besten Repräsentierende aus. Besonders vor dem Hintergrund, dass jede Teilskala aus lediglich 6 Items bestand, um ein möglichst ökonomisches Inventar zu entwickeln, waren diese ersten Ergebnisse sehr motivierend. Ein innovativer Aspekt des neuen Instruments bestand darin, aus Gruppen- und Selbstforschung bekannte negative Eigenschaften in Bezug auf das geschlechtsbezogene Selbstkonzept zu berücksichtigen. Die grundsätzliche Annahme bestand darin, dass sich Individuen auch hinsichtlich normativ unerwünschter Attribute klar positionieren können, was durch die Ergebnisse bestätigt wurde: geschlechtstypische negative Eigenschaften wurden eher zur Selbstbeschreibung herangezogen, und im stärkeren Maße, als die dem anderen Geschlecht entsprechenden negativen Eigenschaften – zumindest zeigte sich dies so bei Frauen.

Die Konstruktvalidität wurde schließlich in zwei weiteren Studien (Studie 5 und Replikationsstudie 6) durch das Einbeziehen einer Vielzahl von psychologischen Konzepten überprüft. Jeder Teilskala wurden Konstrukte zugeordnet, denen bereits durch vorliegende Forschungsergebnisse relevante Beziehungen zum biologischen Geschlecht und dem Geschlechtsrollenselbstkonzept nachgewiesen wurden. So gelten beispielsweise Aggressivität und Macht als Eigenschaften, die Männern im Allgemeinen und der maskulinen (negativen) Geschlechtsrolle im Speziellen zugeordnet werden können. Gleichermaßen sind leistungsbezogene Parameter wie beruflicher Ehrgeiz, Hoffen auf Erfolg und das Leistungsstreben ebenfalls mit dem maskulinen Geschlechtsrollenkonzept assoziiert, jedoch mit den eher positiven Eigenschaften. Ähnliches lässt sich für die Femininität finden. So gelten beispielsweise Empathie und sozial ausgerichtete Eigenschaften (Werte und Orientierungen) als primär mit der femininen Geschlechtsrolle verbundene Konstrukte und im Speziellen dabei mit deren positiven Aspekten. Demgegenüber können als Vulnerabilitätsfaktoren geltende Konzepte wie ängstliche Depressivität und problematische emotionale Selbstwertschätzung sehr eindeutig der negativen femininen Teilskala zugewiesen werden. Diese hypothetisierten Zusammenhänge ließen sich für alle berichteten Konzepte und ihren respektiven Teilskalen in Regressionsanalysen bestätigen. Darüber hinaus konnte der Ansatz des destruktiven Testens (Anderson & Anderson, 1996), der die Stärke einer solchen Verbindung zwischen zwei Konstrukten durch das sukzessive Hinzufügen möglicher konkurrierender Konzepte prüft, die Relevanz der neuen negativen Teilskalen zusätzlich bestätigen. Die Ergebnisse der Analysen

belegen eindeutig, dass es sich bei den negativen Teilskalen nicht einfach um eine Umkehr ihrer respektiven positiven Gegenstücke handelt, wodurch deren Erfassung redundant wäre. Stattdessen zeigte das destruktive Testen die Unterschiedlichkeit der vier Teilskalen sowohl bezüglich der Typizität als auch ihrer Valenz. Alle vier Teilskalen konnten erfolgreich mit sehr unterschiedlichen Konzepten in Verbindung gebracht werden, was ihre inhaltliche und methodische Vielfalt unterstreicht.

Im Folgenden werden die Befunde im Einzelnen einer kritischen Diskussion unterzogen:

- 4.1 Notwendigkeit eines neuen Instruments: Die Grenzen des BSRI
- 4.2 Entwicklung und Validierung des PoNeGRI - Befunde zur Neukonstruktion
- 4.3 Generalisierbarkeit
- 4.4 Einschränkungen, methodenkritische Aspekte und Ausblick

#### **4.1 Notwendigkeit eines neuen Instruments: Die Grenzen des BSRI**

Der BSRI hat den scheinbaren Vorteil, dass dieses Instrument auch Dekaden nach der Entwicklung gute interne Konsistenzen aufweist und zwischen den Geschlechtern zu trennen vermag; auch hinsichtlich spezifischer Validierungskonstrukte. Im Zuge dieser Argumentation werden jedoch wichtige Aspekte außer Acht gelassen. So umspannt der BSRI durch die große Itemanzahl eine Vielzahl unterschiedlicher Eigenschaften, womit statistisch per se die Beziehung zu mehreren Validierungskonstrukten leichter hergestellt und auch überschätzt werden kann. Jackson (1985) argumentierte, dass überlappende Iteminhalte aufgrund der daraus resultierenden gemeinsamen Varianz eine Validierung problematisch gestalten könnten. Die durch die schiere Menge der BSRI-Eigenschaften leichter hergestellten Beziehungen zu Validierungskonstrukten verschleiern wiederum die eigentlichen Mängel dieses Instruments und erschweren nicht nur deren Nachweis, sondern vor allem die Erkenntnis der Notwendigkeit eines solchen Nachweises.

Auch die guten Reliabilitäten sind nicht zuletzt der großen Anzahl der Items geschuldet, wodurch vernachlässigt wird, dass einige Items nicht mehr den ursprünglichen Konstruktionsansprüchen genügen. So wiesen viele Studien auf Itemebene Veränderungen dessen nach, was als Maskulinität und Femininität definiert wird. Durch die anhaltende Erhebung mittels des BSRI mit zwei resultierenden Skalenwerten wird dieser Umstand jedoch ignoriert. In der vorliegenden Arbeit wurden nicht nur Selbstbeschreibungsmaße mit dem BSRI erhoben sondern auch der Konstruktionsprozess repliziert.

In einem ersten Teil der Arbeit wurde der Frage nachgegangen, inwieweit die Items des BSRI auch unter heutigen Gesichtspunkten hinsichtlich statistischer Gütekriterien wie Reliabilität und Validität noch Teil eines Instrumentes zur Erfassung des Geschlechtsrollenselbstkonzepts sein können. Vor dem Hintergrund vielfach diskutierter Studien aus der Vergangenheit wurde erwartet, dass einige Items die ursprünglich definierten Voraussetzungen für eine Skalenzugehörigkeit nicht mehr erfüllen und aufgrund ihrer Typizitäts- und Erwünschtheitseinschätzungen aus der Skala ausgeschlossen werden müssten. In zwei Studien wurden mehr als 2.000 Probanden gebeten, die Originalitems einer Überprüfung zu

unterziehen. In Anlehnung an die Entwicklung von Bem (1974) und der deutschen Version durch Schneider-Düker und Kohler (1988) wurde in Studie 1 das gleiche Vorgehen (Skalen, Instruktion) zur Erfassung der sozialen Erwünschtheit angewandt. So galt ein Item dann als für die entsprechende Skala geeignet, wenn es unabhängig vom Geschlecht des Beurteilers als signifikant erwünschter für das eine als das andere Geschlecht beurteilt werden würde. Dazu sollten Probanden zunächst die dargebotenen Eigenschaften dahingehend einschätzen, wie sozial (un-)erwünscht diese für einen Mann oder eine Frau in unserer Gesellschaft sind. Zusätzlich wurden in dieser ersten Studie weitere Personen um Bewertungen der Typizität (ebenfalls für jeweils Mann und Frau getrennt befragt) besagter Eigenschaften gebeten. In einer zweiten Studie wurden Männer und Frauen schließlich hinsichtlich ihrer Selbstbeschreibungen mittels dieser Items befragt.

Wie aus anderen Studien bekannt, warf die Überprüfung des BSRI Zweifel an der Eignung einzelner Eigenschaften und damit ihrer respektiven Skalen zur Erfassung des GSK auf. Generell zeigte zunächst der querschnittliche Vergleich mit den Daten von Bem (1974) und Schneider-Düker und Kohler (1988) was auch Twenge (1997; 2001) oder Holt und Ellis (1998) bereits kritisierten: während die soziale Erwünschtheit der maskulinen Eigenschaften für Männer und der femininen Eigenschaften für Frauen auch mehr als 30 Jahre nach der ersten Entwicklung nahezu konstant blieb, stiegen die Einschätzungen der Eigenschaften für die jeweiligen anderen Geschlechter; der Unterschied in den Erwünschtheitsratings wird kleiner. So zeigte sich, dass sowohl die femininen Eigenschaften für Männer als auch die maskulinen Eigenschaften für Frauen annähernd gleichermaßen erwünscht waren wie für die jeweils andere Gruppe. Generell muss bei den Schlussfolgerungen bezüglich der gesellschaftlichen Veränderungen jedoch berücksichtigt werden, dass die vorliegenden Ergebnisse nicht aus einer wiederholten Überprüfung der Stichproben aus den jeweiligen Entwicklungsstudien stammen, sondern lediglich einen querschnittlichen Vergleich erlauben, für den unterschiedliche demografische Charakteristiken der jeweiligen Stichproben ebenfalls in Betracht gezogen werden sollten.

Des Weiteren bestätigte sich ebenfalls, dass die femininen Eigenschaften in ihren Bewertungen insgesamt als weniger sozial erwünscht eingeschätzt wurden als die maskulinen Items und dies von beiden Geschlechtern gleichermaßen. Auch bemerkenswert sind jedoch in diesem Zusammenhang die Beurteilungen der Frauen, die die maskulinen Eigenschaften als annähernd gleich sozial erwünscht für Frauen einschätzen wie für Männer. Frauen, noch eher als Männer, spiegeln in ihren Ratings die bereits von Twenge (1997; 2001) festgestellte Annäherung der sozialen Erwünschtheit femininer und maskuliner Eigenschaften wider, da sie sowohl die maskulinen Eigenschaften für sich selbst als auch die femininen Eigenschaften für Männer für sozial erwünschter halten als die Männer dies taten. Unterstützt wurde dieses Ergebnis durch die Einschätzungen hinsichtlich der Typizität. So wurde Maskulinität insgesamt als typischer für Frauen eingeschätzt als für Männer. Der von Twenge (1997; 2001) beschriebene gesellschaftliche Wandel schien sich zu bestätigen: Maskulinität, so wie durch den BSRI erhoben, ist mittlerweile auch für Frauen attraktiv.

Auffällig war die fehlende Korrelation zwischen Maskulinität und Femininität, was trotz andersartiger Ergebnisse und diesbezüglicher Kritik (Blanchard-Fields et al., 1994) doch für die von Bem postulierte Unabhängigkeit der Skalen spricht. Als bemerkenswert dagegen erwiesen sich Veränderungen im Ausmaß der Erwünschtheitszuschreibungen von Maskulinität und Femininität. In früheren Studien von Bem (1974) und Schneider-Düker und Kohler (1988) waren Mittelwertsunterschiede von mindestens zwei Skalenpunkten in den Einschätzungen der Erwünschtheit für Männer und Frauen festzustellen. In der aktuellen Stichprobe hatte sich diese Differenz auf Unterschiede von kleiner als eins reduziert, was eine verringerte Geschlechterstereotypisierung andeutet. Holt und Ellis (1998) hatten diesen Effekt nicht nur ebenfalls bereits nachweisen können, sondern zudem spekuliert, inwieweit eine Replikation in 20 Jahren zu dem Ergebnis kommen könnte, dass der BSRI nicht länger als valides Messinstrument gelten kann.

Die Analysen auf Itemebene brachten hervor, dass einige der jeweiligen skalenspezifischen Attribute nach diagnostischen Gesichtspunkten gar aus den Skalen ausgeschlossen werden müssten. Die Probleme zeigten sich bei Items, die nicht den erwarteten Geschlechtsunterschied in der Beurteilung evozierten (z. B. maskuline Items *trete bestimmt auf, ehrgeizig, sicher*), die unzureichende Faktorladungen auf ihren respektablen Teilskalen (feminine Items *nachgiebig, bescheiden und liebe Sicherheit*) oder Ladungen auf der jeweils anderen Teilskala (*glücklich*) aufwiesen, Mittelwerte beider Geschlechter unterhalb des Skalenmittelwerts aufzeigten sowie über mehrere Kriterien hinweg den Auswahlprozess nicht bestanden (*abhängig; ein ursprünglich feminines Item, welches als typischer und sozial erwünschter für Männer gehalten wurde*). Verschiedene Autoren (Auster & Ohm, 2000; Sethi & Allen, 1984) hatten ähnlich detaillierte Itemanalysen vorgenommen und so Items identifiziert, die keine andauernde Berechtigung in den finalen Skalen hätten.

In einer zweiten Studie im Rahmen der Überprüfung des BSRI wurden Probanden gebeten, sich auf den Items des BSRI per Selbstbeschreibung zu positionieren. Erwartet wurde, dass im Zuge des interindividuellen Vergleichs Frauen höhere Femininitätswerte und Männer höhere Maskulinitätswerte im Vergleich zur jeweils anderen Gruppe zeigen. Im intraindividuellen Vergleich sollten Frauen dann höhere Selbsteinschätzungen auf der Femininitätsskala im Vergleich zur Maskulinitätsskala vornehmen, während sich Männer höher auf der Maskulinitätsskala im Vergleich zur Femininitätsskala einschätzen sollten. Wie bei den Einschätzungen zu der sozialen Erwünschtheit und der Typizität – entsprechend der eigentlichen Voraussetzungsprüfung – offenbarten auch die Ergebnisse der Selbsteinschätzungen von Frauen und Männern weitere Mängel. Zunächst ließ sich erwartungsgemäß feststellen, dass sich Männer maskuline Eigenschaften stärker zuschrieben als Frauen dies taten, während Frauen feminine Eigenschaften im stärkeren Maße für sich beanspruchten. Im gewissen Sinne lässt sich somit durchaus ein kontinuierlicher Einfluss traditioneller Geschlechtsrollen nachweisen, was die Relevanz der Erfassung des geschlechtsbezogenen Selbst für die Forschung unterstreicht. Auch Auster und Ohm (2000) mittels Typizitätseinschätzungen und Blanchard-Fields et al. (1994) mittels Selbstbeschreibungsmaße hatten bereits den anhaltenden Einfluss des GSK – trotz nachgewiesener Mängel des BSRI – für das eigene Selbst diskutiert.

Bedeutsam sind vor allem auch die Ergebnisse von Auster und Ohm (2000) bezüglich der Frage, wie Probanden gern wären, die eine annähernd gleiche Bedeutsamkeit maskuliner und femininer Eigenschaften aufzeigten. Während anhand dieser Resultate zwar keine Aussagen darüber getroffen werden können, inwieweit Probanden diese Eigenschaften auch tatsächlich zeigen, demonstrieren sie dennoch die Selbstverständlichkeit, mit der Männer und Frauen geschlechtsbezogene Eigenschaften zur Selbstbeschreibung heranziehen. Trotz der scheinbaren gesellschaftlichen Veränderungen demonstrieren diese Ergebnisse die Beständigkeit traditioneller Geschlechterstereotype.

Während grundsätzlich die Erhebung des GSK nicht in Frage gestellt werden muss, zeigten die Analysen der Selbsteinschätzungen dennoch Unzulänglichkeiten der einzelnen Skalen auf. So war festzustellen, dass sich die feminine Skala im Vergleich zu den maskulinen Items als weniger homogen erwies. Gleichzeitig wurde sie jedoch am stärksten zur Selbstbeschreibung herangezogen, was die ursprünglich deklarierte unabhängige Gleichberechtigung beider Teilskalen in Frage stellt. Bemerkenswert war vor allem, dass der von Twenge (1997) berichtete Anstieg der Maskulinitätswerte von Frauen, der sich in den Einschätzungen der sozialen Erwünschtheit noch hatte demonstrieren lassen, in den Selbstbeschreibungen nicht in der Form bestätigt werden konnte. So lag der durchschnittliche Wert zwar oberhalb des Skalenmittelwerts, stellte insgesamt aber die am geringsten ausgeprägte Selbstzuschreibung dar. Auch die bereits aus den Voraussetzungsprüfungen auffällige Eigenschaft *abhängig* zeigte bei den Analysen auf Itemebene weitere Schwächen, da es im Zuge der Selbstbeschreibung nicht als ausschließlich positiv deklariert werden konnte. Eine durchgeführte konfirmatorische Faktorenanalyse bestätigte überdies, was die Überprüfung der Auswahlkriterien bereits angedeutet hatte. Die Daten erlaubten in dieser Form keinen Modell-Fit. Vor allem die feminine Teilskala fiel durch unzureichende Faktorladungen auf. An dieser Stelle lässt sich ihr Anteil am unzureichenden Modell-Fit nur spekulieren.

Obwohl die Ergebnisse der Selbstbeschreibung nicht gegen einen grundsätzlichen Einsatz des BSRI sprachen, so legten sie doch eine klare Überarbeitung nahe. Zudem wiesen die Items klare Mängel hinsichtlich ihrer sozialen Erwünschtheit auf und die Typizität, die Bem nicht einmal als klares eigenes Kriterium festgelegt hatte, kann nachweislich nicht per se als gegeben vorausgesetzt werden. Aufgrund der vorliegenden ausführlichen Datenbasis musste der Schluss gezogen werden, dass die deutsche Übersetzung des BSRI dem heutigen Zeitgeist und dem, was gesellschaftlich als maskuline und feminine Werte definiert ist, nicht mehr entspricht. Für zukünftige Einsätze eines solchen Fragebogens kann das einfache Eliminieren einiger Items nicht die angestrebte Lösung sein. Eine Neuentwicklung wurde somit nicht nur wünschenswert sondern notwendig. Vor allem auch vor dem Hintergrund, dass selbst die Erfassung positiver Aspekte nicht zufriedenstellend erfolgte, und der Nicht-Berücksichtigung jeglicher negativer Attribute, deren Anteil am Selbst jedoch ausgiebig nachgewiesen wurde, mussten Items gesucht werden, die aktuelle Vorstellungen von Maskulinität und Femininität widerspiegeln. Mit Bezug auf aktuelle Forschungsergebnisse und unter Berücksichtigung der Konstruktion des geschlechtlichen Selbst mussten im Rahmen dieser Neukonstruktion Items

gefunden werden, die sowohl die positiven als auch negativen Aspekte der Geschlechtsrolle abzubilden vermögen, um so die Vorhersagefähigkeit dieses Teils der Selbstbeschreibung zu maximieren.

## **4.2 Entwicklung und Validierung des PoNeGRI - Befunde zur Neukonstruktion**

### **4.2.1 Identifizierung positiver und negativer Attribute im GSK**

Die erste Phase der Skalenkonstruktion, in der Probanden gebeten wurden, Attribute und Beschreibungen von Männern und Frauen frei zu produzieren, sollte nicht nur sicherstellen, dass ausschließlich aktuelle, im zeitgemäßen Sprachgebrauch repräsentierte Eigenschaften das Ausgangsmaterial bilden, sondern dass darüber hinaus lediglich die Items ausgewählt wurden, die von beiden Geschlechtern gleichermaßen genannt wurden. So sollte bereits auf dieser Prozessstufe das Kriterium der Typizität als Maßstab angelegt werden. Gerade der Aktualität und dem Verständnis der Eigenschaften wurde große Bedeutung zuteil, hatten doch Antill et al. (1981) in ihren Studien feststellen müssen, dass Schülerinnen und Schüler 25% der 512 ihnen präsentierten Items nicht verstanden hatten, was für Selbstbeschreibungsmaße problematisch ist. Auch Perkins und Forehand (2006) hatten die Relevanz selbst gewählter gegenüber einer vorgegebenen Gruppe von Eigenschaften bereits in den Fokus gerückt. So konnten sie die Tendenz einer stärkeren Zuschreibung von Eigenschaften, die den selbst generierten semantisch ähneln, nachweisen.

Somit stand am Ende dieser Phase eine Menge von ca. jeweils 20 positiven und negativen sowohl maskulinen wie auch femininen Eigenschaften, die das Ausgangsmaterial für die Neukonstruktion der Skala zur Erfassung der positiven und negativen Konstrukte des GSK darstellen. In Anlehnung an die soziale Identitätstheorie (Tajfel & Turner, 1979; 1986) sowie Studien mit ähnlicher freier Produktion von Beschreibungen (Helgeson, 1994a) erwies sich auch in der vorliegenden Studie das Generieren positiver Eigenschaften als grundsätzlich einfacher im Vergleich zu negativen Eigenschaften. Dabei zeigten sich Frauen insgesamt als kreativer und produktiver in ihren Beschreibungen als Männer, was die höhere Anzahl maskuliner Items erklären könnte. An dieser Stelle muss auch die Frage gestellt werden, inwieweit Frauen selbstkritischer sind, da sie nahezu allein für die Menge der generierten typisch negativen Eigenschaften für ihre Gruppe verantwortlich waren.

Dem Anliegen, die Teilskalen so divers wie möglich zu gestalten und sie dennoch auf wenige Items beschränken zu können, wurde entsprochen, indem Probanden selbst Eigenschaften generierten, die als typisch für Männer und Frauen in dieser Gesellschaft gelten. Während man durch dieses Vorgehen in der Tat sicherstellen konnte, dass gegebene Attribute sowohl aktuell als auch Bestandteil des aktiven Sprachschatzes sind, schränkte es auf der anderen Seite die Gesamtmenge der Items von vornherein ein. Zum Einen konnte bei der freien Generierung die im späteren Verlauf zu vermeiden versuchte Vergleichsperspektive nicht verhindert werden: Probanden sahen auf dem Bildschirm die freien Felder für die Geschlechter, die sie jeweils mit positiven und negativen typischen Beschreibungen füllen sollten. Dies könnte



Erwartungen und direkte Vergleiche evoziert haben. Als eine Schwierigkeit der Itemsauswahl zeigten sich daher bereits an dieser Stelle unerwünschte inhaltliche Überschneidungen der Teilskalen. Als eine weitere Herausforderung galten in diesem Prozess die durch die gewünschte Attributen-basierte Skala verbundenen Anforderungen und Beschränkungen an die deutsche Sprache. So mussten beispielsweise „wahre“ Adjektive vs. Partizipien und substantivierte Adjektiven unterschieden werden (z. B. *intrigierend* vs. *intrigant*). Weiterhin wurden in der freien Produktion Items genannt, deren Endsilbe „-los“ (z. B. *teilnahmslos*, *verantwortungslos*, *emotionslos*) oder das Präfix „un-“ (z. B. *ungeduldig*, *unberechenbar*, *undiszipliniert*) bereits prinzipiell aufzeigen, dass es eigentlich um das Fehlen einer Eigenschaft geht, die ihrerseits wieder Bestandteil einer anderen Teilskala sein könnte. In diesem Zusammenhang stellte sich die Frage, inwieweit negative Eigenschaften mitunter lediglich das Fehlen der Eigenschaft einer anderen positiven Teilskala bedeuten könnte.

#### *Soziale Erwünschtheit und Typizität*

Um das Evozieren einer Vergleichsperspektive, die daraus resultierende Akzentuierung der Geschlechterdifferenzen und damit eine Überschätzung der Geschlechterunterschiede zu vermeiden (Doise, Deschamps, & Meyer, 1978), schätzten Probanden die Typizität von Eigenschaften für jeweils nur ein Geschlecht ein; obwohl es nachweislich schwierig ist, die Salienz weiblicher und männlicher Kategorien zu vermeiden (Hardie & McMurray, 1992). Die Frage, die sich hier auch stellte, hing direkt mit der Fragestellung nach der Typizität selbst zusammen. In der vorliegenden Studie wurden die Versuchspersonen instruiert, Eigenschaften dahingehend einzuschätzen, wie typisch diese für einen Mann oder eine Frau in unserer Gesellschaft sind: „*Wie typisch ist es für einen Mann (eine Frau) ... zu sein?*“. Entscheidend sind in diesem Fall die Formulierung und das Vorgeben des Rahmens innerhalb dessen die Probanden ihre Einschätzungen vornehmen sollen. So bleibt offen, inwieweit bestimmte Formen der Einschätzungen von der gefühlten Kompatibilität mit der eigenen Geschlechtergruppe abhängen, was sowohl für Typizität als auch soziale Erwünschtheit gilt. Möglich wären dazu Studien, die genau dies in Beziehung zueinander setzen: die Einschätzung der eigenen Geschlechtsrolle sowie das Beurteilen der sozialen Erwünschtheit derselben Eigenschaften. Während sich durch diese Art der Erfassung das Heranziehen einer Vergleichsperspektive nicht vermeiden lässt, gibt diese Strategie genau die Möglichkeit mit einer bereits vorliegenden Versuchsgruppe die Einschätzungen der sozialen Erwünschtheit von eigenen Ratings unbeeinflusst mit jenen zu vergleichen, zu denen man die zusätzlichen Informationen der Selbsteinschätzung hat. So könnte der Anteil dessen, was man in den allgemeinen Einschätzungen der sozialen Erwünschtheit zu projizieren vermag, aus der eigenen Position dieser Items heraus gerechnet werden. Bedeutsam hierzu sind die Ergebnisse aus dem Versuch, eine individuelle Wertekomponente zu integrieren und der Erkenntnis, dass die individuellen Wichtigkeitszuschreibungen denen der sozialen Erwünschtheit ähneln. Obwohl diese Ergebnisse aus unterschiedlichen, jedoch vergleichbaren, Stichproben stammen, lässt sich diese Ähnlichkeit nicht verleugnen.

In Konstruktionsprozessen wie der vorliegenden Studie wird in der Regel eine Reihe von Auswahlkriterien angelegt, um Items zu finden, die dem Forschungsvorhaben entsprechen. Gerade bezüglich von Instrumenten, die attributbasiert sind, birgt dieses Vorgehen einige Gefahren. Zieht man beispielsweise die Trennschärfe als Kriterium heran, bedeutet dies lediglich, dass für eine homogene Skala gesorgt wird. Betrachtet man vor diesem Hintergrund die femininen Teilskalen aus anderen Instrumenten genauer und insbesondere die positive Femininität, fällt auf, was man bereits aus anderen Instrumenten weiß – die Skala beinhaltet eine Vielzahl emotional besetzter Items. In der vorliegenden Studie wurden viele Kriterien formuliert, die sowohl Ergebnisse anderer Studien berücksichtigten, als auch gleichzeitig für ein maximal heterogenes Instrument sorgen sollten und dennoch ließ sich diese Emotionalität als vorherrschendes latentes Thema nicht gänzlich vermeiden. Vor diesem Hintergrund stellt sich daher die Frage, was mit der Teilskala tatsächlich gemessen wird und inwieweit, wenn nicht gänzlich ein anderes Konstrukt, dann doch wenigstens maximal ein Ausschnitt des Geschlechtsrollenselbstkonzepts abgebildet wird. Diese Argumentation ist nicht neu und gliedert sich in die Debatte um die Mehrdimensionalität des Geschlechtsrollenselbstkonzepts (Ashmore, 1990; Athenstaedt, 2003; Baber & Tucker, 2006; Connell & Messerschmidt, 2005; Deaux & LaFrance, 1998; Deaux & Lewis, 1984; Deaux & Major, 1987; Eagly & Kite, 1987; Feather, 1984; Freeman, 1987; Helgeson, 1994a; Huston, 1983; Jackson & Cash, 1985; Koestner & Aube, 1995; Marsh & Myers, 1986; Myers & Gonda, 1982a; Oswald & Lindstedt, 2006; Ricciardelli & Williams, 1995; Ruble & Martin, 1998; Spence, 1993; Spence, 1999; Spence & Buckner, 2000; Spence & Sawin, 1985; Taylor & Hall, 1982; Twenge, 1999).

Während diese Diskussion weniger als eine Einschränkung der Teilskalen selbst zu verstehen ist, muss anerkannt sein, dass ein attributbasiertes Instrument, sofern es den grundlegenden diagnostischen Kriterien der Testkonstruktion entsprechen soll ohne deren Voraussetzungen zu verletzen, generell nur jeweils Ausschnitte des GSK berücksichtigen kann. Dennoch ist festzuhalten, dass am Ende des Prozesses zur Einschätzung der Typizität, Erwünschtheit und erster Ergebnisse zur Selbstbeschreibung ein Instrument mit 24 Eigenschaften steht, die mit hohem Konsens über beide Geschlechter als jeweils typisch für Männer oder für Frauen eingeschätzt wurden, und entsprechend der Vorgabe hinsichtlich ihrer Valenz, eindeutig ober- (positive Items) oder unterhalb (negative Items) des Skalenmittelwertes lagen. Somit war die Voraussetzung für den Einsatz des neuen Instruments zur Messung des Geschlechtsrollenselbstkonzepts in einer Validierungsstudie erfüllt.

#### **4.2.2 Struktur des neuen Instruments**

Die konfirmatorische Faktorenanalyse bestätigte einen für die entwickelten Teilskalen äußerst zufriedenstellenden 4-Faktoren-Modell-Fit. Bezüglich der Verhältnisse der Teilskalen untereinander zeigt sich weitgehend das hypothetisierte Bild. So erwiesen sich jeweils die Beziehungen der gegensätzlichen Teilskalen (Mas+ und Fem- sowie Mas- und Fem+) als die am stärksten negativen im Modell. Eine Aufgabe im Rahmen der Entwicklung der Teilskalen bestand vor allem darin die entgegengesetzten Teilskalen nicht lediglich nur eine Umkehr ihrer Gegenstücke sein zu lassen; negative feminine Items dürften demnach weder eine negativ

umformulierte Variante der positiven maskulinen, noch ein Fehlen eben solcher positiver maskuliner Eigenschaften darstellen. Mit nachgewiesenen Korrelationskoeffizienten von  $-.23$  und  $-.24$  ist dies geglückt. So zeigten die gegensätzlichen Teilskalen insgesamt im Modell zwar die stärksten und auch vermuteten negativen Beziehungen zueinander auf, was jedoch zum einen im Einklang mit früheren Arbeiten ist und zum anderen im Rahmen bleibt, da weniger als 6% der Unterschiede in der einen Merkmalsausprägung durch die Unterschiede in den anderen Merkmalen erklärt werden können. Somit leisten alle vier Teilskalen ihren eigenen erklärenden Beitrag innerhalb der Geschlechtsrolle einer Person. Hinzu kommt, dass die Höhe der Koeffizienten bisweilen sogar geringer ausfiel als die aus anderen Arbeiten: so fanden Choi und Kollegen (Choi et al., 2007) beispielsweise eine Korrelation von  $r = .34$  zwischen Maskulinität und Femininität, was insgesamt schon immer kritisch als lediglicher Nachweis der gemeinsamen sozialen Erwünschtheit diskutiert wurde.

Die Interkorrelationen gerade der gegensätzlichen Teilskalen lagen damit über denen, wie sie beispielsweise schon mit den EPAQ berichtet wurden. So fanden Evans und Dinning (1982) Korrelationskoeffizienten zwischen der positiven maskulinen und den negativen femininen Teilskalen von  $F_{va}$   $r = -.07$  und  $F_c$   $r = -.05$  sowie zwischen der negativen maskulinen und der positiven femininen Teilskala von  $r = -.02$ , die die Unabhängigkeit der Teilskalen eindrucksvoll untermauert. Nimmt man jedoch eine Vergleich mit den Originaldaten von Spence und Helmreich (Spence et al., 1979) vor, stellt man fest, dass deren berichtete Interkorrelationen (M+ mit  $F_{va}$   $r = -.28$  und M+ mit  $F_c$   $r = -.38$ ) viel eher denen der vorliegenden Studie ähneln. Während bisweilen die Klarheit der orthogonalen Abgrenzung der einzelnen Faktoren in Abhängigkeit von Zeit und Stichprobe Schwankungen unterliegen mag, wird dennoch deutlich, dass die Beziehung der jeweils umgekehrten Skalen zueinander niemals eine perfekte sein wird, was die Bedeutung der negativen Teilskalen als eigenständige Informationsquellen hervorhebt.

Im Zuge der Selbstbeschreibung wird zudem die Stärke der negativen Eigenschaften unterstrichen. Zwar zeigte sich, wie auch schon bei Krahe et al. (2007) für Jugendliche als auch Athenstaedt (2003) mit dem EPAQ für Erwachsene, ein grundsätzlicher Effekt der Valenz: positive Eigenschaften wurden als typischer für das eigene Geschlecht betrachtet als negative Attribute. Gleichzeitig stellte sich jedoch auch heraus, dass bei der Einschätzung negativer Items geschlechtskongruente Eigenschaften stärker zur Selbstbeschreibung herangezogen werden als negative Eigenschaften, die dem anderen Geschlecht entsprechen. So schrieben sich Frauen negative feminine Eigenschaften im stärkeren Maße zu als negative maskuline. Lediglich bei den Männern wich dieses Bild von der Erwartung ab. In beiden Validierungsstudien beanspruchten die Männer die negativen femininen Eigenschaften im stärkeren Maße zur Selbstbeschreibung als die negativen maskulinen. Während dieser Effekt nur in einer der beiden Studien signifikant war, deutet er einen wichtigen Aspekt der Femininität an. Obwohl insgesamt betrachtet maskuline Eigenschaften häufig positiver eingeschätzt und favorisiert werden (Long, 1991), zeigte sich ähnlich wie in anderen Studien (Schwartz et al., 1997) in der vorliegenden Arbeit, dass die femininen Eigenschaften insgesamt sozial erwünschter

erscheinen und womöglich weniger bedrohlich für das eigene Selbstkonzept sind, so dass sich Männer selbst die negativen femininen Eigenschaften eher zuschrieben als die negativen Charakteristiken, die typischer für das eigene Geschlecht sind.

Dass Frauen gleichzeitig wiederum grundsätzlich geschlechtskongruente Eigenschaften unabhängig von ihrer Valenz stärker zur Selbstbeschreibung heranzogen, könnte für eine unbeliebtere Position maskuliner negativer Eigenschaften sprechen, die zwar typisch für das männliche Geschlecht sind, gleichermaßen aber am wenigsten sozial erwünscht und daher nicht individuell präferiert – selbst nicht von der eigenen Gruppe der Männer. Unterstützt werden die Ergebnisse aus den Analysen zur sozialen Erwünschtheit, die die geringere Erwünschtheit der sechs maskulinen Items im Vergleich zu den femininen nachweist. In diese Ergebnisse reihen sich auch die Resultate der Typizitätseinschätzungen der jeweils 6 Items beider Teilskalen ein. Verglichen zu den negativen maskulinen Eigenschaften liegen die Mittelwerte der Typizitätsbewertung der negativen femininen Items bemerkenswerterweise nicht unterhalb, sondern im Durchschnitt über beide Geschlechter oberhalb des Skalenmittelwertes, was an dieser Stelle die Frage offen lässt, inwieweit hier bereits ein gewisser Anteil der Itemvalenz diese Einschätzungen beeinflusste. So könnte sich die verringerte soziale Erwünschtheit der maskulinen Attribute womöglich bereits in den geringen Typizitätseinschätzungen widerspiegeln, auch wenn diese eindeutig als typischer für Männer als für Frauen gelten, während die unerwünschten femininen Eigenschaften zwar auch charakteristischer für Frauen sind, insgesamt jedoch womöglich wesentlich weniger sozial unerwünscht sind als die maskulinen negativen Items.

Ähnlich wie auch Bems – ursprünglich ausschließlich positive – Skalen dahingehend kritisiert wurden, dass sie nicht gleichermaßen sozial erwünscht sind, erfahren auch die Ergebnisse der neuen Teilskalen eine kleine Einschränkung dahingehend, dass es nicht gänzlich geglückt ist für eine gleich ausgeprägte Erwünschtheit der Teilskalen zu sorgen. Teils kann diese Einschränkung jedoch dahingehend wieder aufgehoben werden, dass es – und dies im Unterschied zu Bem – nun vier Teilskalen gibt, die eine variable Positionierung der Individuen in Abhängigkeit von Typizität und Erwünschtheit ermöglichen. Jackson (1985) hatte vor dem Hintergrund ihrer eigenen Studie und der Herstellung einer Beziehung zwischen Geschlechtsrolle und salienten Dimensionen des eigenen Selbst bereits argumentiert, dass überlappende Iteminhalte aufgrund der daraus resultierenden gemeinsamen Varianz zu einem Problem werden könnten. Während dieser Aspekt vor allem bei der Validierung und den dort zum Einsatz gekommenen Konstrukten Berücksichtigung finden muss, ist auch ein Bezug zu den vier Teilskalen des GSK unerlässlich und es war festzustellen, dass der individuelle Beitrag aller vier Teilskalen eindrucksvoll nachgewiesen werden konnte.

Die geringen Korrelationen zwischen den positiven und negativen Teilskalen innerhalb einer Dimension (z. B. Fem+ und Fem-) zeigen dann auch, dass die negativen Teilskalen eigenständige Informationen des individuellen GSK erfassen und dass die Ausprägung der negativen Selbstbeschreibungen nicht aus den Zuschreibungen der positiven per Umkehrschluss abgeleitet werden können. Diese Interpretation wird unterstrichen durch

Befunde aus den Validierungsanalysen, die die Bedeutsamkeit aller vier Teilskalen bezüglich ganz eigener Validierungskonstrukte eindrucksvoll unter Beweis stellen.

### *Negative Femininität*

Betrachtet man besonders die negative feminine Teilskala, muss man anerkennen, dass sie die am wenigsten homogene der vier Teilskalen ist. Während dies zum einen zugunsten der inhaltlich vielschichtigeren Ausgestaltung der Skala geht und damit die Konzentration auf lediglich eine, womöglich durch verschiedene Formulierungen unterschiedlich ausgedrückte Eigenschaftsgruppe (z. B. ängstlich, wachsam, besorgt, furchtvoll, vorsichtig) verhindert werden sollte, überraschen die schwachen Skalenkennwerte vor dem Hintergrund früherer Arbeiten wenig. In einer Vielzahl von Studien (Athenstaedt et al., 2004; Athenstaedt et al., 2008; Diekman & Eagly, 2000; Helmreich et al., 1981; McCreary & Korabik, 1994; Runge et al., 1981; Russell & Antill, 1984; Wilde & Diekman, 2005) wurden der (zumeist positiven) Femininität wiederholt schlechte oder zumindest die schwächsten Reliabilitäten im Vergleich zu den anderen gemessenen Teilskalen bescheinigt. Andere Studien konnten diese Schwächen auch der negativen Femininität nachweisen. So berichteten Helmreich und Kollegen (Helmreich et al., 1981) Cronbachs Alphas der femininen Teilskala für Männer und Frauen zwischen lediglich .41 und .46, McCreary und Korabik (1994) berechneten ein Cronbachs Alpha von .42, Wilde und Diekman (2005) kamen auf .65 im Vergleich zu .71, .74 und .82 der anderen Spence-Teilskalen und auch Helgeson und Fritz (1999) fanden im Vergleich zu den anderen Teilskalen für Unmitigated Communion nur die schlechtesten Reliabilitäten: .69 bis .76. Ricciardelli und Williams (1995) diskutierten zwar in ihrem Artikel die Validität und Reliabilität der femininen negativen Teilskala, geben selbst jedoch die interne Konsistenz dieser Teilskala in ihrer Stichprobe nicht an. Gleichzeitig müssen sie jedoch einschränken, dass der Zusammenhang der sozial unerwünschten femininen Attribute zu Verhaltenskonstrukten kein klares Bild zu zeichnen vermag, was zumindest im weitesten Sinne Einschränkungen dieser Teilskala vermuten lässt. Auch in Bereichen, in denen andere als reine Selbstbeschreibungsmaße zum Einsatz kamen, stellt sich das Feminine als insgesamt statistisch und methodisch weniger eindeutig und deutlich heraus: so konnte Lippa durch seinen speziellen diagnostischen Ansatz (Lippa & Connelly, 1990) beispielsweise 92% der Männer und dennoch „nur“ 88% der Frauen korrekt klassifizieren.

Weiterhin ist vermutlich die soziale Erwünschtheit der Eigenschaften ein Kriterium, welches nicht notwendigerweise für eine starke negative feminine Teilskala sorgt. Zwar gilt dieses auch für die negativen maskulinen Eigenschaften, die ebenso sozial unerwünscht sein sollen. Es muss jedoch berücksichtigt werden, dass die Maskulinität als solche das attraktivere Konzept zu sein scheint. Twenge (Twenge, 1997; Twenge & Im, 2007) hat nicht nur als eine der ersten auf die veränderten Geschlechtsrollen über die Zeit hinweg aufmerksam gemacht, sondern auch nachweisen können, dass es vor allem Veränderungen hinsichtlich der Maskulinität sind, die zur Annäherung von Frauen und Männern in den Ausprägungen ihres Geschlechtsrollenselbstkonzepts beitragen. In der heutigen Gesellschaft ist es zeitgemäß und

notwendig, maskuline Eigenschaften ins Selbst integriert zu haben. Dies gilt nicht nur für Männer und Frauen, sondern vor allem auch unabhängig davon, ob diese Eigenschaften positiv oder maskulin sind. Mag es vor dem Hintergrund einer möglichen gesellschaftlichen Bewertung sicherlich vorteilhaft sein, über positive maskuline Eigenschaften zu verfügen, sind es vermutlich nicht zuletzt auch gerade die negativen maskulinen Eigenschaften, die zum eigenen Erfolg beitragen mögen. So ist bisweilen ein gewisses Maß an Aggressivität unerlässlich, um in der beruflichen Welt voranzukommen und Erfolg zu haben.

Ein anderes Bild ergibt sich jedoch hinsichtlich der negativen femininen Eigenschaften. Sie mögen gleichermaßen sozial unerwünscht sein wie die negativen maskulinen – auf einer ausschließlich methodischen Ebene im Rahmen der Itemauswahl. Geht es jedoch darum, sich diese Eigenschaften aktiv selbst zuzuschreiben, mag erst das wahre Ausmaß an sozialer Unerwünschtheit zum Vorschein kommen. Die Theorie zur Sozialen Identität postuliert, dass Annahmen und Überzeugungen bezüglich einer Gruppe (z. B. Männer und Frauen) in Abhängigkeit von der Assoziation mit eben einer dieser Gruppen in das eigene Selbstkonzept integriert werden. Dabei sind diese Überzeugungen in nur geringer Weise objektiven Gesichtspunkten zugrunde gelegt, sondern es hängt viel stärker von subjektiven Einschätzungen ab, inwieweit beispielsweise die Zuschreibung zu bestimmten Gruppen das eigene Selbst auch aufwerten kann. So besteht zwar für Männer und Frauen die Tendenz, sich vorwiegend stereotype Eigenschaften des eigenen Geschlechts zuzuschreiben, nicht jedoch die Notwendigkeit. Ausgehend von Studien, die diskutieren, wie sich soziale Erwünschtheit von maskulinen Eigenschaften in den vergangenen Jahren verändert hat (Twenge, 1997), wird die Misere der negativen femininen Eigenschaften deutlich. Sie entsprechen sozusagen in zwei Dimensionen nicht dem gesellschaftlichen Ideal: sie sind nicht maskulin und nicht positiv. Sich diese Eigenschaften verstärkt zuzuschreiben, ist nicht ohne Risiko. Männer können auf ihren respektablen Geschlechtsrollen (Mas+, Mas-) Stellung beziehen, ohne das Gesicht zu verlieren oder gesellschaftliche Abwertung zu erfahren. Für Frauen gestaltet sich dies etwas schwieriger. Während Sie auf der einen Seite eine Positionierung auf der negativen femininen Teilskala ohne Weiteres vornehmen können, gilt es gleichzeitig im gewissen Maße, das eigene Selbst (und den Selbstwert) zu schützen.

So könnten negative maskuline Eigenschaften Personen womöglich widerstandsfähiger für Herausforderungen machen und unter Umständen gar ein Schutzschild gegen Angriffe darstellen (z. B. *rücksichtslos*, *angeberisch*, *überheblich*), so dass diese nicht unvermittelt wirken können. Im Vergleich dazu bleiben Nutzen und Gewinn negativer femininer Eigenschaften fragwürdig. *Orientierungslos*, *ängstlich* und *zweifelnd* beispielsweise bereiten niemanden auf schwierige Situationen vor, vielmehr machen sie einen Menschen angreifbar und schwach. Dies erklärt zum einen, warum vor allem Zusammenhänge zu Validierungskonstrukten aus dem psychischen Gesundheitsbereich gefunden werden und dies nahezu ausschließlich. Zum anderen könnte es erklären, warum Skalen mit Items dieser Art so viel weniger reliabel sind als vergleichbare andere Skalen – sie sind in einem stärkeren Maße von der Situation abhängig. Betrachtet man die negativen maskulinen Eigenschaften kommt man

verhältnismäßig einfach zu dem Schluss: man hat sie oder man hat sie nicht. Auch Frauen wurde ein Anstieg in den selbst zugeschrieben maskulinen Maßen nachgewiesen (Diekmann & Eagly, 2000; Twenge, 1997), auch wenn ihnen von der Gesellschaft eine Abwertung droht, beispielsweise im Führungsbereich (Carli & Eagly, 1999; Eagly, Makhijani, & Klonsky, 1992).

Einen aufschlussreichen Ansatz zur Klärung der Frage, was das Diagnostische im femininen Bereich anscheinend so viel anspruchsvoller und inkonsistenter macht, lieferte Helgeson (Helgeson, 1994a). In ihrer Studie fand sie auf Nachfrage, an wen Probanden bei den Einschätzungen von Männern und Frauen dachten, beachtenswerte Ergebnisse. So zogen 79% der Versuchspersonen bei der Bewertung maskuliner Personen auch tatsächlich einen prototypischen Mann zum Vergleich heran, während nur 47% der Versuchspersonen bei der Evaluation femininer Personen auch wirklich an eine Frau als Maßstab dachten. Überhaupt scheint die Prototypikalität von Männern und damit Maskulinität – berücksichtigt man, dass die biologische Unterscheidung zwischen Menschen die Definitionsgrundlage der Geschlechtsrolle ist – klarer und damit leichter abrufbar zu sein. Die Studie von Williams und Best (1990) offenbarte zwischen verschiedensten Ländern mehr Ähnlichkeiten als Unterschiede bei der Beschreibung des Typischen von Männern und Frauen. Das Bedeutsame an den Ergebnissen ist jedoch, dass in allen Kulturen sechs Attribute als zutreffend für Männer (z. B. *dominant, independent, strong*) identifiziert wurden, während es lediglich drei für die Frauen waren (*sentimental, submissive, superstitious*). Während zugegebenermaßen der Unterschied in der Anzahl nicht überwältigend ist, gilt es hier jedoch zu berücksichtigen, dass dies für alle befragten Kulturen gilt und als Hinweis gelten könnte, wie unterschiedlich stark das typisch Männliche und damit Maskulinität in den Kognitionen projiziert ist und damit leichter abrufbar im Vergleich zur Femininität.

Zusammenfassend bleibt festzuhalten, dass Femininität insgesamt und vor allem die neue Erfassung ihrer negativen Aspekte, komplex in Aufbau, Selbstzuschreibung und Interpretation ist. Für zukünftige Studien gilt es dies bei Erhebung und Auswertung nach wie vor zu berücksichtigen, um Fehlinterpretationen und Ableitungen bezüglich dieser Teilskala zu vermeiden.

#### 4.2.3 Konstruktvalidierung

Ein zentrales Ziel dieser Studie bestand darin, die Notwendigkeit einer Diskriminierung der Geschlechtsrollenmessung in positive wie auch negative Teilskalen zu demonstrieren. Es galt zu zeigen, dass negative Aspekte des GSK eine eigenständige Beziehung zu verschiedenen psychologischen Konstrukten aufwiesen, die über eine bloße Umkehrung der positiven Aspekte hinausgingen. So mussten Russell und Antill (1984) feststellen, dass beispielsweise das Konzept Selbstbewusstsein zwar positiv mit der hypothetisierten positiven maskulinen Teilskala korrelierte, gleichzeitig und nahezu gleichermaßen negativ mit der negativen femininen Teilskala. Solch ein Ergebnis deutet die Austauschbarkeit und damit Redundanz der negativen Teilskalen an und stellt ein Problem für die Forschung dar, in der es schlussendlich darum geht, mit möglichst wenigen Mitteln Untersuchungen durchzuführen. Das „Mitziehen“ unnötiger

negativer Eigenschaften, die keinen eigenen varianzaufklärenden Beitrag leisten, gehört sicher nicht dazu.

Während bereits die Ergebnisse aus der Selbstbeschreibung für die eigenständige Bedeutung negativ bewerteter Aspekte des GSK sprachen, erfuhr diese Interpretation durch die Validierungsanalysen Bestätigung. Die positiven und negativen Teilskalen von Maskulinität und Femininität waren nachweislich in sehr unterschiedlicher Weise mit den Validierungskonstrukten verbunden. In zwei Studien zeigte sich, wie vorhergesagt, dass die negativen maskulinen Eigenschaften in engerer Beziehung zu Aggressivität (erhoben über zwei unterschiedliche Maße) und machtbezogenen Werten standen als die positiven Aspekte der Maskulinität. Ob sich eine Person als typisch maskulin im negativen Sinne betrachtet, spielt für die Ausprägung der eigenen Aggressivität eine größere Rolle als vergleichbar ausgeprägte positive maskuline Eigenschaften. Dementsprechend war die Beziehung negativer femininer Ausprägung zu Ängstlicher Depressivität, Trait-Angst und Problematik der emotionalen Selbstwertschätzung enger als die positiver Femininität. Analog zu der Validierung der negativen Teilskalen wurde auch der individuelle Zusammenhang positiver Maskulinität und Femininität zu spezifischen psychologischen Konstrukten überprüft und nachgewiesen. So zeigte sich ebenfalls, dass die positiven Attribute der Maskulinität in engerer Beziehung zu beruflichem Ehrgeiz, Hoffen auf Erfolg und Leistungsorientierung standen als die negativen Aspekte. Für die positive feminine Teilskala ließ sich dieser eigenständige Zusammenhang zu den Konzepten Empathie im Sinne von Sorge, Ausprägung sozialer Werte und Fürsorglichkeit nachweisen.

Somit lassen sich die vier neu entwickelten Teilskalen als zuverlässiges, valides und vor allem kurzes Instrument zur Erfassung des Geschlechtsrollenselbstkonzepts bei Erwachsenen betrachten, welches in vielen Studien einsetzbar sein dürfte. Gerade die zusätzliche Berücksichtigung negativer Aspekte der Maskulinität und Femininität eröffnet eine Vielzahl von Möglichkeiten, auch den Bezug zu negativ ausgeprägten psychologischen Konzepten herzustellen.

#### *Wertkomponente*

Ein zusätzlicher Aspekt der vorliegenden Arbeit bestand darin, die potentielle Bedeutung der subjektiven Wichtigkeit, die einem Attribut jenseits der Selbstzuschreibung beigemessen wird, zu untersuchen. Der allgemeinen Forschung bezüglich der Wichtigkeit ist insgesamt nur wenig Aufmerksamkeit zuteil geworden und dennoch ist sie weder gänzlich neu noch unumstritten (Ajzen, 2001; Fishbein & Ajzen, 1975; Fishbein & Middlestadt, 1995; Jaccard, Brinberg, & Ackerman, 1986; Palmgreen & Rayburn, 1982; Rosenberg, 1956; van der Pligt & de Vries, 1998). Die Ergebnisse der vorliegenden Arbeit sprechen nicht dafür, die Annahme einer Wichtigkeitskomponente im Geschlechtsrollenselbstkonzept zu stärken. Aus den vorliegenden Daten muss die Schlussfolgerung gezogen werden, dass die Erfassung der individuellen Wertekomponente die Erwartungen bezüglich eines eigenen Beitrags zum geschlechtsbezogenen Selbstkonzept nicht erfüllen kann. Die gemittelten Produkte aus Selbstbeschreibung und Wichtigkeitszuschreibung pro Teilskala konnten keine nennenswerten



Korrelationen mit den entsprechenden Validierungskonstrukten aufweisen und klärten nichts an zusätzlicher Varianz.

Damit reihen sich die Ergebnisse in eine Liste von Studien ein, die ebenfalls Schwierigkeiten mit der Messung von dem Konzept Wichtigkeit hatten. So berichteten Fishbein und Ajzen (1975) von Studien, in denen die Zunahme von Wichtigkeit als zusätzlicher Faktor im Einstellungsmodell die Vorhersage von Verhalten nicht verbesserte, sondern sogar unter Umständen die Vorhersagekraft der Einstellung einschränkte. Auch Palmgreen und Rayburn (1982) kamen in ihrer Studie zur Gratifikation durch Fernsehkonsum in Abhängigkeit der Bewertung und der Wichtigkeit zu dem Schluss, dass die zusätzliche Aufnahme von Wichtigkeit in ihr Analysemodell wenig zur Vorhersageleistung beizutragen vermochte. Fishbein und Ajzen (1975) argumentierten diesbezüglich, dass die durch Individuen eingeschätzte wahrgenommene Wichtigkeit vermutlich wenig Ähnlichkeit mit empirisch abgeleiteten Gewichtungen in einer so klar definierten Gleichung wie die der Einstellungsmessung hat. Beide gingen davon aus, dass die relative Wichtigkeit von Attributen sich wenig bis gar nicht in Gewichten einer multiplen Regressionsrechnung widerspiegeln werden, da diese einen objektiven Wichtigkeitsindex darstellen, der jedoch wenig über die Kausalität aussagt. Demnach wäre es beispielsweise unangemessen, im Fall einer hohen Korrelation ohne weiteres auf eine wichtige Determinante rückzuschließen und bei einer niedrigen Korrelation davon auszugehen, dass diese Erwartung keine wichtige Determinante sei. Fishbein und Ajzen (1975) lehnten daher auch die statistische Methode (Korrelation von Erwartungseinschätzung eines jeden Attributs mit der allgemeinen Einstellung) zur Gewinnung von Wichtigkeitswerten ab.

Diese Einwände schwächen die Position der Wichtigkeit in der Geschlechtsrollenerfassung als eigenen Faktor, da anscheinend ähnliche Wirkmechanismen der Faktoren zueinander zu vermuten sind. Eine Ursache in der wenig erfolgreichen Integration der Wichtigkeitskomponente als eigenen varianzaufklärenden Faktor im geschlechtsbezogenen Selbstkonzept könnte im Umgang mit eben dieser Wichtigkeit liegen. Während ein Großteil der Studien ein globales Maß von Wichtigkeit in das Modell aufnimmt, liegt der Schlüssel dieser Variable womöglich weniger in der Eigenständigkeit als Faktor, sondern vielmehr in der Möglichkeit die wichtigen Einstellungen von den weniger wichtigen zu separieren und sie dann getrennt in Bezug auf das Verhalten zu betrachten. Van der Pligt und de Vries (1998) hatten genau dies versucht, indem sie die Wichtigkeitskomponente in einem Einstellungsmodell bezüglich des Rauchens implementiert haben. Neben einem Einstellungsmaß sowie der Bewertung einiger Aussagen über negative Konsequenzen des Rauchens sollten die jungen Probanden zusätzlich die drei wichtigsten Aussagen auswählen. Die Autoren konnten zeigen, dass die Vorhersagekraft der als am wichtigsten erachteten Einstellungen höher ist als die weniger wichtigen Einstellungen, was die separate Erfassung der Wichtigkeit im Rahmen der Einstellungs- und vor allem auch der Interventionsforschung durchaus rechtfertigt. Darüber hinaus waren durch diese Schwerpunktsetzung in den Einstellungen die Unterschiede zwischen den Gruppen der Raucher und Nichtraucher stärker ausgeprägt, während diese gänzlich verschwanden, sobald die weniger wichtigen Einstellungen berücksichtigt wurden.

Eine Möglichkeit, diesen Ansatz auf das GSK zu übertragen und die Wichtigkeitskomponente noch nicht aufzugeben, könnte womöglich auch darin liegen, aus den Teilskalen die jeweils für das Selbstkonzept wichtigsten Items durch die Probanden identifizieren zu lassen und deren Beitrag hinsichtlich möglicher Validierungskonstrukte zu überprüfen.

Darüber hinaus bleibt jedoch grundlegend die Frage, inwieweit der fehlende Nachweis für den Einfluss der Wichtigkeit der Art der Erfassung geschuldet sein mag. Eine abschließende Betrachtung der gemittelten Itemparameter (soziale Erwünschtheit, Typizität, Selbstbeschreibung, Wichtigkeit) pro Skala lieferte dabei einen entscheidenden Hinweis. Die Erfassung und Mittelwerte der Wichtigkeit pro Skala – so wie in dieser Studie umgesetzt – sind denen der sozialen Erwünschtheit aus der Konstruktionsstudie sehr ähnlich und dadurch möglicherweise in den Teilskalen und deren ausgewählten Items per se schon repräsentiert. Die vorgenommene Selbst-Positionierung könnte diese Wichtigkeit bereits beinhalten, weshalb die zusätzliche, losgelöste Erfassung der Wichtigkeitskomponente keinerlei neue Information des Selbstkonzepts zu erfassen vermochte. Auch Jackson (1985) hatte in ihrer Studie hohe Korrelationen zwischen Maskulinität und Femininität mit der jeweils eingeschätzten Wichtigkeit dieser Charakteristiken nachgewiesen und damit die Debatte initiiert, inwieweit Aspekte der Selbstbeschreibung und Wichtigkeit überhaupt unabhängig voneinander existieren und erfasst werden können.

Des Weiteren wäre ebenfalls möglich, dass die gemessene Wichtigkeit prinzipiell einen individuellen Valenzaspekt erfasst, der aber womöglich mit den Validierungskonstrukten insgesamt wenig in Zusammenhang steht, da in diesen Werten keine Valenzaspekte stecken. Als Beispiel bedeutet dies, eine Person kann emotional sein und das auch wichtig finden, aber es beeinflusst nicht notwendigerweise ihre Ausprägung auf Empathie – es sei denn man würde auch hier den Valenzaspekt kennen und einschätzen lassen können.

Zusammenfassend bleibt die Implementierung von Produktberechnungen in der Geschlechtsrollenmessung nicht unmöglich, sehr wohl aber problematisch. Zukünftige Studien mit möglichen unterschiedlichen methodischen Ansätzen könnten untersuchen, ob die Vorhersage der Valenzkonstrukte möglicherweise erhöht werden kann, wenn gezielter zwischen den Gruppen niedriger und hoher Wichtigkeitszuschreibungen unterschieden wird, wie auch in anderen Studien bereits umgesetzt (Pahl & Eiser, 2005). Erste Analysen diesbezüglich in der vorliegenden Studie bestätigen solche Überlegungen jedoch nicht.

## **4.3 Generalisierbarkeit**

### **4.3.1 Onlineerhebung**

Methodisch stellt sich die zentrale Frage, ob es eine Grundlage für den Verdacht gibt, dass ein mithilfe von Onlineerhebungen entwickeltes und erprobtes Inventar auf diesen Erhebungsmodus und die Stichprobe beschränkt und für Offlineerhebungen nicht ohne weiteres einsetzbar sei.

Unabhängig von ihren unterschiedlichen Umsetzungen und Darstellungsformen werden online erhobene Daten im Vergleich zu Offlinestudien ständig einer kritischen Überprüfung unterzogen (Buchanan, 2001; Buchanan, 2002; Knapp & Kirk, 2003; Rietz & Wahl, 2002; Tuten

et al., 2002). Zusammenfassend kommen diese Studien zum Urteil, dass Onlineforschung dem Vergleich mit offline erhobenen Daten sehr wohl standhält und keine nennenswerten Unterschiede in den Ergebnissen hervorbringt (Erdman et al., 1983; Gosling, Vazire, Srivastava, & John, 2004; Knapp & Kirk, 2003; Morrison et al., 1999; Rietz & Wahl, 2002; Schillewaert & Meulemeester, 2005; Skinner & Allen, 1983). Vorteile finden sich vor allem in größeren, demografisch heterogenen Online-Stichproben (Buchanan, 2001; Musch & Reips, 2000) sowie in technischen Aspekten, wie *Automatisierung* von Prozeduren und Analysen, oder der *Flexibilität der Darstellung*. Gefundene Ergebnisse aus Studien zu Persönlichkeitskonstrukten erwiesen sich in ihren statistischen Kennwerten als vergleichbar und zufriedenstellend, sowohl bezogen auf interne Konsistenzen oder Faktorenstrukturen (Pasveer & Ellard, 1998), Ähnlichkeiten in den gefundenen Konstrukten (Schwarzer et al., 1999) als auch Reliabilitätsindizes (Davis, 1999).

Methodisch gesehen sind zwischen online und offline erhobenen Daten vor allem die größere Varianz in den Onlinedaten sowie die Selbstselektion in Onlinestudien (Bandilla, 1999; Reips, 1999; Reips, 2000) zu beachten. Bei dieser passiven Auswahl von Versuchspersonen durch überrepräsentierte geübte und häufig teilnehmende Probanden besteht die Gefahr der Verzerrung von Daten (Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute, 2001), ihretwegen die Onlineforschung abzuwerten, ist jedoch unnötig. Zum einen sind die Vorteile der Onlineforschung (Zeit, Kosten- und Materialeinsparung) in vielen Studien notwendig, zum anderen ist die Umsetzung einer Studie offline nicht automatisch gleichbedeutend mit einem Zuwachs an Qualität bezüglich der Stichprobe, denn auch diese Methode muss sich mit dem Problem der Selbstselektion auseinandersetzen. Daher erfordert es vor allem die Anerkennung der besonderen Umstände in der Form der Datenerhebung sowie Berücksichtigung dieser bei Analysen und Interpretationen der Ergebnisse.

Die Abhängigkeit von der technischen Ausstattung, die Abwesenheit eines Versuchsleiters sowie die hohe Anonymität der Stichprobe oder Teilnahmeabbrüche (= *Drop-out*) müssen sicherlich als Nachteile der Onlineforschung aufgeführt werden (Batinic & Bosnjak, 1997; Bosnjak, 2001; Reips, 1997; Reips, 1999). Da jedoch eine Vielzahl der „Nachteile“ von Studien im WWW durchaus auch bezüglich traditioneller Datenerhebungen diskutiert werden müssen, erfordern die Nachteile nicht etwa eine Abkehr von Onlinestudien als Methode, sondern legen vielmehr eine Berücksichtigung derer in Planung und Programmierung nahe.

In der Online-Umsetzung kann es schlussendlich nur darum gehen, geeignete Vorkehrungen zu treffen, um möglichst viele Personen zu erreichen und die Selbstselektion so gering wie möglich zu halten, z. B. durch Replikationen mit anderen Stichproben, geeigneter Platzierung von Werbelinks, „multiple site entry“-Technik<sup>40</sup> (Reips, 1997; 2000). Im Großen und Ganzen sind online und offline erhobene Daten in ihrer Qualität äquivalent in dem, was sie messen und wie gut sie dies tun.

---

<sup>40</sup> Bei dieser Technik wird eine Onlinestudie auf verschiedenen Webseiten verlinkt, die unterschiedliches Publikum ansprechen. Die so erreichten vielfältigen Substichproben werden – basierend auf ihrem unterschiedlichen Zugang zur Studie - in ihren Ergebnissen verglichen, wodurch Aussagen bezüglich eventueller Verzerrungen schon innerhalb des Mediums Internet möglich sind.

In den hier vorgestellten Studien wurden alle technischen und methodischen Möglichkeiten umgesetzt, um maximale Generalisierbarkeit der Ergebnisse zu gewährleisten. Die umfangreiche Rekrutierung der Stichproben über einen längeren Zeitraum (die sich von Offline-Stichproben in der demografischen Charakteristik nicht unterscheiden); das Ausschließen intimer Fragen, die durch sozial erwünschte Antworten verzerrt werden können; das Überprüfen der Datenqualität durch klassische Methoden; die Replikation der Validierung sowie die Wiederholungsmessungen unterstreichen die Generalisierbarkeit des Instruments über die Erhebungsmodalitäten – im Sinne der Nutzung im Offline-Modus – hinweg.

#### 4.3.2 Stichprobenwahl

Eine wichtige Frage, die sich ganz allgemein auch bei der vorliegenden Arbeit stellt, betrifft die Stichprobenwahl und inwieweit die gefundenen Ergebnisse auf die Gruppe der befragten Studierenden limitiert ist.

Gerade in Bezug zur Zusammensetzung der Stichprobe ist die Kritik groß, wenn diese auf Studierende beschränkt ist, gelten sie doch nur als bedingt geeignet um Aussagen über kulturelle Aspekte einer Gesellschaft zu machen, die allgemeinen Wert haben sollen. Gleichzeitig sind Studierende bevorzugt befragte Probanden und dies auch, wenn es um Einschätzungen der in der Kultur vorherrschenden Werte und Normen geht (Auster & Ohm, 2000; Bem, 1974; Birnbaum, 1999; Harris, 1994; Holt & Ellis, 1998). Birnbaum (1999) hatte es wohl am ehesten auf den Punkt gebracht, als er argumentierte, dass Studierende vor allem aufgrund ihrer leichten Erreichbarkeit zu Befragungen dieser Art herangezogen werden („[...] *and college students because they are available.*“ S. 399).

Verschiedene Forscher hatten zudem bereits angemahnt, dass die Konzepte der Maskulinität und Femininität in Abhängigkeit des sozioökonomischen Status und der ethnischen Zugehörigkeit variieren können (Helgeson, 1994a; Tolson, 1977). Demnach würden Personen aus der arbeitenden Bevölkerung Maskulinität eher mit aggressiven Verhaltensweisen assoziieren, während die Mittelklasse den Schwerpunkt auf Erfolge im Sport oder im Beruf legen würde. Diese Überzeugung gewinnt an Bedeutung, berücksichtigt man die überwiegend studentischen Stichproben aus der vorliegenden Arbeit.

Vor diesem Hintergrund der möglichen Einschränkung wurde großes Bemühen darauf verwendet, die Grundgesamtheit der Studierenden insgesamt divers zu halten, wodurch ein repräsentativer Ausschnitt der Population maximal ermöglicht werden sollte. So wurde beispielsweise deutschlandweit über die Onlineplattform studiVZ rekrutiert und sichergestellt, dass nicht ausschließlich Psychologiestudierende befragt wurden, sondern Studierende aus Fachrichtungen aller Art wie beispielsweise Anglistik, BWL, Sprachen, Erziehungs-, Rechts-, Wirtschafts- oder Ingenieurwissenschaften berücksichtigt wurden. Zusätzlich zu der Studienfachvielfalt der Stichprobe zeigte sich ebenfalls große Varianz in weiteren demografischen Faktoren wie Studiendauer, Alter, Familien- oder Wohnsituationen. Zudem demonstrierten Ergebnisse aus anderen Studien, dass studentische und nichtstudentische Stichproben durchaus zu vergleichbaren Einschätzungen kommen könnten. So hatten Diekman

und Eagly (2000) einen gesellschaftlichen Konsensus bezüglich der Stereotype von Frauen, wie die Zunahme maskuliner Eigenschaften, nachgewiesen – in sehr unterschiedlichen Gruppen von Studierenden, jüngeren und älteren Probanden. Es sind Ergebnisse wie diese, die für eine Generalisierbarkeit von Studien mit Studierenden sprechen. An dieser Stelle sei die mögliche Selbstselektion von Probanden aus Onlinestudien erwähnt (Bandilla, 1999; Reips, 1999; 2000), die umso mehr Gewicht erhält, zieht man die unbekannte Grundgesamtheit der Internetnutzer in Betracht (Bandilla, 1999). Durch die dynamische Veränderung der Gesamtheit der Internetnutzer gilt diese Gruppe als äußerst schwierig zu definieren und repräsentativ abzubilden. Dies ist in der Tat eine erhebliche Einschränkung für die Verallgemeinerung von Onlineergebnissen und dennoch nicht nur auf diese Methode beschränkt. So ist es offline zwar möglich jeden Haushalt zu erreichen, eine Garantie für eine repräsentative Stichprobe ist dadurch jedoch noch nicht gegeben. Rücklaufquoten postalischer Befragungen von mitunter weniger als 10% (Bortz & Döring, 2006; Porst, 2001) lassen auch hier berechnete Zweifel hinsichtlich der Selbstselektion von Probanden zu.

Darüber hinaus war die Stichprobenwahl und -zusammensetzung vor allem für die Konstruktionsphase und dabei die Generierung der typischen Eigenschaften für Männer und Frauen relevant. Im Zuge dieses Prozesses war es wichtig, Eigenschaften zu finden, die nicht nur dem aktuellen Sprachgebrauch und den vorherrschenden Geschlechtstypizitäten entsprechen, sondern auch Beschreibungen, die über die Gruppe von Studierenden hinaus Bestand haben würden. Daher wurde gerade bei der Generierung der Eigenschaften dafür Sorge getragen, eine große Stichprobe anzusprechen und auch Probanden zu rekrutieren, die nicht notwendigerweise studiert hatten und im Durchschnitt älter waren. Mit diesem Vorgehen sollten die generierten Beschreibungen eine entsprechende Vielfältigkeit widerspiegeln, die in der folgenden Einschätzungsphase hinsichtlich ihrer Eignung überprüft wurden.

Auch in der zweiten Phase des Konstruktionsprozesses zur Beurteilung der Typizität und sozialen Erwünschtheit wurde durch das methodische Vorgehen sichergestellt, dass die Ergebnisse nicht ausschließlich auf Studierende beschränkt bleiben. So wurden beispielsweise die Instruktionsbestandteile für alle Probanden nicht nur gleich gehalten, sondern zudem in einer Weise formuliert, die deutlich machte, dass es gerade in der Konstruktionsphase nicht um ihre eigene Meinung geht, sondern ihre Beurteilungen die soziale Erwünschtheit in dieser Gesellschaft reflektieren sollten. Dieser Aspekt ist auch vor dem Hintergrund des sozial erwünschten Antwortens nicht unbedeutend. Paulhus (2002) hatte beispielsweise in einer seiner Studien zeigen können, wie sehr die Einschätzungen von den Instruktionen abhängig sind und inwieweit manipulierte Anweisungen in eher communale oder agentische Formulierungen zu Unterschieden in der Selbstdarstellung führten. Um zu verhindern, dass bereits in der Konstruktionsphase individuell wertungsbezogene Antworten gegeben werden, wurden Probanden in Anlehnung an andere Arbeiten aus diesem Bereich (Auster & Ohm, 2000;

Bem, 1974; Harris, 1994; Holt & Ellis, 1998)<sup>41</sup> gebeten, Einschätzungen bezüglich der Typizität für eine Frau bzw. einen Mann sowie der sozialen Erwünschtheit vorzunehmen.

Ähnlich wie auch Prentice und Carranza (2002) anmerkten, ist zusammenfassend kritisch zu hinterfragen, inwieweit man ähnliche Resultate auch mit einer repräsentativen Stichprobe hätte erreichen können. Zwei Aspekte müssen in diesem Zusammenhang immer berücksichtigt werden: die Zusammensetzung der Stichprobe und die Möglichkeit, die Ergebnisse zu verallgemeinern.

Grundsätzlich muss festgestellt werden, dass nicht der Anspruch erhoben wird, dass es sich hierbei um eine repräsentative Auswahl aus der Grundpopulation handelt. Gleichzeitig stellen das sorgfältige methodische Vorgehen sowie der aufwendige Prozess der Konstruktion unter Berücksichtigung einer Vielzahl von Itemparametern sicher, dass das neue Instrument ein Abbild aktueller maskuliner und femininer Eigenschaften darstellt, welches auch für die allgemeine Bevölkerung gilt.

Für die vorliegende Arbeit bedeutet dies zusammenfassend, dass das mithilfe von Onlineerhebungen entwickelte und erprobte Inventar nicht auf seinen Erhebungsmodus und die Stichprobe beschränkt ist. Sattdessen ist das neue Instrument sowohl auf andere Gruppen als die im Konstruktionsprozess untersuchten als auch in der Nutzung im Offline-Modus anwendbar.

#### 4.4 Einschränkungen, methodenkritische Aspekte und Ausblick

In der vorliegenden Arbeit wurde ein Instrument zur Messung des Geschlechtsrollenselbstkonzepts entwickelt, das sowohl positive als auch negative maskuline und feminine Eigenschaften erfasst. Während die Ergebnisse klar demonstrieren, dass die Entwicklung eines reliablen und validen Messinstruments geglückt ist, welches aufgrund der Aktualität und Widerspiegelung des aktuellen Sprachgebrauchs einen Zugewinn in diesem Forschungsbereich darstellt, bleiben einige wenige Fragen noch offen.

Obwohl die Bedeutung der vier neuen Teilskalen durch eine Vielzahl von Validierungskonstrukten eindeutig nachgewiesen werden konnte, handelte es sich bei den Konstrukten ausschließlich um Selbstberichtsmaße, wodurch die Objektivität der gefundenen Ergebnisse einschränkt wird. In zukünftigen Studien wäre daher der Einsatz strengerer Außenkriterien, wie beispielsweise objektiver Testdaten, oder die Berücksichtigung von Fremdeinschätzungen wünschenswert. In experimentellen Studien wäre eine Überprüfung der vier Teilskalen im Hinblick auf konkrete Verhaltenskonstrukte in spezifisch definierten Situationen möglich.

---

<sup>41</sup> eingesetzte Instruktionen: "How desirable is it in American society for a man (woman) to possess each of the characteristics?" (Holt & Ellis, 1998; S. 932); "We would like you to indicate how desirable it is in American society for a man/woman to possess each of these characteristics. Note: We are not interested in your personal opinion of how desirable each of these characteristics is. Rather, we want your judgment of how our society evaluates each of these characteristics in a man/woman." (Auster & Ohm, 2000; S. 504); "In American culture, how desirable is it for a woman to be assertive?" (Harris, 1994; S. 246)

Des Weiteren erfolgte die Überprüfung der Zusammenhänge zwischen den Teilskalen und den Validierungskonstrukten nur im querschnittlichen Design. Künftig wäre der prädiktive Wert der Skalen auch im Längsschnitt abzusichern.

Die hier dargestellten Ergebnisse und Diskussionen verdeutlichen, wie schwierig und ungenau das Konzept der eigenschaftsbasierten Geschlechtsrolle bisweilen bleibt und welche Besonderheiten, wie beispielsweise die Zuschreibung der unbeliebteren negativen femininen und maskulinen Geschlechtereigenschaften, bestehen. Es bleiben Fragen offen, wie genau diese Zuschreibungen vorgenommen werden. Vom Ansatz der Selbstwerterhöhung sollten diese Zuschreibungen der positiven Eigenschaften verhältnismäßig einfach und schnell gelingen. Wie dies bei negativen Eigenschaften genau vollzogen wird, bleibt zunächst jedoch unklar. Eigenschaften, die von der Typizität zwar der eigenen Geschlechtsrolle zugeordnet wurden, deren negative Valenz man jedoch nicht im eigenen Selbst wiederfinden möchte, fordern beispielsweise die Aufrechterhaltung des Selbstwertes heraus. So mag eine Tendenz bestehen, diese Eigenschaften per se abzulehnen, gleichzeitig jedoch anzuerkennen, dass sie Teil der geschlechtsspezifischen Rolle sind. Experimente, die diesen Zuschreibungsprozess genauer nachvollziehen, könnten diesen Prozess transparenter machen. Die Aufzeichnung von Reaktionszeiten würde es erlauben, die mögliche Initialtendenz, negative Eigenschaften zunächst abzulehnen um sie schließlich sich doch selbst zuzuschreiben, abzubilden. Zusätzlich wäre in diesem Zusammenhang eine Manipulation der Salienz der Geschlechtsrolle möglich, um womöglich in der Lage zu sein, selbstwertdienliche und eher sozial erwünschte Antworten zu vermeiden. So wäre denkbar, die jeweils eigene Geschlechtszugehörigkeit salient zu machen und explizit Vergleiche mit der anderen Gruppe anzustreben (z. B. „*Wie sehen Sie sich als Frau im Vergleich zu Männern?*“) oder dies bewusst zu unterbinden (z. B. „*Wie sehen Sie sich als Frau?*“) und die Ausprägungen auf dem GSK in Abhängigkeit unterschiedlicher Referenzrahmen zu überprüfen.

Trotz dieser Einschränkung liegt mit dem Inventar *PoNeGRI* ein zuverlässiges und valides Instrument zur Erfassung des Geschlechtsrollenselbstkonzepts bei Erwachsenen vor. Aufgrund der theoretischen Verankerung und des mehrstufigen Auswahlprozesses bildet es sowohl aktuelle Vorstellungen als auch das aktuelle Vokabular dessen ab, was im deutschsprachigen Raum als typisch für Männer und Frauen gilt. Durch die Berücksichtigung der negativen Aspekte von Maskulinität und Femininität bietet sich die Möglichkeit, einen Bezug zu negativen oder als problematisch geltenden psychologischen Korrelativen der Selbstkonstruktion herzustellen. Nicht zuletzt durch seine Kürze dürfte das Instrument in einer Vielzahl von Anwendungskontexten einsetzbar sein, um den Einfluss geschlechtstypischer Merkmale genauer zu analysieren.

## 5 BIBLIOGRAFIE

- Abele, A. E. (1997). Geschlechtsrollen, Geschlechtsrollenorientierungen und Geschlechtsstereotype im Wandel. In E. Liebau (Hrsg.), *Das Generationsverhältnis. Über das Zusammenleben in Familie und Gesellschaft* (S. 123-139). Weinheim: Juventa Verlag.
- Abele, A. E. (2003). The Dynamics of Masculine-Agentive and Feminine-Communal Traits: Findings from a Prospective Study. *Journal of Personality and Social Psychology*, *85*, 768-776.
- Abele, A. E. (2007). Geschlechterdifferenz aus sozialpsychologischer Sicht. In U. Krebs, & J. Forster (Hrsg.), *Sie und Er interdisziplinär - Aktuelle Themen interdisziplinär* (S. 165-180). Berlin: LIT Verlag Dr. W. Hopf.
- Abrams, D., & Hogg, M. A. (1990). Social identification, self- categorization, and social influence. In W. Stroebe, & M. Hewstone (Hrsg.), *European Review of Social Psychology* (S. 195-228). Chichester: Wiley.
- Abrams, D., & Hogg, M. A. (2001). Comments on the motivational status of self-esteem in social identity and intergroup discrimination. In M. A. Hogg, & D. Abrams (Hrsg.), *Intergroup relations: Essential readings* (S. 232-244). New York: Psychology Press.
- Abrams, D., & Hogg, M. A. (2004). Collective Identity: Group Membership and Self-conception. In M. B. Brewer, & M. Hewstone (Hrsg.), *Self and social identity* (S. 147-181). Malden: Blackwell Publishing.
- Acuña, M., & Bruner, C. A. (1986). Una Aproximacion a la Validacion del Inventario de Roles Sexuales de Bem (BSRI) en Mexico. *Revista de Psicologia Social y Personalidad*, *2*, 43-64.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park: Sage.
- Ajzen, I. (1985). From Intentions to actions: A theory of planned behavior. In J. Kuhl, & J. Beckmann (Hrsg.), *Actions control: From cognition to behavior* (S. 11-39). Heidelberg: Springer.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179-211.
- Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*, 27-58.
- Ajzen, I., & Fishbein, M. (2005). The Influence of Attitudes on Behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 173-221). Mahwah: Lawrence Erlbaum Associates Publishers.
- Ajzen, I., & Fishbein, M. (2008). Scaling and Testing Multiplicative Combinations in the Expectancy-Value Model of Attitudes. *Journal of Applied Social Psychology*, *38*, 2222-2247.
- Albert, A. A., & Porter, J. R. (1983). Age patterns in the development of children's gender-role stereotypes. *Sex Roles*, *9*, 59-67.
- Alfermann, D. (1996). *Geschlechterrollen und geschlechtstypisches Verhalten*. Stuttgart: Kohlhammer.
- Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology*, *49*, 1621-1630.
- Alicke, M. D., Klotz, M. L., Breitenbecher, D. L., Yurak, T. J., & Vredenburg, D. S. (1995). Personal contact, individuation, and the better-than-average effect. *Journal of Personality and Social Psychology*, *68*, 804-825.
- Allport, G. W. (1954). *The nature of prejudice*. Oxford England: Addison-Wesley.
- Altermatt, T. W., DeWall, C. N., & Leskinen, E. (2003). Agency and Virtue: Dimensions Underlying Subgroups of Women. *Sex Roles*, *49*, 631-641.
- Altstötter-Gleich, C. (2004). Expressivität, Instrumentalität und psychische Gesundheit: Ein Beitrag zur Validierung einer Skala zur Erfassung des geschlechtsrollenbezogenen Selbstkonzepts. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *25*, 123-139.
- Altstötter-Gleich, C., Eglau, B., & Kramer, J. (2000). *Möglichkeiten der Operationalisierung von Expressivität und Instrumentalität: Entwicklung von Skalen zur Erfassung der Geschlechtstypizität (GTS)*. Forschungsbericht des Fachbereichs Psychologie, Nr.23: Universität Koblenz-Landau, Abteilung Landau.
- Amelang, M., & Borkenau, P. (1981). Untersuchungen zur Validität von Kontroll-Skalen für Soziale Erwünschtheit und Akquieszenz. *Diagnostica*, *27*, 295-312.
- 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.
- Anderson, N. H. (1968). Likableness Ratings of 555 Personality - Trait Words. *Journal of Personality and Social Psychology*, *9*, 272-279.



- Angleitner, A., & Ostendorf, F. (1994). Von aalglatt bis zynisch: Merkmale persönlichkeitsbeschreibender Begriffe. In W. Hager, & M. Hasselhorn (Hrsg.), *Handbuch deutschsprachiger Wortnormen* (S. 340-381). Göttingen: Hogrefe.
- Antill, J. K. (1983). Sex role complementarity versus similarity in married couples. *Journal of Personality and Social Psychology, 45*, 145-155.
- Antill, J. K., Cunningham, J. D., Russell, G., & Thompson, N. L. (1981). An Australian sex-role scale. *Australian Journal of Psychology, 33*, 169-183.
- Antill, J. K., & Russell, G. (1982). The factor structure of the Bem Sex-Role Inventory: Method and sample comparisons. *Australian Journal of Psychology, 34*, 183-193.
- Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute (2001). Standards zur Qualitätssicherung für Online-Befragungen. *Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.(ADM)*. Retrieved 11.06.2009 from [http://www.adm-ev.de/pdf/Onlinestandards\\_D.PDF](http://www.adm-ev.de/pdf/Onlinestandards_D.PDF).
- Archer, J. (2006). The Importance of Theory for Evaluating Evidence on Sex Differences. *American Psychologist, 61*, 638-639.
- Archer, J., & Lloyd, B. (2002). *Sex and Gender*. Cambridge: Cambridge University Press.
- Archer, J., Smith, J., & Kilpatrick, G. (1995). The association between gender scale measures and gender clustering in recall. *Sex Roles, 33*, 299-308.
- Asendorpf, J. B. (1992). Beyond stability: Predicting inter-individual differences in intra-individual change. *European Journal of Personality, 6*, 103-117.
- Ashmore, R. D. (1981). Sex stereotypes and implicit personality theory. In D. L. Hamilton (Hrsg.), *Cognitive processes in stereotyping and intergroup behaviour* (S. 37-82). Hillsdale: Erlbaum.
- Ashmore, R. D. (1990). Sex, gender, and the individual. In L. A. Pervin (Hrsg.), *Handbook of personality: Theory and research* (S. 486-526). New York: Guilford Press.
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Hrsg.), *Cognitive processes in stereotyping and intergroup behavior* (S. 1-36). Hillsdale: Lawrence Erlbaum.
- Ashmore, R. D., & del Boca, F. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive-social psychological conceptualization. *Sex Roles, 5*, 219-248.
- Athenstaedt, U. (2000). Normative Geschlechtsrollenorientierung: Entwicklung und Validierung eines Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie, 21*, 91-104.
- Athenstaedt, U. (2003). On the Content and Structure of the Gender Role Self-Concept: Including Gender-Stereotypical Behaviors in Addition to Traits. *Psychology of Women Quarterly, 27*, 309-318.
- Athenstaedt, U., & Altstötter-Gleich, C. (2009). Gender role attribute self-concept: Conceptualization, assessment, and correlates (submitted).
- Athenstaedt, U., Altstötter-Gleich, C., & Hillberger, P. (2007). Negative and Positive Gender Traits as Dimensions of the Gender Role Self-Concept: Validation of New Gender Trait Self-Concept Scales (GTSS) By Relating Them To Psychological Health Related Variables and Social Desirability Orientation.
- Athenstaedt, U., Haas, E., & Schwab, S. (2004). Gender role self-concept and gender-typed communication behavior in mixed-sex and same-sex dyads. *Sex Roles, 50*, 37-52.
- Athenstaedt, U., Heinzle, C., & Lerchbaumer, G. (2008). Gender subgroup self-categorization and gender role self-concept. *Sex Roles, 58*, 266-278.
- Aubé, J., & Koestner, R. (1992). Gender characteristics and adjustment: A longitudinal study. *Journal of Personality and Social Psychology, 63*, 485-493.
- Augoustinos, M., Ahrens, C., & Innes, J. M. (1994). Stereotypes and prejudice: The Australian experience. *British Journal of Social Psychology, 33*, 125-141.
- Augoustinos, M., & Rosewarne, D. L. (2001). Stereotype knowledge and prejudice in children. *British Journal of Developmental Psychology, 19*, 143-156.
- Auster, C. J., & Ohm, S. C. (2000). Masculinity and femininity in contemporary American society: A reevaluation using the Bem Sex-Role Inventory. *Sex Roles, 43*, 499-528.
- Baber, K. M., & Tucker, C. J. (2006). The Social Roles Questionnaire: A New Approach to Measuring Attitudes Toward Gender. *Sex Roles, 54*, 459-467.

- Bailey, R. D., Foote, W. E., & Throckmorton, B. (2000). Human Sexual Behavior: A Comparison of College and Internet Surveys. In M. H. Birnbaum (Hrsg.), *Psychological Experiments of the Internet* (S. 141-168). London: Academic Press.
- Baird, B. M., Le, K., & Lucas, R. E. (2006). On the Nature of Intraindividual Personality Variability: Reliability, Validity, and Associations With Well-Being. *Journal of Personality and Social Psychology, 90*, 512-527.
- Bakan, D. (1966). *The duality of human existence*. Chicago: Rand McNally.
- Ballard-Reisch, D., & Elton, M. (1992). Gender orientation and the Bem Sex Role Inventory: A psychological construct revisited. *Sex Roles, 27*, 291-306.
- Bandilla, W., & Hauptmanns, P. (1998). Internetbasierte Umfragen als Datenerhebungstechnik für die empirische Sozialforschung? *ZUMA-Nachrichten, 43*, 36-53.
- Bandilla, W. (1999). WWW-Umfragen - Eine alternative Datenerhebungstechnik für die empirische Sozialforschung. In B. Batinic, A. Werner, L. Gräf, & W. Bandilla (Hrsg.), *Online Research* (S. 9-19). Göttingen: Hogrefe.
- Bandura, A. (1979). *Sozial-kognitive Lerntheorie*. Stuttgart: Klett-Cotta.
- Bandura, A., Ross, D., & Ross, S. A. (1963). A comparative test of the status envy, social power, and secondary reinforcement theories of identificatory learning. *Journal of Abnormal & Social Psychology, 67*, 527-534.
- Barry, H. I., Bacon, M. K., & Child, I. L. (1957). A cross-cultural survey of some sex differences in socialization. *Journal of Abnormal & Social Psychology, 55*, 327-332.
- Bason, J. J. (2000). Comparison of telephone, mail, web and IVR surveys of drug and alcohol use among University of Georgia students. *Proceedings of the Survey Research Methods Section, American Statistical Association*. Retrieved 01.08.2008 from [http://www.amstat.org/sections/SRMS/proceedings/papers/2000\\_184.pdf](http://www.amstat.org/sections/SRMS/proceedings/papers/2000_184.pdf).
- Basow, S. A. (1984). Cultural variations in sex-typing. *Sex Roles, 10*, 577-585.
- Basow, S. A. (1992). *Gender: Stereotypes and roles (3rd ed.)*. Belmont: Thomson Brooks/Cole Publishing Co.
- Batinic, B. (2001). *Fragebogenuntersuchungen im Internet*. Aachen: Shaker-Verlag.
- Batinic, B., & Bosnjak, M. (1997). Fragebogenuntersuchung im Internet. In B. Batinic (Hrsg.), *Internet für Psychologen* (S. 221-244). Göttingen: Hogrefe.
- Baumeister, R. F. (1999). *The self in social psychology*. New York: Psychology Press.
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5*, 323-370.
- Beck, A. T., Steer, R. A., & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review, 8*, 77-100.
- Beckwith, J. B. (1993). Gender stereotypes and mental health revisited. *Social Behavior and Personality, 21*, 85-88.
- Beere, C. A. (1979). *Women and women's issues: A handbook of tests and measures*. San Francisco: Jossey-Bass.
- Beere, C. A. (1990). *Gender roles: A handbook of tests and measures*. New York: Greenwood Press.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162.
- Bem, S. L. (1977). On the Utility of Alternative Procedures for Assessing Psychological Androgyny. *Journal of Consulting and Clinical Psychology, 45*, 196-205.
- Bem, S. L. (1979). Theory and measurement of androgyny: A reply to the Pedhazur-Tetenbaum and Locksley-Colten critiques. *Journal of Personality and Social Psychology, 37*, 1047-1054.
- Bem, S. L. (1981a). *Bem Sex-Role Inventory - Professional Manual*. Palo Alto: Consulting Psychologists Press, Inc.
- Bem, S. L. (1981b). Gender Schema Theory: A Cognitive Account of Sex Typing. *Psychological Review, 88*, 354-364.
- Bem, S. L. (1984). Androgyny and gender schema theory: A conceptual and empirical integration. *Nebraska Symposium on Motivation, 32*, 179-226.
- Bem, S. L. (1987a). Gender schema theory and the romantic tradition. In P. Shaver, & C. Hendrick (Hrsg.), *Sex and gender* (S. 251-271). Thousand Oaks: Sage Publications, Inc.
- Bem, S. L. (1987b). No: Gender schema theory and its implications for child development: Raising gender-aschematic children in a gender-schematic society. In M. R. Walsh (Hrsg.), *The psychology of women: Ongoing debates* (S. 226-245). New Haven: Yale University Press.
- Bem, S. L. (1993). *The lenses of gender: Transforming the debate on sexual inequality*. New Haven: Yale University Press.

- Bem, S. L., & Lenney, E. (1976). Sex Typing and the Avoidance of Cross-Sex Behavior. *Journal of Personality and Social Psychology*, 33, 48-54.
- Bem, S. L., & Lewis, S. A. (1975). Sex Role Adaptability: One Consequence of Psychological Androgyny. *Journal of Personality and Social Psychology*, 31, 634-643.
- Bem, S. L., Martyna, W., & Watson, C. (1976). Sex typing and androgyny: Further explorations of the expressive domain. *Journal of Personality and Social Psychology*, 34, 1016-1023.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Berenbaum, S. A. (1998). How hormones affect behavioral and neural development: Introduction to the special issue on 'Gonadal hormones and sex differences in behavior.'. *Developmental Neuropsychology*, 14, 175-196.
- Berzins, J. I., Welling, M. A., & Wetter, R. E. (1978). A New Measure of Psychological Androgyny Based on the Personality Research Form. *Journal of Consulting and Clinical Psychology*, 46, 126-138.
- Best, D. L. (2001). Gender Concepts: Convergence in Cross-Cultural Research and Methodologies. *Cross-Cultural Research: The Journal of Comparative Social Science*, 35, 23-43.
- Bierhoff-Alfermann, D. (1989). *Androgynie - Möglichkeit und Grenzen der Geschlechterrollen*. Opladen: Westdeutscher Verlag.
- Biernat, M., & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66, 5-20.
- Biernat, M., Vescio, T. K., & Green, M. L. (1996). Selective Self-Stereotyping. *Journal of Personality and Social Psychology*, 71, 1194-1209.
- Birnbaum, M. H. (1999). Testing critical properties of decision making on the Internet. *Psychological Science*, 10, 399-407.
- Bischof-Köhler, D. (2006). *Von Natur aus anders. Die Psychologie der Geschlechtsunterschiede*. Stuttgart: Kohlhammer.
- BITKOM. (2008). Fast jeder fünfte Mensch ist online. *Bundesverband Informationswirtschaft Telekommunikation und neue Medien e.V.* Retrieved 24.06.2008 from [http://www.bitkom.org/46074\\_46069.aspx](http://www.bitkom.org/46074_46069.aspx).
- Blanchard-Fields, F., Suhrer-Roussel, L., & Hertzog, C. (1994). A confirmatory factor analysis of the Bem Sex Role Inventory: Old questions, new answers. *Sex Roles*, 30, 423-457.
- Blanton, H., & Jaccard, J. (2006). Tests of multiplicative models in psychology: A case study using the unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, 113, 155-165.
- Blatt, S. J., & Shichman, S. (1983). Two primary configurations of psychopathology. *Psychoanalysis and Contemporary Thought*, 6, 187-254.
- Block, J. H. (1973). Conceptions of Sex Role: Some Cross-Cultural and Longitudinal Perspectives. *American Psychologist*, 512-526.
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler (4. Auflage)*. Berlin: Springer.
- Bosnjak, M. (2001). Participation in Non-Restricted Web Surveys: A Typology and Explanatory Model for Item Non-Response. In U.-D. Reips, & M. Bosnjak (Hrsg.), *Dimensions of Internet Science* (S. 193-207). Berlin: Pabst.
- Bosnjak, M., & Tuten, T. L. (2001). Classifying Response Behaviors in Web-based Surveys. *Journal of Computer-Mediated Communication*. Retrieved 11.06.2009 from <http://jcmc.indiana.edu/vol6/issue3/bosnjak.html>.
- Bowker, D. K. (2000). Minimizing Survey Error Establishing Protocols for Web Survey Design. *Paper presented at the International Field Director's and Technologies Conference in Portland, Oregon*.
- Bozionelos, N., & Bozionelos, G. (2003). Instrumental and Expressive Traits: Their Relationship and their Association with Biological Sex. *Social Behavior and Personality*, 31, 423-430.
- Brandtstädter, J., & Greve, W. (1994). The aging self: Stabilizing and protective processes. *Developmental Review*, 14, 52-80.
- Brannon, L. (2005). *Gender: Psychological perspectives*. Needham Heights: Allyn and Bacon.
- Branscombe, N. R., & Smith, E. R. (1990). Gender and racial stereotypes in impression formation and social decision-making processes. *Sex Roles*, 22, 627-647.

- Brems, C., & Johnson, M. E. (1990). Reexamination of the Bem Sex-Role Inventory: The Interpersonal BSRI. *Journal of Personality Assessment*, 55(3-4), 484-498.
- Brewer, M. B., & Gardner, W. (1996). Who is this 'We'? Levels of collective identity and self representations. *Journal of Personality and Social Psychology*, 71, 83-93.
- Brickenkamp, R. (1962). *Aufmerksamkeits-Belastungs-Test Handanweisung d-2*. Oxford England: C. J. Hogrefe.
- Brigham, J. (1986). *Social Psychology*. Toronto: Little & Brown.
- Broverman, I. K., Broverman, D. M., Clarkson, F. E., Rosenkrantz, P. S., & Vogel, S. R. (1970). Sex-role stereotypes and clinical judgments of mental health. *Journal of Consulting and Clinical Psychology*, 34, 1-7.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues*, 28, 59-78.
- Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1994). Sex-role stereotypes: A current appraisal. In B. Puka (Hrsg.), *Caring voices and women's moral frames: Gilligan's view* (S. 191-210). New York: Garland Publishing.
- Brown, J. (1998). Comparing charges: The experience of discrimination and harassment among women police officers serving in Australia, the British Isles and the United States of America. *International Journal of Police Science & Management*, 1, 227-240.
- Bruch, M. A. (2002). The relevance of mitigated and unmitigated agency and communion for depression vulnerabilities and dysphoria. *Journal of Counseling Psychology*, 49, 449-459.
- Brunstein, J. C., Lautenschlager, U., Nawroth, B., & Pöhlmann, K. (1995). Persönliche Anliegen, soziale Motive und emotionales Wohlbefinden. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 16, 1-10.
- Bryson, J. B., & Corey, D. M. (1977). Sex-diagnosticsity in 100 personality trait adjectives. *Personality and Social Psychology Bulletin*, 3, 301-304.
- Buchanan, T. (2001). Online Personality Assessment. In U.D. Reips & M. Bosnjak (Eds.). *Dimensions of Internet Science*. In (S. 57-74). Berlin: Pabst.
- Buchanan, T. (2002). Online assessment: Desirable or dangerous? *Professional Psychology: Research and Practice*, 33, 148-154.
- Buchanan, T., & Smith, J. L. (1999). Research on the internet: Validation of a World-Wide Web mediated personality scale. *Behavior Research Methods, Instruments and Computers*, 31, 565-571.
- Bundesministerium für Familie - Senioren - Frauen und Jugend (2007). 20-jährigen Frauen und Männern heute - Lebensentwürfe, Rollenbilder, Einstellungen zur Gleichstellung. Retrieved 20.06.2009 from <http://www.bmfsfj.de/bmfsfj/generator/RedaktionBMFSFJ/Broschuerenstelle/Pdf-Anlagen/sinus.property=pdf,bereich=bmfsfj,sprache=de,rwb=true.pdf>.
- Burchardt, C. J., & Serbin, L. A. (1982). Psychological androgyny and personality adjustment in college and psychiatric populations. *Sex Roles*, 8, 835-851.
- Buss, A. H., & Warren, W. L. (2000). *Aggression Questionnaire: Manual*. Los Angeles: Western Psychological Services.
- Buss, D. M. (1990). Unmitigated agency and unmitigated communion: An analysis of the negative components of masculinity and femininity. *Sex Roles*, 22, 555-568.
- Buss, D. M., & Kenrick, D. T. (1998). Evolutionary social psychology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Hrsg.), *The Handbook Of Social Psychology (4th ed.)* (S. 982-1026). New York: McGraw-Hill.
- Butler, D., & Geis, F. L. (1990). Nonverbal affect responses to male and female leaders: Implications for leadership evaluations. *Journal of Personality and Social Psychology*, 58, 48-59.
- Cadinu, M., Maass, A., Rosabianca, A., & Kiesner, J. (2005). Why Do Women Underperform Under Stereotype Threat? Evidence for the Role of Negative Thinking. *Psychological Science*, 16, 572-578.
- Campbell, M., Steffen, J. J., & Langmeyer, D. (1981). Psychological androgyny and social competence. *Psychological Reports*, 48, 611-614.
- Campbell, T., Gillaspay, J. A. J., & Thompson, B. (1997). The factor structure of the Bem Sex-Role Inventory (BSRI): Confirmatory analysis of long and short forms. *Educational and Psychological Measurement*, 57, 118-124.
- Carli, L. L. (1990). Gender, language, and influence. *Journal of Personality and Social Psychology*, 59, 941-951.
- Carli, L. L., & Eagly, A. H. (1999). Gender effects on social influence and emergent leadership. In G. N. Powell, & G. N. Powell (Hrsg.), *Handbook of gender and work* (S. 203-222). Thousand Oaks, CA US: Sage Publications, Inc.

- Caspi, A., & Roberts, B. W. (2001). Target Articles: Personality Development Across the Life Course: The Argument for Change and Continuity. *Psychological Inquiry*, 12, 49-66.
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality Development: Stability and Change. *Annual Review of Psychology*, 56, 453-484.
- Cejka, M. A., & Eagly, A. H. (1999). Gender-stereotypic images of occupations correspond to the sex segregation of employment. *Personality and Social Psychology Bulletin*, 25, 413-423.
- Chassin, L., & Stager, S. F. (1984). Determinants of self-esteem among incarcerated delinquents. *Social Psychology Quarterly*, 47, 382-390.
- Cheng, C. (1999). Gender-role differences in susceptibility to the influence of support availability on depression. *Journal of Personality*, 67, 439-467.
- Chia, R. C., Moore, J. L., Lam, K. N., & Chuang, C. J. (1994). Cultural differences in gender role attitudes between Chinese and American students. *Sex Roles*, 31, 23-30.
- Choi, N., Fuqua, D. R., & Newman, J. L. (2007). Hierarchical confirmatory factor analysis of the Bem Sex Role Inventory. *Educational and Psychological Measurement*, 67, 818-832.
- Coats, S., & Smith, E. R. (1999). Perceptions of gender subtypes: Sensitivity to recent exemplar activation and in-group/out-group differences. *Personality and Social Psychology Bulletin*, 25, 515-526.
- Cohen, J. B., Fishbein, M., & Ahtola, O. T. (1972). *The Nature and Uses of Expectancy-Value Models in Consumer Attitude Research*. (9 ed.).
- Coleman, J. M., & Hong, Y. y. (2008). Beyond nature and nurture: The influence of lay gender theories on self-stereotyping. *Self and Identity*, 7, 34-53.
- Connell, R. W., & Messerschmidt, J. W. (2005). Hegemonic Masculinity: Rethinking the Concept. *Gender & Society*, 19, 829-859.
- Constantinople, A. (1973). Masculinity-femininity: An exception to a famous dictum? *Psychological Bulletin*, 80, 389-407.
- Converse, P. D., Peterson, M. H., & Griffith, R. L. (2009). Faking on personality measures: Implications for selection involving multiple predictors. *International Journal of Selection and Assessment*, 17, 47-60.
- Conway, M., Pizzamiglio, M. T., & Mount, L. (1996). Status, communality, and agency: Implications for stereotypes of gender and other groups. *Journal of Personality and Social Psychology*, 71, 25-38.
- Cook, E. P. (1985). *Psychological androgyny*. New York: Pergamon.
- Costa, P. T. J., & McCrae, R. R. (1988). Personality in adulthood: A six-year longitudinal study of self-reports and spouse ratings on the NEO Personality Inventory. *Journal of Personality and Social Psychology*, 54, 853-863.
- Costrich, N., Feinstein, J., Kidder, L., Marecek, J., & Pascale, L. (1975). When stereotypes hurt: Three studies of penalties for sex-role reversals. *Journal of Experimental Social Psychology*, 11, 520-530.
- Council of American Survey Research Organizations. (1998). Frequently asked questions about conducting online research "New methodologies for traditional techniques". *Council of American Survey Research Organizations*. Retrieved 24.06.2008 from <http://www.casro.org/faq.cfm>.
- Craighead, L. W., & Green, B. J. (1989). Relationship between depressed mood and sex-typed personality characteristics in adolescents. *Journal of Youth and Adolescence*, 18, 467-474.
- Crocker, J., & Major, B. (1989). Social Stigma and Self-Esteem: The Self-Protective Properties of Stigma. *Psychological Review*, 96, 608-630.
- Cross, S. E., & Madson, L. (1997). Models of the self: Self-construals and gender. *Psychological Bulletin*, 122, 5-37.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Dalgard, O. S., Dowrick, C., Lehtinen, V., Vazquez-Barquero, J. L., Casey, P., Wilkinson, G. et al. (2006). Negative life events, social support and gender difference in depression. *Social Psychiatry and Psychiatric Epidemiology*, 41, 444-451.
- Danoff-Burg, S., Mosher, C. E., & Grant, C. A. (2006). Relations of agentic and communal personality traits to health behavior and substance use among college students. *Personality and Individual Differences*, 40, 353-363.
- Danoff-Burg, S., Revenson, T. A., Trudeau, K. J., & Paget, S. A. (2004). Unmitigated Communion, Social Constraints, and Psychological Distress Among Women With Rheumatoid Arthritis. *Journal of Personality*, 72, 29-46.
- Davies, A. P. C., & Shackelford, T. K. (2006). An Evolutionary Psychological Perspective on Gender Similarities and Differences. *American Psychologist*, 61, 640-641.

- Davis, R. N. (1999). Web-based administration of a personality questionnaire: Comparison with traditional methods. *Behavior Research Methods, Instruments and Computers*, 31, 572-577.
- De Leo, D., Villa, A., Magni, G., Andreatta, A., & Gagliardi, A. (1986). Presentazione della versione italiana e contributo alla taratura del Bem Sex-Role Inventory. *Bollettino di Psicologia Applicata*, 175, 21-28.
- Dear, G. E., & Roberts, C. M. (2002). The relationships between codependency and femininity and masculinity. *Sex Roles*, 46, 159-165.
- DeArmond, S., Tye, M., Chen, P. Y., Krauss, A., Rogers, D. A., & Sintek, E. (2006). Age and Gender Stereotypes: New Challenges in a Changing Workplace and Workforce. *Journal of Applied Social Psychology*, 36, 2184-2214.
- Deaux, K. (1984). From individual differences to social categories: Analysis of a decade's research on gender. *American Psychologist*, 39, 105-116.
- Deaux, K. (1985). Sex and gender. *Annual Review of Psychology*, 36, 49-81.
- Deaux, K., & Kite, M. E. (1987). Thinking about gender. In B. B. Hess, & M. M. Ferree (Hrsg.), *Analyzing gender: A handbook of social science research* (S. 92-117). Thousand Oaks: Sage Publications, Inc.
- Deaux, K., Kite, M. E., & Lewis, L. L. (1985). Clustering and gender schemata: An uncertain link. *Personality and Social Psychology Bulletin*, 11, 387-397.
- Deaux, K., & LaFrance, M. (1998). Gender. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Hrsg.), *The handbook of social psychology* (4th ed.) (S. 788-827). New York: McGraw-Hill.
- Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology*, 46, 991-1004.
- Deaux, K., & Major, B. (1987). Putting gender into context: An interactive model of gender-related behavior. *Psychological Review*, 94, 369-389.
- Deaux, K., Winton, W., Crowley, M., & Lewis, L. L. (1985). Level of categorization and content of gender stereotypes. *Social Cognition*, 3, 145-167.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5-18.
- Devine, P. G., & Elliot, A. J. (2000). Are racial stereotypes really fading? The Princeton trilogy revisited. In C. Stangor, & C. Stangor (Hrsg.), *Stereotypes and prejudice: Essential readings* (S. 86-99). New York, NY US: Psychology Press.
- Diekmann, A. B., & Eagly, A. H. (2000). Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Personality and Social Psychology Bulletin*, 26, 1171-1188.
- Diekmann, A. B., Eagly, A. H., Mladinic, A., & Ferreira, M. C. (2005). Dynamic Stereotypes About Women and Men in Latin America and the United States. *Journal of Cross-Cultural Psychology*, 36, 209-226.
- Dillman, D. A. (2000). *Procedures for conducting government-sponsored establishment surveys: Comparisons of the Total Design Method (TDM), a traditional cost-compensation Model, and tailored design*. Proceedings of Second International Conference on Establishment Surveys (ICESII), Buffalo, NY.
- Dillman, D. A., & Bowker, D. K. (2001). The Web Questionnaire Challenge to Survey Methodologists. In U.-D. Reips, & M. Bosnjak (Hrsg.), *Dimensions of Internet Science*. (S. 159-178). Berlin: Pabst.
- Dillman, D. A., Tortora, R. D., Conradt, J., & Bowker, D. (1998). *Influence of plain vs. fancy design on response rates for web surveys*. Presented at Joint Statistical Meetings, Dallas, Texas.
- Dillman, D. A., Tortora, R. D., & Bowker, D. K. (1999). Principles for Constructing Web Surveys. *SESRC Technical Report 98-50, Pullman, Washington*. Retrieved 02.12.2001 from <http://survey.sesrc.wsu.edu/dillman/papers/websurveyppr.pdf>.
- Dohi, I., & Hirokawa, K. (2004). Development of CAS (Communion-Agency Scale): Measurement of positive and negative aspects of gender personality. *Japanese Journal of Psychology*, 75, 420-427.
- Doise, W., Deschamps, J. C., & Meyer, G. (1978). The accentuation of intra-category similarities. In H. Tajfel (Hrsg.), *Differentiation between social groups* (S. 159-168). London: Academic Press.
- Doss, B. D., & Hopkins, J. R. (1998). The Multicultural Masculinity Ideology Scale: Validation From Three Cultural Perspectives. *Sex Roles*, 38, 719-742.
- Doyle, J. A., & Paludi, M. A. (1995). *Sex and gender: The human experience* (3rd ed.). Madison: Brown & Benchmark.
- Eagly, A. H. (1983). Gender and social influence: A social psychological analysis. *American Psychologist*, 38, 971-981.

- Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social Role Interpretation*. Hillsdale: Lawrence Erlbaum Associates.
- Eagly, A. H., & Diekmann, A. B. (1997). The accuracy of gender stereotypes: A dilemma for feminism. *Revue Internationale de Psychologie Sociale/International Review of Social Psychology*, 10, 11-30.
- Eagly, A. H., & Kite, M. E. (1987). Are stereotypes of nationalities applied to both women and men? *Journal of Personality and Social Psychology*, 53, 451-462.
- Eagly, A. H., Makhijani, M. G., & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111, 3-22.
- Eagly, A. H., & Mladinic, A. (1989). Gender stereotypes and attitudes toward women and men. *Personality and Social Psychology Bulletin*, 15, 543-558.
- Eagly, A. H., & Mladinic, A. (1994). Are People Prejudiced Against Women? Some Answers From Research on Attitudes, Gender Stereotypes, and Judgments of Competence. *European Review of Social Psychology*, 5, 1-35.
- Eagly, A. H., Mladinic, A., & Otto, S. (1994). Cognitive and affective bases of attitudes toward social groups and social policies. *Journal of Experimental Social Psychology*, 30, 113-137.
- Eagly, A. H., & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46, 735-754.
- Eagly, A. H., & Steffen, V. J. (1986). Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin*, 100, 309-330.
- Eagly, A. H., & Steffen, V. J. (2000). Gender stereotypes stem from the distribution of women and men into social roles. In C. Stangor (Hrsg.), *Stereotypes and prejudice: Essential readings* (S. 142-160). New York: Psychology Press.
- Eagly, A. H., & Wood, W. (1982). Inferred sex differences in status as a determinant of gender stereotypes about social influence. *Journal of Personality and Social Psychology*, 43, 915-928.
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, 54, 408-423.
- Eagly, A. H., Wood, W., & Diekmann, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 123-174). Mahwah: Lawrence Erlbaum Associates Publishers.
- Eagly, A. H., Wood, W., & Johannesen-Schmidt, M. C. (2004). Social Role Theory of Sex Differences and Similarities: Implications for the Partner Preferences of Women and Men. In A. H. Eagly, A. E. Beall, R. J. Sternberg, A. H. Eagly, A. E. Beall, & R. J. Sternberg (Hrsg.), *The psychology of gender (2nd ed.)* (S. 269-295). New York: Guilford Press.
- Eaton, W. O. (1983). Gender understanding and sex-role stereotyping in preschoolers: Implications for caregivers. *Child Care Quarterly*, 12, 28-35.
- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L. et al. (1983). In J. T. Spence (Hrsg.), *Achievement and achievement motivation* (S. 75-146). San Francisco: Freeman.
- Eckes, T. (1994a). Explorations in gender cognition: Content and structure of female and male subtypes. *Social Cognition*, 12, 37-60.
- Eckes, T. (1994b). Features of men, features of women: Assessing stereotypic beliefs about gender subtypes. *British Journal of Social Psychology*, 33, 107-123.
- Eckes, T., Trautner, H. M., & Behrendt, R. (2005). Gender Subgroups and Intergroup Perception: Adolescents' Views of Own-Gender and Other-Gender Groups. *Journal of Social Psychology*, 145, 85-111.
- Edwards, A. L. (1953). The relationship between the judged desirability of a trait and the probability that the trait will be endorsed. *Journal of Applied Psychology*, 37, 90-93.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. New York: Dryden.
- Edwards, A. L., & Ashworth, C. D. (1977). A replication study of item selection for the Bem Sex Role Inventory. *Applied Psychological Measurement*, 1, 501-507.
- Edwards, V. J., & Spence, J. T. (1987). Gender-related traits, stereotypes, and schemata. *Journal of Personality and Social Psychology*, 53, 146-154.
- Ehrenberg, K., Cataldegirmen, H., & Klauer, K. C. (2001). Valenz und Geschlechtstypikalität von 330 Verhaltensbeschreibungen: Eine Normierung für studentische Stichproben. *Zeitschrift für Sozialpsychologie*, 32, 13-28.

- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, *94*, 100-131.
- Ellis, L. J., & Bentler, P. M. (1973). Traditional sex-determined role standards and sex stereotypes. *Journal of Personality and Social Psychology*, *25*, 28-34.
- Engeser, S. (2005). Messung des expliziten Leistungsmotivs: Kurzform der Achievement Motives Scale. Retrieved 10.08.2010 from [http://www.uni-trier.de/fileadmin/fb1/prof/PSY/PGA/bilder/Engeser\\_2005\\_Kurzform\\_der\\_AMS.pdf](http://www.uni-trier.de/fileadmin/fb1/prof/PSY/PGA/bilder/Engeser_2005_Kurzform_der_AMS.pdf).
- Erdman, H., Klein, M. H., & Greist, J. H. (1983). The reliability of a computer interview for drug use/abuse information. *Behavior Research Methods and Instrumentation*, *15*, 66-68.
- Erikson, E. H. (1950). *Childhood and society*. Oxford: Norton & Co.
- Evans, R. G., & Dinning, W. D. (1982). MMPI correlates of the Bem Sex Role Inventory and Extended Personal Attributes Questionnaire in a male psychiatric sample. *Journal of Clinical Psychology*, *38*, 811-815.
- Eysenck, H. J. (1988). The biological basis of intelligence. In S. H. Irvine, J. W. Berry, S. H. Irvine, & J. W. Berry (Hrsg.), *Human abilities in cultural context* (S. 87-104). New York: Cambridge University Press.
- Fagot, B. I., Rodgers, C. S., & Leinbach, M. D. (2000). Theories of gender socialization. In T. Eckes, & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 65-89). Mahwah: Lawrence Erlbaum Associates Publishers.
- Fahrenberg, J., Hampel, R., & Selg, H. (2001). *Das Freiburger Persönlichkeitsinventar FPI. Revidierte Fassung FPI-R und teilweise geänderte Fassung FPI-A1. Handanweisung. 7.Aufl.* Göttingen: Hogrefe.
- Feather, N. T. (1978). Factor structure of the BEM Sex-Role Inventory: Implications for the study of masculinity, femininity, and androgyny. *Australian Journal of Psychology*, *30*, 241-254.
- Feather, N. T. (1982). Actions in relation to expected consequences: An overview of a research program. In N. T. Feather (Hrsg.), *Expectations and actions: Expectancy-value models in psychology* (S. 53-95). Hillsdale: Erlbaum.
- Feather, N. T. (1984). Masculinity, femininity, psychological androgyny, and the structure of values. *Journal of Personality and Social Psychology*, *47*, 604-620.
- Feather, N. T., & O'Brien, G. E. (1987). Looking for employment: An expectancy-valence analysis of job-seeking behaviour among young people. *British Journal of Psychology*, *78*, 251-272.
- Feinman, S. (1981). Why is cross-sex-role behavior more approved for girls than for boys? A status characteristic approach. *Sex Roles*, *7*, 289-300.
- Feinman, S. (1984). A status theory of the evaluation of sex-role and age-role behavior. *Sex Roles*, *10*, 445-456.
- Ferring, D., Boll, T., & Neumann, B. (2002). Soziale Vergleiche zwischen Geschwistern im frühen Erwachsenenalter: Eine explorative Studie. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *34*, 174-183.
- Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*, 117-140.
- Fiebert, M. S., & Meyer, M. W. (1997). Gender stereotypes: A bias against men. *Journal of Psychology: Interdisciplinary and Applied*, *131*, 407-410.
- Filipp, S. H. (1980). Development of self-concepts. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *12*, 105-125.
- Fischer, J., & Leonard, R. N. (1981). Sex-Role Development in Late Adolescence and Adulthood. *Sex Roles*, *7*(2), 97-106.
- Fishbein, M. (1963). An investigation of the relationship between beliefs about an object and the attitude toward that object. *Human Relations*, *16*, 233-239.
- Fishbein, M. (1967). A Behavioral Theory Approach to the Relations between Beliefs about an Object and the Attitude Toward the Object. In M. Fishbein (Hrsg.), *Readings in Attitude Theory and Measurement* (S. 389-400). New York: Wiley.
- Fishbein, M. (2007). A reasoned action approach: Some issues, questions, and clarifications. In I. Ajzen, D. Albarracín, & R. Hornik (Hrsg.), *Prediction and change of health behavior: Applying the reasoned action approach* (S. 281-295). Mahwah: Lawrence Erlbaum Associates Publishers.
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading: Addison-Wesley.
- Fishbein, M., & Middlestadt, S. (1995). Noncognitive effects on attitude formation and change: Fact or artifact? *Journal of Consumer Psychology*, *4*, 181-202.



- Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889-906.
- Fiske, S. T., & Stevens, L. E. (1993). What's so special about sex? Gender stereotyping and discrimination. In S. Oskamp, & M. Costanzo (Hrsg.), *Gender issues in contemporary society* (S. 173-196). Thousand Oaks: Sage Publications, Inc.
- Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis)respecting versus (Dis)liking: Status and Interdependence Predict Ambivalent Stereotypes of Competence and Warmth. *Journal of Social Issues*, 55, 473-489.
- Fleming, J. S., & Courtney, B. E. (1984). The dimensionality of self-esteem: II. Hierarchical facet model for revised measurement scales. *Journal of Personality and Social Psychology*, 46, 404-421.
- Fling, S., & Manosevitz, M. (1972). Sex typing in nursery school children's play interests. *Developmental Psychology*, 7, 146-152.
- Fontayne, P., Sarrazin, P., & Famose, J. P. (2000). The Bem Sex-Role Inventory: Validation of a short version for French teenagers. *European Review of Applied Psychology/Revue Europeenne de Psychologie Appliquee*, 50, 405-416.
- Forschungsgruppe Wahlen Online (2008). *Internet-Strukturdaten*.
- Freeman, H. R. (1987). Structure and Content of Gender Stereotypes: Effects of Somatic Appearance and Trait Information. *Psychology of Women Quarterly*, 11, 59-68.
- Freud, S. (1927). Some psychological consequences of the anatomical distinction between the sexes. *International Journal of Psycho-Analysis*, 8, 133-142.
- Fritz, H. L. (2000). Gender-linked personality traits predict mental health and functional status following a first coronary event. *Health Psychology*, 19, 420-428.
- Fritz, H. L., & Helgeson, V. S. (1998). Distinctions of Unmitigated Communion From Communion: Self-Neglect and Overinvolvement With Others. *Journal of Personality and Social Psychology*, 75, 121-140.
- Funder, D. C., & Colvin, C. R. (1991). Explorations in behavioral consistency: Properties of persons, situations, and behaviors. *Journal of Personality and Social Psychology*, 60, 773-794.
- Gabriel, S., & Gardner, W. L. (1999). Are there 'his' and 'hers' types of interdependence? The implications of gender differences in collective versus relational interdependence for affect, behavior, and cognition. *Journal of Personality and Social Psychology*, 77, 642-655.
- Gackenbach, J. (1981). Sex-role identity across two cultures. *Psychological Reports*, 49, 677-678.
- Gage, N. L., & Berliner, D. C. (1986). *Pädagogische Psychologie (4. Auflage)*. Weinheim und München: Psychologie Verlags Union.
- Gage, N. L., & Berliner, D. C. (1992). *Educational psychology (5th ed.)*. Boston: Houghton, Mifflin and Company.
- Gall, M. D. (1969). The relationship between masculinity-femininity and manifest anxiety. *Journal of Clinical Psychology*, 25, 294-295.
- Gana, K. (1995). Androgynie psychologique et valeurs socio-cognitives des dimensions du concept de soi. *Cahiers Internationaux de Psychologie Sociale*, 25, 27-43.
- Gaudreau, P. (1977). Factor analysis of the Bem Sex-Role Inventory. *Journal of Consulting and Clinical Psychology*, 45, 299-302.
- Ghaed, S. G., & Gallo, L. C. (2006). Distinctions Among Agency, Communion, and Unmitigated Agency and Communion According to the Interpersonal Circumplex, Five-Factor Model, and Social-Emotional Correlates. *Journal of Personality Assessment*, 86, 77-88.
- Gjesme, T., & Nygard, R. (1970). *Achievement-related motives: Theoretical considerations and construction of a measuring instrument*. Unpublished manuscript: University of Oslo.
- Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should We Trust Web-Based Studies? A Comparative Analysis of Six Preconceptions About Internet Questionnaires. *American Psychologist*, 59, 93-104.
- Götttert, R., & Kuhl, J. (1980). *LM-Fragebogen: Deutsche Übersetzung der AMS-Scale von Gjesme und Nygard. Unveröffentlichtes Manuskript*. Bochum, Psychologisches Institut der Ruhr-Universität.
- Gough, H. G. (1957). *Manual for the California Psychological Inventory*. Palo Alto: Consulting Psychologists Press.
- Green, B. L., & Kenrick, D. T. (1994). The attractiveness of gender-typed traits at different relationship levels: Androgynous characteristics may be desirable after all. *Personality and Social Psychology Bulletin*, 20, 244-253.

- Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual Review of Psychology, 54*, 461-490.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25.
- Grimmel, D. W., & Stern, G. S. (1992). The relationship between gender role ideals and psychological well-being. *Sex Roles, 27*, 487-497.
- Gross, N., Mason, W. S., & McEachern, A. W. (1958). *Explorations in role analysis: Studies of the school superintendency role*. Oxford: Wiley.
- Gross, R., Batlis, N. C., Small, A. C., & Erdwins, C. (1979). Factor structure of the Bem Sex-Role Inventory and the Personal Attributes Questionnaire. *Journal of Consulting and Clinical Psychology, 47*, 1122-1124.
- Guilford, J. P., & Zimmerman, W. S. (1956). Fourteen dimensions of temperament. *Psychological Monographs, 70*, 11-24.
- Gurman, E. B., Wekselberg, V., & Long, K. (1993). Work Role Orientation: Conceptualization and Measurement. *Journal of Social Behavior and Personality, 8*, 621-634.
- Hager, W., & Hasselhorn, M. (1994). *Handbuch deutschsprachiger Wortnormen*. Göttingen: Hogrefe.
- Hager, W., Schmuck, P., & Mecklenbräuker, S. (1996). Soziale Erwünschtheit von Eigenschaften bei Frauen und bei Männern. *Zeitschrift für Experimentelle Psychologie, 43*, 22-39.
- Hannover, B. (1997). Zur Entwicklung des geschlechtsrollenbezogenen Selbstkonzepts: Der Einfluß 'maskuliner' and 'femininer Tätigkeiten' auf die Selbstbeschreibung mit instrumentellen und expressiven Personeneigenschaften. *Zeitschrift für Sozialpsychologie, 28*, 60-75.
- Hannover, B. (2000). Development of the Self in Gendered Contexts. In T. Eckes, & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 177-206). Mahwah: Lawrence Erlbaum Associates Publishers.
- Hardie, E. A., & McMurray, N. E. (1992). Self-stereotyping, sex role ideology, and menstrual attitudes: A social identity approach. *Sex Roles, 27*, 17-37.
- Hargreaves, D. J. (1976). What are little boys and girls made of? *New Society, 37*, 542-544.
- Hargreaves, D. J. (1986). Psychological Theories Of Sex-Role Stereotyping. In D. J. Hargreaves, & A. M. Colley (Hrsg.), *The Psychology of Sex Roles* (S. 27-44). London: Harper & Row, Publishers.
- Hargreaves, D. J., & Colley, A. M. (1986). *The Psychology of Sex Roles*. London: Harper & Row, Publishers.
- Harris, A. C. (1994). Ethnicity as a determinant of sex role identity: A replication study of item selection for the Bem Sex Role Inventory. *Sex Roles, 31*, 241-273.
- Harris, M. (1993). The Evolution of Human Gender Hierarchies: A Trial Formulation. In B. Miller (Hrsg.), *Sex and Gender Hierarchies* (S. 57-79). Cambridge: Cambridge University Press.
- Harris, T. L., & Schwab, R. (1979). Personality characteristics of androgynous and sex-typed females. *Journal of Personality Assessment, 43*, 614-616.
- Harris, T. L., & Schwab, R. (1990). Sex-role orientation and personal adjustment. *Journal of Social Behavior and Personality, 5*, 473-479.
- Hartung, C. M., & Widiger, T. A. (1998). Gender differences in the diagnosis of mental disorders: Conclusions and controversies of the DSM-IV. *Psychological Bulletin, 123*, 260-278.
- Hathaway, S. R., & McKinley, J. C. (1951). *MMPI manual*. New York: Psychological Corporation.
- Havighurst, R. J. (1983). Sex role development. *Journal of Research & Development in Education, 16*, 60-65.
- Hecheltjen, K.-G., & Mertesdorf, F. (1973). Entwicklung eines mehrdimensionalen Stimmungsfragebogens. *Gruppendynamik, 40*, 110-122.
- Heckhausen, H. (1977a). Achievement motivation and its constructs: A cognitive model. *Motivation and Emotion, 1*, 283-329.
- Heckhausen, H. (1977b). Kognitionspsychologische Aufspaltung eines summarischen Konstrukts. *Psychologische Rundschau, 28*, 175-189.
- Heckhausen, H. (1989). *Motivation und Handeln (2.Aufl.)*. Berlin: Springer.
- Heilbrun, A. B. (1976). Measurement of masculine and feminine sex role identities as independent dimensions. *Journal of Consulting and Clinical Psychology, 44*, 183-190.

- Heilbrun, A. B. (1981). *Human sex-role behavior*. New York: Pergamon Press.
- Helgeson, V. S. (1993). Implications Of Agency And Communion For Patient And Spouse Adjustment To A First Coronary Event. *Journal of Personality and Social Psychology*, *64*, 807-816.
- Helgeson, V. S. (1994a). Prototypes and dimensions of masculinity and femininity. *Sex Roles*, *31*, 653-682.
- Helgeson, V. S. (1994b). Relation of agency and communion to well-being: Evidence and potential explanations. *Psychological Bulletin*, *116*, 412-428.
- Helgeson, V. S. (2003). Unmitigated Communion and Adjustment to Breast Cancer: Associations and Explanations. *Journal of Applied Social Psychology*, *33*, 1643-1661.
- Helgeson, V. S., Escobar, O., Siminerio, L., & Becker, D. (2007). Unmitigated Communion and Health Among Adolescents With and Without Diabetes: The Mediating Role of Eating Disturbances. *Personality and Social Psychology Bulletin*, *33*, 519-536.
- Helgeson, V. S., & Fritz, H. L. (1996). Implications of unmitigated communion and communion for adolescent adjustment to Type I diabetes. *Women's Health: Research on Gender, Behavior, and Policy*, *2*, 163-188.
- Helgeson, V. S., & Fritz, H. L. (1998). A Theory of Unmitigated Communion. *Personality and Social Psychology Review*, *2*, 173-183.
- Helgeson, V. S., & Fritz, H. L. (1999). Unmitigated agency and unmitigated communion: Distinctions from agency and communion. *Journal of Research in Personality*, *33*, 131-158.
- Helgeson, V. S., & Fritz, H. L. (2000). The Implications of Unmitigated Agency and Unmitigated Communion for Domains of Problem Behavior. *Journal of Personality*, *68*, 1031-1057.
- Helgeson, V. S., & Lepore, S. J. (1997). Men's adjustment to prostate cancer: The role of agency and unmitigated agency. *Sex Roles*, *37*, 251-267.
- Helmreich, R. L. (1982). On the distinction between sex-role attitudes and sex-linked traits and their stability. In P. W. Berman, & E. R. Ramey (Hrsg.), *Women: A Developmental Perspective* (S. 157-163). Washington: GPO.
- Helmreich, R. L., Spence, J. T., & Wilhelm, J. A. (1981). A psychometric analysis of the Personal Attributes Questionnaire. *Sex Roles*, *7*, 1097-1108.
- Helson, R., Kwan, V. S. Y., John, O. P., & Jones, C. (2002). The growing evidence for personality change in adulthood: Findings from research with personality inventories. *Journal of Research in Personality*, *36*, 287-306.
- Henning-Thurau, T., & Dallwitz-Wegner, D. (2002). Online-Befragungen - Ein neues Instrument für die Marktforschung. *Wirtschaftswissenschaftliches Studium*, *31*, 309-314.
- Hewson, C. M., Laurent, D., & Vogel, C. M. (1996). Proper methodologies for psychological and sociological studies conducted via the Internet. *Behavior Research Methods, Instruments and Computers*, *28*, 186-191.
- Hewstone, M., & Young, L. (1988). Expectancy-value models of attitude: Measurement and combination of evaluations and beliefs. *Journal of Applied Social Psychology*, *18*, 958-971.
- Heyman, G. D., & Giles, J. W. (2004). Valence Effects in Reasoning About Evaluative Traits. *Merrill-Palmer Quarterly*, *50*, 86-109.
- Higgins, E. T. (1990). Personality, social psychology, and person-situation relations: Standards and knowledge activation as a common language. In L. A. Pervin, & L. A. Pervin (Hrsg.), *Handbook of personality: Theory and research* (S. 301-338). New York: Guilford Press.
- Hines, M. (2004). *Brain gender*. New York: Oxford University Press.
- Hoffman, R. M. (2001). The Measurement of Masculinity and Femininity: Historical Perspective and Implications for Counseling. *Journal of Counseling & Development*, *79*, 472-485.
- Hofstede, G. (1980). *Culture Consequences: International Differences in Work-related Values*. London: Sage Publications.
- Hofstede, G. (1994). *Uncommon sense about organisations, cases, studies and field observations*. London: Sage.
- Hofstede, G. (1998). *Masculinity & Femininity: The Taboo of National Cultures*. Thousand Oaks: Sage Publications.
- Hogan, H. W. (1979). German and American responses to the Bem Sex Role Inventory. *Journal of Social Psychology*, *109*, 141-142.
- Hogg, M. A. (2003). *Social identity*. New York: The Guilford Press.
- Hogg, M. A., & Abrams, D. (1988). *Social identifications: A social psychology of intergroup relations and group processes*. Florence: Taylor & Frances/Routledge.

- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology, 26*, 325-340.
- Hogg, M. A., & Vaughan, G. M. (2008). *Social psychology (5th edition)*. London: Prentice Hall.
- Holahan, C. K., & Spence, J. T. (1980). Desirable and undesirable masculine and feminine traits in counseling clients and unselected students. *Journal of Consulting and Clinical Psychology, 48*, 300-302.
- Holt, C. L., & Ellis, J. B. (1998). Assessing the current validity of the Bem Sex-Role Inventory. *Sex Roles, 39*, 929-941.
- Hoorens, V. (1995). Self-favoring biases, self-presentation, and the self-other asymmetry in social comparison. *Journal of Personality, 63*, 793-817.
- Hörhold, M., & Klapp, B. F. (1993). Testung der Invarianz und der Hierarchie eines mehrdimensionalen Stimmungsmodells auf der Basis von Zweipunkterhebungen an Patienten- und Studentenstichproben. *Zeitschrift für medizinische Psychologie, 1*, 27-35.
- Hort, B. E., Fagot, B. I., & Leinbach, M. D. (1990). Are people's notions of maleness more stereotypically framed than their notions of femaleness? *Sex Roles, 23*, 197-212.
- Hosoda, M., & Stone, D. L. (2000). Current gender stereotypes and their evaluative content. *Perceptual and Motor Skills, 90*, 1283-1294.
- Hoyt, C. L., Simon, S., & Reid, L. (2009). Choosing the best (wo)man for the job: The effects of mortality salience, sex, and gender stereotypes on leader evaluations. *Leadership Quarterly, 20*, 233-246.
- Hu, L. t., & Bentler, P. M. (1999). Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria Versus. *Structural Equation Modeling, 6*, 1.
- Hunt, M. G. (1993). Expressiveness does predict well-being. *Sex Roles, 29*, 147-169.
- Huston, A. C. (1983). Sex-typing. In P. H. Mussen, & E. M. Hetherington (Hrsg.), *Handbook of Child Psychology (Vol. 4: Socialization, personality and social development)* (S. 387-468). New York: Wiley.
- Huston, A. C., & Alvarez, M. M. (1990). The socialization context of gender role development in early adolescence. In R. Montemayor, G. R. Adams, & T. P. Gulotta (Hrsg.), *From Childhood to Adolescence: A Transitional Period? Advances in adolescent development* (S. 156-179). Thousand Oaks: Sage Publications, Inc.
- Huyck, M. H. (1996). Continuities and discontinuities in gender identity. In V. L. Bengtson (Hrsg.), *Adulthood and aging: Research on continuities and discontinuities* (S. 98-121). New York: Springer.
- Hyde, J. S. (2005). The Gender Similarities Hypothesis. *American Psychologist, 60*, 581-592.
- Hyde, J. S., Krajinik, M., & Skuldt-Niederberger, K. (1991). Androgyny across the life span: A replication and longitudinal followup. *Developmental Psychology, 27*, 516-519.
- Hyde, J. S., & Rosenberg, B. G. (1980). *Half the human experience: The psychology of women (2nd Ed.)*. Lexington: D.C. Heath & Co.
- i2010 (2008). IKT-Fortschrittsbericht der Kommission. *Preparing Europe's digital future i2010 Mid-Term Review*. Retrieved 24.06.2008 from [http://ec.europa.eu/information\\_society/eeurope/i2010/mid\\_term\\_review\\_2008/index\\_en.htm](http://ec.europa.eu/information_society/eeurope/i2010/mid_term_review_2008/index_en.htm).
- Idle, T., Wood, E., & Desmarais, S. (1993). Gender role socialization in toy play situations: Mothers and fathers with their sons and daughters. *Sex Roles, 28*, 679-691.
- Jaccard, J., Brinberg, D., & Ackerman, L. J. (1986). Assessing attribute importance: A comparison of six methods. *Journal of Consumer Research, 12*, 463-468.
- Jackson, D. N. (1974). *Personality research form manual*. Goshen: Research Psychologists Press.
- Jackson, D. N., & Rushton, J. P. (2006). Males have greater g: Sex differences in general mental ability from 100,000 17- to 18-year-olds on the Scholastic Assessment Test. *Intelligence, 34*, 479-486.
- Jackson, L. A. (1985). Self-conceptions and gender role: The correspondence between gender-role categorization and open-ended self-descriptions. *Sex Roles, 13*, 549-566.
- Jackson, L. A., & Cash, T. F. (1985). Components of gender stereotypes: Their implications for inferences on stereotypic and nonstereotypic dimensions. *Personality and Social Psychology Bulletin, 11*, 326-344.
- Jacobson, L. I., Kellogg, R. W., Cauce, A., & Slavin, R. S. (1977). A multidimensional social desirability inventory. *Bulletin of the Psychonomic Society, 9*, 109-110.
- Jansen, W. H. M. (1997). *The Dynamics of Gender: Transformation context and design*. University of Nijmegen: Centre for Women's Studies.

- Joinson, A. N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. *European Journal of Social Psychology, 31*, 177-192.
- Kagan, J. (1964). Acquisition and Significance of Sex Typing and Sex Role Identity. In M. L. Hoffman, & L. W. Hoffman (Hrsg.), *Review of child development research* (S. 137-167). New York: Russell Sage.
- Kanter, R. M. (1997). *Frontiers of Management*. Boston: Harvard Business School Press.
- Karlin, M., Coffman, T. L., & Walters, G. (1969). On the fading of social stereotypes: Studies in three generations of college students. *Journal of Personality and Social Psychology, 13*, 1-16.
- Karniol, R., Gabay, R., Ochion, Y., & Harari, Y. (1998). Is gender or gender-role orientation a better predictor of empathy in adolescence? *Sex Roles, 39*, 45-59.
- Katenbrink, J. (2006). Translation and standardization of the Sex-Role Egalitarianism Scale (SRES-B) on a German sample. *Sex Roles, 54*, 485-493.
- Katsurada, E., & Sugihara, Y. (1999). A preliminary validation of the Bem Sex Role Inventory in Japanese culture. *Journal of Cross-Cultural Psychology, 30*, 641-645.
- Kelly, J. A., Caudill, M. S., Hathorn, S., & O'Brien, C. G. (1977). Socially undesirable sex-correlated characteristics: Implications for androgyny and adjustment. *Journal of Consulting and Clinical Psychology, 45*, 1185-1186.
- Kelly, J. A., & Worell, J. (1977). New formulations of sex roles and androgyny: A critical review. *Journal of Consulting and Clinical Psychology, 45*, 1101-1115.
- Kent, R. L., & Moss, S. E. (1994). Effects of sex and gender role on leader emergence. *Academy of Management Journal, 37*, 1335-1346.
- Kimura, D. (1992). Weibliches und männliches Gehirn. *Spektrum der Wissenschaft, 11/92*, 104-113.
- Kirchmeyer, C. (1996). Gender roles and decision-making in demographically diverse groups: A case for reviving androgyny. *Sex Roles, 34*, 649-663.
- Kite, M. E., & Deaux, K. (1986). Gender versus category clustering in free recall: A test of gender schema theory. *Representative Research in Social Psychology, 16*, 38-43.
- Kite, M. E., Deaux, K., & Haines, E. L. (2008). Gender stereotypes. In F. L. Denmark, & M. A. Paludi (Hrsg.), *Psychology of women: A handbook of issues and theories (2nd ed.)* (S. 205-236). Westport: Praeger Publishers/Greenwood Publishing Group.
- Klapprott, J. (1994). Erwünschtheit und Bedeutung von 338 alltagspsychologischen Eigenschaftsbegriffen. In W. Hager, & M. Hasselhorn (Hrsg.), *Handbuch deutschsprachiger Wortnormen* (S. 292-309). Göttingen: Hogrefe.
- Klingenspor, B. (1994a). Gender identity and bulimic eating behavior. *Sex Roles, 31*, 407-431.
- Klingenspor, B. (1994b). Geschlecht, soziale Identität und bulimisches Essverhalten. *Zeitschrift für Sozialpsychologie, 25*, 108-125.
- Klingenspor, B. (2002). Gender-related self-discrepancies and bulimic eating behavior. *Sex Roles, 47*, 51-64.
- Klingenspor, B., & Rastetter, D. (2004). Geschlechtsspezifische Identitätsentwicklung und bulimisches Essverhalten bei Jugendlichen. *Zeitschrift für Sozialpsychologie, 35*, 67-82.
- Knapp, F., & Heidlingsfelder, M. (2001). Drop-Out Analysis: Effects of the Survey Design. In U.-D. Reips, & M. Bosnjak (Hrsg.), *Dimensions of Internet Science* (S. 221-230). Berlin: Pabst.
- Knapp, H., & Kirk, S. A. (2003). Using pencil and paper, Internet and touch-tone phones for self-administered survey: does methodology matter? *Computers in Human Behavior, 19*, 117-134.
- Koestner, R., & Aube, J. (1995). A multifactorial approach to the study of gender characteristics. *Journal of Personality, 63*, 681-710.
- Kohlberg, L. (1966). A Cognitive-Developmental Analysis of Children's Sex Role Concepts and Attitudes. In E. E. Maccoby (Hrsg.), *The Development of Sex Differences* (S. 82-173). Stanford: Stanford University Press.
- Kohlberg, L., & Ullian, D. Z. (1974). Stages in the development of psychosexual concepts and attitudes. In R. C. Friedman, R. M. Richart, R. L. Vande Wiele, L. O. Stern, R. C. Friedman, R. M. Richart, R. L. Vande Wiele, & L. O. Stern (Hrsg.), *Sex differences in behavior* (S. 209-222). Oxford: John Wiley & Sons.
- Konrad, A. M., & Harris, C. (2002). Desirability of the Bem Sex-Role Inventory items for women and men: A comparison between African Americans and European Americans. *Sex Roles, 47*, 259-271.
- Korabik, K. (1999). Sex and gender in the new millennium. In G. N. Powell, & G. N. Powell (Hrsg.), *Handbook of gender and work* (S. 3-16). Thousand Oaks: Sage Publications, Inc.

- Krahé, B., Berger, A., & Möller, I. (2007). Entwicklung und Validierung eines Inventars zur Erfassung des Geschlechtsrollen-Selbstkonzepts im Jugendalter. *Zeitschrift für Sozialpsychologie*, *38*, 195-208.
- Krahé, B., & Herrmann, J. (2003). Verfälschungstendenzen im NEO-FFI: Eine experimentelle Überprüfung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *24*, 105-117.
- Krampen, G. (1979). Eine Skala zur Messung der normativen Geschlechtsrollenorientierung (GRO-Skala). *Zeitschrift für Soziologie*, *8*, 254-266.
- Krampen, G. (1983). Eine Kurzform der Skala zur Messung normativer Geschlechtsrollen-Orientierungen. *Zeitschrift für Soziologie*, *12*, 152-156.
- Krantz, J. H., & Dalal, R. (2000). Validity of Web-based Psychological Research. In M. H. Birnbaum (Hrsg.), *Psychological Experiments on the Internet* (S. 35-60). San Diego: Academic Press.
- Kühnen, U., & Hannover, B. (2003). Kultur, Selbstkonzept und Kognition. *Zeitschrift für Psychologie mit Zeitschrift für angewandte Psychologie und Sprache & Kognition*, *211*, 212-224.
- LaBrie, R. A., Shaffer, H. J., LaPlante, D. A., & Wechsler, H. (2003). Correlates of College Student Gambling in the United States. *Journal of American College Health*, *52*, 53-62.
- Landesbeauftragte für den Datenschutz und für das Recht auf Akteneinsicht Brandenburg (2005). Nutzung der Initialen zum Zweck der Pseudonymisierung oder Anonymisierung. *13. Tätigkeitsbericht 2004/2005*. Retrieved 22.06.2008 from [http://www.lida.brandenburg.de/sixcms/detail.php?qsid=lbm1.c.336188.de&template=allgemeintb13\\_lida](http://www.lida.brandenburg.de/sixcms/detail.php?qsid=lbm1.c.336188.de&template=allgemeintb13_lida).
- LaPiere, R. T. (1934). Attitudes vs. Actions. *Social Forces*. *Social Forces*, *13*, 230-237.
- Lara-Cantu, M. A., & Navarro-Arias, R. (1986). Positive and negative factors in the measurement of sex roles: Findings from a Mexican sample. *Hispanic Journal of Behavioral Sciences*, *8*, 143-155.
- Lara-Cantu, M. A., & Navarro-Arias, R. (1987). Self-descriptions of Mexican college students in response to the Bem Sex Role Inventory and other sex role items. *Journal of Cross-Cultural Psychology*, *18*, 331-344.
- Larsen, R. J., & Seidman, E. (1986). Gender schema theory and sex role inventories: Some conceptual and psychometric considerations. *Journal of Personality and Social Psychology*, *50*, 205-211.
- Laux, L., Glanzmann, P., Schaffner, P., & Spielberger, C. D. (1981). *State-Trait-Angstinventar (STAI)*. Weinheim: Beltz.
- Leibetseder, M., Mackinger, H., Hammerer, A., Staudinger, E., & Vogl, B. (1996). Geschlechtsunterschiede bei der Bewertung von Substantiva, Adjektiva und Verba. *Zeitschrift für Experimentelle Psychologie*, *43*, 585-599.
- Lewis, M., & Weinraub, M. (1979). Origins of Early Sex-Role Development. *Sex Roles*, *5*(2), 135-153.
- Li, H. Z. (2002). Culture, gender and self-close-other(s) connectedness in Canadian and Chinese samples. *European Journal of Social Psychology*, *32*, 93-104.
- Linville, P. W., & Carlston, D. E. (1994). Social cognition of the self. In P. G. Devine, D. L. Hamilton, T. M. Ostrom, P. G. Devine, D. L. Hamilton, & T. M. Ostrom (Hrsg.), *Social cognition: Impact on social psychology* (S. 143-193). San Diego: Academic Press.
- Lippa, R. A. (1991). Some Psychometric Characteristics of Gender Diagnosticity Measures: Reliability, Validity, Consistency Across Domains, and Relationship to the Big Five. *Journal of Personality and Social Psychology*, *6*, 1000-1011.
- Lippa, R. A. (1995). Gender-related Individual Differences and Psychological Adjustment in Terms of the Big Five and Circumplex Models. *Journal of Personality and Social Psychology*, *69*, 1184-1202.
- Lippa, R. A. (1997). The Display of Masculinity, Femininity, and Gender Diagnosticity in Self-Descriptive Photo Essays. *Journal of Personality*, *65*, 137-169.
- Lippa, R. A. (1998). Gender-related individual differences and the structure of vocational interests: The importance of the people-things dimension. *Journal of Personality and Social Psychology*, *74*, 996-1009.
- Lippa, R. A. (2001). On deconstructing and reconstructing masculinity-femininity. *Journal of Research in Personality*, *35*, 168-207.
- Lippa, R. A. (2002). *Gender, nature, and nurture*. Mahwah: Lawrence Erlbaum Associates.
- Lippa, R. A. (2005). Subdomains of Gender-Related Occupational Interests: Do They Form a Cohesive Bipolar M-F Dimension? *Journal of Personality*, *73*, 693-729.
- Lippa, R. A. (2006). The Gender Reality Hypothesis. *American Psychologist*, *61*, 639-640.
- Lippa, R. A., & Connelly, S. (1990). Gender diagnosticity: A new Bayesian approach to gender-related individual differences. *Journal of Personality and Social Psychology*, *59*, 1051-1065.

- Lippa, R. A., & Hershberger, S. (1999). Genetic and Environmental Influences on Individual Differences in Masculinity, Femininity, and Gender Diagnosticity: Analyzing Data From a Classic Twin Study. *Journal of Personality, 67*, 127-155.
- Lippa, R. A., & Tan, F. D. (2001). Does culture moderate the relationship between sexual orientation and gender-related personality traits? *Cross-Cultural Research: The Journal of Comparative Social Science, 35*, 65-87.
- Lobel, T. E., Slone, M., & Winch, G. (1997). Masculinity, popularity, and self-esteem among Israeli preadolescent girls. *Sex Roles, 36*, 395-408.
- Lobel, T. E., & Teiber, A. (1994). Effects of self-esteem and need for approval on affective and cognitive reactions: Defensive and true self-esteem. *Personality and Individual Differences, 16*, 315-321.
- Lockheed, M. E. (1985). Women, girls, and computers: A first look at the evidence. *Sex Roles, 13*, 115-122.
- Locksley, A., & Colten, M. E. (1979). Psychological androgyny: A case of mistaken identity? *Journal of Personality and Social Psychology, 37*, 1017-1031.
- Long, V. O. (1991). Gender role conditioning and women's self-concept. *Journal of Humanistic Counseling, Education and Development, 30*, 19-29.
- Lönqvist, J. E., Verkasalo, M., & Bezmenova, I. (2007). Agentic and communal bias in socially desirable responding. *European Journal of Personality, 21*, 853-868.
- Looy, H. (2001). Sex differences: Evolved, constructed, and designed. *Journal of Psychology & Theology, 29*, 301-313.
- Lorenzi-Cioldi, F. (1991). Self-stereotyping and self-enhancement in gender groups. *European Journal of Social Psychology, 21*, 403-417.
- Lubinski, D., Tellegen, A., & Butcher, J. N. (1981). The relationship between androgyny and subjective indicators of emotional well-being. *Journal of Personality and Social Psychology, 40*, 722-730.
- Lubinski, D., Tellegen, A., & Butcher, J. N. (1983). Masculinity, femininity, and androgyny viewed and assessed as distinct concepts. *Journal of Personality and Social Psychology, 44*, 428-439.
- Lück, H. E., & Timaeus, E. (1969). Skalen zur Messung Manifester Angst (MAS) und sozialer Wünschbarkeit (SDS-E und SDS-CM). *Diagnostica, 15*, 134-141.
- Maccoby, E. E. (1966). *The Development of Sex Differences*. Stanford: Stanford University Press.
- Maccoby, E. E., & Jacklin, C. N. (1974). *The Psychology of Sex Differences*. Stanford University Press.
- MacKenzie, S. B., Podsakoff, P. M., & Jarvis, C. B. (2005). The Problem of Measurement Model Misspecification in Behavioral and Organizational Research and Some Recommended Solutions. *Journal of Applied Psychology, 90*, 710-730.
- Maes, J., Schmitt, M., & Schmal, A. (1995). *Gerechtigkeit als innerdeutsches Problem: Werthaltungen, Kontrollüberzeugungen, Freiheitsbezeugungen, Drakonität, Soziale Einstellungen, Empathie und Protestantische Arbeitsethik als Kovariate*. Berichtaus der Arbeitsgruppe „Verantwortung, Gerechtigkeit, Moral“ Nr.85.Trier: Universität Trier, Fachbereich I - Psychologie.
- Marcell, M. M., & Falls, A. L. (2001). Online data collection with special populations over the World Wide Web. *Down Syndrome: Research and Practice, 7*, 106-123.
- Markus, H. R. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35*, 63-78.
- Markus, H. R. (1983). Self-knowledge: An expanded view. *Journal of Personality, 51*, 543-565.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review, 98*, 224-253.
- Markus, H. R., & Kitayama, S. (1998). The cultural psychology of personality. *Journal of Cross-Cultural Psychology, 29*, 63-87.
- Markus, H. R., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. In M. R. Rosenzweig, L. W. Porter, M. R. Rosenzweig, & L. W. Porter (Hrsg.), *Annual review of psychology* (S. 299-337). Palo Alto: Annual Reviews.
- Marsh, H. W. (1985). The Structure of Masculinity/Femininity: An Application of Confirmatory Factor Analysis to Higher-Order Factor Structures and Factorial Invariance. *Multivariate Behavioral Research, 20*, 427-449.
- Marsh, H. W., Antill, J. K., & Cunningham, J. D. (1987). Masculinity, Femininity, and Androgyny: Relations to Self-Esteem and Social Desirability. *Journal of Personality, 55*, 661-685.

- Marsh, H. W., & Byrne, B. M. (1991). Differentiated additive androgyny model: Relations between masculinity, femininity, and multiple dimensions of self-concept. *Journal of Personality and Social Psychology*, *61*, 811-828.
- Marsh, H. W., & Myers, M. (1986). Masculinity, femininity, and androgyny: A methodological and theoretical critique. *Sex Roles*, *14*, 397-430.
- Marshall, S. J., & Wijting, J. P. (1980). Relationships of achievement motivation and sex-role identity to college women's career orientation. *Journal of Vocational Behavior*, *16*, 299-311.
- Martin, C. L. (1991). The role of cognition in understanding gender effects. In H. W. Reese, & H. W. Reese (Hrsg.), *Advances in child development and behavior* (S. 113-149). San Diego: Academic Press.
- Martin, C. L. (1993). New directions for investigating children's gender knowledge. *Developmental Review*, *13*, 184-204.
- Martin, C. L. (2000). Cognitive theories of gender development. In T. Eckes, & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 91-121). Mahwah: Lawrence Erlbaum Associates Publishers.
- Martin, C. L., Wood, C. H., & Little, J. K. (1990). The Development Of Gender Stereotype Components. *Child Development*, *61*, 1891-1904.
- Martin, H. J., & Ramanaiah, N. V. (1988). Confirmatory factor analysis of the Bem Sex-role Inventory. *Psychological Reports*, *62*, 343-350.
- Marx, D. M., Stapel, D. A., & Muller, D. (2005). We Can Do It: The Interplay of Construal Orientation and Social Comparisons Under Threat. *Journal of Personality and Social Psychology*, *88*, 432-446.
- May, A., & Spangenberg, J. J. (1997). Androgyny and coping in men with a managerial orientation. *South African Journal of Psychology*, *27*, 244-249.
- Mazis, M. B., Ahtola, O. T., & Klippel, R. E. (1975). A comparison of four multi-attribute models in the prediction of consumer attitudes. *Journal of Consumer Research*, *2*, 38-52.
- Maznah, I., & Choo, P. F. (1986). The Factor Structure of the Bem Sex-Role Inventory (BSRI). *International Journal of Psychology*, *21*, 31-41.
- McAdams, D. P., & Vaillant, G. E. (1982). Intimacy motivation and psychosocial adjustment: A longitudinal study. *Journal of Personality Assessment*, *46*, 586-593.
- McCabe, S. E. (2002). Gender differences in collegiate risk factors for heavy episodic drinking. *Journal of Studies on Alcohol*, *63*, 49-56.
- McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. (1976). *The Achievement Motive*. New York: Irvington.
- McCrae, R. R., & Costa, P. T. (1991). The NEO Personality Inventory: Using the Five-Factor Model in counseling. *Journal of Counseling & Development*, *69*, 367-372.
- McCrae, R. R., & Costa, P. T. (1994). The stability of personality: Observation and evaluations. *Current Directions in Psychological Science*, *3*, 173-175.
- McCreary, D. R., & Korabik, K. (1994). Examining the relationships between the socially desirable and undesirable aspects of agency and communion. *Sex Roles*, *31*, 637-651.
- McGuinness, C. (1998). Cognition. In K. Trew, & J. Kremer (Hrsg.), *Gender & Psychology* (S. 66-81). London: Arnold.
- Mead, M. (1935). *Sex and Temperament in Three Primitive Societies*. Oxford: William Morrow.
- Meade, A. W., & Kroustalis, C. M. (2006). Problems With Item Parceling for Confirmatory Factor Analytic Tests of Measurement Invariance. *Organizational Research Methods*, *9*, 369-403.
- Mecklenbräuker, S., Hager, W., & Möller, H. (1994). Erwünschtheit von 908 Eigenschaften bei Frauen und Männern aus der Sicht von Männern und Frauen. In W. Hager, & M. Hasselhorn (Hrsg.), *Handbuch deutschsprachiger Wortnormen* (S. 310-328). Göttingen: Hogrefe.
- Meeker, B. F., & Weitzel-O'Neill, P. A. (1977). Sex roles and interpersonal behavior in task-oriented groups. *American Sociological Review*, *42*, 91-105.
- Miller, J. B., Lewy, J., & Peckham, E. (1997). Context effects on self-perceptions of feminine and masculine qualities. *Sex Roles*, *37*, 723-751.
- Millham, J., & Jacobson, L. I. (1978). The need for approval. In H. London, & J. J. Exner (Hrsg.), *Dimensions of Personality* (S. 365-390). New York: Wiley.
- Mischel, W. (1966). A Social-Learning View of Sex Differences in Behaviour. In Maccoby, E.E. (Ed.) *The Development of Sex Differences*. In (S. 56-81). Stanford: Stanford University Press.



- Möller, I. (2006). Mediengewalt und Aggression: Eine längsschnittliche Betrachtung des Kausalzusammenhangs am Beispiel des Konsums gewalthaltiger Bildschirmspiele. *Dissertation, Universität Potsdam*, <http://opus.kobv.de/ubp/volltexte/2006/773>.
- Moore, S. M., & Rosenthal, D. A. (1980). Sex-roles: Gender, generation, and self-esteem. *Australian Psychologist*, *15*, 467-477.
- Moosbrugger, H., & Schermelleh-Engel, K. (2007). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 307-324). Heidelberg: Springer.
- Moreland, J. R., Gulanick, N., Montague, E. K., & Harren, V. A. (1978). Some psychometric properties of the Bem Sex-Role Inventory. *Applied Psychological Measurement*, *2*, 249-256.
- Morris, M. W., Kwok, L., & Iyengar, S. S. (2004). Person perception in the heat of conflict: Negative trait attributions affect procedural preferences and account for situational and cultural differences. *Asian Journal of Social Psychology*, *7*, 127-147.
- Morrison, D. M., Leigh, B. C., & Rogers Gillmore, M. (1999). Daily Data Collection: A Comparison of Three Methods. *Journal of Sex Research*, *36*, 76-81.
- Moskowitz, D. S., Suh, E. J., & Desaulniers, J. (1994). Situational influences on gender differences in agency and communion. *Journal of Personality and Social Psychology*, *66*, 753-761.
- Muldoon, O., & Reilly, J. (1998). Biology. In K. Trew, & J. Kremer (Hrsg.), *Gender & Psychology* (S. 55-65). London: Arnold.
- Mummendey, A., & Otten, S. (1998). Positive-negative asymmetry in social discrimination. In W. Stroebe, & M. Hewstone (Hrsg.), *European Review Of Social Psychology* (S. 107-143). Chichester: Wiley.
- Mummendey, A., Otten, S., Berger, U., & Kessler, T. (2000). Positive-negative asymmetry in social discrimination: Valence of evaluation and salience of categorization. *Personality and Social Psychology Bulletin*, *26*, 1258-1270.
- Mummendey, H. D. (1981). Methoden und Probleme der Kontrolle sozialer Erwünschtheit (Social Desirability). *Zeitschrift für Differentielle und Diagnostische Psychologie*, *2*, 199-218.
- Musch, J., & Reips, U.-D. (2000). A Brief History of Web Experimenting. In M. H. Birnbaum (Hrsg.), *Psychological Experiments on the Internet* (S. 61-87). San Diego: Academic Press.
- Myers, A. M., & Gonda, G. (1982a). Empirical validation of the Bem Sex-Role Inventory. *Journal of Personality and Social Psychology*, *43*, 304-318.
- Myers, A. M., & Gonda, G. (1982b). Utility of the masculinity-femininity construct: Comparison of traditional and androgyny approaches. *Journal of Personality and Social Psychology*, *43*, 514-523.
- Nagurney, A. J. (2007). The effects of relationship stress and unmitigated communion on physical and mental health outcomes. *Stress and Health: Journal of the International Society for the Investigation of Stress*, *23*, 267-273.
- Nagurney, A. J. (2008). The effects of unmitigated communion and life events among women with fibromyalgia syndrome. *Journal of Health Psychology*, *13*, 520-528.
- Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.
- Nguyen, H. H., & Ryan, A. M. (2008). Does stereotype threat affect test performance of minorities and women? A meta-analysis of experimental evidence. *Journal of Applied Psychology*, *93*, 1314-1334.
- Novakovic, T., & Kidd, A. H. (1988). Gender self-concepts in the USA and Yugoslavia. *Psychological Reports*, *62*, 611-617.
- O'Heron, C. A., & Orlofsky, J. L. (1990). Stereotypic and nonstereotypic sex role trait and behavior orientations, gender identity, and psychological adjustment. *Journal of Personality and Social Psychology*, *58*, 134-143.
- Oakes, P. J., Haslam, S. A., & Turner, J. C. (1994). *Stereotyping and social reality*. Malden: Blackwell Publishing.
- Orlofsky, J. L., Ramsden, M. W., & Cohen, R. S. (1982). Development of the Revised Sex-Role Behavior Scale. *Journal of Personality Assessment*, *46*, 632-638.
- Orlofsky, J. L., & Stake, J. E. (1981). Psychological Masculinity and Femininity: Relationship to Striving and Self-Concept in the Achievement and Interpersonal Domains. *Psychology of Women Quarterly*, *6*, 218-233.
- Oswald, D. L., & Lindstedt, K. (2006). The content and function of gender self-stereotypes: An exploratory investigation. *Sex Roles*, *54*, 447-458.
- Pahl, S., & Eiser, J. R. (2005). Valence, Comparison Focus and Self-Positivity Biases: Does it Matter Whether People Judge Positive or Negative Traits? *Experimental Psychology*, *52*, 303-310.

- Palmgreen, P., & Rayburn, J. D. (1982). Gratifications sought and media exposure: An expectancy value model. *Communication Research, 9*, 561-580.
- Parsons, T., & Bales, R. F. (1955). *Family, Socialization and Interaction Processes*. New York: Free Press.
- Pasveer, K. A., & Ellard, J. H. (1998). The making of a personality inventory: Help from the WWW. *Behavior Research Methods, Instruments and Computers, 30*, 309-313.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.
- Paulhus, D. L. (1986). Self-description and impression management in test responses. In A. Angleitner, & J. S. Wiggins (Hrsg.), *Personality assessment via questionnaires: Current issues in theory and measurement* (S. 143-165). Berlin: Springer.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Hrsg.), *The role of constructs in psychological and educational measurement* (S. 49-69). Mahwah: Lawrence Erlbaum Associates Publishers.
- Paulhus, D. L., & Martin, C. L. (1988). Functional flexibility: A new approach to interpersonal flexibility. *Journal of Personality and Social Psychology, 55*, 88-101.
- Payne, T. J., Connor, J. M., & Colletti, G. (1987). Gender-based schematic processing: An empirical investigation and reevaluation. *Journal of Personality and Social Psychology, 52*, 937-945.
- Pedhazur, E. J., & Tetenbaum, T. J. (1979). Bem Sex Role Inventory: A theoretical and methodological critique. *Journal of Personality and Social Psychology, 37*, 996-1016.
- Peng, T. K. (2006). Construct Validation of the Bem Sex Role Inventory in Taiwan. *Sex Roles, 55*, 843-851.
- Perkins, A. W., & Forehand, M. R. (2006). Decomposing the Implicit Self-Concept: The Relative Influence of Semantic Meaning and Valence on Attribute Self-Association. *Social Cognition, 24*, 387-408.
- Perry, D. G., & Bussey, K. (1984). *Social development*. Englewood Cliffs: Prentice Hall.
- Persson, R. S. (1999). *Exploring the meaning of gender: Evaluating and revising the Bem Sex-Role Inventory (BSRI) for a Swedish research context (BSRI-SE)*. Jönköping: School of Education & Communication, Jönköping University.
- Petersen, L. E., Stahlberg, D., & Dauheimer, D. (2000). Effects of Self-Schema Elaboration on Affective and Cognitive Reactions to Self-Relevant Information. *Genetic, Social and General Psychology Monographs, 126*, 25-42.
- Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology, 29*, 137-165.
- Phoenix, J. R. (1982). Positive and negative masculinity and femininity: Relationship to self-actualization, attitudes about change as rated by self and close others. *Dissertation Abstracts International, 42*(11), 4587.
- Pickett, C. L., Bonner, B. L., & Coleman, J. M. (2002). Motivated self-stereotyping: Heightened assimilation and differentiation needs result in increased levels of positive and negative self-stereotyping. *Journal of Personality and Social Psychology, 82*, 543-562.
- Pierce, W. D., Sydie, R. A., Stratkotter, R., & Krull, C. (2003). Social Concepts and Judgments: A Semantic Differential Analysis of the Concepts Feminist, Man, and Woman. *Psychology of Women Quarterly, 27*, 338-346.
- Pleck, J. H. (1975). Masculinity-Femininity: Current and Alternative Paradigms. *Sex Roles, 1*, 161-178.
- Porst, R. (2001). Wie man die Rücklaufquote bei postalischen Befragungen erhöht. *ZUMA How-to-Reihe Nr.9*. Retrieved 17.07.2009 from [http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis\\_reihen/howto/how-to9rp.pdf](http://www.gesis.org/fileadmin/upload/forschung/publikationen/gesis_reihen/howto/how-to9rp.pdf).
- Portman, T. A. A. (2001). Sex role attributions of American-Indian women. *Journal of Mental Health Counseling, 23*, 72-84.
- Powlisha, K. K. (1995). Gender bias in children's perceptions of personality traits. *Sex Roles, 32*, 17-28.
- Prentice, D. A., & Carranza, E. (2002). What Women and Men Should Be, Shouldn't Be, Are Allowed to Be, and Don't Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly, 26*, 269-281.
- Pronin, E., Steele, C. M., & Ross, L. (2004). Identity bifurcation in response to stereotype threat: Women and mathematics. *Journal of Experimental Social Psychology, 40*, 152-168.
- Qin, L., & Yanjie, S. (2003). Revision of Bem Sex Role Inventory. *Chinese Mental Health Journal, 17*, 550-553.
- Reips, U.-D. (1997). Das psychologische Experimentieren im Internet. In B. Batinic (Hrsg.), *Internet für Psychologen* (S. 245-265). Göttingen: Hogrefe.

- Reips, U.-D. (1999). Theorie und Techniken des Web-Experimentierens. In B. Batinic, A. Werner, L. Gräf, & W. Bandilla (Hrsg.), *Online Research* (S. 277-295). Göttingen: Hogrefe.
- Reips, U.-D. (2000). The Web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Hrsg.), *Psychological experiments on the Internet* (S. 89-117). San Diego: Academic Press.
- Rhoads, K. W. (2004a). Adolescent substance use and violence: A cross-cultural comparison. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 64(9), 4631.
- Rhoads, S. E. (2004b). *Taking Sex Differences Seriously*. San Francisco: Encounter Books.
- Ricciardelli, L. A., & Williams, R. J. (1995). Desirable and undesirable gender traits in three behavioral domains. *Sex Roles*, 33, 637-655.
- Rice, T. W., & Coates, D. L. (1995). Gender Role Attitudes in the Southern United States. *Gender & Society*, 9, 744-756.
- Rietz, I., & Wahl, S. (2002). Comparison of psychologists' self image and their image in the Internet and in print. In B. Batinic, U. D. Reips, M. Bosnjak, B. Batinic, U. D. Reips, & M. Bosnjak (Hrsg.), *Online social sciences* (S. 135-149). Ashland: Hogrefe & Huber Publishers.
- Rinfret, N., & Lévesque, M. F. (2000). Comparaison transculturelle d'une analyse psychométrique du Personal Attributes Questionnaire (PAQ). *Science et Comportement*, 28, 185-202.
- Roberts, B. W., & DelVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126, 3-25.
- Roberts, B. W., & Donahue, E. M. (1994). One Personality, Multiple Selves: Integrating Personality and Social Roles. *Journal of Personality*, 62, 199-218.
- Robson, S. M., Jones, A., & Abraham, J. (2008). Personality, faking, and convergent validity: A warning concerning warning statements. *Human Performance*, 21, 89-106.
- Rose, A. J., & Montemayor, R. (1994). The relationship between gender role orientation and perceived self-competency in male and female adolescents. *Sex Roles*, 31, 579-595.
- Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *The Journal of Abnormal and Social Psychology*, 53, 367-372.
- Rosenblatt, P. C., & Cunningham, M. R. (1976). Sex Differences in Cross-Cultural Perspectives in Lloyd, B. (1976). *Exploring Sex Differences*. In B. Lloyd, & J. Archer (Hrsg.), *Exploring Sex Differences* (S. 71-94). Oxford: Academic Press.
- Rosenkrantz, P., Bee, H., Vogel, S. R., & Broverman, I. K. (1968). Sex-Role Stereotypes and Self-Concepts in College Students. *Journal of Consulting and Clinical Psychology*, 32, 287-295.
- Roth, D. L., Snyder, C. R., & Pace, L. M. (1986). Dimensions of favorable self-presentation. *Journal of Personality and Social Psychology*, 51, 867-874.
- Rotter, J. B. (1954). *Social Learning and Clinical Psychology*. New York: Prentice-Hall.
- Ruble, D. N. (1988). Sex-Role Development. In M. H. Bornstein, & M. E. Lamb (Hrsg.), *Developmental Psychology* (S. 411-460). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Ruble, D. N., & Martin, C. L. (1998). Gender development. In N. Eisenberg, & N. Eisenberg (Hrsg.), *Handbook of child psychology (5th ed.)*, *Social, emotional, and personality development* (S. 933-1016). Hoboken: John Wiley & Sons Inc.
- Ruble, D. N., & Stangor, C. (1986). Stalking the elusive schema: Insights from developmental and social-psychological analyses of gender schemas. *Social Cognition*, 4, 227-261.
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74, 629-645.
- Rudman, L. A., & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77, 1004-1010.
- Rudman, L. A., & Glick, P. (2001). Prescriptive Gender Stereotypes and Backlash Toward Agentic Women. *Journal of Social Issues*, 57, 743-762.
- Rudman, L. A., & Glick, P. (2008). *The social psychology of gender: How power and intimacy shape gender relations*. New York: Guilford Press.

- Runge, T. E. (1981). Masculine (instrumental) and feminine (expressive) traits: A comparison between students in the United States and West Germany. *Journal of Cross-Cultural Psychology, 12*, 142-162.
- Runge, T. E., Frey, J., Gollwitzer, P. M., Helmreich, R. L., & Spence, J. T. (1981). Cross-cultural stability of masculine (instrumental) and feminine (expressive) traits. *Journal of Cross-Cultural Psychology, 12*, 142-162.
- Russell, G., & Antill, J. (1984). An Australian Sex-Role Scale: Additional psychometric data and correlations with self-esteem. *Australian Psychologist, 19*, 13-18.
- Rustemeyer, R., & Fischer, N. (2007). Paradigms of self-concept research with focus on gender identity. In A. J. Lauber (Hrsg.), *Gender identity, psychology and lifestyle* (S. 1-35). Hauppauge: Nova Science Publishers.
- Sabini, J. (1992). *Social Psychology*. New York: W W Norton & Co.
- Saragovi, C., Koestner, R., Di Dio, L., & Aubé, J. (1997). Agency, communion, and well-being: Extending Helgeson's (1994) model. *Journal of Personality and Social Psychology, 73*, 593-609.
- Sassenberg, K., Kessler, T., & Mummendey, A. (2003). Less negative = more positive? Social discrimination as avoidance or approach. *Journal of Experimental Social Psychology, 39*, 48-58.
- Sawrie, S. M., Watson, P. J., & Biderman, M. D. (1991). Aggression, sex role measures, and Kohut's psychology of the self. *Sex Roles, 25*, 141-161.
- Schaarschmidt, U., & Fischer, A. (1996). *AVEM - Arbeitsbezogenes Verhaltens- und Erlebensmuster. Handanweisung*. Frankfurt: Swets & Zeitlinger.
- Schahn, J. (2002). Aufwand und soziale Erwünschtheit als Prädiktoren von Itemmittelwerten: Analysen auf Aggregat- und Individualdatenebene im Bereich Umwelteinstellung und -verhalten. *Zeitschrift für Differentielle und Diagnostische Psychologie, 23*, 45-54.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the Fit of Structural Equation Models: Tests of Significance and Descriptive Goodness-of-Fit Measures. *Methods of Psychological Research, 8*, 23-74.
- Schillewaert, N., & Meulemeester, P. (2005). Comparing response distributions of offline and online data collection methods. *International Journal of Market Research, 47*, 163-178.
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology, 38*, 194-201.
- Schmidt, W. C. (1997). World-Wide Web Survey Research: Benefits, potential problems, and solutions. *Behavior Research Methods, Instruments and Computers, 29*, 274-279.
- Schmitt, D. P., Realo, A., Voracek, M., & Allik, J. (2008). Why can't a man be more like a woman? Sex differences in Big Five personality traits across 55 cultures. *Journal of Personality and Social Psychology, 94*, 168-182.
- Schmitt, M. J., Altstötter-Gleich, C., Hinz, A., Maes, J., & Brähler, E. (2006). Normwerte für das Vereinfachte Beck-Depressions-Inventar(BDI-V) in der Allgemeinbevölkerung. *Diagnostica, 52*, 51-59.
- Schmitt, M. J., Beckmann, M., Dusi, D., Maes, J., Schiller, A., & Schonauer, K. (2003). Messgüte des vereinfachten Beck-Depressions-Inventars (BDI-V). *Diagnostica, 49*, 147-156.
- Schmitt, M. J., & Maes, J. (2000). *Vorschlag zur Vereinfachung des Beck-Depressions-Inventars (BDI)*. (46 ed.).
- Schmitt, M. J., & Steyer, R. (1993). A latent state-trait model (not only) for social desirability. *Personality and Individual Differences, 14*, 519-529.
- Schmitz, K., & Diefentahler, S. (1998). An Examination of Traditional Gender Roles Among Men and Women in Mexico and the United States. Retrieved 12.07.2008 from <http://murphylibrary.uwlax.edu/digital/jur/1998/schmitz.pdf>.
- Schneider-Düker, M. (1978). *Deutsche Neukonstruktion des Bem-Sex-Role-Inventary*. Arbeiten der Fachrichtung Psychologie Nr.51, Universität des Saarlandes, Saarbrücken.
- Schneider-Düker, M., & Kohler, A. (1988). Die Erfassung von Geschlechtsrollen: Ergebnisse zur deutschen Neukonstruktion des Bem Sex-Role Inventory. *Diagnostica, 34*, 256-270.
- Schröder, E. (2005). Berufliche Selbständigkeit als Ziel ?! : Entwicklung und Evaluation eines "Life-Skills"-basierten Trainingsprogramms für Jugendliche. *Dissertation, Universität Jena, [http://www.db-thueringen.de/servlets/DocumentServlet?name=FSU\\_Jena\\_Schroeder\\_Elke](http://www.db-thueringen.de/servlets/DocumentServlet?name=FSU_Jena_Schroeder_Elke)*.
- Schultz, D. P. (1972). The human subject in psychological research. In C. L. Sheridan (Hrsg.), *Readings for Experimental Psychology* (S. 263-282). New York: Holt.
- Schütz, A., & Sellin, I. (2006). *Multidimensionale Selbstwertskala*. Göttingen: Hogrefe.

- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Hrsg.), *Advances in Experimental Social Psychology* (25 Ed., S. 1-65). San Diego: Academic Press.
- Schwartz, S. H., Verkasalo, M., Antonovsky, A., & Sagiv, L. (1997). Value priorities and social desirability: Much substance, some style. *British Journal of Social Psychology*, *36*, 3-18.
- Schwarzer, R., Mueller, J., & Greenglass, E. (1999). Assessment of perceived general self-efficacy on the Internet : Data collection in cyberspace. *Anxiety, Stress & Coping*, *12*, 145-161.
- Sebastián, J. (1990). Las escalas de masculinidad y feminidad: presupuestos subyacentes al modelo clásico y actual. Segunda parte: el modelo actual. *Evaluación Psicológica*, *6*, 327-367.
- Sedney, M. A. (1985). Growing more complex: Conceptions of sex roles across adulthood. *International Journal of Aging and Human Development*, *22*, 15-29.
- Sethi, R. R., & Allen, M. J. (1984). Sex-role Stereotypes in Northern India and the United States. *Sex Roles*, *11*, 615-626.
- Shapiro, D. H. (1994). *Manual for the Shapiro Control Inventory*. Palo Alto: Behaviordyne.
- Shaver, P. R., Papalia, D., Clark, C. L., & Koski, L. R. (1996). Androgyny and attachment security: Two related models of optimal personality. *Personality and Social Psychology Bulletin*, *22*, 582-597.
- Shimonaka, Y., Nakazato, K., Kawaai, C., & Sato, S. (1997). Androgyny and succesful adaptation across the life span among.. *Journal of Genetic Psychology*, *158*, 389-400.
- Sieverding, M. (1999). Weiblichkeit - Männlichkeit und psychische Gesundheit. In E. Brähler, & H. Felder (Hrsg.), *Weiblichkeit, Männlichkeit und Gesundheit* (S. 31-57). Opladen: Westdeutscher Verlag.
- Sieverding, M. (2002). Gender and healthrelated attitudes: The role of a "macho" selfconcept. In G. Weidner, Kopp S.M., & M. Kristenson (Hrsg.), *Heart Disease: Environment, Stress and Gender. NATO Science Series, Series I: Life and Behavioural Sciences* (327 Ed., S. 237-250). Amsterdam: IOS Press.
- Sieverding, M., & Alfermann, D. (1992). Instrumentelles (maskulines) und expressives (feminines) Selbstkonzept: Ihre Bedeutung für die Geschlechtsrollenforschung. *Zeitschrift für Sozialpsychologie*, *23*, 6-15.
- Simon, B., & Trötschel, R. (2007). Das Selbst und die soziale Identität. In K. Jonas, W. Stroebe, & M. Hewstone (Hrsg.), *Sozialpsychologie: Eine Einführung* (S. 147-185). Heidelberg: Springer.
- Sinclair, S. A., Hardin, C. D., & Lowery, B. S. (2006). Self-Stereotyping in the Context of Multiple Social Identities. *Journal of Personality and Social Psychology*, *90*, 529-542.
- Singelis, T. M. (1994). The measurement of independent and interdependent self-construals. *Personality and Social Psychology Bulletin*, *20*, 580-591.
- Singleton, C. H. (1986). Biological and social explanations of sex-role stereotyping. In D. J. Hargreaves, & A. M. Colley (Hrsg.), *The Psychology Of Sex Roles* (S. 3-26). London: Harper & Row, Publishers.
- Skinner, H. A., & Allen, B. A. (1983). Does the computer make a difference? Computerized versus face-to-face versus self-report assessment of alcohol, drug, and tobacco use. *Journal of Consulting and Clinical Psychology*, *51*, 267-275.
- Skoe, E. E. (1995). Sex role orientation and its relationship to the development of identity and moral thought. *Scandinavian Journal of Psychology*, *36*, 235-245.
- Slaby, R. G., & Frey, K. S. (1976). Development of gender constancy and selective attention to same-sex models. *Child Development*, *46*, 849-856.
- Smart, R. (1966). Subject selection bias in psychological research. *Canadian Psychologist*, *7*, 115-121.
- Smith, C. J., Noll, J. A., & Bryant, J. B. (1999). The effect of social context on gender self-concept. *Sex Roles*, *40*, 499-512.
- Smith, E. R., & Mackie, D. M. (2000). *Social psychology (2nd ed.)*. New York: Psychology Press.
- Smith, J. L. (1996). Expectancy, value, and attitudinal semantics. *European Journal of Social Psychology*, *26*, 501-506.
- Snell, W. E., Belk, S. S., & Hawkins, R. C. (1987). Alcohol and drug use in stressful times: The influence of the masculine role and sex-related personality attributes. *Sex Roles*, *16*, 359-373.
- Sparks, P., Hedderley, D., & Shepherd, R. (1991). Expectancy-value models of attitudes: A note on the relationship between theory and methodology. *European Journal of Social Psychology*, *21*, 261-271.
- Spence, J. T. (1980). The psychological dimensions of masculinity and femininity and achievement motivation. *Annals of the New York Academy of Sciences*, *340*, 88-101.

- Spence, J. T. (1984). Masculinity, femininity and gender-related traits: A conceptual analysis and critique of current research. In B. A. Maher, & W. B. Maher (Hrsg.), *Progress in experimental research* (S. 2-66).
- Spence, J. T. (1985). *Gender identity and its implications for the concepts of masculinity and femininity*. Nebraska Symposium on Motivation 1984, Vol. 32: Psychology and Gender, Lincoln: University of Nebraska Press, 1985.
- Spence, J. T. (1991). Do the BSRI and PAQ measure the same or different concepts? *Psychology of Women Quarterly*, 15, 141-165.
- Spence, J. T. (1993). Gender-related traits and gender ideology: Evidence for a multifactorial theory. *Journal of Personality and Social Psychology*, 64, 624-635.
- Spence, J. T. (1999). Thirty years of gender research: A personal chronicle. In W. B. J. Swann, J. H. Langlois, L. A. Gilbert, W. B. J. Swann, J. H. Langlois, & L. A. Gilbert (Hrsg.), *Sexism and stereotypes in modern society: The gender science of Janet Taylor Spence* (S. 255-289). Washington: American Psychological Association.
- Spence, J. T., & Buckner, C. (1995). Masculinity and femininity: Defining the undefinable. In P. J. Kalbfleisch, M. J. Cody, P. J. Kalbfleisch, & M. J. Cody (Hrsg.), *Gender, power, and communication in human relationships* (S. 105-138). Hillsdale: Lawrence Erlbaum Associates, Inc.
- Spence, J. T., & Buckner, C. E. (2000). Instrumental and expressive traits, trait stereotypes, and sexist attitudes. *Psychology of Women Quarterly*, 24, 44-62.
- Spence, J. T., Deaux, K., & Helmreich, R. L. (1985). Sex roles in contemporary American society. In G. Lindzey, & E. Aronson (Hrsg.), *Handbook of Social Psychology* (S. 149-178). New York: Random House.
- Spence, J. T., & Helmreich, R. (1972). The Attitudes toward Women Scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. *Catalog of Selected Documents in Psychology*, 66-67.
- Spence, J. T., Helmreich, R., & Stapp, J. (1974). The Personal Attributes Questionnaire: A measure of sex role stereotypes and masculinity-femininity. *Catalog of Selected Documents in Psychology*, 4, 43-44.
- Spence, J. T., Helmreich, R., & Stapp, J. (1975). Ratings of self and peers on sex role attributes and their relation to self-esteem and conceptions of masculinity and femininity. *Journal of Personality and Social Psychology*, 32, 29-39.
- Spence, J. T., & Helmreich, R. L. (1978). *Masculinity & Femininity. Their Psychological Dimensions, Correlates & Antecedents*. University of Texas Press: Austin.
- Spence, J. T., & Helmreich, R. L. (1980). Masculine Instrumentality and Feminine Expressiveness: Their Relationships With Sex Role Attitudes and Behaviors. *Psychology of Women Quarterly*, 5, 147-163.
- Spence, J. T., & Helmreich, R. L. (1981). Androgyny versus gender schema: A comment on Bem's gender schema theory. *Psychological Review*, 88, 365-368.
- Spence, J. T., Helmreich, R. L., & Holahan, C. K. (1979). Negative and Positive Components of Psychological Masculinity and Femininity and Their Relationships to Self-Reports of Neurotic and Acting out Behaviors. *Journal of Personality and Social Psychology*, 37, 1673-1682.
- Spence, J. T., & Sawin, L. L. (1985). Images of masculinity and femininity: A reconceptualization. In V. E. O'Leary, R. K. Unger, & B. X. Wallston (Hrsg.), *Women, Gender and Social Psychology* (S. 35-66). Hillsdale: Erlbaum.
- Spörrle, M., Gerber-Braun, B., & Försterling, F. (2007). The influence of response lines on response behavior in the context of open-question formats. *Swiss Journal of Psychology/Schweizerische Zeitschrift für Psychologie/Revue Suisse de Psychologie*, 66, 103-107.
- Srocke, B. (1989). *Mädchen und Mathematik: historisch-systematische Untersuchung der unterschiedlichen Bedingungen des Mathematiklernens von Mädchen und Jungen (Habilitation)*. Wiesbaden: Deutscher Universitätsverlag.
- Stake, J. E. (1997). Integrating expressiveness and instrumentality in real-life settings: A new perspective on the benefits of androgyny. *Sex Roles*, 37, 541-564.
- Stangor, C., & Ruble, D. N. (1987). Development of gender role knowledge and gender constancy. *New Directions for Child Development*, 38, 5-22.
- Stapleton, C. D. (1997). *Basic Concepts and Procedures of Confirmatory Factor Analysis*. Paper presented at the Annual Meeting of the Southwest Educational Research Association (Austin, TX, January 23-25, 1997).
- Statistisches Bundesamt Deutschland (2008). Studierende an Hochschulen Wintersemester 2007/2008. *Vorbericht Fachserie*, 11.

- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, *52*, 613-629.
- Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, *69*, 797-811.
- Steele, C. M., & Aronson, J. A. (2004). Stereotype Threat Does Not Live by Steele and Aronson (1995) Alone. *American Psychologist*, *59*, 47-48.
- Steiner, I. D. (1972). *Group Process and Productivity*. New York: Academic Press.
- Stewart, A. J., & Healy, J. M. (1989). Linking individual development and social changes. *American Psychologist*, *44*, 30-42.
- Stöber, J. (1999). Die Soziale-Erwünschtheits-Skala-17 (SES-17): Entwicklung und erste Befunde zu Reliabilität und Validität. *Diagnostica*, *45*, 173-177.
- Stocké, V. (2004). Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit: Ein Vergleich der Prognosen der Rational-Choice Theorie und des Modells der Frame-Selektion. *Zeitschrift für Soziologie*, *33*, 303-320.
- Stoppard, J. M., & Kalin, R. (1978). Can gender stereotypes and sex-role conceptions be distinguished? *British Journal of Social & Clinical Psychology*, *17*, 211-217.
- Storms, M. D. (1979). Sex role identity and its relationships to sex role attributes and sex role stereotypes. *Journal of Personality and Social Psychology*, *37*, 1779-1789.
- Strauss, B., Köller, O., & Möller, J. (1996). Geschlechtsrollentypologien - eine empirische Prüfung des additiven und des balancierten Modells. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *17*, 67-83.
- Strickland, B. R. (1977). Approval motivation. In T. Blass (Hrsg.), *Personality variables in social behavior* (S. 315-356). Hillsdale: Erlbaum.
- Strong, E. K. J. (1943). *Vocational interests of men and women*. Stanford: University Press.
- Strough, J., Leszczynski, J. P., Neely, T. L., Flinn, J. A., & Margrett, J. (2007). From adolescence to later adulthood: Femininity, masculinity, and androgyny in six age groups. *Sex Roles*, *57*, 385-396.
- Stryker, S., & Statham, A. (1985). Symbolic interaction and role theory. In G. Lindzey, & E. Aronson (Hrsg.), *Handbook of Social Psychology* (3rd. Ed.) (S. 311-378). New York: Random House.
- Stuckless, N., & Goranson, R. (1992). The Vengeance Scale: Development of a measure of attitudes toward revenge. *Journal of Social Behavior and Personality*, *7*, 25-42.
- Stumpf, H., Angleitner, A., Wieck, T., Jackson, D. N., & Beloch-Till, H. (1985). *Deutsche Personality Research Form (PRF)*. Göttingen: Hogrefe.
- Sugihara, Y., & Katsurada, E. (2000). Gender-role Personality Traits in Japanese Culture. *Psychology of Women Quarterly*, *24*, 309-318.
- Sugihara, Y., & Warner, J. A. (1999). Endorsements by Mexican-Americans of the Bem Sex-Role Inventory: Cross-ethnic comparison. *Psychological Reports*, *85*, 201-211.
- Suh, E. J., Moskowitz, D. S., Fournier, M. A., & Zuroff, D. C. (2004). Gender and Relationships: Influences on Agentic and Communal Behaviors. *Personal Relationships*, *11*, 41-59.
- Swan, S., & Wyer, R. S. Jr. (1997). Gender stereotypes and social identity: How being in the minority affects judgment of self and others. *Personality and Social Psychology Bulletin*, *23*, 1265-1276.
- Swazina, K. R., Waldherr, K., & Maier, K. (2004). Geschlechtsspezifische Ideale im Wandel der Zeit. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *25*, 165-176.
- Swim, J. K. (1994). Perceived versus meta-analytic effect sizes: An assessment of the accuracy of gender stereotypes. *Journal of Personality and Social Psychology*, *66*, 21-36.
- Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues*, *25*, 79-97.
- Tajfel, H. (1972). La categorization sociale (English trans.). In S. Moscovici (Hrsg.), *Introduction a la Psychologie Sociale* (S. 272-302). Paris: Larousse.
- Tajfel, H. (1978). *Differentiation between social groups: Studies in the social psychology of intergroup relations*. Oxford: Academic Press.
- Tajfel, H. (1979). Individuals and groups in social psychology. *British Journal of Social & Clinical Psychology*, *18*, 183-190.

- Tajfel, H. (2001). Social stereotypes and social groups. In M. A. Hogg, & D. Abrams (Hrsg.), *Intergroup relations: Essential readings* (S. 132-145). New York: Psychology Press.
- Tajfel, H., & Forgas, J. P. (2000). Social categorization: Cognitions, values and groups. In C. Stangor (Hrsg.), *Stereotypes and prejudice: Essential readings* (S. 49-63). New York: Psychology Press.
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin, & S. Worchel (Hrsg.), *The social psychology of intergroup relations* (S. 33-48). Monterey: Brooks/Cole.
- Tajfel, H., & Turner, J. (1986). The social identity of intergroup behavior. In W. G. Austin, & S. Worchel (Hrsg.), *The social psychology of intergroup relations* (S. 7-24). Monterey: Brooks/Cole.
- Taylor, M. C., & Hall, J. A. (1982). Psychological androgyny: Theories, methods, and conclusions. *Psychological Bulletin*, 92, 347-366.
- Tepper, B. J., Brown, S. J., & Hunt, M. D. (1993). Strength of subordinates' upward influence tactics and gender congruency effects. *Journal of Applied Social Psychology*, 23, 1903-1919.
- Terman, L. M., & Miles, C. C. (1936). *Sex and Personality*. New Haven: Yale University Press.
- Thompson, B. (1989). Meta-analysis of factor structure studies: A case example with Bem's androgyny measure. *Journal of Experimental Education*, 57, 187-197.
- Thompson, E. H., Pleck, J. H., & Ferrera, D. L. (1992). Men and masculinities: Scales for masculinity ideology and masculinity-related constructs. *Sex Roles*, 27, 573-607.
- Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. New York: Appleton-Century-Crofts.
- Tolson, A. (1977). *The Limits of Masculinity: Male Identity and Women's Liberation*. New York: Harper & Row Publishers.
- Trew, K. (1998). Identity and the Self. In K. Trew, & J. Kremer (Hrsg.), *Gender & Psychology* (S. 3-14). London: Arnold.
- Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, 96, 506-520.
- Triandis, H. C. (1995). The self and social behavior in differing cultural contexts. In N. R. Goldberger, J. B. Veroff, N. R. Goldberger, & J. B. Veroff (Hrsg.), *The culture and psychology reader* (S. 326-365). New York: New York University Press.
- Triandis, H. C., & Gelfand, M. J. (1998). Converging measurement of horizontal and vertical individualism and collectivism. *Journal of Personality and Social Psychology*, 74, 118-128.
- Troche, S. J., Weber, N., Hennigs, K., Andresen, C. R., & Rammsayer, T. H. (2007). The relationship of digit ratio (2D:4D) and gender-role orientation in four national samples. *Journal of Individual Differences*, 28, 78-87.
- Trost, G. (1992). Test für medizinische Studiengänge (TMS): Studien zur Evaluation. *16.Arbeitsbericht, Kurzfassung.Bonn: Institut für Test- und Begabungsforschung der Studienstiftung des deutschen Volkes*.
- Turner, J. C. (1981). Towards a cognitive redefinition of the social group. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, 1, 93-118.
- Turner, J. C. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Hrsg.), *Social identity and intergroup relations* (S. 15-40). Cambridge: University Press.
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Cambridge: Basil Blackwell.
- Turner, J. C., & Onorato, R. S. (1999). Social identity, personality, and the self-concept: A self-categorizing perspective. In T. R. Tyler, R. M. Kramer, & O. P. John (Hrsg.), *The psychology of the social self* (S. 11-46). Mahwah: Lawrence Erlbaum Associates Publishers.
- Turner, J. C. (1978a). Social categorization and social discrimination in the minimal group paradigm. In H. Tajfel (Hrsg.), *Differentiation between social groups: studies in the social psychology of intergroup relations* (S. 101-140). London: Academic Press.
- Turner, J. C. (1978b). Social comparison, similarity and ingroup favouritism. In H. Tajfel (Hrsg.), *Differentiation between social groups: studies in the social psychology of intergroup relations* (S. 235-250). London: Academic Press.
- Tuten, T. L., Urban, D. J., & Bosnjak, M. (2002). Internet Surveys and Data Quality - A Review. In B. Batinic, U. D. Reips, & M. Bosnjak (Hrsg.), *Online Social Sciences* (S. 7-26). Seattle: Hogrefe & Huber Publishers.
- Twenge, J. M. (1997). Changes in masculine and feminine traits over time: A meta-analysis. *Sex Roles*, 36, 305-325.
- Twenge, J. M. (1999). Mapping gender: The multifactorial approach and the organization of gender-related attributes. *Psychology of Women Quarterly*, 23, 485-502.



- Twenge, J. M. (2001). Changes in women's assertiveness in response to status and roles: A cross-temporal meta-analysis, 1931-1993. *Journal of Personality and Social Psychology, 81*, 133-145.
- Twenge, J. M., & Im, C. (2007). Changes in the need for social approval, 1958-2001. *Journal of Research in Personality, 41*, 171-189.
- Ullian, D. Z. (1976). The Development of Conceptions Of Masculinity & Femininity. In B. Lloyd (Hrsg.), *Exploring Sex Differences* (S. 25-47). Oxford: Academic Press.
- van der Pligt, J., & de Vries, N. K. (1998). Belief importance in expectancy-value models of attitudes. *Journal of Applied Social Psychology, 28*, 1339-1354.
- van Wersch, A. (1998). Health and illness. In K. Trew, & J. Kremer (Hrsg.), *Gender & Psychology* (S. 167-178). London: Arnold.
- Viklund, M., & Sjöberg, L. (2008). An expectancy-value approach to determinants of trust. *Journal of Applied Social Psychology, 38*, 294-313.
- Vonk, R., & Olde-Monnikhof, M. (1998). Gender subgroups: Intergroup bias within the sexes. *European Journal of Social Psychology, 28*, 37-47.
- Vroom, V. H. (1964). *Work and Motivation*. New York: John Wiley & Sons.
- Walkup, H., & Abbott, R. D. (1978). Cross-validation of item selection on the Bem Sex Role Inventory. *Applied Psychological Measurement, 2*, 63-71.
- Wänke, M., Bless, H., & Wortberg, S. (2003). Der Einfluss von "Karrierefrauen" auf das Frauenstereotyp: Die Auswirkungen von Inklusion und Exklusion. *Zeitschrift für Sozialpsychologie, 34*, 187-196.
- Ward, C. A., & Sethi, R. R. (1986). Cross-cultural validation of the Bem Sex-Role Inventory: Malaysian and South Indian research. *Journal of Cross-Cultural Psychology, 17*, 300-314.
- Waters, C. W., Waters, L. K., & Pincus, S. (1977). Factor analysis of masculine and feminine sex-typed items from the Bem Sex-Role Inventory. *Psychological Reports, 40*, 567-570.
- Wechsler, H., Kuo, M., Lee, H., & Dowdall, G. W. (2000). Environmental correlates of underage alcohol use and related problems of college students. *American Journal of Preventive Medicine, 19*, 24-29.
- Wechsler, H., & Nelson, T. F. (2001). Binge drinking and the American college students: What's five drinks? *Psychology of Addictive Behaviors, 15*, 287-291.
- Weston, R., & Gore, P. A. J. (2006). A Brief Guide to Structural Equation Modeling. *Counseling Psychologist, 34*, 719-751.
- Wiggins, J. S. (1979). A Psychological Taxonomy Of Trait Descriptive Terms: The Interpersonal Domain. *Journal of Personality and Social Psychology, 37*, 395-412.
- Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In W. M. Grove, & D. Cicchetti (Hrsg.), *Thinking clearly about psychology: Vol. 2. Personality and Psychopathology* (S. 89-113). Minneapolis: University of Minnesota Press.
- Wiggins, J. S., & Holzmüller, A. (1978). Psychological Androgyny and Interpersonal Behavior. *Journal of Consulting and Clinical Psychology, 46*, 40-52.
- Wiggins, J. S., & Holzmüller, A. (1981). Further evidence on androgyny and interpersonal flexibility. *Journal of Research in Personality, 15*, 67-80.
- Wiggins, J. S., & Trapnell, P. D. (1996). A dyadic-interactional perspective on the Five Factor model. In J. S. Wiggins (Hrsg.), *The Five Factor model of personality: Theoretical perspectives* (S. 88-162). New York: Guilford Press.
- Wilcox, C., & Francis, L. J. (1997a). The Relationship Between Neuroticism and the Perceived Social Desirability of Feminine Characteristics among 16-19 Year Olds Females. *Social Behavior and Personality, 25*, 291-294.
- Wilcox, C., & Francis, L. J. (1997b). Beyond gender stereotyping: Examining the validity of the BEM Sex-Role Inventory among 16- to 19-year old females in England. *Personality and Individual Differences, 23*, 9-13.
- Wilde, A., & Diekmann, A. B. (2005). Cross-Cultural Similarities and Differences in Dynamic Stereotypes: A Comparison between Germany and the United States. *Psychology of Women Quarterly, 29*, 188-196.
- Williams, J. E. (1977). Sex-trait stereotypes in England, Ireland and the United States. *British Journal of Social & Clinical Psychology, 16*, 303-309.
- Williams, J. E. (1979). Sex-trait stereotypes in France, Germany, and Norway. *Journal of Cross-Cultural Psychology, 10*, 133-156.

- Williams, J. E. (1981). Traits associated with men and women: Attribution by young children in France, Germany, Norway, the Netherlands, and Italy. *Journal of Cross-Cultural Psychology, 12*, 327-346.
- Williams, J. E., & Bennett, S. (1975). The definition of sex stereotypes via the Adjective Check List. *Sex Roles, 1*, 327-337.
- Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multinational study (rev. ed.)*. Thousand Oaks: Sage Publications, Inc.
- Williams, R. J., & Ricciardelli, L. A. (2000). Relationship Between Restrained Eating and Drinking, Sex-Role Stereotypes, and Self-Control. *Journal of Social Behavior and Personality, 15*, 259-268.
- Williams, R. J., & Ricciardelli, L. A. (2001). Sex-role traits and the comorbidity of symptoms of disordered eating and problem drinking. *Eating Behaviors, 2*, 67-77.
- Wilson, D., McMaster, J., Greenspan, R., & Mboyi, L. (1990). Cross-cultural validation of the Bem Sex Role Inventory in Zimbabwe. *Personality and Individual Differences, 11*, 651-656.
- Wolff, L., & Taylor, S. E. (1979). Sex, sex-role identification, and awareness of sex-role stereotypes. *Journal of Personality, 47*, 177-184.
- Wong, F. Y., McCreary, D. R., & Duffy, K. G. (1990). A further validation of the Bem Sex Role Inventory: A multitrait-multimethod study. *Sex Roles, 22*, 249-259.
- Wood, J. V. (1989). Theory and research concerning social comparisons of personal attributes. *Psychological Bulletin, 106*, 231-248.
- Wood, W., & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin, 128*, 699-727.
- Woodhill, B. M., & Samuels, C. A. (2003). Positive and Negative Androgyny and Their Relationship with Psychological Health and Well-Being. *Sex Roles, 48*, 555-565.
- Wubbenhorst, T. M. (1994). Personality characteristics of music educators and performers. *Psychology of Music, 22*, 63-74.
- Wulff, M. B., & Steitz, J. A. (1999). A path model of the relationship between career indecision, androgyny, self-efficacy and self-esteem. *Perceptual and Motor Skills, 88*, 935-940.
- Zimet, C. N., & Zimet, S. G. (1978). Educators view people: Ethnic group stereotyping. *Journal of Community Psychology, 6*, 189-193.

## DANKSAGUNG

Mit der Fertigstellung der Dissertationsschrift ist es an der Zeit, nochmals denjenigen zu danken, die mich begleitet und unterstützt haben und zum Gelingen dieser Arbeit einen Beitrag leisteten.

An erster Stelle sei daher Frau Prof. B. Krahe für die langjährige und zielgerichtete Betreuung dieser Arbeit herzlichst gedankt. Sie hat nicht nur die Entwicklung meines Dissertationsprojektes durch kritische wie inspirierende Diskussionen bereichert, sondern mich auch im Hinblick auf meine fachliche, berufliche und persönliche Weiterentwicklung stets gefördert. Ihre wertvollen Anregungen und Ratschläge habe ich immer sehr geschätzt.

Darüber hinaus bedanke ich mich ebenfalls bei allen Mitarbeitern aus der Sozialpsychologie, die mich über die lange Zeit begleiteten und mir mit ihrem Fachwissen, ihrer konstruktiven Kritik und vielen Ideen sowie der produktiven, wissenschaftlichen Zusammenarbeit auch immer wieder den nötigen Schwung gegeben haben.

Ein besonderes Wort des Dankes gilt schließlich auch meinen fleißigen Korrekturlesern und sehr guten Freunden: Claudi, Franzi, Mario und Jule, die mich nicht nur tatkräftig unterstützt haben, sondern mich stets aufbauten und für die erforderliche Abwechslung sorgten. In denkbar kurzer Zeit haben sie zudem am Ende noch die Fehler ausgeräumt, die sich im Zuge einer langen Arbeit einschleichen und oftmals vom eigenen Auge unentdeckt bleiben. Durch ihre sehr genauen und kritischen Fragen bekam die Arbeit ihren letzten und notwendigen Schliff. So wurden aus Stolpersteinen und unverständlichen Erklärungen geeignete Überleitungen und Zusammenfassungen.

Meiner Mum bin ich für die unentwegte und gleichermaßen moralische wie tatkräftige Unterstützung sowohl während meines Studiums als auch während der Promotion zutiefst dankbar. Obgleich die Anfertigung einer Dissertationsschrift ein hohes Maß an Fleiß und Ausdauer abverlangt, war ihre Hilfe vor allem im persönlichen Bereich unersetzlich.

Schlussendlich möchte ich Michael von ganzem Herzen für seine unermüdliche Unterstützung, seine Geduld, seine Liebe und Motivation sowie Aufmunterung in den schwierigen Phasen danken.

## ERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbst verfasst habe und keine anderen Hilfsmittel und Quellen als angegeben verwendet habe.

Ich erkläre, dass die von mir vorgelegte Dissertation nicht in dieser oder einer anderen Form bereits anderweitig als Prüfungsarbeit verwendet wurde oder einer anderen Fakultät als Dissertation vorgelegen hat.

Potsdam, den 22.09.2010

\_\_\_\_\_