

# **FEINDSELIG- UND WOHLWOLLEND-SEXISTISCHE EINSTELLUNGEN IN BEZUG AUF MÜTTER:**

**ENTWICKLUNG UND VALIDIERUNG  
EINER AMBIVALENTEN SEXISMUS-SKALA  
MIT HILFE KOGNITIVER INTERVIEWS**

## **Dissertation**

zur Erlangung des Doktorinnengrades  
der Sozialwissenschaften (Dr. rer. soc.)

des Fachbereichs Sozial- und Kulturwissenschaften  
der Justus-Liebig-Universität Gießen

vorgelegt von:

**DIPL.-SOZIALWISS. ASTRID SCHÜßLER**  
geboren in Wetzlar

Gießen, Januar 2013

Erstgutachter: Prof. Dr. Peter Schmidt (Justus-Liebig-Universität Gießen)  
Zweitgutachter: Prof. Dr. Gerd Bohner (Universität Bielefeld)



## Danksagung

Ich möchte mich herzlich bei allen Personen bedanken, die mich auf dem Weg zur Fertigstellung dieser Dissertation begleitet und unterstützt haben. Mein besonderer Dank gilt...

... meinen beiden Betreuern Peter Schmidt und Gerd Bohner, die mich nicht nur durch hilfreiche Anmerkungen und Kommentare in Bezug auf meine Dissertation, sondern auch im Hinblick auf meine berufliche Weiterentwicklung unterstützt und manche Tür geöffnet haben. Vielen Dank dafür!

... allen übrigen Mitgliedern des Graduiertenkollegs, für die tolle gemeinsame Zeit. Bei Julia Becker, Frank Asbrock und Friederike Eyssel möchte ich mich zudem für den inhaltlichen Austausch sowie bei Jost Reinecke für seine hilfreichen Tipps bedanken, die er mir bei Zentralen Kolloquien und Summer Schools gegeben hat.

... Peter Glick, Rebeca Raijman und Janet Swim für ihre Hinweise und Anregungen bezüglich meiner Überlegungen zur Entwicklung der FWSEM-Skala in der Anfangsphase.

... den vier Interviewerinnen Marie-Luise Adler, Sarah Franziska Dietz, Anna Mlynarska und Stephanie Pape für ihre Unterstützung bei der Datenerhebung.

... meinen beiden studentischen Hilfskräften Stefanie Gosen und Stefan Thörner, die mir Stunden um Stunden beim Transkribieren der kognitiven Interviews geholfen haben. Ihr wart mir eine große Hilfe.

... dem Team der Arbeitsstelle Gender Studies der Justus-Liebig-Universität Gießen. Insbesondere Katharina Kirschbaum und Aylin Balsoy möchte ich für ihre Unterstützung beim Einpflegen der Literatur in die Datenbank danken.

... Julia und Christian Klewer sowie Stefanie und Jan Grote für ihre Aufmunterungen und Zerstreuungen bei einem Leben abseits der Dissertation.

... Veronika Schmid, Gunnar Lemmer, Corinna Zakikhany, Judith Lanphen und Stephan Sielschott für die gegenseitige Unterstützung und die schönen gemeinsamen Feiern, auf denen wir viel gelacht haben und sogar über das böse „D-Wort“ scherzen konnten.

... Diana Auth, die mir gleichzeitig Mentorin und Freundin ist. Danke auch für die hilfreichen Anmerkungen zu meiner Dissertation und die aufbauenden Worte zur rechten Zeit.

... meinem Bruder Lennart Schüßler, auf den ich mich immer verlassen kann – ich bin froh, dass es Dich gibt.

... meinen Eltern Ulrike Schüßler und Karl-Heinz Schüßler-Walter, denen ich dieses Buch von Herzen widmen möchte. Danke, dass Ihr meine Arbeit so sorgfältig Korrektur gelesen habt, dass Ihr immer für mich da seid, wenn ich Euch brauche und mich stets mit allen Kräften unterstützt.

... meinem Freund Stefan Kindermann, der immer an mich geglaubt, mich stets aufgebaut und mit allen Kräften unterstützt hat und der es immer wieder verstand, mich zur rechten Zeit vom Schreibtisch wegzulocken, um frische Luft zu tanken und den Kopf wieder frei zu bekommen. Danke für Deine Liebe und ich freue mich unbeschreiblich auf unseren nächsten gemeinsamen Lebensabschnitt, der bald beginnt!

Fernwald, Januar 2013

Astrid Schüßler

## Zusammenfassung

Die Gleichberechtigung zwischen Frauen und Männern scheint im 21. Jahrhundert fast erreicht zu sein. Nach der Geburt eines Kindes greifen viele Paare jedoch nach wie vor auf traditionelle Muster der Arbeitsteilung zurück. Während Väter den Umfang ihrer Erwerbstätigkeit sogar erhöhen, reduzieren Mütter ihr Arbeitspensum oder steigen (vorübergehend) aus dem Berufsleben aus, um die Kinderbetreuung übernehmen zu können. Diese Brüche im Erwerbverlauf sowie Teilzeitbeschäftigungsverhältnisse erschweren den Wiedereintritt in den Arbeitsmarkt und führen häufig zu einem „Karriereknick“ oder sogar einem „Karriereaus“ von Müttern. Sie mindern die Chancen auf eine Führungsposition und tragen bedeutend zu den Einkommensunterschieden zwischen den Geschlechtern bei. Mit dem geringeren Verdienst der Frau geht eine finanzielle Abhängigkeit vom (Ehe-)Mann einher, zudem reduzieren sich ihre Ansprüche auf Sozialleistungen des Staates (z.B. bezüglich der Rente), die Macht- und Statusunterschiede zwischen den Geschlechtern werden gefestigt. Folglich sind insbesondere Mütter in unserer Gesellschaft von Benachteiligungen betroffen.

Ziel dieser Arbeit ist es daher, eine Skala zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter (kurz: FWSEM-Skala) zu entwickeln und zu validieren. Dabei werden zwei zentrale Anliegen verfolgt: ein inhaltliches und ein methodisches. *Inhaltlich* liegt bei der Skalenentwicklung der Fokus auf einer Subgruppe von Frauen, die in bisherigen Messinstrumenten vernachlässigt wurde: die Gruppe der Mütter. Um die Ansichten der Befragten differenziert erfassen zu können, werden der Theorie des ambivalenten Sexismus (Glick & Fiske, 1996, 1997) folgend neben offen feindseligen auch (scheinbar) wohlwollende Einstellungen berücksichtigt. Der *feindselige Sexismus* erfasst Einstellungen gegenüber berufstätigen Müttern (vermeintlichen „Rabenmüttern“). Er zeichnet ein *negatives* Frauenbild und wertet Mütter ab, die aus der traditionellen Rolle ausbrechen und ihre Kinder bereits nach kurzer Zeit zur Betreuung weggeben. Der *wohlwollende Sexismus* bezieht sich auf Hausfrauen und Vollzeitmütter. Er nimmt scheinbar *positive* Merkmalszuschreibungen vor, die Frauen für die Übernahme der Kinderbetreuung unersetzbar erscheinen lassen, sie dadurch jedoch gleichzeitig auf ihre traditionelle Geschlechterrolle als sorgende Mutter festlegen.

In *methodischer Hinsicht* kommt bei der Skalentestung ein Pretest-Verfahren zum Einsatz, welches in den vergangenen Jahren international zunehmend an Bedeutung gewonnen hat, aber in der deutschen Forschungslandschaft bislang noch zu selten verwendet wird: die Technik der kognitiven Interviews. Um die Güte der Skala zu analysieren, wird ein Drei-Phasen-Pretest entwickelt, das qualitative und quantitative Methoden kombiniert. In der ersten Phase, die aus quantitativen Erhebungen besteht, wird die Itemanzahl des ursprünglichen Itempools reduziert. Oftmals endet eine Skalenkonstruktion an diesem Punkt. Diejenigen Items, die die besten Kennwerte aufweisen, werden in die Skala aufgenommen, ohne zu wissen, ob die Formulierungen

der Items von den Befragten auch so verstanden werden, wie es intendiert ist. Um dies zu überprüfen, werden die verbliebenen Items in Phase 2 mit Hilfe kognitiver Interviews getestet und im Falle von Verständnisschwierigkeiten umformuliert. In der dritten Phase, die erneut quantitativ angelegt ist, werden schließlich die Items für die endgültige FWSEM-Skala ausgewählt und ihre Güte bestimmt.

In insgesamt acht Untersuchungen ( $N = 1317$ ) erweist sich die Skala als zweidimensional und hoch reliabel. Geschlechtsdifferenzen in den Skalenwerten werden nicht gefunden, allerdings zeichnen sich Unterschiede zwischen Befragten aus Ost- und Westdeutschland ab. Für die konvergente Validität der FWSEM-Skalen sprechen signifikante, positive Korrelationen mit der Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999) sowie mit drei weiteren Sexismus-Skalen (MSS, GRB und AMMSA), mit einer Autoritarismus-Skala (RWA) und einer Skala zur Sozialen Dominanzorientierung (SDO). Auch mit einer Reihe anderer Vorurteils-Skalen, die gemeinsam das „Syndrom Gruppenbezogener Menschenfeindlichkeit“ bilden (Heitmeyer, 2002), bestehen signifikante, positive Korrelationen. Auf die Übereinstimmungsvalidität verweisen Korrelationen mit negativen Stereotypen über Karrierefrauen und mit positiven Stereotypen über Hausfrauen. Im Hinblick auf die Tendenz von Befragten, ein sozial erwünschtes Antwortverhalten zu zeigen, konnte keine systematische Verzerrung im Zusammenhang mit der FWSEM-Skala festgestellt werden.

Alles in allem sprechen die Ergebnisse dafür, dass es gelungen ist, mit der FWSEM-Skala ein zweidimensionales, hoch reliables und valides Messinstrument zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter zu entwickeln. Die Skala hat sich bei verschiedenen Erhebungsmodalitäten (schriftlich, telefonisch, online) bewährt und ist nicht nur in Studierenden-, sondern auch in Bevölkerungsstichproben einsetzbar. Dabei zeigte die Skala durchweg ein hohes Maß an Stabilität bezüglich der ermittelten Item- und Skalenkennwerte.

Schlüsselwörter: Ambivalenter Sexismus, Geschlechterrollen, Maternal Wall, Skalenentwicklung, kognitive Interviews, Pretest-Verfahren

## Inhaltsverzeichnis

<b>1 EINLEITUNG</b> .....	<b>7</b>
1.1 ZUM STAND DER GESCHLECHTERGERECHTIGKEIT.....	7
1.2 UNTERSUCHUNGSZIEL UND METHODISCHE VORGEHENSWEISE.....	13
1.3 AUFBAU DER ARBEIT.....	15
<b>2 THEORETISCHE GRUNDLAGEN</b> .....	<b>18</b>
2.1 GESCHLECHTSSPEZIFISCHE ARBEITSTEILUNG.....	18
2.1.1 Institutionelle Rahmenbedingungen und kulturelle Leitbilder als Einflussfaktoren.....	18
2.1.1.1 Institutionelle Rahmenbedingungen.....	19
2.1.1.2 Der Muttermythos.....	23
2.1.1.3 Frauen- und Mütterleitbilder in West- und Ostdeutschland.....	26
2.1.2 Ausgestaltung der Haus- und Familienarbeiten.....	30
2.1.3 Folgen der geschlechtsspezifischen Arbeitsteilung für die Stellung der Frauen im Erwerbsleben.....	33
2.2 GESCHLECHTERVORURTEILE UND GESCHLECHTERSTEREOTYPE.....	40
2.2.1 Einstellungen, Vorurteile, Stereotype und Diskriminierungen.....	40
2.2.2 Geschlechtervorurteile - reine Antipathie?.....	43
2.2.3 Geschlechterstereotype.....	44
2.2.3.1 Definition.....	44
2.2.3.2 Inhalte von Geschlechterstereotypen.....	46
2.2.3.3 Substereotype.....	47
2.2.3.4 Erklärungsansätze.....	47
2.2.3.5 Geschlechterstereotype – zutreffend oder sexistisch?.....	54
2.3 GESCHLECHTERDISKRIMINIERUNG.....	55
2.3.1 Definition.....	55
2.3.2 Maternal wall.....	57
2.3.2.1 Diskriminierung qua Geschlecht(errolle).....	57
2.3.2.2 Muster der Maternal Wall.....	62
2.4 SEXISMUS: THEORIEN UND MESSINSTRUMENTE.....	66
2.4.1 Allgemeine Definition.....	66
2.4.2 Facetten von Sexismus.....	67
2.4.2.1 Traditioneller Sexismus.....	67
2.4.2.2 Moderner Sexismus.....	72
2.4.2.3 Ambivalenter Sexismus.....	75
2.4.2.4 Zusammenfassung.....	81
2.5 KOGNITIVE UND KOMMUNIKATIVE PROZESSE IN BEFRAGUNGSSITUATIONEN.....	84
2.5.1 Frage-Antwort-Prozess in standardisierten Befragungen.....	84
2.5.1.1 Verstehen der Frage.....	85
2.5.1.2 Bildung oder Abruf eines Urteils.....	89
2.5.1.3 Formatierung der Antwort.....	91

2.5.1.4	<i>Editieren der Antwort</i> .....	92
2.5.1.5	<i>Zusammenfassung</i> .....	92
2.5.2	Unterschiede zwischen einer Konversation im Alltag und in standardisierten Befragungen.....	93
2.5.2.1	<i>Sinngebung bei Fragen in Alltagskonversationen</i> .....	93
2.5.2.2	<i>Sinngebung bei Fragen in standardisierten Befragungen</i> .....	96
2.5.2.3	<i>Mögliche negative Folgen der Standardisierung für die Datenqualität</i> .....	99
2.5.2.4	<i>Implikationen für die Umfrageforschung</i> .....	99
2.6	PRETEST-VERFAHREN ZUR EVALUIERUNG VON FRAGEBÖGEN.....	102
2.6.1	Kognitive Interviews.....	105
2.6.1.1	<i>Historische Entwicklung</i> .....	105
2.6.1.2	<i>Ziele</i> .....	106
2.6.1.3	<i>Empfehlungen zur Durchführung</i> .....	107
2.6.1.4	<i>Ablauf</i> .....	110
2.6.1.5	<i>Kognitive Techniken</i> .....	111
2.6.1.5	<i>Auswertung</i> .....	117
2.6.1.6	<i>Stärken und Schwächen des Verfahrens</i> .....	120
2.6.2	Fokusgruppen.....	123
2.6.3	ExpertInnen-Review (Checklisten-Systeme).....	124
2.6.4	Beobachtungspretest mit InterviewerInnen-Debriefing.....	125
2.6.5	Behavior Coding.....	127
2.6.6	Befragten-Debriefing.....	128
2.6.7	Analyse von Pretest-Daten.....	129
2.6.9	Response Latency.....	130
2.6.8	Reinterviews.....	131
2.6.10	Split-Ballot.....	131
2.6.11	Fazit Pretest-Verfahren.....	132
<b>3</b>	<b>EMPIRISCHER TEIL</b> .....	<b>138</b>
3.1	ÜBERBLICK ENTWICKLUNG UND TESTUNG DER FWSEM-SKALA.....	138
3.2	ENTWICKLUNGS-PHASE: GENERIERUNG DES URSPRÜNGLICHEN ITEMPOOLS.....	145
3.2.1	<i>Feindseliger Sexismus</i> .....	145
3.2.2	<i>Wohlwollender Sexismus</i> .....	148
3.3	PRETEST-PHASE 1: REDUZIERUNG DES URSPRÜNGLICHEN ITEMPOOLS.....	153
3.3.1	<i>Stichprobe 1 und 2</i> .....	153
3.3.2	<i>Erhebungsinstrumente in Stichprobe 1 und 2</i> .....	153
3.3.3	<i>Ergebnisse aus Studie 1 und 2</i> .....	154
3.4	PRETEST-PHASE 2: ÜBERPRÜFUNG DES ITEMVERSTÄNDNISSES MIT HILFE KOGNITIVER INTERVIEWS.....	155
3.4.1	<i>Überblick</i> .....	155
3.4.2	<i>Stichprobe 3, 4 und 5</i> .....	155
3.4.3	<i>Erhebungsinstrumente und Durchführung in Stichprobe 3, 4 und 5</i> .....	156
3.4.4	<i>Auswertung der kognitiven Interviews</i> .....	159

3.4.5 Ergebnisse der kognitiven Interviews aus Runde 1 bis 3.....	160
3.4.5.1 <i>Feindseliger Sexismus</i> .....	160
3.4.5.2 <i>Wohlwollender Sexismus</i> .....	179
3.4.5.3 <i>Zusammenfassung</i> .....	193
3.5 PRETEST-PHASE 3: AUSWAHL DER ITEMS UND EINSATZ DER FWSEM-SKALA .....	194
3.5.1 Überblick.....	194
3.5.2 Stichprobe 6 und 7.....	194
3.5.3 Stichprobe 8.....	195
3.5.4 Erhebungsinstrumente in Stichprobe 6.....	196
3.5.5 Erhebungsinstrumente in Stichprobe 7 .....	198
3.5.6 Erhebungsinstrumente in Stichprobe 8.....	199
3.5.7 Ergebnisse aus Studie 6, 7 und 8.....	201
3.5.7.1 <i>Endgültige Auswahl der Items für die FWSEM-Skala</i> .....	201
3.5.7.2 <i>Itemkennwerte der FWSEM-Skala</i> .....	203
3.5.7.3 <i>Reliabilitäten und Itemwert-Skalenwert-Korrelationen der FWSEM-Skalen</i> .....	208
3.5.7.4 <i>Explorative Faktorenanalyse der FWSEM-Skala</i> .....	210
3.5.7.5 <i>Konfirmatorische Faktorenanalyse der FWSEM-Skala</i> .....	211
3.5.7.6 <i>Explorative Faktorenanalyse über alle Sexismus-Items in Stichprobe 6</i> .....	215
3.5.7.7 <i>Explorative Faktorenanalyse über alle Sexismus-Items in Stichprobe 8</i> .....	216
3.5.7.8 <i>Beziehung zwischen der FS- und der WS-Subskala</i> .....	219
3.5.7.9 <i>Geschlechtsunterschiede in den Mittelwerten</i> .....	220
3.5.7.10 <i>Ost-/West-Unterschiede in den Mittelwerten</i> .....	221
3.5.7.11 <i>Korrelationen der FWSEM-Skala mit der ASS</i> .....	222
3.5.7.12 <i>Korrelationen der FWSEM-Skala mit anderen Sexismus-Skalen</i> .....	224
3.5.7.13 <i>Korrelationen der FWSEM-Skalen mit Autoritarismus und Sozialer Dominanzorientierung</i> .....	226
3.5.7.14 <i>Korrelationen der FWSEM-Skalen mit dem GMF-Syndrom</i> .....	228
3.5.7.15 <i>Korrelationen der FWSEM-Skalen mit Sozialer Erwünschtheit</i> .....	231
3.5.7.16 <i>Übereinstimmungsvalidität der FWSEM-Skalen im Vergleich zur ASS</i> .....	231
<b>4 FAZIT UND DISKUSSION .....</b>	<b>235</b>
4.1 FAZIT ZUR METHODISCHEN VORGEHENSWEISE.....	236
4.2 FAZIT ZU DEN INHALTLICHEN ERGEBNISSEN .....	238
4.3 FORSCHUNGSPERSPEKTIVEN UND AUSBLICK .....	247
<b>5 LITERATUR .....</b>	<b>250</b>
<b>ANHANG .....</b>	<b>268</b>
<b>ERKLÄRUNG .....</b>	<b>327</b>

## Tabellenverzeichnis

Tabelle 1: Zeitaufwendungen von Paaren mit und ohne Kindern für unbezahlte und bezahlte Arbeiten im Jahre 2001/02.....	31
Tabelle 2: Theorie des Stereotype Content Model.....	51
Tabelle 3: Nutzen verschiedener Pretest-Verfahren zur Evaluierung der Fragen im Hinblick auf die zentralen Standards .....	136
Tabelle 4: Überblick über die einzelnen Phasen der Entwicklung und Testung der FWSEM-Skala.....	140
Tabelle 5: Überblick über die vorgenommenen Veränderungen am Itempool während der einzelnen Phasen der Entwicklung und Testung der FWSEM-Skala...	141
Tabelle 6: Zusammensetzung der Stichproben in Studie 1 bis 8 .....	143
Tabelle 7: Itemformulierungen der FS-Items, die in Runde 1 der kognitiven Interviews getestet wurden.....	161
Tabelle 8: Itemformulierungen der WS-Items, die in Runde 1 der kognitiven Interviews getestet wurden .....	179
Tabelle 9: Itembeispiele und Reliabilitäten der Skalen, die zur Konstruktvalidierung der FWSEM-Skala in Stichprobe 6 eingesetzt wurden.....	197
Tabelle 10: Itembeispiele und Reliabilitäten der Skalen, die zur Konstruktvalidierung der FWSEM-Skala in Stichprobe 8 eingesetzt wurden.....	200
Tabelle 11: Mittelwerte und Standardabweichungen der FWSEM-Items in drei Stichproben.....	203
Tabelle 12: Mittelwerte, Standardabweichungen, Schiefen und Kurtoses der FWSEM-Skalen in drei Stichproben .....	207
Tabelle 13: Reliabilitäten (Cronbachs Alpha) der FWSEM-Gesamtskala und der beiden Subskalen in drei Stichproben.....	208
Tabelle 14: Korrigierte Itemwert-Skalenwert-Korrelationen der FWSEM-Items in drei Stichproben.....	209
Tabelle 15: Faktorladungen der FWSEM-Items in drei Stichproben .....	210
Tabelle 16: Faktorladungen der FWSEM-Items für die Gesamtstichprobe und die beiden Teilstichproben in Studie 8 .....	211
Tabelle 17: Fitmaße des Ein- und des Zwei-Faktoren-Modells in drei Stichproben ....	213
Tabelle 18: Faktorladungen aller Sexismus-Items in Stichprobe 8.....	217



---

Tabelle 19: Korrelationen zwischen der FS- und der WS-Subskala in drei Stichproben .....	219
Tabelle 20: Mittelwerte der FWSEM-Skalen für Frauen und Männer in drei Stichproben .....	221
Tabelle 21: Mittelwerte der FWSEM-Skalen für West- und Ostdeutsche in Stichprobe 8 .....	222
Tabelle 22: Korrelationen zwischen den FWSEM-Skalen und den ASS-Skalen in Stichprobe 6 und 7 .....	223
Tabelle 23: Korrelationen zwischen den FWSEM-Skalen und drei anderen Sexismus-Skalen in Stichprobe 6 .....	224
Tabelle 24: Korrelationen zwischen den FWSEM-Skalen und drei anderen Sexismus-Skalen in Stichprobe 8 .....	225
Tabelle 25: Korrelationen zwischen den FWSEM-Skalen und der RWA- bzw. SDO-Skala in Stichprobe 6 und 8.....	228
Tabelle 26: Korrelationen zwischen den FWSEM-Skalen und dem GMF-Syndrom in Stichprobe 8 .....	230
Tabelle 27: Korrelationen zwischen den FWSEM-Skalen und einer Skala zur sozialen Erwünschtheit in Stichprobe 6 .....	231
Tabelle 28: Korrelationen zwischen den FWSEM-Skalen und den Stereotypen über Karrierefrauen und Hausfrauen in Stichprobe 7 .....	234
Tabelle 29: Korrelationen zwischen den ASS-Skalen und den Stereotypen über Karrierefrauen und Hausfrauen in Stichprobe 7 .....	234

## Abbildungsverzeichnis

Abbildung 1. Einstellungsmodell nach Eagly und Chaiken (1993) mit drei Reaktionsklassen. ....	41
Abbildung 2. Idealtypische Zuordnung der beiden weiblichen Subgruppen „Hausfrau“ und „Karrierefrau“ in ihr entsprechendes Wärme-Kompetenz-Cluster. 50	
Abbildung 3. Modell des Frage-Antwort-Prozesses in einer standardisierten Umfragesituation. ....	85
Abbildung 4. Häufigkeitsverteilungen der FWSEM-Skalen in Stichprobe 8. ....	206
Abbildung 5. Ursprüngliches Zwei-Faktoren-Modell der FWSEM-Skala in Stichprobe 8 mit den entsprechenden Gütemaßen. ....	214
Abbildung 6. Modifiziertes Zwei-Faktoren-Modell der FWSEM-Skala in Stichprobe 8 mit den entsprechenden Gütemaßen. ....	214

*„Es gibt einen Satz, den kann ich nicht mehr hören: Wir haben doch jetzt eine Bundeskanzlerin! Dieser Satz muss neuerdings als Entschuldigung für so ziemlich alles herhalten, wenn es um eine mögliche Benachteiligung von Frauen geht. Viele denken: Der Feminismus hat sich erledigt, wir sind doch alle, Frauen und Männer, komplett durchemanzipiert. Es herrscht die Vorstellung, man müsse nicht mehr kämpfen um die Gleichberechtigung. Das Gegenteil stimmt. Wir erleben einen ziemlichen Rückfall.“*

(Göring-Eckardt, 2006, o.S.)

## 1 Einleitung

### 1.1 Zum Stand der Geschlechtergerechtigkeit

Die Antwort auf die Frage, ob das Glas halb leer oder halb voll ist, hängt wohl vom Standpunkt des Betrachters bzw. der Betrachterin ab. Dies gilt auch in Bezug auf die Frage nach der Geschlechtergerechtigkeit. Für viele scheint die Gleichstellung der Geschlechter im 21. Jahrhundert (nahezu) erreicht und die Diskriminierung von Frauen kein Problem mehr zu sein. Seit mehr als 60 Jahren ist die Gleichberechtigung von Frauen und Männern im Grundgesetz (Art. 3 Abs. 2) verankert. Das Bürgerliche Gesetzbuch von 1900 wurde reformiert und die dem Gleichberechtigungsgebot entgegenstehenden Regelungen angepasst (vgl. Hoecker, 1999, S. 34 ff.; Reich-Hilweg, 1979, S. 17 ff.). Mittlerweile scheint die formale Gleichstellung der Geschlechter erreicht. So lassen sich im Grundgesetz (so gut wie) keine Rechtsnormen mehr finden, die Frauen unmittelbar diskriminieren (vgl. Cordes, 2008, S. 916). Ganz im Gegenteil wurden seit der Gründung der Bundesrepublik immer mehr Schritte unternommen, um die Diskriminierung aufgrund des Geschlechts gesetzlich zu unterbinden. Mit dem Konzept des Gender Mainstreamings wurde die Geschlechterperspektive „zur Leitlinie des politischen Handelns und politischer Entscheidungen“ (Cordes, 2008, S. 920). Es wurden gleichstellungspolitische Maßnahmen auf den Weg gebracht, um eine bessere Vereinbarkeit von Beruf und Familie zu ermöglichen, Frauen zu fördern und die Chancengleichheit der Geschlechter voranzutreiben (vgl. BMFSFJ, 2004; BMFSFJ, BMBF, BMVBS, BMAS, & BMWI, 2006). Das neue Elterngeldgesetz von 2007 etwa soll durch den Einkommensersatzcharakter des Bezuges sowie die beiden Partnermonate u.a. Anreize für Väter schaffen, sich stärker an der Kinderbetreuung zu beteiligen (vgl. Gerlach, 2010, S. 232 f.; Walther & Schaeffer-Hegel, 2007, S. 15).

Auch die Ergebnisse von Einstellungsumfragen haben sich aufgrund der gesellschaftlichen Diskurse über die Gleichstellung der Geschlechter in den letzten Jahrzehnten gewandelt. Immer seltener werden offen sexistische Ansichten geäußert; eine Befürwortung traditioneller Geschlechterrollen nimmt ab (vgl. Campbell, Schellenberg, & Senn, 1997, S. 89; Eckes & Six-Materna, 1998, S. 224; McHugh & Frieze, 1997, S. 5). In Untersuchungen zu Geschlechterstereotypen werden Frauen sogar in einem positiveren Licht beschrieben als Männer („women are wonderful“ effect) (vgl. Eagly & Mladinic, 1994). In

Umfragen geben junge Männer vermehrt an, dass sie sich gerne aktiv an der Betreuung der Kinder beteiligen möchten (vgl. Zerle & Krok, o.J.).

In Bezug auf die tatsächliche Gleichstellung der Geschlechter lassen sich ebenso positive Entwicklungen konstatieren. Im Bereich der Bildung haben Mädchen heute sogar größere Chancen als Jungen. Sie werden früher eingeschult, erzielen bessere Noten, wiederholen seltener ein Schuljahr und stellen den Hauptteil an AbiturientInnen (vgl. Stürzer, 2005, S. 28 ff.).

Im Hinblick auf die kulturellen Leitbilder scheint sich ebenfalls ein Wandel abzuzeichnen. Alte Rollenbilder brechen (scheinbar) auf: Frauen gehen zur Bundeswehr oder sind Pilotinnen und Männer arbeiten in Kindertagesstätten oder als Altenpfleger. Die Zahl der Väter, die Elternzeit beantragen, ist im Jahr 2009 auf 23.6% angewachsen; „neue Väter“ erobern das Land (vgl. Statistisches Bundesamt, 2011). Das männliche Allein-Verzorger-Modell bröckelt (vgl. Sachverständigenkommission, 2011, S. 27).

Gleichzeitig wird die Erwerbstätigkeit von Frauen nicht mehr in Frage gestellt und die Frauenerwerbsquote ist in den letzten Jahrzehnten angestiegen. Frauenquoten und -quoten führten zu einem höheren Frauenanteil in der Politik (vgl. Hoecker, 2008, S. 13). Auch an den Hochschulen wirken sich Gleichstellungskonzepte und Frauenfördermaßnahmen positiv aus. Der weibliche Anteil innerhalb der ProfessorInnenschaft ist in den vergangenen Jahren stetig gestiegen und hat sich im Zeitraum von 1999 bis 2009 fast verdoppelt (vgl. Destatis, 2010). In der Wirtschaft tut sich ebenfalls etwas: Die Telekom hat sich als erstes Unternehmen freiwillig dazu verpflichtet, einen Frauenanteil von 30% in den Führungsetagen weltweit bis zum Ende des Jahres 2015 zu erreichen (vgl. Deutsche Telekom, 2010). Im Oktober 2011 erklärten sich schließlich auch die übrigen der 30 größten DAX-Unternehmen zur Förderung von Frauen in Spitzenpositionen mit einer selbst festgelegten Quote bereit (vgl. Caspari, 2011).

Diese gleichstellungspolitisch durchaus positiv zu bewertenden Entwicklungen lassen hoffen. Gleichwohl sehen auch heute noch viele Frauen und Männer Handlungsbedarf – für sie ist die Gleichstellung der Geschlechter noch längst nicht erreicht.

Dass die Verwirklichung der Geschlechtergerechtigkeit nach wie vor noch nicht verwirklicht ist und auch weiterhin Handlungsbedarf in dieser Hinsicht besteht, wird im Folgenden deutlich: Im Zuge der Wiedervereinigung wurde Artikel 3 Absatz 2 im Grundgesetz angesichts fehlender Chancengleichheit 1994 um folgenden Nachsatz ergänzt: „Der Staat fördert die tatsächliche Durchsetzung der Gleichberechtigung von Frauen und Männern und wirkt auf die Beseitigung bestehender Nachteile hin“ (vgl. Deutsches Jugendinstitut, 2006). Auch wenn sich in gesetzlicher Hinsicht keine Rechtsnormen mehr im Grundgesetz finden lassen, die Frauen *direkt* diskriminieren, existieren immer noch Normen, die Frauen *indirekt* diskriminieren (vgl. Cordes, 2008, S. 916; siehe Unterkapitel 2.3). Das Konzept des Gender Mainstreamings scheint ebenfalls auf Grenzen zu stoßen: Ohne ein wirkliches Interesse und ein Gespür für Gender-Problematiken auf Seiten der politischen Akteure läuft das Konzept ins Leere und wird unwirksam (vgl. Cordes, 2008, S. 921).

Hinsichtlich der Einstellungsebene ist es ebenfalls nötig, eine differenzierte Betrachtungsweise vorzunehmen: So sind sexistische Einstellungen in der Bevölkerung nicht etwa vollkommen verschwunden, es findet lediglich eine Verlagerung in Richtung subtilerer, scheinbar positiver (paternalistischer und wohlwollender) Formen, die eine höhere soziale Akzeptanz genießen, statt (vgl. Eckes & Six-Materna, 1998, S. 224 f.). Selbst die vermeintlich positiven Merkmalszuschreibungen von Frauen als empathiefähig, fürsorglich oder aufopferungsvoll tragen dazu bei, die bestehenden Geschlechterhierarchien zu festigen, da sie Eigenschaften ansprechen, die Frauen für die traditionelle Geschlechterrolle prädestinieren (vgl. Eckes & Six-Materna, 1999, S. 213 f.; Glick & Fiske, 1996, S. 492 f., Glick & Fiske, 2001a, S. 110; siehe Unterkapitel 2.2.2). Männern hingegen werden solche Merkmale attestiert, die in der Berufswelt gefragt sind („think manager – think male“-Phänomen) (vgl. z.B. Sczesny & Stahlberg, 2002, S. 36).

Die Erfolge im Bildungsbereich werden ebenfalls getrübt, da die besseren Schulleistungen der Mädchen und ihre höheren Bildungsabschlüsse nicht automatisch auch zu besseren Berufschancen führen (sogenanntes „Verwertungsdefizit“). Bei der Leistungskurs-, Studienfach- und Berufswahl bestehen – trotz verschiedener Berufsorientierungsinitiativen wie dem „Girls’ day“ und dem „Boys’ day“ – nach wie vor deutliche geschlechtstypische Unterschiede (vgl. Sachverständigenkommission, 2011, S. 78 f.). Dementsprechend sind Frauen und Männer auch auf dem Arbeitsmarkt in unterschiedlichen Berufsfeldern und Wirtschaftsbereichen vertreten (horizontale Segregation des Arbeitsmarktes). Dies ist insofern bedeutsam, als dass „typisch weibliche“ Berufsfelder schlechter entlohnt werden und zudem ein geringeres Ansehen genießen (vgl. Sachverständigenkommission, 2011, S. 117 ff.). Tendenziell ziehen Frauen solche Arbeitsplätze vor, die eine Vereinbarkeit von Beruf und Familie ermöglichen und geben sich dafür auch mit geringeren Aufstiegsmöglichkeiten und einem geringeren Einkommen zufrieden (vgl. Hoecker, 1999, S. 67).

Im Hinblick auf die Familienmodelle und deren zugrundeliegenden kulturellen Leitbilder lässt sich folgendes festhalten: Das „Modell der Hausfrauenehe“ ist im Westen Deutschlands lediglich durch ein „Vereinbarkeitsmodell der männlichen Versorgerehe“ (Pfau-Effinger, 2005, S. 4) abgelöst worden.<sup>1</sup> Dieses Modell sieht auch weiterhin in Zeiten der „aktiven Elternschaft“ eine geschlechtsspezifische Arbeitsteilung und somit eine Zuständigkeit der Frau für die Kinderbetreuung vor. Diesem Modell zufolge untergliedert sich die Erwerbstätigkeit der Frau in die drei folgenden Phasen: So lange die Frau noch keine Kinder hat, ist generell eine Vollzeitenerwerbstätigkeit für sie vorgesehen. Mit der Geburt des ersten Kindes erfolgt eine Phase der Erwerbsunterbrechung für einige Jahre. Wenn die Kinder etwas älter sind, kehrt die Mutter in ihren Beruf zurück. Dabei nimmt

---

<sup>1</sup> Während im Westen Deutschlands das Zuverdiener-Modell (bzw. „Vereinbarkeitsmodell der männlichen Versorgerehe“) am häufigsten praktiziert wird (kinderlose Paare: 36.7%; Paare mit Kindern: 47.5%), trifft dies im Osten Deutschlands auf das Zweiverdiener-Modell zu (kinderlose Paare: 37.8%; Paare mit Kindern: 37%) (vgl. Sachverständigenkommission, 2011, S. 27).

sie ihre Erwerbstätigkeit zunächst in Teilzeit wieder auf, so lange die Kinder noch als betreuungsbedürftig gelten, um auf diese Weise Familie und Erwerbsarbeit vereinbaren zu können (vgl. Pfau-Effinger, 2001, S. 496; Pfau-Effinger, 2005, S. 4). Dieses Arrangement verdeutlicht, dass „die ‚private Kindheit‘, die auf gemeinsamer Zeit von Eltern mit ihren Kindern und der zeitweiligen Betreuung von Kindern im eigenen Haushalt beruht, (...) aus der Zeit der Hausfrauenehe ansatzweise als wichtiger Wert erhalten“ bleibt (Pfau-Effinger, 2005, S. 4). Demgemäß wird die Erwerbstätigkeit von Frauen ohne Kinder auch als Normalität angesehen und ist gesellschaftlich akzeptiert, auf eine Erwerbstätigkeit von Müttern kleiner Kinder trifft dies hingegen im Westen Deutschlands nicht ausnahmslos zu (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 297).

Im Osten Deutschlands herrscht, dem kulturellen Leitbild aus Zeiten der DDR entsprechend, als Ideal das „Doppelversorger-Modell mit außerhäuslicher Kinderbetreuung“ vor (Pfau-Effinger, 2005, S. 4). Es sieht für Männer *und* für Frauen – auch nach dem Übergang zur Elternschaft – gleichermaßen eine Erwerbstätigkeit in Vollzeit vor. Die Kinderbetreuung wird in externe, außerfamiliäre Hände übergeben (vgl. Pfau-Effinger, 2001, S. 497; Pfau-Effinger, 2005, S. 5). Dementsprechend leichter fällt es hier auch Müttern mit kleinen Kindern einer Erwerbstätigkeit nachzugehen, da die staatliche Form der Betreuung von jeher üblich war und auch heute noch auf hohe gesellschaftliche Akzeptanz stößt (vgl. Schneider & Dorbritz, 2011, S. 32 f.; siehe Unterkapitel 2.1.1.3). An der grundsätzlichen Zuständigkeit der Frauen für die Haus- und Familienarbeiten wurde allerdings auch hier nicht gerüttelt (vgl. Schäffgen, 2000, S. 93).

Vor diesem Hintergrund fällt das Aufbrechen traditioneller Geschlechterarrangements schwer. Auch die Einführung des Elterngeldes im Jahr 2007 konnte – trotz stärkerer Beteiligung der Männer an der Kinderbetreuung – keinen grundlegenden Wandel herbeiführen. So ist absolut gesehen zwar der Anteil der Väter, die Elterngeld beziehen, im Jahr 2009 auf 23.6% gestiegen, doch betrifft dies vornehmlich (in 75% der Fälle) die beiden zusätzlichen Partnermonate. Von den Müttern beanspruchten hingegen 96% das Elterngeld. Zudem bezogen sie es in der überwiegend Mehrzahl der Fälle (neun von zehn Müttern) für einen Zeitraum von 12 Monaten (vgl. Statistisches Bundesamt, 2011).

Bezüglich der Arbeitsteilung in Paarbeziehungen besteht dementsprechend ebenfalls noch keine Geschlechtergerechtigkeit. So haben sich die Relationen der geleisteten (Re-)Produktionsarbeiten beider Geschlechter zwar angenähert, dennoch übernehmen Frauen nach wie vor den Großteil der unbezahlten Haus- und Familienarbeiten (siehe Unterkapitel 2.1.2), während Männer mehr bezahlte Erwerbstätigkeit leisten (siehe Unterkapitel 2.1.3) (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 304).

Den drei Phasen des „Vereinbarkeitsmodells der männlichen Versorgerehe“ entsprechend, setzt nach der Geburt des ersten Kindes eine (Re-)Traditionalisierungstendenz der Arbeitsteilung ein: So entscheiden sich selbst viele Paare, die zuvor die Aufgaben egalitär geteilt haben, mit dem Übergang zur Elternschaft für eine traditionelle Rollenaufteilung. Der Mann verstärkt sogar sein berufliches Arbeitspensum, während die Frau vorübergehend aus ihrem Beruf aussteigt, um sich der Betreuung der Kinder widmen zu können

(vgl. BMFSFJ, 2008a, S. 7; Kortendiek, 2008, S. 438). Einerseits hängt dies mit den eben beschriebenen kulturellen Normen und Leitbildern hinsichtlich der Mutterschaft zusammen, die in Westdeutschland vorherrschen, andererseits wird die Vereinbarkeit von Beruf und Familie auch durch strukturelle Rahmenbedingungen wie z.B. mangelnde Betreuungsmöglichkeiten für unter Dreijährige erschwert. Hinzu kommt, dass das deutsche Steuer- und Sozialsystem nach wie vor monetäre Anreize für eine männliche Versorger-ehe offeriert (vgl. Hummelsheim, 2008; siehe Unterkapitel 2.1).

Auch in beruflicher Hinsicht wurde noch keine Geschlechtergerechtigkeit erreicht. Obwohl die Erwerbstätigenquote der Frauen in den letzten Jahren gestiegen ist, liegt sie nach wie vor deutlich unter der Rate der Männer (vgl. Allmendinger, Leuze & Blank 2008, S. 20). Besonders gravierend fällt der Unterschied aus, wenn man die Vollzeitäquivalente betrachtet: Im Jahr 2009 betrug die Geschlechterlücke 21.5%. Der zuvor beschriebene Anstieg der Zahl erwerbstätiger Frauen ist vor allem auf einen Zuwachs bei den teilzeitbeschäftigten Frauen und „Minijobberinnen“ zurückzuführen. Während Männer zumeist eine Vollzeitstellung haben, üben Frauen häufig eine Teilzeiterwerbstätigkeit aus. (Im Jahr 2008 hatten 46% aller abhängig beschäftigten Frauen eine Teilzeitstelle, bei Männern belief sich die Quote lediglich auf 9%) (vgl. Sachverständigenkommission, 2011, S. 212, S. 90 ff.). Zudem unterbrechen Frauen häufiger familienbedingt ihre Erwerbstätigkeit. Beides wirkt sich wiederum negativ auf ihr Einkommen und ihre beruflichen Erfolgchancen aus (vgl. Dressel, 2005, S. 105 ff.). Im Jahr 2008 lag der Bruttostundenverdienst vollzeiterwerbstätiger Frauen im Durchschnitt 23% unter dem ihrer männlichen Kollegen (Verdienstunterschied Westdeutschland: 25%; Ostdeutschland: 5%) (vgl. Fuchs & Bothfeld, 2011, S.15; Statistisches Bundesamt, 2009a).

Obgleich Frauenfördermaßnahmen positive Effekte erzielen konnten, ist immer noch eine vertikale Segregation auf dem deutschen Arbeitsmarkt zu beobachten. In der Politik, der Wissenschaft oder der Wirtschaft ist das Verhältnis von Frauen und Männern in höheren Positionen längst nicht ausgewogen. So entfielen beispielsweise 2011 ein Drittel der Sitze (32.9%) im Deutschen Bundestag auf Frauen (vgl. Deutscher Bundestag, 2011) – eine Parität ist somit noch nicht erreicht. An den Hochschulen ist der Frauenanteil in den höheren Stadien der akademischen Laufbahn noch geringer: Ende 2009 besetzten Frauen nur 18.3% der Professuren (vgl. Statistisches Bundesamt Deutschland, 2010a). Die Unterrepräsentanz von Frauen ist in der Privatwirtschaft am stärksten ausgeprägt: Je höher die Führungsebene und je größer das Unternehmen, desto geringer ist der Frauenanteil. So waren Frauen im Jahr 2010 auf der ersten und zweiten Führungsebene privatwirtschaftlicher Unternehmen mit 19.6% vertreten. In großen Firmen, deren Umsatz über 20 Millionen Euro beträgt, liegt der Anteil an weiblichen Topmanagerinnen allerdings nur noch bei 5.9% und in DAX-Konzernen lediglich bei 3.2% (vgl. Sachverständigenkommission, 2011, S. 214).

Noch immer stoßen Frauen auf dem Weg nach oben an die sogenannte „gläserne Decke“ bzw. werden sie durch immer neue Sackgassen im „Labyrinth“ der Berufswelt (vgl.

Eagly & Carli, 2007) aufgehalten. So kommt es trotz des Allgemeinen Gleichbehandlungsgesetzes (AGG) von 2006 immer wieder zu Diskriminierungen aufgrund des Geschlechts: In vielen Fällen sind es schwangere Frauen bzw. Mütter, die bei Einstellungen, Weiterbildungsmaßnahmen oder Beförderungen übergangen werden (vgl. z.B. Obermeier, 2006; Täubner, 2008, S.16), da die Personalverantwortlichen zusätzliche Arbeit und Kosten aufgrund familienbedingter Auszeiten von Müttern vermeiden wollen bzw. sie eine Förderung nicht als lohnend ansehen, da eine Rückkehr nach der Elternzeit ungewiss ist (vgl. Sachverständigenkommission, 2011, S. 104). Auch andere Diskriminierungen aufgrund von Geschlechtervorurteilen sind denkbar, wenn Frauen beispielsweise aufgrund ihrer Mutterschaft im Vergleich zu ihren männlichen Kollegen als weniger kompetent, flexibel und/oder engagiert angesehen werden (siehe Unterkapitel 2.3.2).

In den USA hat sich für derartige Formen der Geschlechterdiskriminierung, bei der ArbeitnehmerInnen aufgrund der Übernahme von familiären Betreuungs- und Pflegearbeiten benachteiligt werden, die Bezeichnung „Family responsibilities discrimination (FRD)“ (Center for Work Life Law, 2006) durchgesetzt. Übernehmen Väter einen aktiven Part bei der Kinderbetreuung und gehen sie für mehrere Monate oder auch Jahre in Elternzeit oder signalisieren sie auch nur, dass sie überlegen, dies zu tun, sehen sie sich ebenfalls Diskriminierungen ausgesetzt (vgl. Center for Work Life Law, 2006, S.1). Folglich ist in diesen Fällen die Diskriminierung an die Übernahme einer bestimmten Geschlechterrolle geknüpft („Diskriminierung qua Mutterschaft bzw. qua Vaterschaft“) – völlig losgelöst von einer „reinen“ Diskriminierung auf der Basis der Zugehörigkeit zur Geschlechtskategorie „Frau“ ist sie jedoch nicht. So sind Frauen aufgrund der vorherrschenden Traditionalisierungstendenz der Arbeitsteilung nach der Geburt eines Kindes de facto in viel stärkerem Maße betroffen als Männer. Zudem haben selbst Frauen, die keine Kinder haben und sich auch keine wünschen, mit den oben beschriebenen Benachteiligungsstrukturen zu kämpfen (vgl. Sachverständigenkommission, 2011, S. 104). Bei Bewerbungen etwa können junge Frauen, die sich in einer festen Beziehung befinden, unter dem „Generalverdacht“ stehen, bald ein Kind zu bekommen und für das Unternehmen familienbedingt auszufallen (vgl. Obermeier, 2006). Derartige Generalisierungseffekte gibt es bei Männern nicht. Dementsprechend wird für die oben beschriebenen Diskriminierungsmechanismen häufig synonym der Begriff „Maternal Wall“ (Williams, 2001) verwendet (siehe Unterkapitel 2.3.2).

Als Resümee lässt sich Folgendes sagen: Auch wenn in einigen Bereichen bereits große Fortschritte erzielt wurden, sehen sich Frauen gleichwohl immer noch Diskriminierungen ausgesetzt; die formal rechtliche Gleichbehandlung der Geschlechter führte nicht automatisch zu einer faktischen Geschlechtergerechtigkeit. Sowohl im öffentlichen als auch im privaten Leben lässt sich eine Benachteiligung von Frauen erkennen, die zu einem Großteil mit den alten Rollenbildern bezüglich der Kinderbetreuung und dem Festhalten an traditionellen Arbeitsteilungsmustern zusammenhängt. Mit dem Übergang zur Elternschaft ist eine (Re-)Traditionalisierung der Arbeitsteilung zu beobachten, bei der die Mutter die Betreuung der Kinder übernimmt – sei es nun aufgrund institutioneller



Rahmenbedingungen, was beispielweise die mangelnden Kinderbetreuungsmöglichkeiten anbelangt, und/oder angesichts gesellschaftlicher Wertvorstellungen und Leitbilder, die die besondere Bedeutung der Mutter für das Wohlergehen des Kleinkindes betonen und gleichzeitig dessen außerhäusliche Betreuung und somit auch die mütterliche Erwerbstätigkeit in den ersten Jahren ablehnen.

Die traditionelle Ausgestaltung der Arbeitsteilung mit den familienbedingten beruflichen Auszeiten bringt zahlreiche negative Konsequenzen hinsichtlich der Erwerbstätigkeit von Müttern mit sich (wie in Unterkapitel 2.1.3 und Unterkapitel 2.3.2 noch aufgezeigt wird), die wiederum die Macht- und Statusunterschiede zwischen den Geschlechtern festigen. So führt die Übernahme der familiären Betreuungsaufgaben zu verschiedenen (unmittelbaren wie mittelbaren) Diskriminierungsmechanismen, die u.a. mit verminderten beruflichen Entwicklungschancen und Gehaltseinbußen für Frauen einhergehen. Auf die eingangs gestellte Frage hinsichtlich des Standes der Gleichberechtigung könnte man daher zusammenfassend festhalten: Es können Fortschritte verzeichnet werden, aber das Ziel ist noch nicht erreicht. Nach wie vor haben Frauen mit Benachteiligungsstrukturen und Diskriminierungsmechanismen in Deutschland zu kämpfen – besonders betroffen sind dabei die Mütter.

## 1.2 Untersuchungsziel und methodische Vorgehensweise

Ziel der vorliegenden Arbeit ist es daher, eine Skala zu entwickeln und zu validieren, die sexistische Einstellungen in Bezug auf Mütter erfasst. Dabei werden zwei zentrale Anliegen verfolgt: ein inhaltliches und ein methodisches. So wird bei der Generierung des Messinstruments thematisch eine Subgruppe von Frauen in den Mittelpunkt gerückt, die in bisherigen Sexismus-Skalen überwiegend unberücksichtigt blieb, nämlich die Gruppe der Mütter. In methodischer Hinsicht kommt bei der Skalentestung ein Pretest-Verfahren zum Einsatz, welches in den vergangenen Jahren international zunehmend an Bedeutung gewonnen hat (vgl. Conrad & Blair, 2009, S. 33), aber insbesondere bei der Skalentwicklung (in Deutschland) noch zu selten eingesetzt wird: die Technik der kognitiven Interviews. Zunächst soll jedoch auf das inhaltliche Anliegen näher eingegangen werden.

*Inhaltliches Anliegen.* Bislang werden in der einschlägigen Literatur drei Arten von Sexismus unterschieden. Der *traditionelle*, der *moderne* und der *ambivalente Sexismus*. In Deutschland etablierten sich vor allem sieben Skalen, die die entsprechenden Einstellungen erfassen. Der traditionelle Sexismus wird durch die „Misogynie-Skalen“ (MS) von Krameyer und Schmidt (1973) abgebildet. Zur Erfassung von Geschlechterrolleneinstellungen sind drei Skalen zu nennen: Die „Skala zur Messung der normativen Geschlechtsrollen-Orientierung“ (GRO) von Krampen (1979, 1983), ein „Fragebogen zur Messung normativer Geschlechtsrollenorientierung“ (NGRO) von Athenstaedt (2000) und die „Geschlechterrollen-Batterie“ (GRB) des International Social Survey Programme (ISSP) 2002 (siehe hierzu Braun, 2006; Scholz, Harkness, & Klein, 2003). Eckes und Six-Materna (1998) generierten mit der „Skala zur Erfassung von Vorurteilen gegenüber Frauen auf

der Basis des Konzepts des modernen Sexismus“ (MSS) ein Messinstrument für den modernen Sexismus. Ambivalent sexistische Einstellungen werden mit Hilfe der „Misogynie-Ritterlichkeits-Skala“ (MC) von Schmerl und Steinbach (1973) sowie der „Skala zur Erfassung des ambivalenten Sexismus“ (ASS) von Eckes und Six-Materna (1999) ermittelt.

In der Geschlechterrollen-Batterie des ISSP spricht die Hälfte der Items Frauen in ihrer Rolle als Mutter an, in drei Skalen (MS, GRO und MC) tauchen nur vereinzelt Items über Mütter auf und die übrigen drei Sexismus-Skalen (NGRO, MSS und ASS) beziehen sich ausschließlich auf die Gesamtgruppe der Frauen (bzw. im Falle der ASS mitunter auch auf Feministinnen). Keine der genannten Skalen konzentriert sich jedoch ausschließlich auf die Subgruppe der Mütter, obwohl – wie in Unterkapitel 2.1.3 und 2.3.2 aufgezeigt wird – in vielen Fällen gerade sie es sind, die mit Diskriminierungen im Berufsleben zu kämpfen haben. Um diese Forschungslücke zu schließen, wird eine Skala zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter (kurz FWSEM-Skala) entwickelt und validiert.

Als Basis hierfür wird die Theorie des ambivalenten Sexismus von Glick und Fiske (1996, 1997) herangezogen, da sie neuere Erkenntnisse der (Geschlechter-) Vorurteilsforschung (Eagly & Mladinic, 1994) aufgreift, nach denen Vorurteile nicht nur negativer, sondern (aus Sicht einer sexistisch eingestellten Person) auch positiver Art sein können (siehe Unterkapitel 2.2.2) (vgl. Glick & Fiske, 1997, S. 119 ff.). Folglich differenzieren Glick und Fiske (1996) zwischen feindseligen und subjektiv positiven, wohlwollenden Einstellungen gegenüber Frauen. Letztere sind umso folgenschwerer, da sie aufgrund ihrer subtilen Art häufig nicht als sexistisch wahrgenommen werden. So werden scheinbar wohlwollende Rechtfertigungen für eine Diskriminierung von Frauen („Eine Frau sollte ihre Karriere hintenanstellen, da sich niemand sonst so gut um das Kind kümmern kann wie sie.“) auch eher gesellschaftlich akzeptiert als feindselige Begründungen („Eine Frau sollte ihre Karriere hintenanstellen, da ihr die nötigen Fähigkeiten dazu fehlen.“) (vgl. Glick & Fiske, 2001a; Social Psychology Network, o.J.). Auch wenn beide Argumentationsmuster in unterschiedliche Richtungen gehen, bleibt das Resultat das gleiche: Frauen werden auf ihre Rolle als Hausfrau und Mutter festgelegt und die gesellschaftliche Vorrangstellung der Männer gefestigt (vgl. Glick & Fiske, 1996, S. 492). Mit der Differenzierung zwischen feindseligen und scheinbar wohlwollenden Einstellungen gegenüber Frauen wird die Theorie des ambivalenten Sexismus einem weiteren relevanten Sachverhalt gerecht: Offen sexistische Einstellungen werden heutzutage immer seltener geäußert, da sie als sozial unerwünscht gelten. Dies ließ die Notwendigkeit entstehen Messinstrumente zu entwickeln, die subtile Vorurteile – wie beispielsweise den wohlwollenden Sexismus – erfassen (siehe Unterkapitel 2.4.2.2) (vgl. Campbell, Schellenberg, & Senn, 1997, S. 89; Eckes & Six-Materna, 1998, S. 226 ff.; Swim, Becker, Lee, & Pruitt, 2010, S. 138).

Dementsprechend wird auch bei der Entwicklung der FWSEM-Skala zwischen feindselig- und wohlwollend-sexistischen Einstellungen in Bezug auf Mütter differenziert. Die

*wohlwollende* Subkomponente der Skala bezieht sich auf scheinbar *positive* Zuschreibungen, die Frauen für die Übernahme der Kinderbetreuung unersetzbar erscheinen lassen, sie dadurch gleichzeitig jedoch auf ihre traditionelle Rolle als Hausfrau und Mutter festlegen. Der *feindselige* Sexismus zeichnet hingegen ein *negatives* Frauenbild und wertet Mütter, die aus der traditionellen Rolle ausbrechen und ihre Kinder bereits nach kurzer Zeit zur Betreuung „weggeben“, ab und stellt sie als vermeintliche „Rabenmütter“ hin.

*Methodisches Anliegen.* Das zweite Anliegen dieser Arbeit betrifft die methodische Vorgehensweise. So kommt bei der Skalenentwicklung (neben der psychometrischen Analyse von Pretest-Daten und einem ExpertInnen-Review; siehe Unterkapitel 2.6) eine Technik zum Einsatz, die über die gängigen Verfahren hinausgeht: Die Methode der kognitiven Interviews. Mit ihrer Hilfe können das intendierte Frageverständnis sowie die mentalen Prozesse, die bei der Beantwortung der Items ablaufen, überprüft werden. Hierfür werden die InterviewteilnehmerInnen unabhängig von ihrem gezeigten Antwortverhalten gezielt um die Erklärung bestimmter Begriffe oder die Offenlegung ihrer Gedankengänge gebeten. Auf diese Art und Weise können mögliche Schwachstellen des Items aufgespürt und neue Formulierungen erarbeitet werden (vgl. Prüfer & Rexroth, 1996, 2000, 2005), was wiederum zu einer besseren Verständlichkeit der Items auf Seiten der Befragten führt, die dementsprechend akkuratere Antworten abgeben können, was schließlich zu einer Reduzierung des Messfehlers beiträgt (vgl. Conrad & Blair, 2009, S. 33; vgl. Willis, 2005, S. 4).

Um die Güte der Skala zu analysieren, wurde ein Drei-Phasen-Verfahren entwickelt, das qualitative und quantitative Methoden kombiniert. In der ersten Phase, die aus quantitativen Erhebungen besteht, wird die Itemanzahl des ursprünglichen Itempools reduziert. Oftmals endet eine Skalenkonstruktion an diesem Punkt. Diejenigen Items, die die besten Kennwerte aufweisen, werden in die Skala aufgenommen, ohne zu wissen, ob die Formulierungen der Items von den Befragten auch so verstanden werden, wie es intendiert ist. Um dies zu überprüfen, werden die verbliebenen Items in Phase 2 mit Hilfe verschiedener kognitiver Methoden getestet und im Falle von Verständnisschwierigkeiten umformuliert. In der dritten Phase, die erneut quantitativ angelegt ist, werden schließlich die Items für die endgültige FWSEM-Skala ausgewählt und ihre Güte bestimmt.

### **1.3 Aufbau der Arbeit**

In der Einleitung wurde zunächst aufgezeigt, dass Sexismus nach wie vor in unserer Gesellschaft existiert und die These aufgestellt, dass Mütter in besonderer Weise von Diskriminierungen betroffen sind, um daraus das Forschungsziel dieser Arbeit abzuleiten, welches in der Entwicklung und Validierung einer ambivalenten Sexismus-Skala in Bezug auf Mütter besteht. In Kapitel 2 folgen nun die theoretischen Grundlagen, auf deren Basis die Skalenkonstruktion vorgenommen wird. Entsprechend den beiden Anliegen dieser

Arbeit erfolgt auch die Unterteilung dieses Kapitels. Während der Schwerpunkt der ersten vier Unterkapitel auf dem inhaltlichen Gegenstand dieser Arbeit, nämlich der traditionellen Arbeitsteilung und der sich daraus ergebenden Diskriminierung von Frauen bzw. Müttern (Sexismus) liegt, widmen sich die letzten beiden Unterkapitel den kognitiven und kommunikativen Prozessen in der Umfrageforschung sowie den methodischen Aspekten der Skalenentwicklung und -testung.

Unterkapitel 2.1 beginnt mit der traditionellen Arbeitsteilung zwischen beiden Geschlechtern. Zunächst wird dargestellt, welche Faktoren das Verharren in geschlechtsspezifischen Arbeitsteilungsmustern begünstigen (Unterkapitel 2.1.1). Neben den institutionellen Rahmenbedingungen in Deutschland werden kulturelle Leitbilder hinsichtlich von Mutterschaft in West- und Ostdeutschland beschrieben. Anschließend wird aufgezeigt, dass nach wie vor eine traditionelle Form der Arbeitsteilung bei Paaren in Deutschland besteht, Frauen deutlich mehr Zeit für unbezahlte Arbeiten aufwenden als Männer und diese Tendenz mit dem Übergang zur Elternschaft weiter zunimmt (Unterkapitel 2.1.2). Im darauffolgenden Kapitel werden die negativen Konsequenzen hinsichtlich des Erwerbslebens von Müttern erläutert, die sich nach der Geburt eines Kindes mit dem Rückfall in traditionelle Arbeitsteilungsmuster und der damit einhergehenden Erwerbsunterbrechung bzw. -reduzierung ergeben und wiederum die Macht- und Statusunterschiede zwischen den Geschlechtern festigen (Unterkapitel 2.1.3).

Unterkapitel 2.2 beschäftigt sich mit Geschlechtervorurteilen und Geschlechterstereotypen, die in zweierlei Hinsicht für diese Arbeit von Bedeutung sind: Zum einen prägen sie u.a. die kulturellen Leitbilder hinsichtlich von Mutterschaft über die Annahme, Mütter seien aufgrund ihres Wesens besonders für die Versorgung und Betreuung des Kindes geeignet (deskriptive Stereotype) – damit können sie Einfluss auf die Entscheidung nehmen, wie innerhalb einer Partnerschaft die Arbeiten aufgeteilt werden und wer beruflich kürzer tritt, um sich der Betreuung der Kinder widmen zu können. Zum anderen ist es möglich, dass es auf der Basis von Geschlechterstereotypen zur Diskriminierung von Müttern im Berufsalltag kommt, wenn sie beispielsweise bei Einstellungen, Weiterbildungen oder Beförderungen übergangen werden, da bestimmte Verhaltenserwartungen (präskriptive Stereotype) an sie gestellt werden.

Zu Beginn des Kapitels wird eine allgemeine Definitionen der Begriffe „Einstellungen“, „Vorurteile“, „Stereotype“ und „Diskriminierung“ gegeben (Unterkapitel 2.2.1), um anschließend den Schwerpunkt auf die jeweiligen Konzepte im Zusammenhang mit der sozialen Kategorie Geschlecht zu richten. In Unterkapitel 2.2.2 wird die klassische Vorurteilsdefinition von Allport erweitert, da Geschlechtervorurteile (wie bereits beschrieben) auch scheinbar positive Formen annehmen können. Unterkapitel 2.2.3 widmet sich den Geschlechterstereotypen en detail. Das Hauptaugenmerk liegt dabei auf den Inhalten der Geschlechterstereotype, die mit verschiedenen Subgruppen von Frauen in Verbindung gebracht werden. Es werden zwei Erklärungsansätze (Social Role Theory und Stereotype Content Model) für die zeitlich und kulturell konsistente Zuschreibung von Geschlechterstereotypen geliefert und dargelegt, welche Rolle Geschlechterstereotype im

Zusammenhang mit Diskriminierungsmechanismen von erwerbstätigen Müttern in der Arbeitswelt spielen („Maternal Wall“) (Unterkapitel 2.3). Im Zentrum der Aufmerksamkeit des folgenden Unterkapitels 2.4 steht die Erfassung sexistischer Einstellungen. Neben der Theorie des traditionellen (Unterkapitel 2.4.2.1) und des modernen Sexismus (Unterkapitel 2.4.2.2) erfährt die Theorie des ambivalenten Sexismus (Unterkapitel 2.4.2.3) von Glick und Fiske (1996, 1997) besondere Aufmerksamkeit, da sie den Ausgangspunkt für die Entwicklung der FWSEM-Skala bildet. Begleitend zu den Theorien werden die in Unterkapitel 1.2 angeführten Messinstrumente näher beschrieben.

Im Anschluss daran werden die theoretischen Grundlagen zur Entwicklung und Testung von Messinstrumenten geliefert. In Unterkapitel 2.5 wird zunächst auf die besonderen kognitiven und kommunikativen Prozesse in Befragungssituationen eingegangen. Zu Beginn wird aufgezeigt, welche Aufgaben Befragte im Rahmen einer standardisierten Befragung zu bewältigen haben (Unterkapitel 2.5.1). Es wird verdeutlicht, welche Unterschiede bei der Sinngebung von Fragen in standardisierten Befragungen im Vergleich zur Alltagskonversation bestehen und welche Probleme sich daraus für die Fragebogenkonstruktion ergeben (Unterkapitel 2.5.2). Um Fehlinterpretationen entgegenzuwirken, gibt es verschiedene Leitlinien zur Itemformulierung, welche zu Beginn des Unterkapitels 2.6 vorgestellt werden. Anschließend werden verschiedene Pretestformen präsentiert sowie deren Stärken und Schwächen aufgezeigt. Besondere Beachtung erfährt dabei die Methode der kognitiven Interviews (Unterkapitel 2.6.7), welche im Rahmen dieser Arbeit eingesetzt wurde, um das Frageverständnis der Befragten zu überprüfen.

Der empirische Teil der Arbeit beginnt in Kapitel 3. Zunächst wird ein kurzer Überblick über den Ablauf der Untersuchungen gegeben und die verschiedenen Phasen dargelegt, die bei der Skalenkonstruktion durchlaufen wurden (Unterkapitel 3.1). Anschließend wird die Generierung des ursprünglichen Itempools erläutert, welcher als Ausgangsbasis für die Skalenkonstruktion dient (Unterkapitel 3.2). Um die Güte der Skala zu analysieren, ist (wie bereits beschrieben) ein Drei-Phasen-Pretest entwickelt worden, der qualitative und quantitative Methoden kombiniert. Jede Phase wird in einem eigenen Unterkapitel (3.3, 3.4 und 3.5) behandelt, wobei zu Beginn jeweils die Methode (Stichprobe, Durchführung und Erhebungsinstrumente) beschrieben und anschließend die Ergebnisse geschildert werden.

In Kapitel 4 werden schließlich die wichtigsten Ergebnisse der Untersuchung zusammengefasst und im Hinblick auf das Forschungsziel der Arbeit diskutiert. Dabei wird neben einem methodischen (Unterkapitel 4.1) auch ein inhaltliches Fazit (Unterkapitel 4.2) gezogen. Zudem werden mögliche Forschungsperspektiven im Zusammenhang mit der neu entwickelten Skala aufgezeigt und verschiedene Forschungsfragen erörtert, die sich im Anschluss an diese Arbeit ergeben haben (Unterkapitel 4.3).

## 2 Theoretische Grundlagen

### 2.1 Geschlechtsspezifische Arbeitsteilung

Wie in der Einleitung ausgeführt wurde, haben Frauen heutzutage sowohl in der privaten als auch in der öffentlichen Sphäre nach wie vor mit Benachteiligungen zu kämpfen, die mit dem Übergang zur Elternschaft (bzw. Mutterschaft) zusammenhängen. „Die Geburt des ersten Kindes erweist sich aus gleichstellungspolitischer Perspektive als konfliktträchtige Statuspassage oder Knotenpunkt (...), die eine Umverteilung von Erwerbsarbeit und generativer Sorgearbeit zwischen den Frauen und Männern zugunsten traditioneller Geschlechterrollen“ nach sich zieht (Sachverständigenkommission, 2011, S. 166). Der Mann übernimmt die Rolle des Versorgers der Familie und steigert in vielen Fällen sogar den Umfang seiner Erwerbstätigkeit, während die Frau sich (vorübergehend) aus dem Erwerbsleben zurückzieht bzw. ihre Erwerbstätigkeit reduziert, um sich der Kinderbetreuung zu widmen (vgl. ebd.). Im Folgenden werden die Faktoren, die einen „Rückfall“ in traditionelle Arbeitsteilungsmuster begünstigen, vorgestellt. Dabei liegen neben den institutionellen Rahmenbedingungen auch die kulturellen Leitbilder in beiden Teilen Deutschland im Fokus des Interesses.

*“(...) [I]n many workplace situations today, women – not women in general, but mothers in particular – are systematically disadvantaged by unexamined assumptions. However, as much as we may profess a belief in gender equality, the persistent notion that a good mother is ‘always there’ for her children serves as an important engine of traditionalist gender roles (...). Women with demanding careers may be judged to be inadequate mothers (...)”*  
(Crosby, Williams, & Biernat, 2004, S. 679).

#### 2.1.1 Institutionelle Rahmenbedingungen und kulturelle Leitbilder als Einflussfaktoren

Pfau-Effingers (2000, S. 68 ff.) Theorie zufolge gehen die Erwerbsbeteiligungsmuster von Frauen und somit auch die Arbeitsteilungsmuster innerhalb von Paarbeziehungen auf zwei wesentliche Faktoren zurück: zum einen auf die sogenannte „Geschlechterkultur“ und zum anderen auf die sogenannte „Geschlechterordnung“. Mit der Geschlechterkultur sind die Werte und Leitbilder innerhalb einer Gesellschaft gemeint. Die Geschlechterordnung bezieht sich auf die jeweiligen Strukturen des Geschlechterverhältnisses eines Landes sowie die verschiedenen gesellschaftlichen Institutionen (v.a. der Arbeitsmarkt, die Familie und der Wohlfahrtsstaat), die relevant für die Arbeitsteilung von Frauen und Männern ist. Als dritten zentralen Begriff nennt Pfau-Effinger den des „Geschlechter-Arrangements“. „Mit dem Geschlechter-Arrangement ist sozusagen das Profil gemeint, das die Geschlechterbeziehungen innerhalb einer Gesellschaft annehmen. (...) Die kulturellen und institutionellen Grundlagen des Arrangements werden auf der Grundlage der

long durée solcher Arrangements sowie von Aushandlungsprozessen und Kompromißbildungen zwischen sozialen Akteuren in der aktuellen historischen Situation gebildet“ (Pfau-Effinger, 2000, S. 71).

Damit stellt die Theorie des Geschlechter-Arrangements von Pfau-Effinger eine Erweiterung von rein wohlfahrtsstaatlichen Theorien dar, die das individuelle Handeln der Paare hinsichtlich ihrer Arbeitsteilung sowie das weibliche Erwerbsverhalten der Mütter durch politische Maßnahmen gelenkt sehen – kulturelle Aspekte werden hierbei vernachlässigt (vgl. Hummelsheim, 2008, S. 34 ff.). Dass kulturelle Leitbilder und Normen jedoch ebenso einen Einfluss auf das Verhalten der Paare ausüben und mit berücksichtigt werden sollten, wird Pfau-Effinger und Geissler (2002) zufolge am Beispiel von Deutschland besonders deutlich. So unterscheiden sich auch nach der Wiedervereinigung beider deutscher Staaten die sozialen Praktiken hinsichtlich der Mutterschaft in West- und Ostdeutschland, obwohl heute für beide Teile die gleichen gesetzlichen Rahmenbedingungen in Bezug auf die Familienpolitik gelten. Trotz des gemeinsamen Rahmens des deutschen Wohlfahrtsstaates herrschen in den alten und den neuen Bundesländern unterschiedliche kulturelle Normen, Werte und Leitbilder hinsichtlich von Mutterschaft vor (siehe Unterkapitel 2.1.1.3), die sich in den 40 Jahren während der Teilung entwickelt und verfestigt haben und einen Einfluss auf die innerfamilialen Arbeitsteilungsmuster und das Erwerbsverhalten von Müttern ausüben (siehe Unterkapitel 2.1.2 und 2.1.3) (vgl. Pfau-Effinger & Geissler, 2002, S. 77 ff.).

Dementsprechend wird im folgenden Kapitel auch auf beide Faktoren Bezug genommen. Zuerst werden in Unterkapitel 2.1.1.1 die institutionellen Rahmenbedingungen vorgestellt. Dabei wird auf drei wesentliche Maßnahmen innerhalb der Familienpolitik eingegangen, die die Arbeitsteilungsmuster der Paare und das weibliche Erwerbsverhalten beeinflussen: Monetäre Unterstützungen, Freistellungsregelungen und öffentliche Kinderbetreuungsmöglichkeiten. An die institutionellen Rahmenbedingungen schließen die Ausführungen in Bezug auf die kulturellen Leitbilder an. Unterkapitel 2.1.1.2 behandelt allgemeine Vorstellungen bezüglich von Mutterschaft – den sogenannten Muttermythos und schildert die erhebliche Verantwortung, die Müttern heutzutage für das (psychologische) Wohlbefinden und die gesunde Entwicklung der Kinder zugeschrieben wird. Daran anknüpfend nimmt Unterkapitel 2.1.1.3 das kulturelle Leitbild der „guten Mutter“ in den Blick und verdeutlicht, welche unterschiedlichen Ansichten hinsichtlich von Mutterschaft in Ost- und Westdeutschland vorherrschen.

### **2.1.1.1 Institutionelle Rahmenbedingungen**

Als wesentliche institutionelle Rahmenbedingungen des Wohlfahrtsstaates, die einen Einfluss auf die innerfamilialen Arbeitsteilungsmuster, die Art und die Dauer der Kinderbetreuung und somit auch auf die Erwerbsbeteiligung der Frauen nehmen, sind nach Hummelsheim (2008) drei familienpolitische Maßnahmenkomplexe zu nennen:

- (a) *Finanzielle Unterstützungen von Familien durch den Staat* (z.B. in Form von Steuervorteilen und Elterngeld),
- (b) *Gewährung von Zeitrechten für die familiäre Kinderbetreuung* (z.B. Mutterschutz und Elternzeit)
- (c) *Bereitstellung von öffentlichen Kinderbetreuungsmöglichkeiten.*

Während die ersten beiden Maßnahmen Erwerbsunterbrechungen ermöglichen und somit die Betreuung der Kinder in den Zuständigkeitsbereich der Familie legen, bietet die dritte Maßnahme Optionen, eine Erwerbstätigkeit wieder aufzunehmen und die Eltern (zeitweise) von der Erziehungsarbeit zu entlasten, indem die Hilfe öffentlicher Betreuungseinrichtungen in Anspruch genommen werden kann (vgl. Hummelsheim, 2008, S. 46 f.).

#### *Steuer- und Sozialsystem.*

Nach wie vor wird die traditionelle Form der Arbeitsteilung einschließlich der familialen Kinderbetreuung (die i.d.R. durch die Mutter übernommen wird) durch das deutsche Steuer- und Sozialsystem gefördert. Zu nennen sind in diesem Zusammenhang monetäre Anreize wie z.B. die beitragsfreie Mitversicherung der Ehefrau in der gesetzlichen Krankenversicherung, die Witwenrente oder das Ehegattensplitting (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 294). Letzteres bietet verheirateten Paaren (auch wenn sie kinderlos sind) die Möglichkeit, ihr Einkommen gemeinsam zu versteuern. Aufgrund des progressiven Steuersatzes fällt der Vorteil am größten aus, wenn nur ein Einkommen vorliegt bzw. die Differenz zwischen den Einkommen beider Ehegatten sehr groß ist. Damit wird das männliche Allein-Versorger-Modell mit einem vollzeiterwerbstätigen Ehemann und einer nicht erwerbstätigen Hausfrau vom Staat gefördert; gleichzeitig werden negative Anreize für eine Erwerbstätigkeit der Frau gesetzt (vgl. Hummelsheim, 2008, S. 65).

#### *Elternzeit und Elterngeld.*

1979 wurde unter der sozial-liberalen Koalition der Mutterschaftsurlaub eingeführt, der berufstätigen Frauen über den Mutterschutz (6 Wochen vor und 8 Wochen nach der Geburt eines Kindes) hinaus die Möglichkeit bot, sich bis zu 6 Monate von ihrer Erwerbstätigkeit freistellen zu lassen. Während dieser Zeit galt ein Kündigungsverbot und sie erhielten vom Staat bis zu 750 DM (vgl. Mühling, Rost, Rupp, & Schulz, 2006, S. 81). 1985 wurde unter der christlich-liberalen Koalition der Mutterschaftsurlaub durch den Erziehungsurlaub und das Erziehungsgeld abgelöst. Prinzipiell waren nun beide Elternteile anspruchsberechtigt und eine vorausgehende Erwerbstätigkeit war für den Erziehungsgeldbezug nicht zwingend erforderlich (so wie dies noch beim Mutterschaftsurlaubsgeld der Fall war, welches nur an erwerbstätige Mütter gezahlt wurde) (vgl. Leitner, 2006, S. 328). Die Regelungen des Erziehungsurlaubes und des Erziehungsgeldes wurden seither mehrfach reformiert. U.a. wurde 1992 der Erziehungsurlaub von ehemals 10 auf 36 Monate ausgeweitet und die Bezugsdauer des Elterngeldes auf zwei Jahre angehoben (vgl.



Mühling et al., 2006, S. 81). Mit dieser (im Vergleich zu anderen europäischen Ländern) relativ langen Zeitspanne sollte den Eltern die Möglichkeit geboten werden, die Kinderbetreuung und -erziehung selbst zu übernehmen (vgl. Hummelsheim, 2008, S. 53). Beide Maßnahmen ließen somit die Wahrscheinlichkeit einer schnellen Rückkehr der Mütter ins Erwerbsleben sinken (vgl. Mühling et al., 2006, S. 81).

2001 wurde unter der rot-grünen Koalition der Erziehungsurlaub in Elternzeit umbenannt und abermals modifiziert. Um nicht einen Elternteil zeitweise auf den Familienbereich zu verweisen, bestand nun die Möglichkeit, dass Mütter und Väter zur gleichen Zeit in Elternzeit gehen. Zudem war es ihnen möglich während ihrer Elternzeit einer Teilzeitarbeit im Umfang von 30 Stunden pro Woche nachzugehen (vgl. Gerlach, 2010, S. 216 f.). Diese Reformen zeugen von einem politischen Leitbildwandel: „Nicht mehr der mehrjährige Berufsausstieg (von Müttern) wurde propagiert, sondern die Parallelität von Kindererziehung und (Teilzeit-) Erwerbstätigkeit für Mütter (und für Väter)“ (Auth, 2007, S. 92).

Auch die Regelungen hinsichtlich des Erziehungsgeldes wurden 2001 flexibilisiert. Eltern konnten sich zwischen einem „Regelbetrag“ (der maximal 300 Euro betrug und über einen Zeitraum von 2 Jahren gezahlt wurde) und einer „Budget-Zahlung“ (von maximal 450 Euro über einen Zeitraum von 1 Jahr) entscheiden (vgl. Hummelsheim, 2008, S. 53). Die Budget-Variante sollte für Mütter (und Väter) einen Anreiz bieten, ihre Erwerbstätigkeit wieder schneller aufzunehmen (vgl. Träger, 2007, S. 157). Das Erziehungsgeld wurde allerdings im ersten halben Jahr nur innerhalb einer engen Einkommensgrenze ausgezahlt und war nach wie vor nicht als Lohnersatzleistung konzipiert. Dies führte dazu, dass fast ausschließlich Frauen die Betreuung der Kinder übernahmen, da auf ihr Einkommen eher verzichtet werden konnte (vgl. Hummelsheim, 2008, S. 53).

Dies änderte sich unter der großen Koalition im Jahr 2007 mit dem neuen Bundeselterngeld- und Elternzeitgesetz. Die Höhe des Elterngeldes beläuft sich seither auf 67% des Nettoeinkommens (mindstens jedoch 300 Euro und höchstens 1.800 Euro) desjenigen Elternteils, welcher seine Erwerbstätigkeit unterbricht bzw. reduziert. Es wird für 12 Monate gezahlt, wenn nur ein Elternteil das Elterngeld in Anspruch nimmt, eine Aufstockung von weiteren 2 Monaten ist möglich, wenn ein Wechsel zwischen Mutter und Vater stattfindet (vgl. BMFSFJ, 2013).

Durch den Einkommensersatzcharakter des Bezuges sowie die beiden Partnermonate sollen Anreize für Väter geschaffen werden, sich stärker an der Kinderbetreuung zu beteiligen. Gleichzeitig soll die neue Regelung zu einer kürzeren Unterbrechung der Erwerbstätigkeit von Müttern führen. Hinzu kommt, dass Frauen durch die Einkommensersatzleistung des Elterngeldes, welches mit einem Mindestbetrag gesockelt ist, eine bessere finanzielle Absicherung zu Gute kommt (vgl. Auth, 2007, S. 97). Aus gleichstellungspolitischer Sicht ist die Regelung des Elterngeldgesetzes nach Einschätzung von Auth (2007) somit positiv zu bewerten. Allerdings sieht sie die Gefahr, dass die genannten positiven Effekte ausbleiben, sollten nicht genügend Kinderbetreuungsplätze bereitgestellt werden.

In diesem Fall könnten die Mütter ihren Beruf nicht wieder aufnehmen, sondern müssten die Kinder zu Hause betreuen (vgl. ebd., S. 97).

### *Kinderbetreuung.*

Im Westen Deutschlands erschweren die strukturellen Bedingungen hinsichtlich der Kinderbetreuung es bis heute den Müttern (und den Vätern), die Familienaufgaben mit einer Erwerbstätigkeit zu vereinbaren. Während in Ostdeutschland der Bedarf an Kinderbetreuungsplätzen für unter Dreijährige als sehr gut zu bezeichnen ist, muss die Betreuungssituation in Westdeutschland als unzureichend eingestuft werden (vgl. Hummelsheim, 2008, S. 72).

Dieses Problem wird seit einigen Jahren von Seiten der Politik verstärkt in Angriff genommen. So wurden mit dem Tagesbetreuungsausbaugesetz (TAG) (1. Januar 2005) einerseits ein Ausbau der Kinderbetreuung für unter Dreijährige und mit dem Kinderförderungsgesetz (KiFöG) (2008) andererseits ein Rechtsanspruch auf einen Betreuungsplatz für jedes Kind mit Vollendung des ersten Lebensjahres ab 2013 beschlossen (vgl. BMFSFJ, 2010a). Erste Erfolge dieser gesetzlichen Regelungen sind zu verzeichnen: Während im Jahr 2006 die Betreuungsquote der unter Dreijährigen im Westen noch bei 12% (Gesamtdeutschland: 13.6%) lag, ist sie 2011 auf knapp 20% (Gesamtdeutschland: 25.2%) gestiegen (vgl. BMFSFJ, 2012, S. 63). Um das gesetzte Ziel des Krippengipfels (2007) von Bund, Ländern und Kommunen für das Jahr 2013 hinsichtlich einer bundesweiten Betreuungsquote von 35% zu erreichen (vgl. BMFSFJ, 2010b), muss der Ausbau an Kinderbetreuungseinrichtungen allerdings noch vorangetrieben werden.

Die Betreuungssituation gestaltet sich darüber hinaus auch aus einem anderen Grund schwierig. Neben der Anzahl der Plätze beeinträchtigen ebenso die zeitlich begrenzten Öffnungszeiten der Betreuungseinrichtungen die Möglichkeiten von Müttern (bzw. Vätern) einer Vollzeiterwerbstätigkeit nachzugehen. Dies gilt sowohl für Kindergartenkinder als auch für Grundschulkinder, da es nach wie vor zu wenige Ganztagsplätze bzw. weiterführende Betreuungsmöglichkeiten in den Nachmittagsstunden gibt. Ohne genügend außerhäusliche Betreuungsmöglichkeiten haben Eltern, denen sich nicht die Möglichkeit informeller Betreuungsarrangements (z.B. durch die Großeltern) bietet, keine andere Wahl als ihre Erwerbstätigkeit zu unterbrechen und die Kinder selbst zu betreuen (i.d.R. durch die Mutter). Das Angebot an öffentlicher Kinderbetreuung übt somit einen wesentlichen Einfluss auf die „generelle Chance“ einer mütterlichen Erwerbstätigkeit aus, die Zeitstrukturen der Einrichtungen bestimmen zudem den möglichen „Umfang“ der Erwerbstätigkeit (vgl. Hummelsheim, 2008, S. 48 ff.).

### *Zusammenfassung*

Insgesamt kann festgehalten werden, dass die deutsche Familienpolitik inkonsistente Anreize bietet: So wurden in den vergangenen Jahren verschiedene Maßnahmen eingeführt,

die ein Aufbrechen der traditionellen Geschlechterarrangements, eine bessere Vereinbarkeit von Beruf und Familie sowie eine erhöhte Erwerbstätigkeit von Frauen zum Ziel haben. Zu nennen sind in diesem Zusammenhang v.a. die (Neu-)Regelungen der Elternzeit und des Elterngeldes (was die größere Flexibilität, die beiden Partnermonate und den Einkommensersatzcharakter anbelangt) sowie der zunehmende Ausbau an öffentlicher Kinderbetreuung.

Institutionelle Rahmenbedingungen, die zum heutigen Zeitpunkt nach wie vor die geschlechtsspezifische Arbeitsteilung mit der Übernahme der Kinderbetreuung durch die Frau und eine lange Unterbrechung ihrer Erwerbstätigkeit stützen bzw. Erwerbstätigkeit nur in Teilzeit nahe legen, sind mitunter:

- (1) die beitragsfreie Mitversicherung der Ehefrau in der gesetzlichen Krankenversicherung (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 294),
- (2) das Ehegattensplitting,
- (3) die Möglichkeit, eine Elternzeit von insgesamt maximal drei Jahren zu nehmen,
- (4) das bislang unzureichende Angebot an öffentlichen Kinderbetreuungseinrichtungen (insbesondere für unter Dreijährige) in Westdeutschland sowie die zeitlich begrenzten Öffnungszeiten der Einrichtungen (vgl. Hummelsheim, 2008, S. 71).

Nach den institutionellen Rahmenbedingungen sollen nun die kulturellen Leitbilder, die in West- und Ostdeutschland vorherrschen, in den Blick genommen werden.

Judith Holofernes, Sängerin von Wir sind Helden:

*„Kinder zu kriegen macht unsichtbare Grenzen sichtbar. Und zeigt, wie unterschiedlich Männer und Frauen ihr Elternsein gespiegelt kriegen. Mein Mann [namens Pola, Anm. der Autorin] und ich sind in der Erziehung so gleichberechtigt wie überhaupt möglich. Trotzdem fällt auf, dass Pola permanent Schulterklopfen für identische Erziehungsarbeiten erntet, die mir entsprechend von meinem Arbeitskonto abgezogen werden, als würde ich mich davor drücken. Wenn ich ein Interview gebe und nebenan schreit unser Baby an Polas Schulter, werde ich oft gefragt, wie ich mich dabei fühle; eine Frage, die ihm sicher nie gestellt wird.*

*Da schwingt viel Rabenmutterdenken mit (...)*“

(Freitag, 2011, S. 26)

### 2.1.1.2 Der Muttermythos

Die „Mutter“ ist nach Crawford und Unger (2004) eines der grundlegendsten Archetypen von Frauen. In den meisten Gesellschaften wird Mutterschaft als wesentlich für die Identität und Erfüllung einer Frau angesehen. Sie stellt ein biologisches Privileg dar, das lediglich Frauen vorbehalten ist. Ihre Bedeutung geht jedoch über den biologischen Prozess der Reproduktion hinaus und umfasst verschiedene Gepflogenheiten, Ansichten,

Einstellungen und Normen. In vielen westlichen Gesellschaften herrschen ähnliche Überzeugungen hinsichtlich der Mutterschaft vor, die als „Ideologie der Mutterschaft“ oder als „Muttermythos“ („motherhood mystique“) bezeichnet werden. Diese Ideologie schließt den Autorinnen zufolge folgende Punkte ein:

- (1) *Mutterschaft als Berufung*. Mutterschaft ist die ultimative Erfüllung für eine Frau. Sie stellt eine natürliche und notwendige Erfahrung für alle Angehörigen des weiblichen Geschlechts dar. Denjenigen Frauen, die nicht Mütter werden möchten, wird Unverständnis entgegengebracht. Diejenigen, die Mütter werden wollen, aber nicht können, werden bemitleidet.
- (2) *Natürliche Prädisposition*. Frauen sind von Natur aus für die Betreuung und Versorgung von Kindern prädestiniert. Gute Mütter erfüllt diese Form der Arbeit, sie genießen sie. Frauen, für die dies nicht zutrifft, sind entweder schlecht organisiert oder von ihrem Wesen her anormal.
- (3) *Mütterliche Hingabe*. Eine Mutter besitzt eine grenzenlose Geduld und zeigt eine grenzenlose Bereitschaft, sich für ihre Kinder aufzuopfern. Wenn sie ihre eigenen Bedürfnisse nicht zurückstellt, ist sie keine gute Mutter („Rabenmutter“).
- (4) *Unabkömmlichkeit der Mutter*. Eine intensive, ganztägige Zuwendung (Rundum-die-Uhr-Betreuung) durch die eigene Mutter ist das Beste für das Kind. Frauen, die einer Berufstätigkeit nachgehen, sind Mütter zweiter Klasse („Rabenmütter“).

In unserer Gesellschaft existieren somit Mythen über beides, die „gute Mutter“ und die „schlechte Mutter“ („Rabenmutter“). Die „gute Mutter“ ist eine endlose Quelle des Hegens und Pflegens, sie weiß von Natur aus wie Kinder aufgezogen werden sollten und stellt ihre persönlichen Bedürfnisse hinter denen des Kindes zurück (vgl. Crawford & Unger, 2004, S. 317 ff.).<sup>2</sup> „Rabenmütter“ hingegen sind solche Mütter, die sich (scheinbar) nicht genügend um ihre Kinder kümmern und dadurch mögliche Fehlentwicklungen des Kindes zu verantworten haben (vgl. Mühling et al., 2006, S. 46); ihr Verhalten wird als unnormales und schädlich für das Kind angesehen. Obwohl diese Ansichten veraltet zu sein scheinen, bestehen diese Mythen heutzutage fort. Sie halten Einzug in Erziehungsratgeber, die Mütter die Hauptverantwortung für die Kinderbetreuung zuschreiben, Prozesse des Bemutterns ausführlich schildern und Selbstaufopferung preisen (vgl. Crawford & Unger, 2004, S. 317 ff.).

Dienel (2003) analysierte die Stile des Gebärens, Ratgeber-Literatur sowie Beiträge in Internetforen und kommt zu ähnlichen Ergebnissen: Das Leitbild der Mutterschaft ist in Deutschland durch eine sanfte, nicht-medikamentisierte Geburt, das Stillen des Kindes,

---

<sup>2</sup> Der Muttermythos kann somit als Form des benevolenten Sexismus (siehe Kapitel 2.4.2.3) angesehen werden. So wird Frauen aufgrund ihres natürlichen Wesens die Mutterrolle zugeschrieben, deren Ausübung sie – scheinbar ebenso von Natur aus – glücklich macht (vgl. Crawford & Unger, 2004, S. 317 ff.).

einer mütterlichen Erwerbsunterbrechung von ein bis drei Jahren sowie ein Zurückstellen der mütterlichen Wünsche und Bedürfnisse gekennzeichnet.

*„Typisch für deutsche Frauen scheint die Angst davor symbolträchtige Augenblicke wie den ersten Schritt oder das erste Wort des Kindes zu verpassen. Das Kind wurde nicht geboren, damit es von anderen erzogen wird. Die Bereitschaft zu Verzicht im finanziellen und beruflichen Bereich und in der persönlichen Freiheit ist sehr hoch, während auf die vollständige Präsenz beim Aufwachsen des Kindes nicht verzichtet wird“ (Dienel, 2003, S. 136).*

In der Tat wird in unserer heutigen Gesellschaft den Müttern eine erhebliche Verantwortung für das (psychologische) Wohlbefinden der Kinder zugeschrieben, wie es im Laufe der Geschichte bisher nur wenige Kulturen getan haben (vgl. Crawford & Unger, 2004, S. 354 ff.). Die (intensive) Betreuung der Kinder durch die Mutter wurde in den letzten Jahren immer bedeutsamer. Folglich ist es nicht verwunderlich, dass Frauen heutzutage nicht nur mehr Zeit als Männer mit der Kinderbetreuung und -versorgung verbringen (siehe Unterkapitel 2.1.2), sondern sie auch mehr Zeit als frühere Generationen von Müttern aufwenden. Selbst berufstätige Mütter verbrachten im Jahr 2000 mehr interaktive Zeit mit ihren Kindern, als dies Mütter ohne außerhäuslichen Beruf im Jahr 1975 taten. Diese waren häufig durch die anfallenden Hausarbeiten wie Kochen, Waschen oder Bügeln so in Anspruch genommen, dass sie keine Zeit hatten, sich bewusst den Kindern zu widmen (vgl. Eagly & Carli, 2007, S. 52 ff.). Folglich waren die Mütter zwar in unmittelbarer Nähe der Kinder, diese wuchsen aber häufig nebenher auf und waren sich selbst überlassen oder wurden in jüngeren Jahren von älteren Geschwistern oder den Großeltern betreut. Folgendes Zitat einer 71-jährigen Großmutter verdeutlicht diese Entwicklung:

*“My daughter spends more time with her son than I did with my kids, and I spent more time with my kids than my mother did. My mother really didn't have the time. She had to work on the farm so she didn't help me with homework, spend time playing with us, or take us places ... Even when they have jobs, a lot of women today are pressured. They have to spend lots of time with their children, to make them happier, smarter ... to give them every advantage” (Eagly & Carli, 2007, S. 53).*

Dementsprechend lastet heutzutage ein außerordentlicher Druck auf Müttern, ihre Kinder bestmöglich zu fördern und ihnen die besten Rahmenbedingungen für ihre Entwicklung zu ermöglichen. Insbesondere gut ausgebildete Mütter sind hinsichtlich ihrer mütterlichen Qualitäten/Fähigkeiten selbstkritisch und haben ein schlechtes Gewissen, wenn sie einer Berufstätigkeit nachgehen. Dabei zeigten Tagebuchstudien, dass gut ausgebildete Mütter mehr „Qualitätszeit“ mit ihren Kindern verbringen als weniger gut ausgebildete. Und trotz ihrer Berufstätigkeit stehen sie den Vollzeitmüttern in fast nichts

nach. Eine Studie aus dem Jahr 2000 belegt, dass Mütter, die keiner Erwerbstätigkeit nachgehen, fünf Stunden mehr pro Woche für die Interaktionen mit ihren Kindern verwenden als erwerbstätige Mütter (vgl. Eagly & Carli, 2007, S. 53 f.).

Dennoch scheinen vor diesem Hintergrund für viele Mütter – gerade in den ersten Lebensjahren des Kindes – eine Erwerbsarbeit und v.a. eine berufliche Karriere mit Mutterschaft unvereinbar zu sein (vgl. Vinken, 2001, S. 7). Die Konsequenzen, die sich daraus ergeben, sind zweierlei: Entweder entscheiden sich junge (insbesondere hochqualifizierte) Frauen bewusst gegen Kinder, da sie sich nicht zutrauen, neben ihrer Berufstätigkeit den hohen Ansprüchen des Mutterideals zu genügen (vgl. Kortendiek, 2008, S. 435 f.), oder sie entscheiden sich für Kinder und entschließen sich zu einer (mehrjährigen) Reduzierung der Erwerbstätigkeit bzw. sogar einer Berufsunterbrechung, womit u.a. eine finanzielle Abhängigkeit vom Ehepartner einhergeht.

Die Anforderungen, die an Väter gestellt werden, sind weniger vereinnahmend und lassen sich auch leichter mit einer Erwerbstätigkeit vereinbaren. Auch wenn der väterlichen Anwesenheit und Zuwendung mittlerweile mehr Bedeutung als in früheren Zeiten beigemessen wird, werden Väter immer noch in erster Linie als Versorger der Familie angesehen (vgl. Mühling et al., 2006, S. 45).

Insgesamt gesehen gilt Mutterschaft als natürliche Berufung und Erfüllung einer Frau. Aufgrund ihres Wesens ist sie dem Muttermythos zufolge für die Betreuung der Kinder prädestiniert, sie sollte grenzenlose Hingabe und Opferbereitschaft bei der Ausübung dieser Aufgabe zeigen. Ihre Anwesenheit wird für das Wohlergehen und die Entwicklung des Kindes als unabdingbar angesehen. Dabei hat ihre Bedeutung für die Kinderbetreuung im Vergleich zu vergangenen Zeiten zugenommen. Dies sieht man u.a. daran, dass Frauen dafür nicht nur mehr Zeit als Männer aufwenden, sondern auch mehr Zeit als frühere Generationen von Frauen. Dementsprechend lastet auf Müttern ein großer Druck, ihre Kinder bestmöglich zu fördern und ihnen die besten Rahmenbedingungen für ihre Entwicklung zu bieten. Inwieweit der Muttermythos sich in den kulturellen Frauen- und Familienleitbildern in beiden Teilen Deutschlands niederschlägt, soll im Folgenden dargelegt werden.

### ***2.1.1.3 Frauen- und Mütterleitbilder in West- und Ostdeutschland***

In Westdeutschland herrschte in den 1950er und 1960er Jahren das kulturelle Leitbild der männlichen Versorgerehe vor, welches auf das Bürgertum des 19. Jahrhunderts zurückgeht. Mit dem Übergang zur industriell-kapitalistischen Gesellschaft vollzog sich eine Trennung der privaten und der öffentlichen Sphäre. Frauen und Männer sollten – ihren jeweiligen natürlichen Anlagen entsprechend – in diesen Bereichen unterschiedliche Rollen übernehmen und sich in ihren Tätigkeiten gegenseitig ergänzen (vgl. Pfau-Effinger & Geissler, 2002, S. 78). Während Frauen als passiv, gefühlsbetont und nachgiebig charakterisiert wurden, schrieb man Männern Rationalität und Aktivität zu (siehe hierzu auch Unterkapitel 2.2.3.2). Aufgrund ihres (vermeintlich) natürlichen

Wesens wurden Frauen für die „personenbezogenen Sorge- und Betreuungstätigkeiten in der Familie“ als besonders geeignet angesehen. Dementsprechend fielen ihnen die familiären Aufgaben der Haushaltsführung und vor allem der Kinderbetreuung zu. Männern hingegen sprach man eine besondere Eignung für „sachbezogene, produktive Tätigkeiten in Wirtschaft, Politik Kultur und Wissenschaft“ zu. Sie übernahmen die Rolle als Versorger der Familie und somit die außerhäusliche Erwerbstätigkeit (vgl. Klement & Rudolph, 2003, S. 27).<sup>3</sup>

Das vorherrschende kulturelle Frauenleitbild in der Nachkriegszeit war das der „guten Hausfrau“. Bis Anfang der 1970er Jahre konnte es mit dem Leitbild der „guten Mutter“ nahezu gleichgesetzt werden, „welches für die ersten Lebensjahre jede dauerhafte Fremdbetreuung ausschließt, (...) eine mehrjährige Berufsunterbrechung [impliziert, d. Verf.] und sich gleichzeitig nur schwer mit dem Leitbild der ‚selbständigen Frau‘ vereinbaren [lässt, d. Verf.]“ (Kortendiek, 2008, S. 435). Im Zuge der Modernisierung der Gesellschaft (was beispielsweise den Bereich der Bildung anbelangte) und dem durch die Neue Frauenbewegung angeregten gesellschaftlichen Diskurs über Gleichberechtigung löste sich diese Verknüpfung jedoch. Frauen stellten zunehmend die traditionelle Arbeitsteilung und ihre primäre Verantwortung für die Haus- und Familienarbeiten in Frage. Die alleinige Hausfrauenrolle (ohne die Verbindung mit der Mutterrolle) wurde vermehrt als überholt angesehen (vgl. Pfau-Effinger & Geissler, 2002, S. 78 f.) und die Erwerbstätigkeit von Frauen zunehmend gesellschaftlich gebilligt. Weiterhin wurde jedoch an der mütterlichen Betreuung der Kinder im eigenen Elternhaus festgehalten (vgl. Auth, 2007, S. 90; Pfau-Effinger, 2005, S. 4).

Bis heute hat sich an dem Leitbild der „guten Mutter“ wenig geändert – vielmehr besteht im Westen Deutschlands nach wie vor ein „konservatives Mutterrollenkonzept“ (Onnen-Isemann, 2007, S. 177). Obwohl für Frauen von heute eine gute Ausbildung sowie eine kontinuierliche Erwerbstätigkeit zunehmend an Bedeutung gewinnt, messen Mütter nach wie vor der Betreuung der Kinder zu Hause während der Phase der aktiven Mutterschaft eine hohe Bedeutung bei. Staatliche, externe Kinderbetreuung wird von den meisten Westdeutschen – sowohl von Vätern als auch von Müttern – als mögliche Ergänzung zur privaten Kinderbetreuung eingestuft, nicht jedoch als ein angemessener Ersatz (vgl. Pfau-Effinger & Geissler, 2002, S. 79).

Während die mütterliche Fürsorge als unabdingbar, gewissermaßen als Pflicht erachtet wird, ist eine väterliche Fürsorge hingegen lediglich wünschenswert. Die Mutter wird folglich als Hauptverantwortliche für die Betreuung der Kinder angesehen. Ihre Anwesenheit – und somit ihre zeitliche und persönliche Verfügbarkeit – ist für eine gesunde Entwicklung des Kindes maßgeblich. Die Basis für das kindliche Wohlergehen

---

<sup>3</sup> In diesem Zusammenhang wird auf verschiedene psychologische (wie z.B. psychoanalytische, lerntheoretische, kognitive, kulturpsychologische und evolutionsbiologische) Ansätze verwiesen, die versuchen Geschlechtsunterschiede in Motivation und Verhalten zu erklären (siehe hierzu beispielsweise Asendorpf & Neyer, 2012).

wird bereits unmittelbar nach der Geburt gelegt, in dieser Zeit sollte die Mutter idealerweise das Kind selbst stillen (vgl. Mühling et al., 2006, S. 46).

*„Die ‚moderne‘ Variante attribuiert dies als ‚Selbstverwirklichung in der Mutterrolle‘ (...) und auch viele junge Mütter betrachten ihre Zuständigkeit für die Babys so: Sie möchten ‚etwas von Ihrem [sic] Kind haben‘ im Sinne von die Mutterfreuden genießen, das Kind heranwachsen sehen etc. (...) Dies wird sehr klar assoziiert mit der Pause im Beruf und dem ausschließlichen Dasein für Kind und Haushalt“ (Mühling et al., 2006, S. 46).*

Sobald ein Kind zur Welt kommt, wird in Westdeutschland demzufolge mehrheitlich erwartet, dass die Mutter den Umfang ihrer Erwerbstätigkeit reduziert bzw. vorübergehend aus dem Erwerbsleben aussteigt. So vertraten im Rahmen des ISSP 2002 etwa 97% der befragten Frauen und Männer im Westen die Ansicht, dass Frauen halbtags bzw. überhaupt nicht erwerbstätig sein sollten, so lange ein Kind noch nicht zur Schule geht. Eine Rückkehr in Form einer Vollzeitbeschäftigung wird von der Mehrheit der Befragten im Westen (rund 70%) für die Mutter erst akzeptiert, wenn die Kinder das Elternhaus verlassen haben (siehe hierzu Cornelißen, 2005, S. 760 ff.; Abbildung A 5.2). Auch die Ergebnisse des Generations and Gender Surveys sowie der European Values Study aus dem Jahr 2008 zeigen, dass eine außerhäusliche Kinderbetreuung in den alten Bundesländern weniger akzeptiert wird als in den neuen. So wird in Westdeutschland signifikant häufiger die Ansicht vertreten, Kinder könnten am besten durch die eigenen Eltern betreut werden, seien sie hingegen den größten Teil der Woche in einer Kinderbetreuungseinrichtung untergebracht, hätten sie vermutlich im späteren Leben mit Problemen zu kämpfen. Gehe die Mutter einer Erwerbstätigkeit nach, werde ein Kind, das sich noch im Vorschulalter befinde, darunter leiden (vgl. Schneider & Dorbritz, 2011, S. 32 f.).

Das sozialistische Frauen- und Mutterleitbild im Osten Deutschlands sah hingegen folgendermaßen aus: Dort sollten Frauen und auch Mütter durchgehend einer Vollzeiterwerbstätigkeit nachgehen (vgl. Kortendiek, 2008, S. 436).

*„In der DDR war diejenige Frau eine ‚gute Mutter‘ – und gehörte zur anerkannten Gruppe ‚unsere Muttis‘ –, die erwerbstätig und somit keine Hausfrau war und ihr Kind in einer öffentlichen Ganztageseinrichtung betreuen ließ. Jedoch ist hier festzuhalten, dass auch in der DDR die Sozialpolitik stärker auf Mutterschaft und Beruf und nicht etwa auf die Vereinbarkeit von Elternschaft und Beruf ausgerichtet war“ (Kortendiek, 2008, S. 436).*

Auch in den neuen Bundesländern hielt man an der Zuständigkeit der Frauen für den familiären Bereich fest – ihre Gleichberechtigung sollte über ihre Einbeziehung in das Erwerbsleben erzielt werden (vgl. Schäfers, 2000, S. 93).



Dementsprechend erfährt auch heute noch in den neuen Bundesländern die durchgängige Erwerbsarbeit von Müttern eine breite Zustimmung – allerdings wird auch hier mehrheitlich eine Teilzeitarbeit präferiert, wenn das Kind noch im Kindergarten- oder im Schulalter ist und die Mutter sich somit in der aktiven Familienphasen befindet (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 298).

Zusammenfassend lässt sich Folgendes sagen: In beiden Teilen Deutschlands wurden und werden Frauen den kulturellen Leitbildern entsprechend als Hauptverantwortliche für den privaten, familiären Bereich angesehen. Unterschiede bestanden und bestehen nach wie vor im Hinblick auf die Akzeptanz der Erwerbstätigkeit von Müttern kleiner Kinder sowie bezüglich der gesellschaftlich gebilligten Betreuungsform. Während im Westen die Betreuung der Kinder im privaten Haushalt durch die Mutter favorisiert wurde, war es im Osten Deutschlands von jeher üblich, dass die Mutter vollzeiterwerbstätig war und die Kinder frühzeitig außerhäuslich betreut wurden. Der Muttermythos entspricht folglich voll und ganz dem Leitbild einer „guten Mutter“ in Westdeutschland. In Ostdeutschland wird hingegen eine intensive, ganztägige Zuwendung durch die eigene Mutter nicht als unabdingbar angesehen.

Der Muttermythos besteht nach Ansicht von Crawford und Unger (2004) weiterhin fort, da dies der einzige Bereich ist, in dem westliche Gesellschaften Verbundenheit und Fürsorge über persönlichen (beruflichen) Erfolg stellen. Die Glorifizierung der Mutterschaft, die Frauen gleichzeitig einem erheblichen Druck aussetzt, und ein schlechtes Gewissen bei Nicht-Gerechtwerten erzeugt, begünstigt traditionelle Geschlechterrollen. Frauen werden ermutigt, andere Bereiche ihres Lebens, wie etwa den eigenen Beruf, zugunsten der Mutterschaft zurückzustellen, was wiederum eine ökonomische Abhängigkeit von (Ehe-)Männern bedingt und Geschlechterhierarchien festigt. Liberale Feministinnen betonen, dass die Institution Mutterschaft dazu verwendet wird, um Frauen vom öffentlichen Leben auszuschließen. Sie haben aufgezeigt, wie der Muttermythos dazu verwendet wird, um Frauen auf ihre traditionelle Rolle zu beschränken („keep in their place“) (vgl. Crawford & Unger, 2004, S. 317 ff.). „The social order that elevates men over women is legitimated by women’s devotion to child care, since it takes them out of the running for top-level jobs and political positions and defuses their consciousness of oppression“ (Lorber, 1994, S. 170).

Entsprechend der eingangs vorgestellten These, dass nach wie vor keine Geschlechtergerechtigkeit im privaten Bereich besteht und Mütter insbesondere von den Ungleichheiten hinsichtlich der Verantwortlichkeit für die Reproduktionsarbeiten betroffen sind, soll nun im Folgenden untersucht werden, inwieweit sich die institutionellen Rahmenbedingungen und die kulturellen Leitbilder einer „guten Mutter“ in den heutigen Arbeitsteilungsmustern der Paare in West- und Ostdeutschland widerspiegeln. Hierfür werden die Zeitaufwendungen von Frauen und Männern für die unbezahlten Familienarbeiten (zu denen u.a. die Kinderbetreuung und die Versorgung des Haushaltes zählen) betrachtet. Dabei wird analysiert, ob sich insbesondere bei Paaren mit Kindern traditionelle Arbeitsteilungsmuster abzeichnen, die Frauen die primäre

Verantwortung für den familiären Bereich zuschreiben und sie somit auf eine traditionelle Rolle verweisen.

### 2.1.2 Ausgestaltung der Haus- und Familienarbeiten

Um zu dokumentieren, wie sich Paare in der heutigen Zeit die Haus- und Familienarbeiten aufteilen, werden im Folgenden vor allem die Ergebnisse der beiden repräsentativen Zeitbudgetstudien des Statistischen Bundesamtes aus den Jahren 1991/92 und 2001/02 herangezogen, die einen zeitlichen Vergleich erlauben (vgl. Sachverständigenkommission, 2011, S. 151).<sup>4</sup> Da sie u.a. auf Tagebuchaufzeichnungen beruhen, geben sie vermutlich eine exakte(re) und weniger fehleranfällige Darstellung der tatsächlich aufgewendeten Zeiten einzelner Haushaltsmitglieder wieder, als dies beispielsweise bei Selbsteinschätzungen der Befragten oder auch insbesondere bei Fremdeinschätzungen darüber, wie viel Zeit der/die PartnerIn investieren, der Fall sein dürfte. Allerdings könnten auch hier Verzerrungen auf Grund eines sozial erwünschten Antwortverhaltens auftreten (siehe Unterkapitel 2.5.1).<sup>5</sup>

Laut Angaben des Statistischen Bundesamtes verbrachten Frauen im Jahr 1991/92 im Westen Deutschlands mit unbezahlter Arbeit<sup>6</sup> 1.8-mal so viel Zeit wie Männer, in Ostdeutschland knapp 1.6-mal so viel. Vergleicht man die Entwicklung über den Zeitverlauf hinweg, zeigt sich folgendes Bild: Für jede Stunde, die ein Mann mit Haus- und Familienarbeiten verbringt, sind es im Jahr 2001/02 hingegen „nur noch“ 1.6 Stunden bei einer Frau in Westdeutschland und 1.4 Stunden bei einer Frau in Ostdeutschland. Folglich hat sich das Verhältnis in den letzten 10 Jahren angeglichen – dabei ist es sowohl in den 1990er Jahren als auch in den 2000er Jahren in den neuen Bundesländern etwas ausgewogener als in den alten Bundesländern – nach wie vor ist es aber in beiden Gebieten Deutschlands nicht egalitär (vgl. BMFSFJ, Statistisches Bundesamt, 2003, S. 14).

Schaut man sich in diesem Zusammenhang nicht die relativen Verteilungen, sondern die absoluten Zahlen an, wird deutlich, dass die zunehmende Parität auf eine Reduzierung

---

<sup>4</sup> Zeitbudgetstudien neueren Datums liegen derzeit nicht vor (vgl. Sachverständigenkommission, 2011, S. 151).

<sup>5</sup> Fthenakis, Kalicki & Peitz (2002, S. 109) verweisen im Zusammenhang mit Vor- und Nachteilen verschiedener methodischer Vorgehensweisen zur Erfassung der Zeitaufwendungen von Frauen und Männern für unbezahlte Arbeit auf folgende Quelle: Twiggs, J. E., McQuillan, J., & Ferree, M. M. (1999). *Meaning and measurement: Reconceptualizing measures of the division of household labor*. *Journal of Marriage and the Family*, 61, 712-724.

<sup>6</sup> Um die unbezahlte Arbeit von etwa Freizeitaktivitäten zu unterscheiden, wurde das sogenannte „Dritt-Personen-Kriterium“ herangezogen. Dementsprechend werden alle Tätigkeiten, die prinzipiell auch gegen ein Entgelt von einer dritten Person geleistet werden könnten, unter dem Oberbegriff „unbezahlten Arbeiten“ subsumiert. Dazu zählen neben den klassischen Haushaltstätigkeiten wie Kochen, Waschen, Putzen auch Handwerksarbeiten, die Betreuung von Kindern, die Pflege von Alten und Kranken sowie ehrenamtliche Tätigkeiten (vgl. Pinl, 2004, S. 20). Die Sorgearbeit für ältere Menschen sowie die ehrenamtlichen Tätigkeiten spielen im Zusammenhang mit einer Traditionalisierung der häuslichen Arbeit nach dem Übergang zur Elternschaft keine Rolle und wären daher bei eigenen Analysen ausgeklammert worden.

der aufgewendeten Zeit von Frauen zurückzuführen ist. Den Zeitbudgetstudien zufolge gingen Frauen im Jahr 1991/92 pro Woche 35 Stunden, Männer hingegen lediglich 19 ½ Stunden unbezahlter Arbeit nach (vgl. BMFSFJ, 2003, S. 130), im Jahr 2001/02 sank die aufgewendete Zeit der Frauen auf 31 Stunden, die der Männer blieb unverändert (vgl. Pinl, 2004, S. 22). So verbrachten Frauen im Jahr 2001/02 knapp 10% weniger Zeit mit Haus- und Familienarbeiten als 10 Jahre zuvor (vgl. BMFSFJ, Statistisches Bundesamt, 2003, S. 14 f.).

Mit der Anwesenheit von betreuungsbedürftigen Kindern im Haushalt fallen die Ungleichheiten hinsichtlich der Verantwortlichkeit für die Reproduktionsarbeiten noch deutlicher zu Ungunsten der Frauen aus: Während bei kinderlosen Paaren die Frauen im Jahr 2001/02 mit gut 4 Stunden pro Tag etwa 1.4-mal (eigene Berechnung) so viel Zeit mit unbezahlter Arbeit verbringen wie Männer (mit gut 2 ¾ Stunden), leisten Mütter mit gut 6 ¼ Stunden fast 2-mal so viel Haus- und Familienarbeit wie Väter (mit knapp 3 ¼ Stunden) (siehe Tabelle 1) (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 306; Gille & Marbach, 2004, S. 92 f.). Selbst bei Konstanzhaltung des stundenmäßigen Umfangs der Erwerbstätigkeit beider Elternteile, zeigt sich, dass Mütter mehr Zeit für Hausarbeit aufwenden als Väter (vgl. Pinl, 2004, S. 24).

*Tabelle 1: Zeitaufwendungen von Paaren mit und ohne Kindern für unbezahlte und bezahlte Arbeiten im Jahre 2001/02*

	Paare ohne Kind(er)		Paare mit Kind(ern)	
	Frauen	Männer	Frauen	Männer
Unbezahlte Arbeit	04:07	02:51	06:16	03:10
Bezahlte Arbeit	03:08	04:33	01:55	05:13
Gesamtarbeit	07:15	07:24	08:11	08:23

*Anmerkungen.* Datenbasis: Zeitbudgetstudie 2001/02; Auswertungen von Gille und Marbach, 2004; Angaben in Stunden und Minuten pro Tag. Dabei handelt es sich um Durchschnittswerte, die sowohl aus den Angaben der Befragten an Wochentagen als auch am Wochenende gemacht wurden. Dementsprechend niedrig(er) fallen auch die Werte für die Erwerbsarbeit aus. Der Begriff „Kinder“ meint betreuungsbedürftige Kinder im Alter unter 15 Jahren. Einbezogen wurden Paare im Alter unter 60 Jahren, die einer Erwerbstätigkeit nachgehen.

Quelle: Tabelle übernommen aus Dressel, Cornelißen und Wolf (2005, S. 306).

Betrachtet man die unbezahlte Arbeit nicht im Gesamten, sondern wirft einen Blick allein auf die Betreuung von kleinen Kindern im Alter unter 6 Jahren, welche immerhin mehr als ein Drittel der gesamten Reproduktionsarbeit beansprucht, zeigt sich folgendes Bild: Sowohl Mütter als auch Väter investieren im Jahr 2001/02 mehr Zeit in die Kinderbetreuung als noch im Jahr 1991/92 (im Osten: fast 1 ¼ Stunden; im Westen: 45

Minuten pro Tag).<sup>7</sup> Eine Parität ist auch in diesem Bereich noch nicht erreicht. Im Vergleich zu Männern, die knapp 1 ¼ Stunden pro Tag mit ihren Kindern verbringen, widmen Frauen ihren Kindern 2 ¾ Stunden pro Tag. Dass Frauen auch für die Betreuung des Nachwuchses mehr Zeit aufbringen, gilt nicht nur für Paarbeziehungen, in denen sich die Frau komplett aus dem Erwerbsleben zurückgezogen hat, sondern auch für Partnerschaften, in denen beide einer außerhäuslichen Arbeit nachgehen: Während nicht erwerbstätige Mütter mit gut 3 ¼ Stunden pro Tag 3-mal so viel Zeit für die Kinderbetreuung aufwenden wie erwerbstätige Väter (1 Stunde und 6 Minuten), ist es bei erwerbstätigen Müttern immerhin noch das Zweifache der Zeit (2 Stunden und 10 Minuten) (vgl. BMFSFJ, 2003, S. 22 ff.).

Die Querschnittsdaten der Zeitbudgetstudien belegen somit eindeutig, dass Frauen nach wie vor mehr Zeit für die unbezahlten Arbeiten aufwenden als Männer. Die Anwesenheit von Kindern im Haushalt übt dabei in der Tat einen Einfluss auf die Arbeitsteilung aus: So zeigt sich, dass bei Paaren mit (betreuungsbedürftigen) Kindern im Haushalt das Ungleichgewicht hinsichtlich der Verantwortlichkeit für die Reproduktionsarbeit noch stärker zu Lasten der Frauen verteilt ist, als dies bei Paaren ohne Kinder der Fall ist. Dass der Übergang zur Elternschaft zu einer (Re-)Traditionalisierung der Arbeitsteilung führt und Männer ihre unbezahlten Arbeiten noch weiter reduzieren, wurde Fthenakis, Kalicki und Peitz (2002) zufolge durch mehrere Längsschnittstudien belegt.

Nicht nur hinsichtlich des zeitlichen Umfangs, sondern auch in der Art der Tätigkeiten zeigen sich sowohl für die Hausarbeit als auch die Kinderbetreuung geschlechtsspezifische Unterschiede. So gibt es Zuständigkeitsbereiche, die überwiegend von Frauen übernommen werden (sogenannte „weibliche Tätigkeiten“) und solche, für die eher Männer verantwortlich sind (sogenannte „männlich Tätigkeiten“). Frauen übernehmen die alltäglichen Routinearbeiten, die häufiger, regelmäßiger und zu festgelegten Zeitpunkten ausgeführt werden müssen, wie etwa das Zubereiten der Mahlzeiten, die Pflege der Wäsche oder das Aufräumen und Reinigen der Wohnung. Gleiches gilt auch für die kindbezogenen Arbeiten: Auch hier sind Frauen für die täglich anfallenden Pflege- und Betreuungsaufgaben verantwortlich, die zeitlich festgelegt und aufwendig sind (z.B. Windeln wechseln, An- und Ausziehen, Füttern und Baden). Häufig übernehmen sie auch die Wege für und mit dem Kind (wie etwa die Fahrten zum Sportverein). Damit sind die Frauen deutlich in ihrer Flexibilität eingeschränkt (vgl. Steinbach, 2004, S. 2 f.). Zudem werden die Arbeiten der Frauen seltener anerkannt. Sie sind „unsichtbar“ und fallen nur negativ auf, wenn sie unzureichend ausgeführt werden oder unerledigt bleiben (z.B. wenn Kinder sich schlecht benehmen oder schmutzige Kleidung tragen) (vgl. Woollett & Mars-

---

<sup>7</sup> In den neuen Bundesländern ist dieser Trend auf zwei Aspekte zurückzuführen: Einerseits ist die außerhäusliche Kinderbetreuung von unter Dreijährigen zurückgegangen und andererseits sank die Erwerbsteilung von Müttern mit Kinder in einem Alter von unter sechs Jahren (vgl. BMFSFJ, 2003, S. 24).

hall, 2001, S. 180). Männer hingegen haben mehr Wahlfreiheiten, was die (zeitliche) Gestaltung ihrer Haus- und Familienaufgaben (wie beispielsweise Reparaturarbeiten, Regelung der finanziellen Angelegenheiten, Behördengänge oder die Pflege des Autos) anbelangt. Ihre Tätigkeitsbereiche sind zeitlich klarer begrenzt, treten seltener und unregelmäßiger auf. Folglich fällt es Männern leichter, sich bei den anfallenden Aufgaben einzubringen, sich ihnen aber auch wieder zu entziehen. Da Männer in der Regel weniger die Grundversorgung des Kindes übernehmen, sondern häufig an Unternehmungen von besonderen Freizeitaktivitäten oder Wochenendausflügen beteiligt sind (Beschäftigung mit dem Kind) (vgl. Steinbach, 2004, S. 3), werden diese Familienarbeiten von den Kindern auch eher bemerkt und positiv gewürdigt (vgl. Woollett & Marshall, 2001, S. 180).

Zusammenfassend kann im Hinblick auf die Verteilung der Haus- und Familienarbeiten folgendes festgehalten werden: Nach wie vor tragen Frauen zu den unbezahlten Arbeiten stärker bei als Männer. In den letzten Jahren hat sich eine zunehmende Parität in der Arbeitsteilung innerhalb der Paarbeziehungen abgezeichnet. Die Angleichung der Verhältnisse ist jedoch weniger auf ein größeres Engagement der Männer als vielmehr auf eine Reduzierung des weiblichen Zeitanteils zurückzuführen. Mit der Geburt des ersten Kindes findet eine zusätzliche Verlagerung der Verantwortlichkeiten für die unbezahlten Arbeiten zu Ungunsten der Frauen statt, da sie den Großteil der kindbezogenen Arbeiten rund um dessen Pflege und Betreuung übernehmen. Männer hingegen gehen v.a. der Beschäftigung mit dem Kind (wie z.B. dem Sport, Spielen und anderen Freizeitaktivitäten) am Wochenende nach.

### **2.1.3 Folgen der geschlechtsspezifischen Arbeitsteilung für die Stellung der Frauen im Erwerbsleben**

Um mehr Zeit für die Haus- und Familienarbeiten und v.a. die Kinderbetreuung aufwenden zu können, reduzieren Mütter den Umfang ihrer Erwerbstätigkeit teilweise beträchtlich, d.h. der Anstieg der unbezahlten Arbeiten geht mit einer Reduzierung der bezahlten Arbeiten einher. Väter hingegen verstärken ihr berufliches Pensum sogar (vgl. Sachverständigenkommission, 2011, S. 166).

Zieht man erneut die Daten der Zeitbudgetstudie aus dem Jahr 2001/02 heran (siehe Tabelle 1), zeigen sich folgende Zeitverwendungsmuster für Paare mit und ohne Kinder: Die *Gesamtarbeitszeiten* fallen sowohl bei Müttern als auch bei Vätern etwa eine Stunde höher aus, als bei Frauen und Männern ohne Kinder. Der Anstieg dieser Gesamtarbeitszeit verteilt sich für Frauen und Männer jedoch nicht gleichermaßen auf unbezahlte und bezahlte Tätigkeiten. Sowohl Mütter als auch Väter wenden mehr Zeit für die *unbezahlten* Arbeiten auf, als dies Frauen und Männer ohne Kinder tun (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 305 f.). Für Mütter fällt jedoch der Anstieg der aufgewendeten Zeit sehr viel deutlicher aus. Während sie 2 Stunden und 9 Minuten mehr investieren als Frauen ohne Kinder, sind es bei den Vätern im Vergleich zu Männern

ohne Kinder nur 19 Minuten.<sup>8</sup> Im Hinblick auf die *bezahlte* Arbeit zeichnet sich Folgendes ab: Während Mütter im Vergleich zu kinderlosen Frauen bezüglich des außerhäuslichen Arbeitspensums fast  $1\frac{1}{4}$  Stunden weniger investieren, wenden Väter im Vergleich zu kinderlosen Männern fast eine  $\frac{3}{4}$  Stunde mehr Zeit auf. Der Anstieg der Gesamtarbeitszeit bei Paaren mit Kindern gegenüber Paaren ohne Kinder lässt sich somit bei Frauen vollständig auf einen Anstieg der unbezahlten Arbeiten zurückführen (bei gleichzeitiger Reduzierung der bezahlten Tätigkeiten). Väter investieren hingegen in beiden Bereichen mehr Zeit, wobei der eindeutig größere Teil der außerhäuslichen Erwerbstätigkeit zufällt.

Nach der Geburt eines Kindes investieren Mütter folglich nicht nur mehr Zeit als Väter in unbezahlte Reproduktionsarbeiten (2-mal so viel), sondern auch mehr Zeit als Frauen ohne Kinder (1.5-mal so viel). Gleichzeitig gehen sie der außerhäuslichen Erwerbstätigkeit in geringerem Maße nach. Verglichen mit Vätern wenden sie nur das 0.7-fache der Zeit auf, im Vergleich zu kinderlosen Frauen das 0.6-fache.

Mit dem Übergang zur Elternschaft „werden [somit] arbeitsteilige Zeitverwendungsmuster zwischen Erwerbsarbeit und generativer Haus- und Sorgearbeit bei beiden Geschlechtern eingeübt, die sich oftmals im weiteren Lebensverlauf verstetigen und zu Nachteilen für die berufliche Entwicklung von Müttern und ihre Alterssicherung führen“ (Sachverständigenkommission, 2011, S. 166).

Auf diese Problematik beziehen sich die nachfolgenden Ausführungen. Zunächst wird beschrieben, wie sich der Übergang zur Elternschaft auf die Erwerbsbeteiligung, den Umfang der Erwerbstätigkeit sowie die Mobilität von west- und ostdeutschen Müttern auswirkt. Anschließend wird auf die negativen Konsequenzen eingegangen, die sich für Mütter aufgrund familienbedingter Erwerbsunterbrechungen und/oder -reduzierungen hinsichtlich ihrer beruflichen Erfolgchancen und ihres Erwerbseinkommens ergeben.

### *Erwerbsbeteiligung*

Betrachtet man die Erwerbsbeteiligung, so lässt sich feststellen, dass die Erwerbsquote der Frauen in den letzten Jahren stetig gestiegen ist. Dennoch lag sie 2005 mit 12.5 Prozentpunkten nach wie vor deutlich unter der Rate der Männer. Dabei beeinträchtigt die Familiengründung – und somit die Anwesenheit von betreuungsbedürftigen Kindern – die Erwerbsbeteiligung von Frauen (vgl. Allmendinger, Leuze & Blank 2008, S. 20). Während Frauen im mittleren Alter ohne Kinder etwa in gleichem Maße einer Erwerbstätigkeit nachgehen wie Männer, trifft dies für Frauen mit Kindern nicht zu. Insbesondere Mütter jüngerer Kinder sind seltener erwerbstätig (Dressel, Cornelißen, & Wolf, 2005, S.281). Im Jahr 2008 gingen lediglich 29% der Mütter mit einem Kind unter drei Jahren aktiv einer Erwerbstätigkeit nach. Erst zwischen dem dritten und fünften Lebensjahr des Kindes (also

---

<sup>8</sup> Anzumerken ist in diesem Zusammenhang, dass sich die Daten der Zeitbudgetstudie auch auf das Wochenende beziehen. Im Falle der Väter dürfte der Anstieg der unbezahlten Arbeiten daher vermutlich zu einem großen Teil auf die Beschäftigung mit dem Kind (wie z.B. diverse Freizeitunternehmungen) zurückzuführen sein (siehe Unterkapitel 2.1.2).

mit Erreichen des Kindergartenalters) nimmt die Erwerbsbeteiligung der Mütter wieder zu und steigt auf 59%. Auf das Erwerbsverhalten der Väter haben die Anwesenheit und das Alter von Kindern im Haushalt weitestgehend keinen Einfluss. Je nach Alter des jüngsten Kindes variierte die Erwerbsbeteiligung der Väter zwischen 83 und 86% (vgl. Statistisches Bundesamt, 2010b, S. 1).

Schaut man sich die alten und die neuen Bundesländer getrennt an, so fällt auf, dass ostdeutsche Mütter im Jahr 2008 über alle Altersgruppen der Kinder hinweg häufiger einer Erwerbstätigkeit nachgingen als westdeutsche Frauen (unter Dreijährige: Ost = 33%, West = 27%; 3- bis 5-Jährige: Ost = 63%, West = 57% und 6- bis 9-Jährige: Ost = 71%, West = 64%). Hauptgrund für diese Unterschiede dürften die bessere Betreuungslage von Kleinkindern im Osten (vgl. Statistisches Bundesamt, 2010b, S. 2 f.) sowie die gesellschaftlich stärker akzeptierte Erwerbstätigkeit der Mütter sein (siehe Unterkapitel 2.1.1.3).

### *Umfang der Erwerbstätigkeit*

Neben der generellen Erwerbsbeteiligung ist auch der zeitliche Umfang der Erwerbstätigkeit von Bedeutung, will man die Gleichstellung von Frauen und Männer bzw. Müttern und Vätern auf dem Arbeitsmarkt beurteilen. So ist die gestiegene Erwerbsbeteiligung von Frauen insbesondere auf eine Ausweitung ihrer Teilzeiterwerbstätigkeit zurückzuführen (vgl. Allmendinger, Leuze & Blank 2008, S. 21 ). Während im Jahr 2001 rund 40% der Frauen teilzeitbeschäftigt waren, stieg die Quote bis 2008 auf 46% an. (Die Teilzeitquote der Männer lag hingegen 2008 lediglich bei 9%) (vgl. Sachverständigenkommission, 2011, S. 92).

Wie in Bezug auf die Erwerbsbeteiligung, spielt auch im Zusammenhang mit dem Arbeitszeitvolumen die Anwesenheit von Kindern im Haushalt eine Rolle. Mit dem Übergang zur Elternschaft reduzieren viele Frauen den Umfang ihrer Erwerbstätigkeit und ziehen sich aus dem Erwerbsleben zurück oder gehen nur noch einer Teilzeitbeschäftigung nach (vgl. Allmendinger, Leuze & Blank 2008, S. 21.) Im Jahr 2008 übten 69% der erwerbstätigen Mütter *mit* (und nur 36% der erwerbstätigen Frauen *ohne*) minderjährige Kinder im Haushalt eine Teilzeitbeschäftigung aus. Bei Vätern waren es hingegen nur 5% (vgl. Statistisches Bundesamt, 2009b). Mit Beginn der Familienphase (ab 25 Jahren) vergrößert sich somit die *Arbeitszeitlücke* der Frauen, d.h. gemessen an ihrem Anteil der Beschäftigten tragen sie einen kleineren Anteil zum gesamtwirtschaftlichen Arbeitsvolumen bei. Männer hingegen steigern sogar ihr Arbeitspensum nach der Geburt des ersten Kindes und erzielen einen *Arbeitszeitüberhang* (vgl. Allmendinger, Leuze & Blank 2008, S. 21).

Vergleicht man abermals die alten und die neuen Bundesländern, zeigt sich, dass im Jahr 2008 deutlich mehr Mütter in Ostdeutschland (53%) mit Kindern unter 15 Jahren einer Vollzeiterwerbstätigkeit nachgingen als Mütter in Westdeutschland (22%). Sind die Mütter teilzeiterwerbstätig, so sind die Motive hierfür unterschiedlicher Natur: Während Mütter im Westen überwiegend (89%) aus persönlichen, familiären Gründen eine Teilzeitbeschäftigung ausübten, traf dies nur auf 52% der ostdeutsche Mütter zu. Weitere

38% gaben in den neuen Bundesländern an, sie seien nur daher teilzeiterwerbstätig, da nicht genügend Vollzeitstellen am Arbeitsmarkt zu Verfügung stünden (vgl. Statistisches Bundesamt, 2010b, S. 2 f.)

### *Mobilität*

Eine repräsentative Studie aus dem Jahr 2007 zum Mobilitätsverhalten von erwerbstätigen Frauen und Männern in sechs europäischen Ländern kam zu folgendem Ergebnis: Solange noch keine Kinder in einer Paarbeziehung vorhanden sind, ist der Anteil mobiler Frauen und Männer gleich hoch (jeweils 27%). Dies ändert sich jedoch mit dem Übergang zur Elternschaft. Es zeigen sich deutliche Geschlechtsunterschiede dahingehend, dass 23% der Väter auch weiterhin mobil sind, bei den Müttern beläuft sich der Anteil hingegen nur noch auf 6%. Während es für Mütter somit schwierig zu sein scheint, nach der Geburt eines Kindes mobil zu bleiben, lässt sich für Väter die Elternschaft mit Mobilität offenbar besser vereinbaren – zwischen ihrem Verhalten und dem kinderloser Männer besteht kaum ein Unterschied (vgl. Sachverständigenkommission, 2011, S. 164 f.).

### *Berufliche Erfolgchancen*

Familienbedingte Erwerbsunterbrechungen, Reduzierungen der Arbeitszeiten und eine eingeschränkte Mobilität vermindern wiederum die beruflichen Erfolgchancen von Müttern. Dies macht sich bereits beim Einstieg ins Berufsleben bemerkbar. In einer Studie der Hans-Böckler-Stiftung aus dem Jahr 2003 begründeten 13% der befragten Personalverantwortlichen die Tatsache, dass sie in den letzten fünf Jahren keine Frau eingestellt hatten, mit dem Risiko einer Schwangerschaft (also einer potentiellen Mutterschaft, die Ausfallzeiten der Angestellten und damit einhergehende Kosten für das Unternehmen nach sich zieht). Weitere 9% führten als Grund die fehlende zeitliche Flexibilität von Frauen an (vgl. Obermeier, 2006).

Ähnliches gilt auch in Bezug auf Weiterbildungsmaßnahmen. So investieren Arbeitgeber aufgrund eventueller familienbedingter Erwerbsunterbrechungen in die Fortbildung junger weiblicher Angestellter weniger als in die Qualifizierung junger männlicher Mitarbeiter (vgl. Sachverständigenkommission, 2011, S. 104).

Würden sich Väter und Mütter gleichermaßen um die Betreuung der Kinder kümmern, sei es generell in den ersten Lebensjahren oder auch im Krankheitsfall des Kindes, könnte hier bei Einstellungen die „potentielle Gefahr des Ausfallens“ nicht mehr am Geschlecht festgemacht werden. Folglich wäre es für Personalverantwortliche egal, ob sie einem Mann oder einer Frau den Job gäben oder in deren Weiterqualifizierung investierten: In beiden Fällen müssten sie damit rechnen, dass es familienbedingte Ausfallzeiten gibt (siehe im Zusammenhang mit Weiterbildungsmaßnahmen auch Sachverständigenkommission, 2011, S. 104).

Aber nicht nur der berufliche Einstieg oder die Weiterqualifizierung, sondern auch der Aufstieg fällt unter diesen Umständen schwer. Die Anforderungen in der Arbeitswelt



sind – und dies gilt im verstärkten Maße für hoch dotierte Positionen – nach wie vor an männlichen Erwerbs- und Lebensmustern ausgerichtet. Damit verbunden sind u.a. ein früher Einstieg ins Erwerbsleben, ein hoher zeitlicher und persönlicher Einsatz, durchgängige Anstellungen ohne längere familienbedingte Erwerbsunterbrechungen (vgl. Crosby, Williams, & Biernat, 2004, S. 677), eine hohe Präsenz am Arbeitsplatz (oftmals bis in die späten Abendstunden) sowie mitunter eine geographische Mobilität und zeitliche Flexibilität.<sup>9</sup>

Mütter können diesen Anforderungen meist nur selten gerecht werden. So wird der Grundstein für eine berufliche Karriere häufig in einer Lebensphase gelegt, die zugleich mit der Familiengründung zusammenfällt (der sogenannten „Rush Hour of Life“) (Sachverständigenkommission, 2011, S. 16) und die folglich oftmals mit einer Erwerbsunterbrechung sowie einer sich anschließenden Reduzierung der Erwerbstätigkeit für Mütter einhergeht (siehe Unterkapitel 2.1.2). Auch die verminderte räumliche Flexibilität dürfte für Frauen berufliche Nachteile mit sich bringen.

*„Eine Karriere, die berufliche Mobilität einschließt, erscheint unter der Bedingung von Mutterschaft deutlich erschwert. Umgekehrt können berufliche Mobilitätsanforderungen die Familienentwicklung für Frauen verzögern oder gänzlich verhindern. Dieser Konflikt wird zum einen von den Frauen selbst antizipiert und in zukünftigen Entscheidungen bei der eigenen Lebensplanung berücksichtigt. Zum anderen dürften diese Zusammenhänge auch für die Entscheidungen anderer Akteure auf dem Arbeitsmarkt (z.B. Personalverantwortliche) eine Rolle spielen und könnten eine weitere Benachteiligung von Frauen auf dem Arbeitsmarkt hervorrufen“* (Sachverständigenkommission, 2011, S. 165).

Folglich haben Frauen es auch wesentlich schwerer, ihre Karriere weiter voranzutreiben und in beruflich leitende Positionen aufzusteigen. Nicht selten müssen Frauen bei ihrer Rückkehr ins Erwerbsleben sogar einen beruflichen Abstieg in Kauf nehmen. Je länger dabei die familienbedingten Auszeiten waren, desto geringer ist auch die Chance auf eine gleichwertige Position bei demselben Arbeitgeber zurückkehren zu können. Bei einer Re-Integration in den Arbeitsmarkt nach fünf Jahren liegt die Chance bei 50% und sinkt bei längeren Erwerbsunterbrechungen bis auf 16% (vgl. Sachverständigenkommission, 2011, S. 104).

Dementsprechend sind Frauen nach wie vor in Führungspositionen in der Politik, der Wissenschaft oder der Wirtschaft unterrepräsentiert (sogenannte vertikale Segregation des deutschen Arbeitsmarktes) (siehe Unterkapitel 1.1). Erneut lohnt ein Blick auf den Familienstand der Frauen. Während Frauen ohne Kinder unter den abhängig erwerbstätigen Angestellten mit umfassenden Führungsaufgaben im Jahr 2004 mit 26% vertreten

---

<sup>9</sup> Folglich kann man bei Führungspositionen auch von „Eineinhalb-Personen-Berufen“ sprechen. Um all diese Anforderungen erfüllen zu können, braucht man eine Partnerin bzw. einen Partner, die bzw. der einem den Rücken im privaten Bereich frei hält (vgl. Hoecker, 1999, S. 68; Weber & Schaeffer-Hegel, 2000).

waren, sinkt ihr Anteil laut Daten des Mikrozensus mit der Anwesenheit von einem Kind im Haushalt auf 19%, bei zwei und mehr Kindern liegt er nur noch bei etwa 10% (vgl. hierzu Abbildung 5.16 in Dressel, Cornelißen, & Wolf, 2005, S. 291). Die Gründung einer Familie und die höheren Teilzeitquoten der Mütter nehmen folglich auf den oberen Ebenen der beruflichen Laufbahn ebenso einen negativen Einfluss: Auch hier nimmt mit steigender Kinderzahl der Anteil der Frauen in Spitzenpositionen ab (vgl. Dressel, Cornelißen, & Wolf, 2005, S. 290). Die Frauen, die den beruflichen Aufstieg geschafft haben, sind häufiger kinderlos als Männer: Während 68% der Frauen in leitenden Positionen keine Kinder unter 18 Jahren haben, die noch im Elternhaushalt leben, trifft dies auf 47% der Männer zu (vgl. BMFSFJ, BMBF, BMVBS, BMAS, & BMWI, 2006, S. 9).

### *Erwerbseinkommen*

Angesichts der institutionellen Rahmenbedingungen (mangelnde außerhäusliche Kinderbetreuungsmöglichkeiten) und des kulturellen Leitbildes der „guten Mutter“ im Westen Deutschlands scheint es nicht verwunderlich, dass junge Frauen häufig (im Gegensatz zu jungen Männern) die Entscheidung für ihren zukünftigen Beruf mit Blick auf seine Vereinbarkeit mit der Familie treffen. So hängt die Wahl des Ausbildungsplatzes mitunter davon ab, ob sich die angestrebte Beschäftigung auch mit einer Teilzeittätigkeit verbinden lässt. Zumeist trifft dies auf sogenannte „Sackgassenberufe“ (wie z.B. Arzthelferin oder Verkäuferin) zu, die gering entlohnt werden und nur wenige Aufstiegsmöglichkeiten bieten. Frauen schränken sich auf diese Weise bereits bei der Berufswahl selbst ein (vgl. Hoecker, 1999, S. 67).

Generell sind Frauen häufiger in schlechter bezahlten Lohngruppen sowie in Branchen mit Niedriglöhnen anzutreffen. So sind Frauen überwiegend in Dienstleistungs- und Pflegeberufen beschäftigt, die laut Tarifvertrag mit einem geringeren Entgelt versehen sind als technische oder naturwissenschaftliche Berufe, die häufig von Männern ausgeübt werden (sogenannte horizontale Segregation des Arbeitsmarktes in weibliche und männliche Tätigkeitsbereiche) (vgl. Allmendinger, Leuze & Blank 2008, S. 24).

Die aus der horizontalen Segregation des Arbeitsmarktes resultierenden Einkommensunterschiede zwischen beiden Geschlechtern haben wiederum einen Einfluss auf die Entscheidungen hinsichtlich der familialen Arbeitsteilung. Wenn es also um die Frage geht, wer zu Hause bleiben und die Betreuung der Kinder übernehmen soll, entscheiden sich viele Paare für die Frau, da ihr Verdienst tendenziell geringer ausfällt (vgl. Dressel & Wanger, 2008, S. 486) und somit die Opportunitätskosten hinsichtlich des Haushaltseinkommens weniger hoch sind.

Wie bereits in den vorangegangenen Abschnitten erwähnt, sind es häufig die Frauen, die aus familiären Gründen ihre Erwerbstätigkeit unterbrechen. Die Rückkehr ins Erwerbsleben erfolgt überwiegend in Teilzeit und ist häufig mit einem beruflichen Abstieg verbunden – ein Aufstieg in eine leitende Position fällt schwer, so dass es zu einer vertikalen Segregation des Arbeitsmarktes und Einkommensverlusten kommt (vgl. Allmendinger, Leuze & Blank 2008, S. 23; Dressel & Wanger, 2008, S. 484).

Die traditionelle Arbeitsteilung hinsichtlich der Kinderbetreuung ist somit mit einer Reihe von Faktoren verbunden (siehe unmittelbare und mittelbare Diskriminierungen in Unterkapitel 2.3.1 und Maternal Wall in Unterkapitel 2.3.2), die zu den Lohnunterschieden zwischen den Geschlechtern („gender wage gap“) beitragen. Aber auch bei einem Vergleich der tatsächlichen Stundenlöhne, bei dem das unterschiedliche Arbeitsvolumen von Frauen und Männern ausgeblendet wird, schneiden Frauen schlechter ab (vgl. Allmendinger, Leuze & Blank 2008, S. 23 f.). Der Bruttostundenlohn vollzeiterwerbstätiger Frauen liegt seit mehreren Jahren im Durchschnitt 23% unter dem ihrer männlichen Kollegen (vgl. Fuchs & Bothfeld, 2011, S.14 f.). (Die Versdienstunterschiede im Jahr 2008 betragen in Westdeutschland 25% und in Ostdeutschland 5%) (vgl. Statistisches Bundesamt, 2009a). Selbst die bereinigte Lohndifferenz, bei der alle Einflussgrößen für die Lohnhöhe wie beispielsweise das Alter, der Ausbildungsgrad oder die Berufserfahrung konstant gehalten werden, beträgt noch 12 Prozent zu Ungunsten der Frauen und offenbart somit eine Diskriminierung aufgrund des Geschlechts (vgl. Allmendinger, Leuze & Blank 2008, S. 23 f.).

Geringere Löhne, familienbedingte Erwerbsunterbrechungen und -reduzierungen weiblicher Arbeitnehmerinnen wirken sich wiederum negativ auf die Höhe ihrer Renteneinkommen aus (vgl. Dressel & Wanger, 2008, S. 487). Betrachtet man die Renteneinkommen, die Frauen aus eigener Erwerbstätigkeit zustehen, so zeigt sich seit Jahrzehnten dasselbe Bild: Frauen steht etwa nur die Hälfte der Altersrente der Männer für ihren Lebensunterhalt zur Verfügung (vgl. Fuchs & Bothfeld, 2011, S. 15).

### *Zusammenfassung*

Mit dem Übergang zur Elternschaft gehen verschiedene Faktoren einher, die für Frauen unmittelbare und mittelbare Diskriminierungen nach sich ziehen (können), welche wiederum soziale Ungleichheiten zwischen den Geschlechtern bedingen, aber auch Mütter im Vergleich zu kinderlosen Frauen schlechter stellen: Die vorangegangenen Ausführungen haben gezeigt, dass Frauen, die keine Kinder haben, die Arbeitsmarktintegration vergleichsweise gut gelingt. Sie sind nahezu genauso häufig erwerbstätig wie Männer und üben deutlich häufiger eine Vollzeitbeschäftigung aus als Frauen mit Kindern. In leitenden Positionen sind Frauen ohne Kinder zwar längst nicht so gut vertreten wie ihre männlichen Kollegen, ihr Anteil liegt aber immer noch höher als der von Müttern.

Nach der Geburt eines Kindes greifen viele Paare auf traditionelle Muster der Arbeitsteilung zurück. Während Väter den Umfang ihrer Erwerbstätigkeit sogar erhöhen, reduzieren Mütter ihr Arbeitspensum oder steigen (vorübergehend) aus dem Berufsleben aus, um die Kinderbetreuung übernehmen zu können. Diese Brüche im Erwerbverlauf sowie Teilzeitbeschäftigungsverhältnisse erschweren den Wiedereintritt in den Arbeitsmarkt und führen häufig zu einem „Karriereknick“ oder sogar einem „Karriereaus“ von Müttern. Sie mindern die Chancen auf eine Führungsposition und tragen bedeutend zu den Einkommensunterschieden zwischen den Geschlechtern bei. Mit dem geringeren Ver-

dienst der Frau geht eine finanzielle Abhängigkeit vom (Ehe-)Mann einher, zudem reduzieren sich ihre Ansprüche auf Sozialleistungen des Staates (z.B. bezüglich der Rente), die Macht- und Statusunterschiede zwischen den Geschlechtern werden gefestigt. Mütter sind also in der Tat in unserer Gesellschaft besonders von Benachteiligungen betroffen.

## 2.2 Geschlechtervorurteile und Geschlechterstereotype

Dieses Kapitel beschäftigt sich mit Geschlechtervorurteilen und Geschlechterstereotypen, die in zweierlei Hinsicht für diese Arbeit von Bedeutung sind. Zum einen sind sie Bestandteil der kulturellen Leitbilder hinsichtlich von Mutterschaft und können somit einen Einfluss auf die Entscheidung von Paaren ausüben, wer die Kinderbetreuung übernimmt.

*„Eine Traditionalisierungsfalle stellen (...) geschlechtsspezifische Deutungen bei Kinderbetreuung und Hausarbeit dar, mit denen die (...) Arbeitsteilung begründet wird. Besonders die Zuschreibung, dass Mütter quasi natürliche Kompetenzen für Kinderbetreuung und Haushalt besäßen, hat nach der Geburt des ersten Kindes auf die Arbeitsteilung zwischen den Partnern häufig starken Einfluss. Einige institutionelle und wohlfahrtsstaatliche Regelungen fördern dieses Verständnis, darunter das Ehegattensplitting“ (Sachverständigenkommission, 2011, S.168).*

Zum anderen ist es möglich, dass es auf der Basis von Geschlechterstereotypen zur Diskriminierung von Müttern im Berufsalltag kommt, indem sie beispielsweise bei Einstellungen, Weiterbildungen oder Beförderungen von Personalverantwortlichen übergangen werden (siehe Unterkapitel 2.3).

### 2.2.1 Einstellungen, Vorurteile, Stereotype und Diskriminierungen

In der psychologischen Forschung lässt sich eine Vielzahl unterschiedlicher Definitionen finden, die überwiegend darin übereinstimmen, dass Einstellungen als eine zusammenfassende Bewertung sozialer „Objekte“ verstanden werden. Exemplarisch soll die Definition von Eagly und Chaiken (1993, p. 1) angeführt werden: Eine *Einstellung* ist “a psychological tendency that is expressed by evaluating a particular entity with some degree of favour or disfavour”. Die beiden zentralen Bestandteile dieser Definition sind zum einen das Einstellungsobjekt und zum anderen die Bewertungstendenz. Als Einstellungsobjekt kommt alles in Frage, was eine Person sich abstrakt vorstellen (z.B. Feminismus) oder konkret wahrnehmen kann, wie etwa Objekte (z.B. eine lila Latzhose), Personen (z.B. Alice Schwarzer) oder auch soziale Gruppen (z.B. Feministinnen, Karrierefrauen) (vgl. Bohner, 2003, S. 267).

Die Bewertungstendenzen in Bezug auf das Einstellungsobjekt selbst können nicht direkt beobachtet werden. Allerdings ruft das Einstellungsobjekt Reaktionen unterschiedlicher Valenz hervor (positive oder negative), die einen Rückschluss auf die Einstellung zulassen (siehe Abbildung 1). Diese Verhaltensreaktionen werden häufig in drei Klassen

untergliedert: *kognitiv*, *affektiv* und *konativ*. Die *kognitive Klasse* bezieht sich auf die Gedanken, Meinungen oder Überzeugungen, die ein Individuum hinsichtlich des Einstellungsobjektes hegt. Die *affektive Klasse* umfasst die Gefühle oder Emotionen im Zusammenhang mit dem Einstellungsobjekt und die *konative Klasse* beinhaltet nicht nur konkrete Handlungen, die sich auf das Einstellungsobjekt beziehen, sondern auch Verhaltensabsichten (vgl. Bohner, 2003, S. 267).

Diese drei Reaktionsklassen sind nicht immer eindeutig voneinander trennbar, so dass es auch empirisch schwierig ist, sie beispielweise in Faktorenanalysen nachzuweisen. Zudem werden nicht alle Klassen auch in jeder Einstellung repräsentiert. So ist es möglich, dass eine Einstellung ausschließlich aus einer kognitiven oder einer affektiven Komponente besteht. Folglich wird häufig lediglich ein eindimensionales Einstellungskonzept verwendet, welches eine zusammenfassende Evaluation darstellt (vgl. Bohner & Wänke, 2002, S. 5).

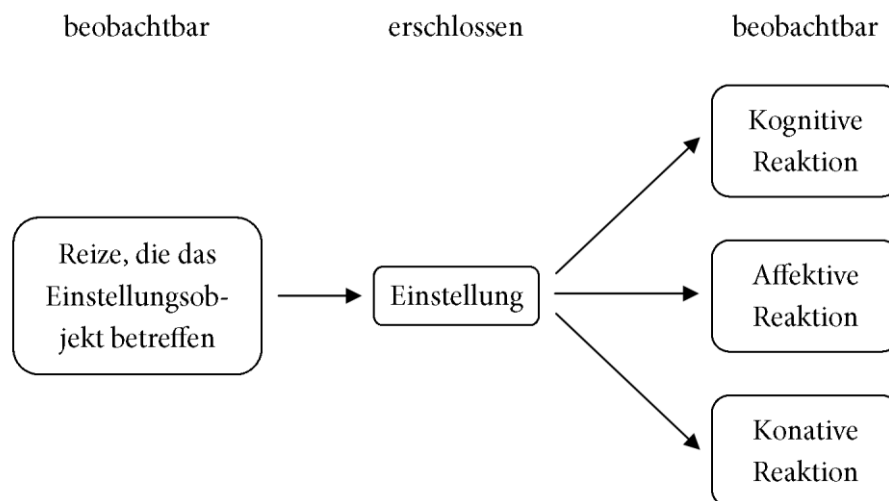


Abbildung 1. Einstellungsmodell nach Eagly und Chaiken (1993) mit drei Reaktionsklassen. Quelle: Abbildung übernommen aus Eagly und Chaiken (1993). Eigen Übersetzung.

Beziehen sich die Einstellungen auf soziale Gruppen und deren Mitglieder so werden sie (vor allem, wenn sie negativ konnotiert sind) als *Vorurteile* bezeichnet (vgl. Bohner, 2003, S. 267). Der klassischen Definition von Allport (1954) zufolge sind Vorurteile eine "antipathy based upon a faulty and inflexible generalization. It may be felt or expressed. It may be directed toward a group as a whole, or toward an individual because he is a member of that group" (S. 9). Unberücksichtigt bleibt in dieser Begriffsbestimmung, dass Vorurteile auch positive Inhalte aufweisen können. Gerade in den letzten Jahrzehnten hat

die Erforschung positiver Vorurteile und deren negativer Folgen an Bedeutung gewonnen (siehe im Zusammenhang mit Geschlechtervorurteilen Unterkapitel 2.2.2).

*Stereotype* wiederum stellen die kognitive Klasse der Reaktionen von Vorurteilen dar. Wenn das Einstellungsobjekt also eine soziale Gruppe ist, kann die kognitive Klasse der Reaktionen (die Gedanken, Überzeugungen, Meinungen und Vorstellungen) mit den Stereotypen über diese Gruppe gleichgesetzt werden (synonymer Gebrauch der Begrifflichkeiten). Stereotype werden dabei als eine Generalisierung von negativen oder auch positiven Merkmalen definiert, die ein Individuum allen Mitgliedern einer sozialen Gruppe zuschreibt (vgl. Eagly & Mladinic, 1989, S. 543 f.).

Mummendey, Kessler und Otten (2009, S.48) zufolge wird in der sozialpsychologischen Forschung nach wie vor im Zusammenhang mit *Diskriminierung* auf die Definition von Allport aus dem Jahre 1954 zurückgegriffen. Diese besagt, dass Diskriminierung ein Verhalten ist, welches auf einer Unterscheidung beruht, die weder mit den Eigenschaften, Leistungen oder dem konkreten Verhalten einer anderen Person zusammenhängen, sondern die auf der Basis natürlicher oder sozialer Kategorien getroffen werden (vgl. Mummendey, Kessler & Otten, 2009, S. 48). Des Weiteren führt er Folgendes aus: "Discrimination comes about only when we deny to individuals or groups of people equality of treatment which they may wish" (Allport, 1954, S. 51).

Allports Definition schließt somit zwei wesentlich Punkte ein: (1) Wie auch im Falle von Vorurteilen und Stereotypen ist die soziale Kategorisierung eine Voraussetzung für Diskriminierungen. (2) Erst wenn die Ungleichbehandlung von Personen aufgrund ihrer Gruppenzugehörigkeit als ungerechtfertigt wahrgenommen wird und ein Dissens besteht, wird von sozialer Diskriminierung gesprochen<sup>10</sup> (vgl. Mummendey, Kessler & Otten, 2009, S. 49 f.).

Zusammenfassend kann folgendes festgehalten werden: Während Vorurteile die Einstellungen gegenüber einer sozialen Gruppe darstellen, sind Stereotype mit der kognitiven Klasse der Reaktionen im Dreikomponentenmodell der Einstellung nach Eagly und Chaiken (1993) gleichzusetzen. Stereotype stellen eine Verallgemeinerung von Eigenschaften dar, die allen Mitgliedern einer sozialen Gruppe zugeschrieben werden. Diskriminierungen sind als benachteiligende Handlungen gegenüber Personen einer sozialen Gruppe zu verstehen, die als ungerechtfertigt wahrgenommen werden. Im Folgenden soll nun auf Vorurteile und Stereotype im Zusammenhang mit der sozialen Kategorie „Geschlecht“ eingegangen werden.

---

<sup>10</sup> Besteht hingegen eine allgemeine Übereinstimmung (von Handelnden, Betroffenen und Außenstehenden) darüber, dass eine unterschiedliche Behandlung von Personen unterschiedlicher Gruppenzugehörigkeit als angemessen gilt, so spricht man von „Sozialer Differenzierung“ (vgl. Mummendey, Kessler, & Otten, 2009, S. 49). Lange Zeit wurde beispielsweise akzeptiert, dass Frauen das Wahlrecht oder der Zugang zu Bildung verwehrt blieb oder sie nur mit Zustimmung ihres Ehemanns einer Erwerbstätigkeit nachgehen durften. Eine derartige Ungleichbehandlung von Frauen und Männer wäre heutzutage unvorstellbar und würde in jedem Fall als Diskriminierung angesehen.

### 2.2.2 Geschlechtstereotypen – reine Antipathie?

In der Sexismusforschung wurde die oben angeführte Definition von Allport im Zusammenhang mit Geschlechtstereotypen aufgegriffen und lange Zeit überwiegend als Feindseligkeit gegenüber Frauen beschrieben. Neuere Forschungen belegen jedoch, dass Stereotypen gegenüber Frauen keineswegs ausschließlich durch eine reine Antipathie gekennzeichnet sind, sondern durchaus auch positive Komponenten aufweisen. Eagly und Mladinic (1994) zeigten, dass Stereotype und Bewertungen in Bezug auf die Gruppe „Frauen“ positiver sind als solche hinsichtlich der Gruppe der „Männer“ („women are wonderful“ effect) (vgl. Glick & Hilt, 2000, S. 244).

Auch wenn Frauen somit oftmals bewundert werden, können sie dennoch gleichzeitig benachteiligt werden. Eagly und Mladinic (1994) lieferten einen Erklärungsansatz hierfür, demzufolge sich die positiven Bewertungen von Frauen insbesondere auf solche Merkmale beziehen, die sie für die häusliche Rolle und typisch „weibliche“ Berufe prädestinieren (siehe feminine Merkmale in Unterkapitel 2.2.3.2). Diese gehen jedoch mit wenig Ansehen und einem geringem Gehalt einher, so dass die niedrigere Statusposition der Frauen durch die Zuschreibung der positiven Eigenschaften sogar gefestigt wird. Die Ansicht beispielsweise, dass Frauen (und Männer nicht) wunderbar darin sind, (kleine) Kinder zu betreuen und zu versorgen, kann somit als Vorurteil angesehen werden, wenn es Bestandteil eines Einstellungssystems ist, das Frauen auf die häusliche Rolle einengt und sie von machtvollen Positionen außerhalb des Hauses fern hält (vgl. Glick & Hilt, 2000, S. 244 f.).

Die Rechtfertigung und letztendliche Aufrechterhaltung der Geschlechterhierarchien muss der Person dabei noch nicht einmal bewusst sein. Folglich kann die vorurteilsgeladene Einstellung von der Person selbst als subjektiv positiv empfunden werden. Tatsächlich stellen subtile sowie scheinbar wohlwollende Formen von Vorurteilen eine Art von Selbsttäuschung dar, die es der Person erlauben ein vorurteilsfreies, egalitäres Selbstbild aufrechtzuerhalten (siehe in Verbindung mit dem modernen Sexismus z.B. Swim, Aikin, Hall, & Hunter, 1995). Im Zusammenhang mit der Sklaverei beispielsweise wurde die Ausbeutung der untergeordneten Gruppe als „white man's burden“, als großmütige Wohltat, dargestellt („Sie sind glücklicher und besser dran, wenn wir sie leiten“). Auch wenn sie nicht auf den ersten Blick als solche zu erkennen sind, können die vorurteilsbehafteten Motive, die hinter diesem Ungleichheitsdenken stehen, durch Zusammenhänge mit offenen Formen von Vorurteilen und dem Wunsch nach sozialer Dominanz aufgezeigt werden (vgl. Glick & Hilt, 2000, S. 245). Benevolenter Sexismus beispielsweise weist in mehreren Studien positive Korrelationen mit Autoritarismus, Sozialer Dominanzorientierung, traditionellem Sexismus und der Akzeptanz von Vergewaltigungsmythen auf (vgl. z.B. Eckes & Six-Materna, 1999; Gerger, Kley, Bohner & Siebler, 2007; Schmidt-Mummendey, Schmerl & Schmidt, 1973).

Allports Definition von Vorurteilen als reine Antipathie ist nach Einschätzung von Rudman (2005) dementsprechend unzureichend, da Geschlechtstereotypen nicht vollständig erklärt werden können – so bleiben die positiven Einstellungen in Bezug auf

Frauen unberücksichtigt. Zudem vernachlässigt die Definition das Zusammenspiel von untergeordneten und dominanten Gruppen der Gesellschaft, die den Status quo aufrechterhalten. So tragen Frauen selbst (mitunter auch unwissentlich) dazu bei, dass die Geschlechterhierarchie fortbesteht, indem sie sich durch scheinbar positive Zuschreibungen „besänftigen“ lassen oder diese nicht als „schädigend“ erkennen und die legitimierende Mythen für sich selbst akzeptieren (siehe Bedeutung ambivalenter Stereotype für das Geschlechterverhältnis in Unterkapitel 2.2.3.4) (vgl. Rudman, 2005, S. 107 ff.).

*“Certainly oppression is more palatable for subordinates when it is ‘sugar-coated,[sic]’ compared with when it is overtly hostile (...). Moreover, women may not recognize that a prescription for female niceness underscores their subordinate status in society. Instead, they may view it as a means of favorably distinguishing themselves from men (...), as well as a means of receiving positive reinforcement”* (Rudman, 2005, S. 113).

Glick und Hilt (2000) schlagen daher eine Definition von Vorurteilen vor, deren Kern in dem Versuch liegt, anderen Gruppen einen Nachteil zuzufügen. Sie definieren Vorurteile als *“the implicit or explicit attitude that a group deserves inferior social status”* (S. 245, Hervorhebung im Original). Diese Definition ist im Hinblick auf die Valenz der Vorurteile weiter gefasst als Allports Begriffbestimmung: Sie schließt sowohl negative, scheinbar positive als auch ambivalente Vorurteile ein. Zudem hält sie die Möglichkeit offen, dass sich die Vorurteile auf die Eigengruppe beziehen (z.B. könnten einige Subgruppen von Frauen ihren Geschlechtsgenossinnen pauschal eine mangelnde Kompetenz im Hinblick auf die Übernahme einer Führungsposition unterstellen). Dabei müssen die Vorteile nicht immer bewusst, sondern sie können auch unbewusst sein (vgl. Glick & Hilt, 2000, S. 245).

## 2.2.3 Geschlechterstereotype

### 2.2.3.1 Definition

Petersen und Six (2008) zufolge werden Stereotype heutzutage üblicherweise „als ‚eine Reihe von Überzeugungen über die Mitglieder einer sozialen Gruppe‘ oder als ‚Assoziation einer Reihe von Merkmalen mit einer Kategorie““ (S. 21) definiert. Grundvoraussetzung für das Abrufen von Stereotypen ist dabei die generelle menschliche Neigung, andere Personen einer bestimmten *sozialen Kategorie* zuzuordnen (vgl. ebd.). Die Einteilung zu einer bestimmten sozialen Kategorie erfolgt auf der Basis von (a) Merkmalen, die äußerlich sofort erkennbar sind – im Falle von Geschlechterstereotypen z.B. an der Kleidung, der Frisur oder der Körperstatur, (b) gemeinsamen Ansichten (z.B. hinsichtlich der Religion) oder (c) der Ähnlichkeit zu einem bestimmten Typus Mensch (z.B. einer „Karrierefrau“) (vgl. Klauer, 2008, S. 23).



Stereotype sind ein kulturell geteiltes Wissen, welches von Kindern im Laufe des Sozialisationsprozesses sehr früh „erlernt“ wird (vgl. Eckes, 2008, S. 171). Bereits ab etwa 12 Monaten beginnen Kinder andere Personen anhand des Geschlechts zu differenzieren, so dass der Prozess der Geschlechterkategorisierung mit fortschreitendem Alter sehr schnell und automatisch abläuft. Es wird fast unmöglich diese Kategorisierung nicht vorzunehmen und das Gegenüber „gender blind“ zu betrachten. Dies wird insbesondere dann deutlich, wenn es doch einmal schwer fallen sollte, das Geschlecht einer anderen Person zu bestimmen. In einer solchen Situation reagiert man mit Unsicherheit und weiß nicht, wie man sich verhalten soll (vgl. Rudman & Glick, 2008, S. 84).

Die reine Kenntnis von Stereotypen bedeutet jedoch nicht, dass sie auch als zutreffend erachtet und auf eine konkrete Person angewendet werden und es zu einer Diskriminierung kommt (vgl. Eckes, 2008, S. 171). So ist es möglich eine Geschlechterkategorisierung ohne weitere Prozesse der Stereotypisierung vorzunehmen. Allerdings ist der Abruf von Stereotypen genau Sinn und Zweck der Kategorisierung. So ordnen wir andere Menschen in eine bestimmte Gruppe ein, um unsere komplexe Welt zu vereinfachen und Rückschlüsse über die Eigenschaften von Personen ziehen zu können. Die Stereotype haben dabei die Funktion, schnell fundamentale Informationen über eine Gruppe bereitzustellen und zugleich eine Orientierung für das eigene Verhalten dieser Gruppe gegenüber zu liefern (vgl. Rudman & Glick, 2008, S. 84). Um den Prozess der Stereotypisierung aufzuhalten, ist folglich ein hohes Maß an bewusster Kontrolle/Reflexion und eine entsprechende Motivation nötig (vgl. Eckes, 2008, S. 171).

Nach Eckes (2008) sind Geschlechterstereotype „kognitive Strukturen, die sozial geteiltes Wissen über die charakteristischen Merkmale von Frauen und Männern enthalten“. Dabei sind sie (genauso wie Stereotype über andere Gruppen von Menschen auch) zum einen Bestandteil des individuellen Wissens und zum anderen „bilden sie den Kern eines konsensuellen, kulturell geteilten Verständnisses von den je typischen Merkmalen der Geschlechter“ (S.171).

Die Besonderheit von Geschlechterstereotypen (verglichen mit Stereotypen im Allgemeinen) besteht darin, dass sie neben den deskriptiven auch präskriptive Anteile besitzen. Während sich die deskriptiven Anteile auf die Annahmen über die Eigenschaften von Frauen und Männern beziehen (Wie „sind“ Frauen und Männer und wie „verhalten“ sie sich), rekurrieren präskriptive Anteile auf normative Verhaltenserwartungen, die an die RollenträgerInnen gestellt werden (Wie „sollen“ Frauen und Männer sein und wie „sollten“ sie sich verhalten). Eine Verletzung der deskriptiven Erwartungen ruft *Überraschung* hervor, eine Abweichung hinsichtlich der präskriptiven Erwartungen hingegen *Ablehnung* oder *Bestrafung* (vgl. Eckes, 2008, S. 171). Wünscht sich eine Frau beispielsweise keine Kinder, so stößt sie in ihrem Umfeld vermutlich auf Verwunderung und Erstaunen. Eine Frau hingegen, die die Mutter eines kleinen Kindes ist, dieses jedoch nicht (aus scheinbar egoistischen Motiven heraus) selbst rund-um-die-Uhr betreut, erfährt vermutlich mitunter Ablehnung (siehe Unterkapitel 2.1.1.2 zum Muttermythos).

Die Inhalte sowohl individueller als auch kultureller Stereotype sind sehr beständig und unterliegen einem sehr langsamen Wandel. Dies hängt u.a. mit der Bildung von Substereotypen zusammen (siehe Unterkapitel 2.2.3.3). Da sich Geschlechterstereotype auf eine sehr große und heterogene Gruppe von Menschen beziehen, wird mit Abweichungen gerechnet. Dennoch werden die Globalstereotype (über Frauen im Allgemeinen) nicht ohne Weiteres aufgegeben. Stattdessen werden Substereotype (beispielsweise über Karrierefrauen) gebildet, die etwa die Verstöße gegen die kategorialen Erwartungen abbilden. Trifft man beispielsweise eine Frau, die sich entgegen den Erwartungen nicht freundlich und liebenswürdig verhält, wird nicht das allgemeine Stereotyp über Frauen revidiert, sondern die entsprechende Frau der Subkategorie „Zicke“ o.ä. zugeordnet (vgl. Rudman & Glick, 2008, S. 83 ff.).

### **2.2.3.2 Inhalte von Geschlechterstereotypen**

Geschlechterstereotype umfassen eine Reihe von Eigenschaften, die mit einem Geschlecht stärker assoziiert werden als mit dem anderen. Dies schließt physische Merkmale (z.B. Männer sind körperlich stärker, Frauen zarter), Vorlieben und Interessen (z.B. Frauen gehen gerne einkaufen und Männer mögen Fußball) sowie soziale Rollen (z.B. Frauen als Hausfrauen und Mutter, Männer als Versorger der Familie) und berufliche Rollen (z.B. Frauen als Krankenschwester oder Grundschullehrerin, Männer als Bauarbeiter oder Pilot) ein (vgl. Rudman & Glick, 2008, S. 85 f.).

Um die Inhalte von Geschlechterstereotypen zu ermitteln, wurden in den letzten Jahrzehnten in zahlreichen Ländern Studien durchgeführt, die zu übereinstimmenden Ergebnissen gelangten. Merkmale, die häufiger mit Frauen in Verbindung gebracht werden, lassen sich unter dem englischen Begriff „communal“ (im Deutschen auch: feminin, gemeinschaftsorientiert, warm oder expressiv) subsumieren. Sie sind auf die Interessen und Bedürfnisse anderer bezogen, spiegeln Selbstlosigkeit und Empathievermögen wider. Männliche Eigenschaften werden hingegen mit dem Begriff „agentic“ (auch: maskulin, selbstorientiert, kompetent oder instrumentell) beschrieben, da sie auf die eigenen Interessen bezogen sind und durch Aufgabenorientierung sowie Aktivität gekennzeichnet sind (vgl. Six-Materna, 2008, S. 122). Dabei gehen die positiven Zuschreibungen häufig mit negativen Charakterisierungen einher (siehe „Stereotype Content Model“ in Unterkapitel 2.2.3.4). Kurz gesagt werden Frauen als „wundervoll“ (warmherzig, gütig, sanftmütig und aufopferungsbereit), aber „schwach“ (abhängig, gehorsam, ängstlich und passiv) und Männer als „böse“ (aggressiv, unhöflich, arrogant und egoistisch), aber „kühn“ (kompetent, rational, abenteuerlustig und aktiv) beschrieben. Obwohl diese Stereotypisierungen nahe legen, dass Frauen eher gemocht werden, zollen sie Männern mehr Respekt und verstärken somit die gesellschaftlichen Machtunterschiede zwischen den Geschlechtern. Die positiven Eigenschaften, die Frauen zugeschrieben werden, lassen sie für pflegende Aufgaben (die Rolle der Mutter oder auch der Beruf als Krankenschwester, Kindergärtnerin) besonders geeignet erscheinen, die negativen Eigenschaften bestärken

ihren niedrigeren Status innerhalb der Hierarchie der Geschlechter. Im Gegensatz dazu prädestinieren die positiven Merkmalszuschreibungen der Männer sie für die Rolle als Beschützer und Versorger der Familie (durch ihre außerhäusliche Erwerbstätigkeit), gleichzeitig festigen ihre negativen Eigenschaften ihre Macht und ihren Status innerhalb der Gesellschaft, da ihnen bescheinigt wird, sie könnten durchgreifen, wenn dies nötig sei (vgl. Rudman & Glick, 2008, S. 81 ff.).

### 2.2.3.3 *Substereotype*

Verschiedene Studien belegen, dass die oben beschriebenen *Globalstereotype* über Frauen häufig zu allgemein und unpräzise sind, um sich als hinreichend nützlich für die eigene Orientierung und die eigene Handlungsplanung im Umgang mit anderen Menschen zu erweisen. Folglich existieren eine Vielzahl von *Substereotypen*, die sich auf bestimmte *Untergruppen* von Frauen beziehen und daher wesentlich spezifischer sind (vgl. Eckes, 2008, S. 174). Stereotypisierungen von Frauen (und auch Männern) finden daher in der Regel auf der Basis von Subgruppen statt, da diese eine exaktere Repräsentation der sozialen Kategorie erlauben.

Betrachtet man die verschiedenen Subgruppen von Frauen, so lassen sich diese thematisch in unterschiedliche Cluster einordnen. So gibt es Frauengruppen, die die traditionelle Rolle erfüllen und andere, die aus ihr ausbrechen. Erstere lassen sich in die Cluster mütterliche, fürsorgliche Frauen (z.B. Mutter, Tante, Großmutter und Hausfrau) oder romantische, attraktive Frauen (z.B. Freundin, Prinzessin und Cheerleaderin) einteilen, letztere in die folgenden Cluster: Frauen, die gegen die präskriptive Norm der weiblichen Freundlichkeit verstoßen (z.B. Zicke, Biest) oder Frauen, die gegen die Norm der weiblichen Sexualität verstoßen, sei es durch Homosexualität (z.B. Lesben) oder Promiskuität (z.B. Schlampen) sowie Frauen, die nicht der traditionellen Geschlechterrolle nachkommen, sondern in beruflicher Hinsicht in „männliche“ Domänen vorstoßen (z.B. Karrierefrauen) (vgl. Rudman & Glick, 2008, S. 101). Den beiden Subtypen „Hausfrauen“ und „Karrierefrauen“ fällt dabei eine prominente Rolle zu, da sie den Studien von Coats und Smith (1999) zufolge diejenigen Subgruppen sind, die am häufigsten von den Befragten genannt werden. Zu den Inhalten der Stereotype lässt sich sagen, dass den weiblichen Subgruppen eine Vielzahl unterschiedlicher Eigenschaften zugeschrieben werden, die nicht alle gleichermaßen den Stereotypen der übergeordneten Kategorie „Frau“ entsprechen. Die stärkste Übereinstimmung mit den Globalstereotypen gibt es für die traditionelle Subgruppe der „Hausfrauen“ (vgl. Wade & Brewer, 2006, S. 754 f.).

### 2.2.3.4 *Erklärungsansätze*

Als Erklärung für die zeitlich und kulturell konsistente Zuschreibung von Geschlechterstereotypen können laut Eckes (2008) insbesondere zwei Theorien herangezogen werden: Eagly's (1987) „Theorie der sozialen Rollen“ („Social Role Theory of Sex Differences and

Similarities“) sowie das „Stereotypinhaltsmodell“ („Stereotype Content Model“) von Fiske, Cuddy, Glick und Xu (2002).

### *Social Role Theory*

Eagly (1987) Social Role Theory (SRT) zufolge werden Frauen und Männern diejenigen Stereotype (Persönlichkeitseigenschaften) zugeschrieben, die mit ihren jeweiligen sozialen Rollen und Berufen in Verbindung gebracht werden. Ausgangspunkt für die Stereotypisierung von Frauen und Männer ist somit die geschlechtsspezifische Arbeitsteilung innerhalb der Gesellschaft und das damit einhergehende Rollenverhalten. Im Laufe der Zeit haben sich geschlechtsspezifische Tätigkeitsfelder im privaten und öffentlichen Bereich entwickelt: Während Männern in der Regel die Versorgerrolle zufällt, die mit einer außerhäuslichen Erwerbstätigkeit einhergeht und maskuline Eigenschaften erfordert, übernehmen Frauen zumeist die Rolle der Hausfrau und Mutter, die mit femininen, gemeinschaftsfördernden Eigenschaften verbunden ist (siehe Unterkapitel 2.2.3.2). Die normativen Erwartungen, die an den/die RollenträgerIn gestellt werden, wirken sich dabei auf das soziale Verhalten beider Geschlechter aus. Die Verhaltensunterschiede von Frauen und Männern (die sie in der Ausübung ihrer jeweiligen sozialen Rolle zeigen) bilden wiederum die Basis der sozialen Wahrnehmung und finden sich folglich in den Geschlechterstereotypen wieder.

Merkmale, die mit Frauen in Verbindung gebracht werden, sind dementsprechend solche, die für die Rolle als Mutter (die traditionell häusliche Rolle der Frau) charakteristisch sind. So werden Frauen „von Natur aus“ als häuslich, familiär angesehen und sie werden mit femininen Eigenschaften (emotionaler Wärme) in Verbindung gebracht. Merkmale, die mit Männern assoziiert werden, stimmen hingegen mit denen der Rolle als Ernährer und Beschützer der Familie (der traditionellen Rolle des Mannes) überein. Folglich werden Männer als dominant wahrgenommen und es werden maskuline Eigenschaften (Kompetenz) mit ihnen assoziiert (vgl. Rudman & Glick, 2008, S. 86). Unterschiede im Verhalten werden somit auf das Geschlecht zurückgeführt, obwohl sie die Anforderungen der unterschiedlichen Rollen widerspiegeln, die Frauen und Männer in der Gesellschaft innehaben (vgl. Eagly, Wood & Diekmann, 2000).

### *Stereotype Content Model*

#### Grundannahmen

Laut der Theorie des Stereotype Content Model (SCM) gehen die Inhalte von Stereotypen (und somit auch die der Geschlechterstereotype) auf die strukturellen Beziehungen einzelner Gruppen innerhalb der Gesellschaft zurück. Fiske, Cuddy, Glick und Xu (2002) nehmen dabei die drei folgenden Kernhypothesen an, die kulturübergreifend (in 17 Ländern) in zahlreichen Studien (z.B. Cuddy, Fiske & Glick, 2004; Eckes, 2002; Fiske, Xu, Cuddy & Glick, 1999) bestätigt werden konnten:

- (1) Die *dimensional hypothesis* besagt, dass *emotionale Wärme* (Ist die Fremdgruppe der Eigengruppe wohlgesinnt oder nicht?) und *Kompetenz* (Wie fähig ist die andere Gruppe der eigenen Gruppe einen Schade zuzufügen bzw. deren Ziele zu gefährden?) die beiden zentralen, universellen Dimensionen der menschlichen sozialen Wahrnehmung darstellen.<sup>11</sup> „Emotionale Wärme bezieht sich auf die wahrgenommene sozio-emotionale Orientierung einer Gruppe gegenüber anderen Personen oder Gruppen, unter Kompetenz wird aufgabenabhängige Kompetenz verstanden (d.h. die wahrgenommene Fähigkeit, Aufgaben, die in der Gesellschaft einen hohen Stellenwert besitzen, zu bewältigen)“ (Eckes, 2001, S. 237 f.).
- (2) Die *social structural hypothesis* postuliert, dass der relative Status einer Gruppe (hoch vs. niedrig) ihre wahrgenommene Kompetenz bestimmt, die Art der Interdependenz im Hinblick auf die Gruppenziele (kooperativ vs. kompetitiv) hingegen übt einen Einfluss auf die wahrgenommene emotionale Wärme aus. So werden Fremdgruppen mit einem hohen Status als kompetent, solche mit einem niedrigen Status als inkompetent wahrgenommen. Kooperative Gruppen, die die gleichen Ziele verfolgen wie die Eigengruppe, werden als warm angesehen, kompetitive Gruppen, deren Ziele den eigenen entgegenstehen, werden als kalt wahrgenommen.
- (3) Der *mixed stereotype hypothesis* zufolge sind die Stereotype vieler Gruppen nicht rein positiv (warmherzig und kompetent) oder rein negativ (kalt und inkompetent), sondern ambivalent. Folglich gehen hohe Einstufungen auf einer der beiden Dimensionen oftmals mit niedrigen Einstufungen auf der anderen Dimension einher (kalt, aber kompetent) oder (warm, aber inkompetent).

Diese drei Hypothesen lassen auch auf die traditionellen Geschlechterstereotype übertragen. Während Männer aufgrund ihres relativ hohen Status in der Gesellschaft als kompetent angesehen werden, wird ihnen gleichzeitig aufgrund ihrer kompetitiven Orientierung gegenüber Frauen im Bereich des Erwerbslebens emotionale Wärme abgesprochen. Somit werden sie zugleich als emotional kalt charakterisiert. Frauen hingegen werden wegen ihres relativ niedrigen gesellschaftlichen Status als inkompetent wahrgenommen, zugleich jedoch wegen ihrer kooperativen Interdependenz in Bezug auf Männer im privaten, familiären Bereich als warmherzig erachtet (vgl. Eckes, 2008, S. 173).

Auch für verschiedene Subgruppen von Frauen konnte eine derartige Klassifizierung gezeigt werden. Während für die traditionelle Subgruppe der Hausfrauen die gleichen Stereotype gelten wie für die Globalkategorie „Frau“ (warmherzig und inkompetent), ste-

---

<sup>11</sup> Zur Unterscheidung vergleichbarer Kerndimensionen („communion“ und „agency“) siehe beispielsweise Abele & Wojciszke, 2007.

hen die Stereotype für den nicht-traditionellen Frauentyp der „Karrierefrauen“ im genauen Gegensatz dazu: Sie werden als kompetent und kalt beschrieben (siehe Abbildung 2) (vgl. Wade & Brewer, 2006, S. 754 f.).<sup>12</sup>

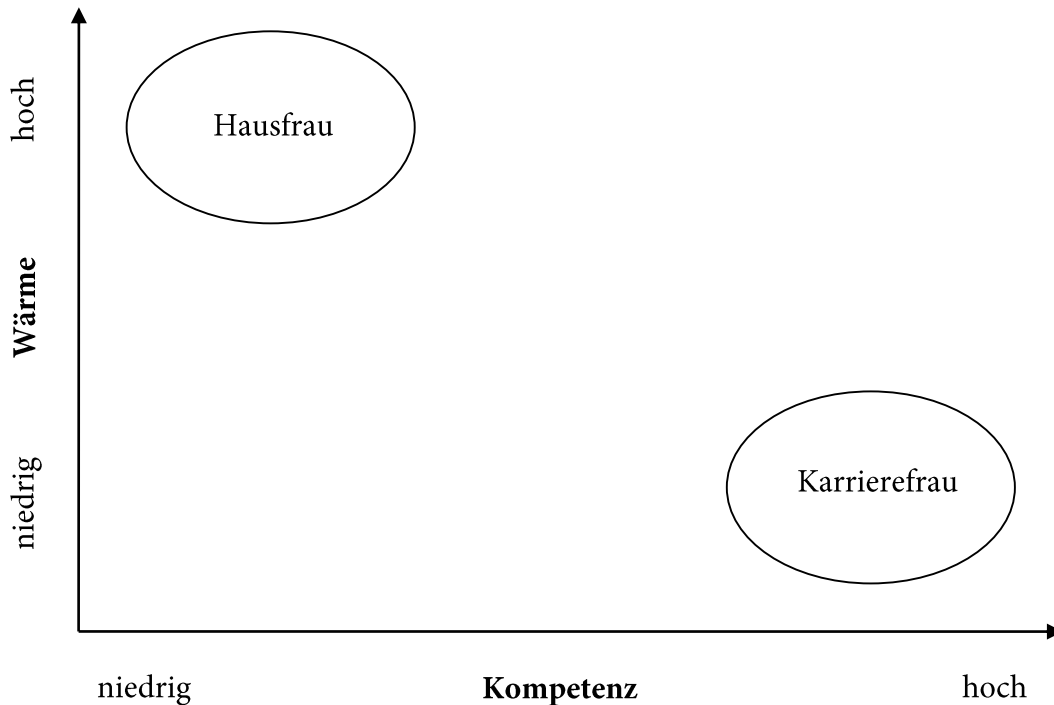


Abbildung 2. Idealtypische Zuordnung der beiden weiblichen Subgruppen „Hausfrau“ und „Karrierefrau“ in ihr entsprechendes Wärme-Kompetenz-Cluster. Quelle: Abbildung übernommen aus Fiske, Cuddy, Glick und Xu (2002). Eigene Übersetzung und Reduzierung auf diese beiden Subgruppen.

Somit ergeben sich vier Kombinationsmöglichkeiten der beiden Dimensionen Wärme und Kompetenz, für die Fiske, Cuddy, Glick und Xu (2002) unterschiedliche Bezeichnungen wählten: Stereotype, die Frauen (oder auch Männer) durchweg positiv – also als warmherzig und kompetent – charakterisieren, werden als *bewundernd* bezeichnet. (Sie beziehen sich in der Regel auf die Eigengruppe.) *Verachtende Stereotype* sind im Gegensatz dazu ausnahmslos negativ (emotional kalt und inkompetent). Die Kombination aus emotionaler Wärme und Inkompetenz wird als *paternalistische Stereotype* bezeichnet. Geht hingegen ein hohes Maß an Kompetenz mit emotionaler Kälte einher, handelt es sich um *neidvolle Stereotype*.

<sup>12</sup> Eckes (2002) konnte zeigen, dass die Stereotype über Hausfrauen und Karrierefrauen die größte Ambivalenz besitzen; so positiv sie auf einer der beiden Dimensionen abschneiden, so negativ schneiden sie auf der anderen Dimension ab. Folglich wurden Hausfrauen im Vergleich zu den anderen Subgruppen am höchsten in Wärme und am niedrigsten in Kompetenz eingestuft. Für Karrierefrauen verhält es sich genau umgekehrt.

Tabelle 2 liefert einen Überblick darüber, welche Stereotypenkombination auf die jeweilige Art der Abhängigkeit und den jeweiligen Status einer Gruppe zurückgehen. Sie führt zudem die Bezeichnung der entsprechenden Stereotype an und liefert einige Beispiele für soziale Subgruppen, auf die sich die Stereotype beziehen könnten

Tabelle 2: Theorie des Stereotype Content Model

Status	Art der Abhängigkeit	
	Kompetitiv	Kooperativ
<b>Niedrig</b>		
Inhalte der Stereotype	kalt und inkompetent	warm, aber inkompetent
Bezeichnung der Stereotype	verachtende Stereotype	Paternalistische Stereotype
Soziale Gruppe	Spießbürgerin, Tussi/Püppchen	Hausfrau und Mutter
<b>Hoch</b>		
Inhalte der Stereotype	kalt, aber kompetent	warm und kompetent
Bezeichnung der Stereotype	neidvolle Stereotype	Bewunderung
Soziale Gruppe	Karrierefrau	selbstbewusste Frau, „Society lady“

Anmerkungen. Quelle: Die Vorlage für die Tabelle wurde aus Glick und Fiske (2001b, S. 173) übernommen; die Beispiele für die sozialen Gruppen stammen aus Eckes (2002). Eigene Übersetzungen.

### Die Bedeutung ambivalenter Stereotype für das Geschlechterverhältnis

Sowohl *neidvolle* als auch *paternalistische Stereotype* gegenüber Karrierefrauen bzw. Hausfrauen sind hinsichtlich des Geschlechterverhältnisses besonders bedeutsam, da sie beide zum Fortbestehen der Geschlechterhierarchie beitragen:

*Neidvolle Stereotype.* Karrierefrauen stellen als sozioökonomisch erfolgreiche Gruppe eine kompetitive Bedrohung für den Status der Männer dar und ihr Erfolg ruft Neid hervor (sie werden als bedrohliche und unfaire Konkurrentinnen wahrgenommen). So werden Karrierefrauen zwar einerseits als kompetent angesehen, was ihren gegenwärtigen Erfolg erklärt und dadurch das System der Leistungsgesellschaft stützt, welches ebenfalls die dominante Eigengruppe der Männer begünstigt. Andererseits werden sie jedoch als emotional kalt stereotypisiert, was (aus männlicher Sicht) Aktionen rechtfertigt, die gegen die beneidete Gruppe der Karrierefrauen gerichtet sind und sie in ihre Schranken verweist (z.B. die fortgesetzte Diskriminierung von Frauen) (vgl. Cuddy, Fiske & Glick, 2008, S. 122 ff.; Eckes, 2008, S. 175 f.; Fiske et al., 2004). Abwertende Bezeichnungen für erfolgreiche Frauen wie „Eiserne Lady“ oder „Mannsweib“ verdeutlichen dies (vgl. Cuddy, Fiske & Glick, 2008, S. 77).

*Paternalistische Stereotype.* Sie sind inhaltlich zwar genau konträr zu den *neidvollen Stereotypen*, ihre Ambivalenz führt jedoch ebenso zur Aufrechterhaltung des Status quo. So werden Hausfrauen als inkompetent angesehen, was ihre untergeordnete Stellung (ihren niedrigen sozialen Status) sowie ihre Bevormundung rechtfertigt. Gleichzeitig wird ihnen emotionale Wärme und somit eine sozial erwünschte, positive Eigenschaft zugesprochen. Diese positiven Zuschreibungen können mitunter dazu führen, dass sich Frauen beschwichtigen lassen und sie die traditionelle Aufteilung der Geschlechterrollen akzeptieren. Gleichzeitig erlaubt diese Art der Stereotypisierung es Männern, sich selbst in ein positives Licht zu rücken, da sie Frauen behüten und beschützen und ihnen positive (wenn auch untergeordnete, nachrangige) Eigenschaften zuschreiben (vgl. Cuddy, Fiske & Glick, 2008, S. 122 ff.; Eckes, 2008, S. 175 f.; Fiske et al., 2004).

Auch Jackman (1994) führt aus, dass Unterordnung und Zuneigung häufig Hand in Hand gehen. Dominante Gruppen neigen dazu, sich warmherzig gegenüber Untergebenen zu zeigen und ihnen bevormundende Zuneigung als Belohnung dafür anzubieten, dass sie „ihren Platz kennen“ und sich nicht dagegen auflehnen. Das wohlwollende Verhalten der untergeordneten Gruppe gegenüber festigt sogar die Vorrangstellung (Position der Überlegenheit und Kontrolle) der dominanten Gruppe. So tragen die Stereotypisierung von Frauen als nett (aber schwach) und ihre Behandlung in einer wohlwollenden (aber bevormundenden) Art in zweifacher Weise dazu bei, die männliche Hegemonie fortzusetzen. Erstens wird dadurch die vermeintlich mangelnde Kompetenz und Autorität der Frauen untermauert, was sie wiederum auf untergeordnete (in der Regel pflegende, erzieherische) Tätigkeiten bzw. Rollen verweist, die ein geringes Ansehen genießen. Zweitens ist es möglich, dass Frauen selbst ihren Anteil zur Aufrechterhaltung des Status quo beisteuern, da sie die „Gefahr“ der Benevolenz nicht als solche erkennen oder sie Diskriminierungen eher tolerieren, wenn sie „nett verpackt“ werden (siehe hierzu Unterkapitel 2.4.2.3 und die Zustimmung zu benevolent sexistischen Einstellungen auf Seiten der Frauen) (vgl. Rudman, 2005, S. 113).

### Kritik am Stereotype Content Model

Das Stereotype Content Model weist zwei Ungenauigkeiten auf: Einerseits müsste der Begriff Kompetenz genauer spezifiziert werden und zum anderen sollte die „perceiver perspective“ berücksichtigt werden. In einer Studie von Cuddy, Fiske und Glick (2007) zeigte sich, dass Hausfrauen nicht immer eindeutig dem Cluster warmherzig und inkompetent zugeordnet werden, sondern auch die Tendenz besteht, sie als emotional warm und kompetent anzusehen. Die AutorInnen stellten die Vermutung an, dass die Kompetenz bei Hausfrauen auf den familiären Bereich bezogen wurde und nicht, wie von den AutorInnen intendiert, auf die Erwerbssphäre (vgl. Cuddy, Fiske & Glick, 2007, S. 638). Um derartigen Missverständnissen vorzubeugen, sollte daher die Kompetenzdimension bei der Operationalisierung explizit auf einen bzw. beide Bereiche verweisen (d.h. es sollte nicht nur erfragt werden für wie „kompetent“, „erfahren“ oder „fähig“ die Befragten bestimmte



soziale Gruppen halten, sondern wie „kompetent in Bezug auf eine außerhäusliche Erwerbstätigkeit“ bzw. wie „kompetent im Zusammenhang mit familiären Aufgaben“ die Gruppen angesehen werden).

Des Weiteren sollte die „perceiver perspective“ berücksichtigt werden. So spiegeln die oben beschriebenen Geschlechterstereotype kulturelle und nicht individuelle Stereotype wieder – je nach Befragtengruppe können diese jedoch differieren. So werden Karrierefrauen den kulturellen Stereotypen zufolge als kompetent, aber emotional kalt und Hausfrauen als warmherzig, aber inkompetent angesehen. Aus der Sicht der Gruppe der Männer dürften die kulturellen sich mit den individuellen Stereotypen decken. Gleiches gilt für Subgruppe der Hausfrauen, wenn sie sich durch die positiven Merkmalszuschreibungen „besänftigen“ lassen und die Zuschreibung der (kulturellen) Stereotype für ihre Eigengruppe akzeptieren (siehe Bedeutung ambivalenter Stereotype für das Geschlechterverhältnis in Unterkapitel 2.2.3.4). Karrierefrauen hingegen werden sich vermutlich selbst als kompetent und warmherzig ansehen, da es sich um die Eigengruppe und folglich keine kompetitive Gruppe handelt (siehe „social structural hypothesis“ in Unterkapitel 2.2.3.4). Die persönlichen Stereotype unterscheiden sich somit in diesem Fall von den kulturellen.

#### *Gemeinsamkeiten des SCM und der SRT*

Während das Stereotype Content Model (SCM) allgemeiner formuliert ist, bietet die Social Role Theory (SRT) eine detaillierte Analyseebene hinsichtlich des Einflusses sozialer Strukturen auf Geschlechterstereotype. Gleichwohl weisen beide Theorien Überschneidungspunkte auf und lassen sich folglich miteinander kombinieren (vgl. Eckes, 2008, S. 173). Zunächst einmal gleichen sich die zentralen Inhalte der Stereotype, auch wenn sie unterschiedlich bezeichnet werden: Während die femininen Eigenschaften der Social Role Theory der Wärmedimension des Stereotype Content Model entsprechen, lassen sich die maskulinen Eigenschaften der SRT der Kompetenzdimension des SCM zuordnen. Zudem ist die Aufteilung der traditionellen Geschlechterrollen durch wechselseitige Interdependenz gekennzeichnet (Frauen sind auf Männer als Ernährer angewiesen, Männer auf Frauen als Mütter und Hausfrauen). Die Inhalte der Geschlechterstereotype lassen sich von sozialen Strukturen (SCM) bzw. der geschlechtsspezifischen Arbeitsaufteilung (SRT) ableiten. So gelten die Hausfrauenrolle oder „typisch-weibliche“ Berufsfelder als status-niedrig, wohingegen die Rolle als Versorger oder „typisch-männliche“ Berufsfelder als status-hoch angesehen werden (vgl. Cuddy, Fiske & Glick, 2008, S. 101 f.). Beide Theorien weisen folglich mehrere Anknüpfungsmöglichkeiten auf.

### 2.2.3.5 Geschlechterstereotype – zutreffend oder sexistisch?

Wenn einige deskriptive Stereotype über Charaktereigenschaften, die typischerweise Frauen und Männern zugeschrieben werden, einen gewissen Grad an Akkuratheit aufweisen, stellt sich die Frage, weshalb solche Zuschreibungen als sexistisch angesehen werden. Swim und Campbell (2003) liefern drei Antworten auf diese Frage:

“First, these [descriptive] beliefs may be misapplied to particular women because any generalization may not be appropriate for any particular woman” (S. 223). Verschiedene Studien zu Geschlechtsdifferenzen kamen zu dem Ergebnis, dass die Variation verschiedener Persönlichkeitsmerkmale innerhalb des Geschlechts (also z.B. zwischen einzelnen Frauen) größer ist als die Differenzen zwischen beiden Geschlechtergruppen (also zwischen Frauen und Männern). Wird die Variation innerhalb einer Gruppe nicht berücksichtigt, so kann es leicht zu einer ungerechten Diskriminierung einzelner Individuen kommen, wenn auf sie Geschlechterstereotype der gesamten Gruppe übertragen werden. Selbst wenn also Frauen *im Durchschnitt* weniger durchsetzungsfähig sein sollten als Männer, wäre es unzutreffend anzunehmen, dass dies auf *jede einzelne* Frau zutrifft. Zu einer Diskriminierung käme es dann, wenn beispielsweise eine weibliche Bewerberin aufgrund dieser Annahme zugunsten eines männlichen Bewerbers bei einer beruflichen Beförderung übergangen wird (vgl. Rudman & Glick, 2008, S. 96 f.).

“Second, although there are likely biological bases for gender differences, the accuracy in people’s beliefs can be attributed in part to these beliefs creating the gender differences or some third variable (such as traditional gender roles) causing both the beliefs and the gender differences” (Swim und Campbell, 2003, S. 223) (siehe hierzu die Social Role Theory in Unterkapitel 2.2.3.4).

“Third, descriptive beliefs can be sexist when they support prescriptive beliefs about women and men that result in gender inequality” (Swim & Campbell, 2003, S. 223). Dabei tendieren Personen, die der Ansicht sind, die Geschlechterdifferenzen hinsichtlich der deskriptiven Stereotype seien in erster Linie auf biologische Faktoren zurückzuführen, in stärkerem Maße dazu, normative Verhaltenserwartungen zu haben (präskriptive Stereotype) als solche Personen, die die gleichen Geschlechtsunterschiede auf Sozialisationsprozesse zurückführen (vgl. ebd.). Ist also eine Person der Überzeugung, dass Frauen von Natur aus einen Mutterinstinkt besitzen und sie einfühlsamer und warmherziger sind als Männer, dann wird vermutlich auch erwartet, dass sich die Mütter (und eben nicht die Väter) um die Betreuung der Kinder kümmern, da sie wesensgemäß am besten dafür geeignet erscheinen.

„Zunächst startete sie [Anke Domscheit, Anm. d. Verf.] bei ihrem ersten Arbeitgeber, einer Unternehmensberatung, richtig durch. Als sie aber ein Kind bekam, änderte sich ihr Standing im Unternehmen radikal. Eigentlich wollte Domscheit ihre Karriere fortsetzen, doch das passte nicht ins Weltbild ihrer Chefs. [Anke Domscheit führt hierzu aus, Anm. d. Verf.] „Bei Beförderungen wurde ich fortan übergangen. Und einer meiner Vorgesetzten schlug mir sogar vor, noch mehr Kinder zu bekommen, anstatt weiter arbeiten zu gehen. (...)“  
(Täubner, 2008, S. 16)

## 2.3 Geschlechterdiskriminierung

Das nachfolgende Kapitel befasst sich mit der Geschlechterdiskriminierung. Vorab wird eine allgemeine Begriffsbestimmung vorgenommen und verschiedene Formen der Diskriminierung werden vorgestellt (Unterkapitel 2.3.1). Besondere Aufmerksamkeit erfahren dann eben jene beschriebenen Diskriminierungen, die auf Geschlechterstereotypen beruhen und sich gegen Mütter richten. So widmet sich Unterkapitel 2.3.2 der „Maternal Wall“. Es wird aufgezeigt, dass in unserer Gesellschaft insbesondere Mütter mit Benachteiligungen im Erwerbsleben zu kämpfen haben, die sich aufgrund der Übernahme von Betreuungs- und Pflegearbeiten für die Familie (u.a. bei der Kinderbetreuung) ergeben.

### 2.3.1 Definition

„Unter Geschlechterdiskriminierung versteht man die ungleiche Behandlung von Personen aufgrund ihres Geschlechts“ (Stahlberg, Dickenberger, & Szillis, 2009, S. 198). In der Literatur (vgl. hierzu Cordes, 2008; Hormel & Scherr, 2010 oder Stahlberg, Dickenberger, & Szillis, 2009) werden verschiedenen Formen der Diskriminierung unterschieden. Auch wenn sie unterschiedliche Bezeichnungen tragen, schließen diese Formen mitunter dieselben Diskriminierungsmechanismen ein (siehe Punkt (a)):

- (a) *unmittelbare* vs. *mittelbare*,  
*direkte* vs. *indirekte*,  
*offene* vs. *verdeckte* und
  - (b) *individuelle* vs. *institutionelle* vs. *strukturelle* Formen.
- (a) Bei der *unmittelbaren* Form der Diskriminierung werden Frauen oder Männer *direkt* und *offensichtlich* aufgrund ihrer Geschlechterzugehörigkeit diskriminiert. Dies kann einerseits auf der Basis von Geschlechterstereotypen geschehen und andererseits aufgrund von (rechtlichen) Normen und Regelungen (vgl. Stahlberg, Dickenberger, & Szillis, 2009, S. 198). Eine direkte, individuelle Form der Diskriminierung wäre es beispielsweise, wenn Frauen (trotz gleicher Qualifikationen) überdurchschnittlich selten im Vergleich zu Männern befördert werden, da das Vorurteil zum Tragen kommt, Frauen seien zu emotional und zu wenig durchsetzungsfähig, um eine Führungsposition einzunehmen. Für eine unmittelbare,

strukturelle Diskriminierung von Frauen, die auf rechtlichen Normen beruht, kann als Beispiel angeführt werden, dass Frauen lange Zeit per Gesetz von Wahlen ausgeschlossen waren und kein Studium aufnehmen durften.

Bei der *mittelbaren Diskriminierung* handelt es sich hingegen um eine *indirekte, verdeckte* Form. Hierbei führen Normen, die an sich geschlechtsneutral sind, zur Benachteiligung einer der beiden Geschlechtergruppen in überproportionaler Weise. Dies gilt beispielsweise für Nachteile, die sich aus einer Teilzeitarbeit ergeben. Da diese fast ausschließlich von Frauen ausgeübt wird, erfolgt eine Schlechterstellung, die das weibliche Geschlecht weitaus häufiger betrifft als das männliche (vgl. Cordes, 2008, S. 916; Stahlberg, Dickenberger, & Szillis, 2009, S. 198; siehe Unterkapitel 2.1.3).

- (b) *Individuelle Diskriminierungen* beziehen sich auf diskriminierendes Verhalten, das von einer einzelnen Person ausgeht. Dabei kann der diskriminierenden Handlung entweder auch eine diskriminierende Absicht zu Grunde liegen oder sie erfolgt unbewusst, auf der Basis persönlicher Einstellungen bzw. Vorurteile bzw. Stereotype (vgl. Hormel & Scherr, 2004, S. 28; siehe Unterkapitel 2.2).

*Institutionelle Diskriminierungen* beruhen hingegen auf den „Normen, Regeln, Vorschriften und Gesetze sozialer Institutionen (Behörden, Verwaltungen, Schulen, Krankenhäuser etc.)“ (Six, 2008, S. 176).

„Im Unterschied zum Vorurteilsansatz (Allport 1954) versteht der Begriff der ‚institutionellen Diskriminierung‘ (...) Sexismus als Ergebnis sozialer Prozesse. Das Wort ‚institutionell‘ lokalisiert die Ursachen von Diskriminierung im organisatorischen Handeln im Netzwerk zentraler gesellschaftlicher Institutionen (z.B. Bildungs- und Ausbildungssektor, Arbeitsmarkt, Wohnungs- und Stadtentwicklungspolitik, Gesundheitswesen und Polizei)“ (Gomolla, o.J., S. 2).

*Strukturelle* Formen der Diskriminierung gehen nicht (direkt) auf eine einzelne ausführende Person und deren Vorurteile oder die Regelungen einzelner sozialer Institutionen zurück, sondern auf die gesamtgesellschaftlichen Strukturen (z.B. die sozialen Normen oder die rechtlichen Rahmenbedingungen in einer Gesellschaft) (vgl. Bielefeldt, 2010, S. 30; Cordes, 2008, S. 916). Beispielsweise wirken sich die unzureichenden strukturellen Rahmenbedingungen hinsichtlich der Kinderbetreuung in Westdeutschland negativ auf die Erwerbstätigkeit von Müttern aus. Gibt es nicht genügend Krippenplätze für die Betreuung kleiner Kinder, führt dies indirekt dazu, dass ein Elternteil – also in der Regel die Mutter – beruflich kürzer treten muss, um die familiären Betreuungsaufgaben zu übernehmen.

### 2.3.2 Maternal wall

Eine Form der Diskriminierung, die auf Geschlechterstereotypen beruht, ist die „Maternal Wall“ bzw. die „family responsibilities discrimination (FRD)“.

*“While the glass ceiling and sexual harassment continue to pose formidable obstacles to women’s advancement in the workplace, many women are also harmed by another form of gender discrimination known as the ‘maternal wall’. Women run up against the maternal wall when they are discriminated against in the workplace because of past, present or future pregnancies or because they have taken one or more maternity leaves. Women also may experience discrimination when they adopt part-time or flexible work schedules” (Williams & Westfall, 2006, S. 31).*

Wie dieses Zitat verdeutlicht, hängt eine Diskriminierung in Form der „Maternal Wall“ mit der (bevorstehenden) Übernahme von Betreuungs- und Pflegearbeiten für die Familie (wie etwa der Kinderbetreuung) zusammen und ist somit an die traditionell weibliche Geschlechterrolle geknüpft („Diskriminierung qua Mutterschaft“). Beteiligen sich Väter aktiv an der Kinderbetreuung und nehmen sich eine familienbedingte Auszeit, sehen sie sich möglicherweise ebenfalls Diskriminierungen ausgesetzt (vgl. Center for Work Life Law, 2006, S. 1). Völlig losgelöst von einer „reinen“ Diskriminierung auf der Basis der Zugehörigkeit zur Geschlechtskategorie „Frau“ ist die Maternal Wall, wie bereits der Name vermuten lässt, jedoch nicht. Wie im Folgenden aufgezeigt wird, sind Frauen in weitaus stärkerem Maße davon betroffen als Männer, da sie als Mütter nach wie vor den Großteil der Kinderbetreuung übernehmen.

#### 2.3.2.1 Diskriminierung qua Geschlecht(errolle)

Diskriminierungen aufgrund der Geschlechterrolle liegen beispielsweise vor, wenn Frauen aufgrund ihrer Schwangerschaft nicht befördert werden oder sie nach der Rückkehr aus der Elternzeit von verantwortungsvollen Aufgaben entbunden oder in ihrer beruflichen Stellung degradiert werden (siehe Unterkapitel 2.1.3). Ebenso zählen stereotype Annahmen darüber, wie sich Arbeitnehmerinnen aufgrund ihrer familiären Betreuungspflichten verhalten werden bzw. verhalten sollten, dazu (siehe Unterkapitel 2.1.1.2 zum Muttermythos und Unterkapitel 2.2.3.4 zu den Verhaltenserwartungen, die auf deskriptiven Geschlechterstereotypen beruhen). So ist es möglich, dass eine Frau für einen Job, der Reisetätigkeiten erfordert, überhaupt nicht in Erwägung gezogen wird, da die Personalverantwortlichen davon ausgehen, dass die familiären Aufgaben einer Mutter schwer mit diesen beruflichen Anforderungen zu vereinbaren seien (vgl. Center for Work Life Law, 2006, S. 1 f.; Williams & Segal, 2003, S. 94 ff.).

Für eine Diskriminierung aufgrund der Geschlechterrolle ist es nicht einmal erforderlich, dass die Frau offensichtlich schwanger ist oder bereits ein Kind hat, dessen Betreu-

ung sie übernimmt. Die Diskriminierungsmechanismen kommen auch in solchen Situationen zum Tragen, in denen eine potentielle Mutterschaft vermutet wird. Bei Bewerbungen etwa können junge Frauen, die sich in einer festen Beziehung befinden, unter dem „Generalverdacht“ stehen, bald ein Kind zu bekommen und für das Unternehmen familienbedingt auszufallen (vgl. hierzu Obermeier, 2006). Dies würde wiederum zusätzliche Kosten und Arbeit für das Unternehmen bedeuten, da eine Vertretung gesucht und diese auch wieder eingearbeitet werden müsste (siehe Unterkapitel 2.1.3 zu den negativen Folgen der geschlechtsspezifischen Arbeitsteilung für die beruflichen Erfolgchancen von Müttern).

Diese Formen der Diskriminierungen, die mit einer vermuteten zukünftigen aktiven Elternschaft zusammenhängen, müssen Männer aufgrund der vorherrschenden traditionellen Arbeitsteilung, was die Kinderbetreuung angeht (siehe Unterkapitel 2.1.2), nicht fürchten. Nur in den seltensten Fällen wird bei personalverantwortlichen Entscheidungen (wie etwa Einstellungen oder Beförderungen) davon ausgegangen, dass ein Mann – selbst wenn er Vater wird – für längere Zeit für das Unternehmen familienbedingt ausfällt. Somit stellt ein Mann kein „Risiko“ dar und ist im Zweifelsfalle die „sichere Wahl“, wenn es um die Entscheidung geht, wer eingestellt oder befördert werden soll. Bei einer Frau wird hingegen vielfach automatisch angenommen, dass sie die Betreuung der Kinder übernimmt und beruflich kürzer tritt, ohne ihre tatsächlichen Absichten und Wünsche zu kennen. Oftmals ist daher – so meine These – die Diskriminierung aufgrund der Geschlechterrolle unmittelbar mit der Diskriminierung aufgrund des Geschlechts verknüpft.

Frauen spüren die Folgen der Maternal Wall im Zusammenhang mit den drei folgenden Lebensphasen:

- (1) Im Zusammenhang mit einer *vermuteten, potentiellen Mutterschaft*. Dies trifft insbesondere auf Frauen im Alter von Anfang 20 bis Mitte 30 zu, die sich in einer festen Beziehung befinden. Ihnen wird am ehesten unterstellt, dass sie in naher Zukunft ein Kind bekommen werden und sich daher eine berufliche Auszeit nehmen möchten (vgl. hierzu Obermeier, 2006).
- (2) Im Zusammenhang mit der *aktuellen Mutterrolle*. Sobald eine Frau bekannt gibt, schwanger zu sein bzw. man ihr die Schwangerschaft ansieht und/oder nachdem sie das Kind zur Welt gebracht hat und zur „Mutter“ geworden ist.
- (3) Im Zusammenhang mit der *Reduktion der Erwerbstätigkeit*. Nachdem sie aus der Elternzeit zurückkehrt und sie beginnt ihre Arbeitszeit zu reduzieren (z.B. auf Teilzeit) oder sie eine flexible Arbeitszeitregelung wünscht (vgl. Williams & Segal, 2003, S. 78).

Für Männer hingegen ergeben sich Diskriminierungen im Zusammenhang mit der Übernahme von familiären Betreuungsaufgaben nur im letzten Falle, d.h. sobald sie tatsächlich beruflich kürzer treten, um sich der Kinderbetreuung zu widmen bzw. sie ihre

Absicht bekunden, dies zu tun. Dann haben sie jedoch Williams & Segal (2003) zufolge mitunter mit stärkeren Sanktionen ihrer Vorgesetzten zu rechnen als Frauen, da sie sich entgegen den Geschlechterrollenerwartungen verhalten, typisch weibliche Aufgaben übernehmen und ihrer traditionellen Rolle als Versorger nicht nachkommen. Wie wenig akzeptiert das Bedürfnis von Vätern ist, sich berufliche Freiräume für die Kinderbetreuung zu nehmen, verdeutlicht das folgende Zitat eines 37-jährigen Referatsleiters in einem Großunternehmen, der Vater eines Kindes ist:

„Im Beruf war es für mich sehr schwierig. Meine Kollegen und das Management konnten meine Wunsch nach mehr Flexibilität zugunsten der Familie nur schwer akzeptieren“ (Walther & Schaeffer-Hegel, 2007, S. 19).

Dem ersten Gleichstellungsbericht der Bundesregierung und der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V. (2011) lässt sich hierzu folgendes entnehmen:

*„Väter, die am Arbeitsplatz auf ihre familialen Verpflichtungen hinweisen, treffen bei Vorgesetzten und Kollegen oft auf Unverständnis und Ablehnung. Von den Vorgesetzten wird ein mangelnder beruflicher Einsatz unterstellt; Kollegen befürchten häufig, dass Mehrarbeit auf sie zukommt und bewerten das Anliegen, das berufliche Engagement zugunsten familialer Verpflichtungen zu reduzieren, als unsolidarische Zumutung“* (Sachverständigenkommission, 2011, S. 154).

Im Allgemeinen ist für einen Mann der Übergang zur Elternschaft jedoch nicht nachteilig. Vorgesetzte, die erfahren, dass ein Mitarbeiter Vater wird, sehen darin nichts Negatives, sondern sogar eher etwas Positives. Dies hängt damit zusammen, dass sich die gesellschaftlichen Erwartungen, die an die Eltern herangetragen werden, komplett unterscheiden: Während die Mutter nach der Geburt des Kindes zuerst einmal zu Hause bleiben und sich um die Betreuung kümmern soll, wird der Vater als Versorger der Familie angesehen (siehe Unterkapitel 2.1.1.2). Dementsprechend wird es auch aus unternehmerischer Perspektive eher begrüßt, wenn ein männlicher Angestellter Nachwuchs bekommt. Schließlich wird angenommen, er werde sich zukünftig noch stärker in seinem Beruf engagieren, um seine Familie finanziell absichern und ihr ein gutes Leben ermöglichen zu können. Elternschaft hat für Frauen und Männer somit unterschiedliche Konsequenzen, was die Erwartungshaltungen, aber auch die Fremdwahrnehmung anbetrifft (vgl. Porter, 2006, S. 62).

Mutterschaft ist ein Schlüsselreiz („key trigger“) für Geschlechterstereotypisierung. So werden Geschlechterstereotype verstärkt aktiviert, wenn die Mutterschaft offensichtlich, salient wird. Dies kann dadurch geschehen, dass eine Frau ihre Schwangerschaft bekannt gibt bzw. diese durch den Babybauch äußerlich sichtbar wird, sie als Mutter aus der Elternzeit zurückkehrt und/oder den Umfang ihrer Stelle reduziert. Unter diesen Umständen werden Angestellte häufig mit beidem konfrontiert: mit Geschlechterstereotypen in Bezug auf die Globalkategorie „Frau“ (z.B. Frauen sind in beruflicher Hinsicht weniger

kompetent als Männer), hinzu kommen Geschlechterstereotype über die Subkategorie „Mutter“ (z.B. Mütter sind – verglichen mit kinderlosen Frauen oder Vätern – weniger engagiert in Bezug auf ihre beruflichen Aufgaben) (vgl. Center for work life law, 2006, S. 2).

Oftmals ist es sogar der Fall, dass die Rolle der Mutter in der Wahrnehmung der KollegInnen und der Vorgesetzten alle anderen Rollen überlagert. Auch in der Arbeitswelt ist sie dann zuallererst eine „Mutter“ und erst danach eine „Angestellte“, was u.a. negative Konsequenzen hinsichtlich ihrer Kompetenzwahrnehmung hat (siehe Unterkapitel 2.3.2.2). Für Männer ist die Vaterrolle hingegen nahezu unsichtbar und gerät eher in Vergessenheit, da sie mit der Ausübung seiner Berufsrolle nicht in Konflikt gerät (vgl. Porter, 2006, S. 62).

Mit der Fremdwahrnehmung der Frauen, die nun in die Subkategorie „Mutter“ eingeordnet werden, ändern sich jedoch nicht nur die deskriptiven Inhalte der Stereotype (z.B. Eine schwangere Frau ist zu emotional, um eine gute Arbeitnehmerin zu sein). Gleichzeitig wandeln sich die präskriptiven Stereotypanteile und somit die Verhaltenserwartungen, die an die Frauen herangetragen werden (siehe Unterkapitel 2.2.3.1). So ist es beispielsweise möglich, dass die Ansicht, eine gute Mutter solle zu Hause bei ihrem Kind bleiben, auf die Arbeitnehmerin übertragen wird, obwohl diese Erwartung eigentlich nichts mit ihrer sozialen Kategorie/Rolle als „Angestellte“ zu tun hat.

Nicht selten hat die unhinterfragte („unexamined“)<sup>13</sup> Anwendung von Geschlechterstereotypen diskriminierende Handlungen zur Folge, wobei Mütter nicht nur im Vergleich zu Vätern, sondern auch im Vergleich zu kinderlosen Frauen benachteiligt werden.

Die Diskriminierung kann dabei unterschiedliche Formen annehmen:

- (a) Sie kann *offen, feindselig* ausfallen, wenn beispielsweise einer Arbeitnehmerin, die aus der Elternzeit zurückkehren möchte, mitgeteilt wird, dass ihr Platz zu Hause bei ihrem Kind sei. Oder es wird ihr nach der Rückkehr der Elternzeit keine gleichwertige Tätigkeit, sondern lediglich eine auf einer niedrigeren Hierarchieebene angeboten.
- (b) Sie kann *indirekte, scheinbar wohlwollende* Formen annehmen, wenn z.B. die Arbeitstätigkeit einer Mutter aus Rücksichtnahme (ohne sie zu fragen) so abgeändert

---

<sup>13</sup> Williams (2006) betont, dass es aus juristischer Sicht notwendig ist, von „unhinterfragten“ („unexamined“) Stereotypen zu sprechen (statt von „subtilen“, „unbewussten“ oder „impliziten“ Stereotypen) da dies die Verantwortung denjenigen zuschreibt, die solche Stereotype vertreten bzw. auch auf deren Basis handeln: “First of all, if this bias is so subtle, is it fair to hold people responsible legally? Secondly, if it’s unconscious, how can it meet the standard for intentional discrimination? ‘Implicit’ doesn’t have those problems, but it’s not sufficiently transparent really for use either in public education or certainly in the courtroom. The new terminology that we have suggested is the terminology of ‘unexamined’ stereotyping. Note how ‘unexamined’ shifts the burden of proof. If it’s unconscious, ‘oh my gosh, I didn’t know.’ But if it’s unexamined and you are clueless, whose fault is that? This new terminology also highlights that although stereotype activation is automatic (...) stereotype application can be controlled” (Williams, 2006, S. 52 f.).



wird, dass sie weniger reisen muss. Selbst wenn dies von den Vorgesetzten gut gemeint ist, kann sich diese Maßnahme letztendlich negativ auf den Verlauf der Karriere der Frau auswirken (vgl. Williams & Segal, 2003, S. 95 ff.; Williams, 2006a, S. 157).

Nicht alle Diskriminierungen sind folglich sofort als solche zu erkennen. Auch Verhaltensweisen, die auf den ersten Blick prosozial erscheinen, können dazu beitragen, die untergeordnete Stellung einer Gruppe aufrechtzuerhalten (vgl. Glick & Hilt, 2000, S. 246).

Auch wenn somit Männer potentiell unter den Folgen der Maternal Wall zu leiden haben, sobald sie eine aktive Rolle bei der Kinderbetreuung einnehmen, sind Frauen dennoch häufiger davon betroffen, da sie nach wie vor den Hauptteil der Familien- und Hausarbeiten einschließlich der Kinderbetreuung übernehmen.

Um etwaige Benachteiligung aufgrund der Maternal Wall zu vermeiden, wird Frauen mitunter sogar empfohlen, ihren Ehering bei Einstellungsgeprächen abzulegen, um die oben beschriebenen Annahmen hinsichtlich einer potentiellen zukünftigen Mutterschaft zu vermeiden (vgl. Williams & Segal, 2003, S. 97). Frauen, die bereits Mütter sind, versuchen hingegen oftmals möglichen Diskriminierungen zu entgehen, indem sie an ihrer Arbeitsstelle verschweigen, dass sie Kinder haben. So auch die Geschäftsführerin in einem Großunternehmen, die Mutter zweier Kinder ist:

*„Die Vorwürfe von außen waren teilweise extrem. ‚Warum schaffst Du Dir Kinder an, wenn Du Dich nicht um sie kümmerst?‘ Das ist so schlimm gewesen, dass ich über Jahre meine Kinder gar nicht erwähnt habe. Viele Kollegen haben nie erfahren, dass ich Kinder habe. Es war eine bewusste Entscheidung, das nicht zu sagen. Ich wollte mich schützen. Denn sonst fragen alle entsetzt, wie ich das hinkriege, und dann kommt man permanent in eine Verteidigungsposition. Das wollte ich vermeiden“ (Walther & Schaeffer-Hegel, 2007, S. 20).*

Eine Untersuchung von Still (2006) bestätigt die stärkere Betroffenheit der Frauen. Sie begutachteten Diskriminierungsklagen der letzten 30 Jahre, die im Zusammenhang mit der Übernahme familiärer Betreuungsaufgaben eingereicht wurden, und stellten fest, dass in 92% der Fälle die KlägerInnen weiblich waren (vgl. Still, 2006, S. 8).

Hinzu kommt, dass für Frauen die Elternschaft im Zusammenhang mit dem Berufsleben andere Auswirkungen hat als für Männer, was die Fremdwahrnehmung (deskriptive Stereotype) und die Erwartungshaltungen (präskriptive Stereotype) anbelangt. Dementsprechend liegt der Fokus der folgenden Ausführungen auch auf den Diskriminierungsmechanismen (den verschiedenen Mustern der Maternal Wall), denen sich Frauen ausgesetzt sehen.

### 2.3.2.2 Muster der Maternal Wall

Bei der Maternal Wall kommen verschiedene Diskriminierungsmechanismen zum Tragen, die sich auf unhinterfragte („unexamined“) stereotype Annahmen stützen. Diese sollen im Folgenden vorgestellt werden:

- (1) *Rollen-Inkongruenz*
- (2) *Feindselige und wohlwollende präskriptive Stereotype*
- (3) *Feindselige und wohlwollende Stereotype*
- (4) *Attribution Bias*
- (5) *Leniency Bias* (vgl. Williams, 2006b, S. 54).

#### (1) *Rollen-Inkongruenz*

Die Ausrichtung des Berufslebens an männlichen Erwerbsmustern/-biographien (siehe Unterkapitel 2.1.3) hat zur Folge, dass die beiden Rollen, die berufstätige Mütter innehaben, als unvereinbar angesehen werden: „Sie kann nicht beides sein: eine gute Mutter und eine gute Arbeitnehmerin.“ Während eine gute Mutter immer für ihre Kinder da sein sollte, wird von einer guten Angestellten der volle Einsatz für ihren Beruf verlangt. Müttern wird daher unterstellt, sie könnten beiden Rollen nicht gleichzeitig gerecht werden (vgl. Williams, 2006a, S. 156; Williams & Segal, 2003, S. 95). Diese angenommene Rollen-Inkongruenz kann zu feindseligen, aber auch zu wohlwollenden präskriptiven Stereotypen führen.

#### (2) *Feindselige und wohlwollende präskriptive Stereotype*

Feindselige Stereotype sind solche Ansichten, dass arbeitende Mütter keiner Rolle wirklich gerecht werden und sie sich daher zwischen ihrer Karriere oder einem Kind entscheiden sollten. Dies kann in der Verhaltenserwartung münden, dass die Mutter ihre Karriere aufgeben und sich lieber auf ihre Rolle als Mutter konzentrieren sollte.

Unter wohlwollenden Stereotypen können hingegen folgende Beispiele subsumiert werden: ArbeitgeberInnen sehen sich möglicherweise selbst als aufmerksam, umsichtig oder besonnen an, wenn sie die mütterlichen Verpflichtungen bei ihren Entscheidungen mit bedenken. So erhält eine Mutter möglicherweise aus Rücksichtnahme einen Führungsposten nicht, da die Personalverantwortlichen davon ausgehen, die familiären Pflichten seien schwer mit den beruflichen Anforderungen zu verbinden und sie wollten die Mutter nicht überlasten. Eine andere Möglichkeit besteht darin, dass ArbeitgeberInnen ihre normativen Erwartungen – wie beispielsweise die Ansicht, Mütter sollten unbegrenzte Zeit aufbringen, um sich den Bedürfnissen der Familie zu widmen – in Annahmen übertragen, die die Wünsche der Mütter selbst betreffen. Einer Mutter wird beispielsweise eine Beförderung nicht angeboten, da angenommen wird, sie sei in ihrer Situation überhaupt nicht an einem Führungsposten interessiert.

Unabhängig davon, ob die Stereotype feindselig oder wohlwollend sind, Frauen werden auf diesem Wege die Entscheidungen darüber, wie die Verpflichtungen einer Mutter zu interpretieren sind, aus der Hand genommen, da die Annahme besteht, dass die Mütter traditionellen Mustern folgen werden bzw. folgen sollten (vgl. Williams & Segal, 2003, S. 94 ff.). Williams und Segal (2003) führen folgendes Beispiel an: Nachdem ein Ehepaar, das im gleichen Unternehmen arbeitete, ein Kind bekam, wurde die Ehefrau stets frühzeitig nach Hause geschickt mit dem besorgten Gefühl, sie solle bei ihrem Kind sein. Den Ehemann hingegen betraute man mit zusätzlichen Aufgaben und es wurde erwartet, dass er Überstunden machte. Beide Verhaltensweisen sollten die Eltern darin unterstützen, ihre Rolle als Mutter bzw. Vater angemessen erfüllen zu können. Während die zusätzliche Arbeit dem Ehemann helfen sollte, seine Familie zu versorgen, sollte die Reduzierung der Arbeitslast der Mutter ihre Anwesenheit beim Kind ermöglichen. Letztendlich drängten die Vorgesetzten die Eltern dadurch jedoch in eine traditionelle Arbeitsteilung und nahmen ihnen die Entscheidung ab, wie sie selbst die Betreuung des Kindes aufteilen möchten.

### (3) *Feindselige deskriptive Stereotype*

Nicht nur in Bezug auf die präskriptiven, sondern auch hinsichtlich der deskriptiven Stereotype wirkt sich die Mutterschaft nachteilig aus. Sobald der Status einer Frau als Mutter salient wird – entweder weil sie schwanger ist, in Elternzeit geht oder anschließend den Umfang ihrer Erwerbstätigkeit reduziert – ändert sich die Fremdwahrnehmung der Frau und somit auch die Stereotype, die ihr zugeschrieben werden. So kann die Rolle der Mutter in der Wahrnehmung der Vorgesetzten oder KollegInnen die der Erwerbstätigen überlagern. Dementsprechend wird sie eher als warmherzige Betreuungsperson wahrgenommen, denn als hoch kompetente Karrierefrau (vgl. Williams & Segal, 2003, S. 90; siehe hierzu auch Unterkapitel 2.2.3.3 und Unterkapitel 2.2.3.4). Die folgenden Untersuchungen bestätigen dies: Eagly und Steffen (1986) fanden heraus, dass Frauen, die teilzeiterwerbstätig sind, in ihren Charaktereinstufungen als warmherziger („communal“), aber gleichzeitig als weniger kompetent („agentic“) angesehen werden als vollzeiterwerbstätige Frauen. Cuddy, Fiske und Glick (2004) kamen zu folgenden Ergebnissen: Berufstätigen Frauen, die ein Kind bekommen, wird zwar eine größere Warmherzigkeit zugesprochen als Frauen ohne Kinder, gleichzeitig werden sie jedoch als weniger kompetent wahrgenommen. Auch im Vergleich zu den berufstätigen Vätern schneiden die Mütter schlechter ab, da diese keine Einbußen hinsichtlich ihrer Kompetenzeinstufungen hinnehmen müssen. So werden Männer nach dem Übergang zur Elternschaft als gleichermaßen kompetent angesehen und gewinnen zudem noch an Wärme hinzu.

Neben der Kompetenzwahrnehmung sinken auch das wahrgenommene Engagement und die wahrgenommene Verfügbarkeit der Mütter. Eine Studie von Fuegen, Biernat, Haines und Deaux (2004) kommt zu dem Ergebnis, dass Mütter als weniger „agentic“, weniger verfügbar und weniger engagiert in Bezug auf ihren Beruf angesehen werden als

Arbeitnehmerinnen ohne Kinder. Im Vergleich zu Vätern schneiden Mütter in der Beurteilung ebenfalls schlechter ab. Es ist anzunehmen, dass der Muttermythos und seine Förderung, Mütter sollten ihren Kindern alle Zeit und Liebe der Welt geben (siehe Unterkapitel 2.1.1.2) mitunter zu dieser Wahrnehmung beitragen.

Derartige Stereotypisierungen sind umso fataler, da sie nur sehr schwer wieder aufgebrochen werden können. Informationen, die den Stereotypen entgegenstehen, werden vom Umfeld der Mutter einfach ignoriert oder verdrängt. So wird ein/e ArbeitgeberIn oder MitarbeiterIn jedes Mal zur Kenntnis nehmen, wenn eine Mutter frühzeitig den Arbeitsplatz verlässt, wohingegen solche Fälle vergessen werden, in denen sie Überstunden macht (vgl. Williams & Segal, 2003, S. 96 f.).

#### (4) *Attribution Bias*

Stereotype können einen Einfluss darauf ausüben, wie uneindeutige Ereignisse bzw. Sachverhalte interpretiert werden. Wenn Eltern zu spät zur Arbeit kommen, wird im Falle einer Frau angenommen, sie hätte Probleme mit den Kindern gehabt, wohingegen dies bei einem Mann nicht getan wird. Hier werden eher familienunabhängige Faktoren, wie z.B. ein unpünktlicher Zug, zur Erklärung herangezogen (Williams & Segal, 2003, S. 96 f.).

Als Folge des attribution bias können auch die Leistungen in ein negatives Licht gerückt werden. Folgende Interviewsequenz einer teilzeiterwerbstätigen Mutter soll dies verdeutlichen:

*“[B]efore I went part-time, when people called and found I was not at my desk, they assumed that I was elsewhere at a business meeting. But after I went part-time, the tendency was to assume that I was not there because of my part-time schedule—even if I was out at a meeting. Also, before I went part-time, people sort of gave me the benefit of the doubt. They assumed that I was giving them as fast a turn-around as was humanly possible. After I went part-time, this stopped, and they assumed that I wasn’t doing things fast enough because of my part-time schedule. As a result, before I went part-time, I was getting top-of-the-scale performance reviews. Now I’m not, though as far as I can tell, the quality of my work has not changed”* (Williams & Segal, 2003, S. 97).

#### (5) *Leniency Bias*

Zahlreiche Untersuchungen belegen, dass die Eigengruppe vorteilhafter behandelt wird als die Fremdgruppe („in-group favoritism“) (vgl. hierzu Brewer, 1979)<sup>14</sup>. Ein bekanntes Muster ist dabei „leniency bias“, bei dem objektive Regeln für die Eigengruppe flexibel eingesetzt werden, bei der Fremdgruppe sich hingegen strikt daran gehalten wird (vgl.

<sup>14</sup> Siehe hierzu auch „Minimal Group Paradigm“ (Tajfel, Billig, Bundy, & Flament, 1971) sowie „Social Identity Theory“ (Tajfel & Turner, 1986).

Williams & Segal, 2003, S. 98). So müssen Mütter beispielsweise viel höheren Pünktlichkeits- und Leistungsstandards genügen als Väter (oder Frauen, die keine Kinder haben) (vgl. Williams, 2006a, S. 157).

### *Zusammenfassung*

Wie im vorangegangenen Abschnitt erläutert, kann die Maternal Wall verschiedene Formen annehmen – alle wirken sich negativ auf die Stellung der Frauen im Erwerbsleben aus: Da die Anforderungen in der Arbeitswelt nach wie vor an männlichen Erwerbs- und Lebensmustern ausgerichtet sind, der Muttermythos jedoch eine ständige Verfügbarkeit der Mutter für ihre Kinder nahe legt, werden beide Rollen als unvereinbar angesehen. Dies kann einerseits in die Erwartungshaltung münden, Mütter sollten sich lieber aus dem Erwerbsleben zurückziehen und sich voll und ganz auf die Familie konzentrieren, da andernfalls beide Bereiche darunter zu leiden hätten (feindselige präskriptive Stereotype). Andererseits ist es auch möglich, dass Personalverantwortliche den Müttern die Erfüllung ihrer „familiären Verpflichtungen“ erleichtern wollen, indem sie ihnen beispielsweise eine Führungsposten überhaupt nicht erst anbieten oder sie von zusätzlichen Reisen und Weiterbildungsmaßnahmen „befreien“ (diskriminierende Handlungen, die auf wohlwollenden präskriptiven Stereotypen beruhen).

Auch in Bezug auf die deskriptiven Geschlechterstereotype ergeben sich Nachteile. So laufen Frauen, deren Mutterrolle salient wird, Gefahr, aus dem Cluster der Karrierefrauen (hochintelligent, ambitioniert, dominant, hart arbeitend, kompetent) in das Cluster der Hausfrauen zu rutschen (nett, ordentlich und warmherzig, aber abhängig, unsicher und inkompetent) (vgl. Williams & Segal, 2003, S. 91). Dies zieht in beruflicher Hinsicht negative Konsequenzen nach sich, da die feminin-positiven Eigenschaften der Hausfrauen in der Arbeitswelt weniger gefragt sind, der Verlust an wahrgenommener Kompetenz und angenommenem Engagement für die Karriere jedoch äußerst nachteilig ist.

Des Weiteren können die präskriptiven Stereotype, die Müttern die primäre Verantwortung für die Kinderbetreuung zuschreiben, einen nachteiligen Einfluss darauf ausüben, wie uneindeutige Ereignisse bzw. Sachverhalte im beruflichen Kontext interpretiert werden (z.B. ein Zuspätkommen) (*attribution bias*). Aufgrund der angenommenen Rollen-Inkongruenz müssen sich erwerbstätige Mütter zudem doppelt bewähren (*leniency bias*).

In Unterkapitel 2.4 wird nun der Bogen zur Einstellungsforschung geschlagen und auf sexistische Einstellungen sowie deren Erfassung eingegangen. Hierfür werden neben den verschiedenen Sexismus-Theorien des traditionellen, modernen und ambivalenten Sexismus etablierte Messinstrumente vorgestellt.

## 2.4 Sexismus: Theorien und Messinstrumente

### 2.4.1 Allgemeine Definition

Als Begriffsbestimmung für „Sexismus“ liegt dieser Arbeit die Definition von Swim und Hyers (2009) zugrunde. Die Autorinnen bezeichnen Sexismus als “individuals’ attitudes, beliefs, and behaviors, and organizational, institutional, and cultural practices that either reflect negative evaluations of individuals based on their gender or support unequal status of women and men” (S. 407). Durch diese Definition wird deutlich, dass beide Geschlechter – sowohl Frauen als auch Männer – mit Sexismus konfrontiert werden können (vgl. Swim & Hyers, 2009, 407 f.). Da in den meisten Gesellschaften jedoch nach wie vor eine Vorrangstellung der Männer hinsichtlich gesellschaftlicher Macht besteht, was beispielsweise das Einkommen oder die Übernahme leitender Positionen in der Wirtschaft anbelangt, (vgl. Glick & Fiske, 1996; siehe Unterkapitel 1.1), konzentriert sich die (Einstellungs-)Forschung überwiegend auf Sexismus in Bezug auf Frauen (vgl. Eckes, 2008, S. 176).<sup>15</sup>

Die Definition von Swim und Hyers (2009) lässt dabei offen, von wem die Diskriminierung ausgeht. Sexistische Einstellungen in Bezug auf Frauen müssen folglich nicht immer zwangsläufig von Männern vertreten werden, sondern können ebenfalls von Frauen gehegt werden. Dies wird v.a. dann zutreffen, wenn es sich um unterschiedliche Subgruppen von Frauen handelt. Eine Hausfrau, die aus Überzeugung ihre Kinder selbst betreut, wird wahrscheinlich eine Karrierefrau abwerten, die ihre Kinder zur Betreuung weggibt („Warum hat sie denn überhaupt Kinder bekommen?“). So geht es darum, die eigenen Handlungen und Lebensweisen der Eigengruppe (in diesem Falle Hausfrauen) gegenüber der Fremdgruppe (in diesem Falle Karrierefrauen) zu rechtfertigen, sie in ein positives Licht zu rücken und gleichzeitig das Verhalten der Fremdgruppe abzuwerten (siehe hierzu „Social Identity Theory“ (Tajfel & Turner, 1986) und “in-group favouritism” und “out-group derogation” (Brewer, 2001).)

Hinsichtlich der Valenz der Einstellungen werden in der Definition von Swim und Hyers zwar explizit nur negative Evaluationen benannt, es wird aber darauf hingewiesen, dass es bei sexistischen Einstellungen oder Verhaltensweisen darauf ankommt, dass sie einen ungleichen Status zwischen Frauen und Männern unterstützen. So können auch (scheinbar) positive Einstellungen als sexistisch angesehen werden, wenn sie dazu beitragen, die sozialen Ungleichheiten zwischen den Geschlechtern zu festigen und die Geschlechterhierarchien fortzusetzen.

Des Weiteren greift die Definition verschiedene Ebenen von Sexismus auf. Neben individuellen Einstellungen, Vorurteilen und Verhaltensweisen können auch organisatorische, institutionelle und kulturelle Praktiken als sexistisch eingestuft werden, insofern

---

<sup>15</sup> Zu sexistischen Einstellungen in Bezug auf Männer siehe beispielsweise das „Ambivalence toward Men Inventory“ (AMI) von Glick und Fiske (1999).

diese eine negative Bewertung von Frauen (oder Männern) widerspiegeln oder soziale Ungleichheiten zwischen den Geschlechtern unterstützen.

Mit der Entwicklung der Skala zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter (kurz FWSEM-Skala) liegt der Schwerpunkt dieser Arbeit auf den individuellen sexistischen Einstellungen (und folglich nicht auf institutionellen oder kulturellen Praktiken). Als Einstellungsobjekt wurden Frauen bzw. Mütter (und nicht Männer) ausgewählt, da sie – wie in Unterkapitel 2.1.3 und 2.3.2 gezeigt werden konnte – in besonderer Weise von mittelbaren und unmittelbaren Diskriminierungen in unserer Gesellschaft betroffen sind. Dies hängt zu einem Großteil mit den alten Rollenbildern bezüglich der Kinderbetreuung und dem Festhalten an traditionellen Arbeitsteilungsmustern zusammen, die zahlreiche negative Konsequenzen hinsichtlich der Erwerbstätigkeit von Müttern mit sich bringen, die wiederum die Macht- und Statusunterschiede zwischen den Geschlechtern festigen. Neben negativen werden subjektiv positiv empfundene Einstellungen erfasst, die dazu beitragen, Frauen auf die traditionelle Geschlechterrolle als Mutter und Hausfrau zu verweisen und somit ihren untergeordneten gesellschaftlichen Status aufrechtzuerhalten.

#### 2.4.2 Facetten von Sexismus

Die Sexismusforschung hat in den vergangenen Jahren eine Reihe von unterschiedlichen Theorien, Messinstrumenten und Methoden hervorgebracht, die dabei helfen die Natur sexistischer Einstellungen zu verstehen und diese zu erfassen. Sowohl die Konzeptualisierung von Sexismus als auch die entsprechenden Messinstrumente wurden dabei immer weiter ausdifferenziert. Während sich die anfängliche Sexismusforschung auf offene Formen von Sexismus konzentrierte, verschob sich der Fokus im Laufe der Zeit auf subtilere Formen (vgl. Goodwin & Fiske, 2001; Swim et al., 2010, S. 138). Man kam von einer reinen Antipathie Frauen gegenüber ab und griff auch scheinbar positive Sexismusvarianten auf.

Derzeit werden drei Arten von Sexismus unterschieden: der *traditionelle*, der *moderne* und der *ambivalente Sexismus*. Um die entsprechenden Einstellungen zu erfassen, wurden mehrere Skalen entwickelt, deren prominenteste Vertreter im Folgenden vorgestellt werden.

##### 2.4.2.1 Traditioneller Sexismus

###### *Begriffsbestimmung*

Cameron (1977) definiert traditionellen Sexismus ("old-fashioned sexism") als "a prejudicial attitude or discriminatory behavior based on the presumed inferiority or difference of women as a group" (S. 340). Der traditionelle Sexismus stellt folglich eine direkte Form der Diskriminierung von Frauen dar und wird häufig als offener Sexismus bezeichnet. Im Wesentlichen ist diese Form des Sexismus durch drei Merkmale gekennzeichnet: (1) Ne-

ben der Befürwortung traditioneller Geschlechterrollen wird (2) auf Geschlechtsunterschiede zwischen Frauen und Männern hingewiesen. Des Weiteren wird (3) die Ansicht vertreten, das männliche Geschlecht sei dem weiblichen überlegen (vgl. Eckes & Six-Materna, 1998, S. 225).

Im Zentrum der Skalen, die traditionell sexistische Einstellungen erfassen sollen, stehen insbesondere die traditionellen Geschlechterrollen, die Frauen und Männern unterschiedliche Verhaltensweisen für den familiären und den beruflichen Bereich vorgeben. Dementsprechend enthalten viele der Items präskriptive Stereotype und somit Aussagen darüber, wie sich Frauen und Männer verhalten *sollten* (z.B. „Frauen sollten sich wieder mehr auf ihre Rolle als Hausfrau und Mutter konzentrieren.“) (vgl. Swim & Campbell, 2003, S. 224). Nicht eindeutig vom traditionellen Sexismus ist daher das Konzept der normativen Geschlechterrollen-Orientierung zu unterscheiden. Da es hierbei ebenfalls um die Befürwortung bzw. die Ablehnung geschlechtsspezifischer Arbeitsteilung geht, wird dieses Konzept in dieser Arbeit unter dem Oberbegriff traditioneller Sexismus subsumiert.

### *Messinstrumente*

Seit Anfang der 1970er Jahre wurden mehrere englischsprachige Skalen zur Messung von traditionellem Sexismus und Geschlechterrolleneinstellungen generiert (vgl. McHugh & Frieze, 1997; Swim et al., 1995): Die „Attitudes Toward Woman Scale“ (AWS, Spence & Helmreich, 1972), die „Sex-Role Orientation Scale“ (SRO, Brogan & Kutner, 1976) und die „Sex-Role Egalitarianism Scale“ (SRES, Beere, King, Beere, & King 1984) (vgl. Braun, 2008, S. 644 f.). Die „Attitudes Toward Women Scale“ ist eine der ersten und am häufigsten angewendeten Skalen im englischsprachigen Raum. Sie erfasst offen sexistische Einstellungen hinsichtlich der Rechte, der Verantwortlichkeiten und der Rolle von Frauen (und somit weniger explizite Einstellungen in Bezug auf Frauen) (vgl. McHugh & Frieze, 1997, S. 6). In ihrer ursprünglichen Fassung umfasste sie 55 Items, daneben wurde eine Kurzform mit 15 Items entwickelt (Spence, Helmreich, & Stapp, 1973), von denen sich zwei auf die Rolle als Mutter beziehen. (Item 1: “Women should worry less about their rights and more about becoming good wives and mothers.” und Item 2: “In general, the father should have greater authority than the mother in the bringing up of the children.”). Die Skala besitzt eine zufriedenstellende interne Konsistenz, weist eine einfaktorielle Struktur auf und zeigt stabile Geschlechtsunterschiede in den Mittelwerten (vgl. Eckes & Six-Materna, 1998, S. 226). Im Laufe der Jahrzehnte offenbarte sich jedoch eine Schwachstelle der AWS: Die sich verändernden Einstellungen hinsichtlich der Rechte, Rollen und Fähigkeiten von Frauen trugen zu Deckeneffekten beim Einsatz der Skala bei, so dass sie nicht mehr hinreichend zwischen den Einstellungen der Befragten differenzieren konnte (vgl. McHugh & Frieze, 1997, S. 7). Spence und Hahn (1997) stellten fest, dass die Mehrzahl der Befragten egalitäre Antwortmuster aufwies und dementsprechend Items wie dem folgenden nicht mehr zustimmten: “Sons in a family should be given more encouragement to go to college than daughters” (vgl. Eckes & Six-Materna, 1998, S. 227). McHugh



und Frieze (1997) verdeutlichen, dass dieses Problem nicht allein für die AWS, sondern im Prinzip für alle Sexismus-Skalen besteht:

*“It may be difficult for any scale to measure gender-related attitudes over an extended period of time. As women successfully challenge restrictions on their rights and roles, new/different issues become the source of controversy and challenge. For example, today few individuals would question women’s ability and right to attend medical school, but there is continuing debate over women as firefighters and combat soldiers” (S. 7).*

Um sexistische Einstellungen zeitgemäß erfassen zu können, muss die Forschung daher den gesellschaftlichen Veränderungen durch die Entwicklung neuer Messinstrumente, aber auch der theoretischen Weiterentwicklung und Ausdifferenzierung des Sexismus-Konzept gerecht werden (siehe Unterkapitel 2.4.2.2 und 2.4.2.3).

Auch in Deutschland kamen mehrere Skalen zum Einsatz, die unter dem Konzept des traditionellen Sexismus subsumiert werden können: Neben den „Misogynie-Skalen“ (MS) von Krameyer und Schmidt (1973) ist die „Skala zur Messung der normativen Geschlechtsrollen-Orientierung“ (GRO) von Krampen (1979), der „Fragebogen zur Messung normativer Geschlechtsrollenorientierung“ (NGRO) von Athenstaedt (2000) sowie die „Geschlechterrollen-Batterie“ (GRB) des International Social Survey Programme (ISSP) 2002 (siehe hierzu Braun, 2006; Scholz, Harkness, & Klein, 2003) zu nennen.

*Misogynie-Skalen.* Für die Entwicklung ihrer Einstellungs-Skalen definierten Krameyer und Schmidt (1973) Frauenfeindlichkeit als eine „Ablehnung der Gleichberechtigung“. Insgesamt generierten sie sieben Subskalen mit je 10 Items zu unterschiedlichen Aspekten bzw. Bereichen der Frauenfeindlichkeit: (1) *Öffentlichkeit* (Bsp.: „Berufsarbeit zerstört die biologischen Fähigkeiten der Frau.“), (2) *Ehe bzw. Paarbeziehungen* (Bsp.: „Es ist die Pflicht einer jeden Frau, für eine gute Ehe ihre Selbstständigkeit aufzugeben.“), (3) *Sexualmoral und Sexualverhalten* (Bsp.: „Frauen dürfen bei weitem nicht so triebhaft sein wie Männer.“), (4) *Beurteilung von Begabung, Fähigkeiten, Leistungen* (Bsp.: „Wenn Frauen nur sorgfältig ausgebildet würden, wären sie in jeder Beziehung den Männern ebenbürtig.“), (5) *Beurteilung von nicht-leistungsbezogenen Eigenschaften* (Bsp.: „Die Frau ist körperlich und seelisch schwächer als der Mann.“), (6) *direkte Stellungnahmen zur Gleichheit und Gleichberechtigung* (Bsp.: „Die völlige Gleichberechtigung von Mann und Frau ist eine der größten Katastrophen der Menschheit.“) und (7) *spezielle Zuweisung von Aufgaben und Pflichten* (Bsp.: „Sobald eine Frau Kinder hat, sollte sie ihre ganze Zeit und Kraft der Welt des Hauses widmen.“). Von den insgesamt 70 Items beziehen sich 6 Items auf Frauen im Zusammenhang mit ihrer Rolle als Mutter. Alle stammen aus der letztgenannten Inhaltskategorie.

Um die „Misogynie-Skalen“ zu konstruieren, entwickelten Krameyer und Schmidt (1973) ein dreistufiges Verfahren, bei dem die nachfolgenden Methoden hintereinander zum Einsatz kamen: (1) Die Methode der gleicherscheinenenden Intervalle, (2) die Methode

der summierten Schätzungen und (3) eine Dimensionalitätsprüfung mit Hilfe einer Skalogramm-Analyse. Die Ergebnisse der Analysen legen eine Eindimensionalität der Skala nahe, Angaben zur Reliabilität und Validität der Skala fehlen. Es wird darauf hingewiesen, dass die Verteilung der Skala linksschief ist (vgl. Krameyer & Schmidt, 1973, S. 41 ff.).

*Skala zur Messung der normativen Geschlechtsrollen-Orientierung.* Die GRO-Skala ist das Ergebnis einer deutschsprachigen Bearbeitung der „Sex-Role Orientation Scale“ von Brogan und Kutner (1976) und liegt sowohl in einer Langfassung (38 Items) als auch in einer Kurzform (elf Items) vor (vgl. Eckes & Six-Materna, 1998, S. 227). Geschlechtsrollen-Orientierungen werden von Krampen (1979) als „normative Einstellungsaspekte bezeichnet, die moralische Werte darüber beinhalten, wie sich Männer und Frauen verhalten sollten und wie sie sein sollten“ definiert (S. 255). Inhaltlich decken die Items der Skala sechs Bereiche ab: (1) Arbeitsteilung in der Ehe, (2) Machtstrukturen in Familie, Politik und der Arbeitswelt, (3) berufliche Beschäftigungen von Frauen und Männern, (4) politischer Status von Frauen, (5) geschlechtsspezifische Sozialisation, (6) weitere Stereotype über Geschlechterrollen-Verhalten. In der Kurz-Skala befassen sich 2 Items mit schwangeren Frauen bzw. der Arbeitsteilung zwischen Mütter und Vätern (Item 1: „Eine Frau, die in der Öffentlichkeit arbeitet (z.B. eine Verkäuferin) sollte nicht mehr arbeiten, wenn zu sehen ist, dass sie schwanger ist.“ und Item 2: „Es ist nicht gut, wenn ein Mann zu Hause bleibt und die Kinder versorgt und seine Frau arbeitet.“). Sowohl die Lang- als auch die Kurzform der Skala erwiesen sich als reliabel und valide (vgl. Krampen, 1979, 1983).

*Fragebogen zur Messung normativer Geschlechtsrollenorientierung.* Athenstaedt entwickelte im Jahr 2000 einen neuen „Fragebogen zur Messung normativer Geschlechtsrollenorientierung“ (NGRO) mit 29 Items, um die „Ausprägung internalisierter, persönlicher Geschlechtsrollennormen“ zeitgemäß erfassen zu können. Wie bereits angedeutet, wiesen ältere Skalen mit der Zeit zunehmend schiefe Verteilungen der Antworten auf, so dass eine Differenzierungsfähigkeit nicht mehr zufriedenstellend gegeben war. Auch die Itemformulierungen erwiesen sich als veraltet. Mit Hilfe der neuen Skala sollte eine Unterscheidung zwischen den einzelnen Befragten innerhalb des gesamten Einstellungsgebietes ermöglicht werden. Für die Ausformulierung der Items dienten vier Inhaltskategorien als Orientierung: (1) Ehe und Familie, (2) Erwerbsleben, (3) Politik sowie (4) Sozialisation. Die Items sollten sich dabei sowohl auf Frauen als auch auf Männer beziehen sowie traditionelle und egalitäre Einstellungen gleichermaßen widerspiegeln (Bsp. 1 (traditionell): „Hemden bügeln ist nicht Sache der Männer.“; Bsp. 2 (egalitär): „Man kann von Frauen nicht fordern, dass sie die Hausarbeit alleine verrichten müssen.“). Keines der Items spricht explizit Mütter an, lediglich die Väter werden im Zusammenhang mit der Kinderbetreuung erwähnt (Bsp.: „Auch Männer sollten nach der Geburt ihres Kindes die Möglichkeit eines Erziehungsurlaubes in Anspruch nehmen können.“). Faktorenanalysen ergaben Athenstaedt (2000) zufolge eine Eindimensionalität der Skala. Die Reliabilität und Validität des Messinstrumentes erwiesen sich als zufriedenstellend. Es zeigte sich zudem eine Normalverteilung der Skala in verschiedenen Alters- und Bildungsgruppen der Befragten.

*Geschlechterrollen-Batterie.* Bereits in den 60er Jahren wurden in Amerika Vorläufer-Items der Geschlechterrollen-Batterie (GRB) des ISSP eingesetzt (siehe hierzu u.a. Mason, Czajka, & Arber, 1976). 1977 wurde ein Teil der Items in den General Social Survey der USA aufgenommen, die wiederum 1982 (neben einigen neuen Items) in der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften in Deutschland berücksichtigt wurden. 1988 schließlich kam die GRB das erste Mal im Rahmen des International Social Survey Programme zum Einsatz (vgl. Braun, 2006, S. 113). Das ISSP ist eine internationale Bevölkerungsumfrage (in mittlerweile 39 Ländern), die seit 1985 mit jährlich wechselnden Themenmodulen durchgeführt wird. In den Jahren 1988, 1994 und 2002 lag der Schwerpunkt auf dem Thema „Familie und sich ändernde Geschlechterrollen“. Im Kern geht es in diesem Modul um die Frage, wie die Rollenverteilung zwischen den Geschlechtern aussehen sollte und ob die Frau einer Erwerbstätigkeit nachgehen sollte oder nicht. Seit dem ersten Einsatz der Geschlechterrollen-Batterie wurde die Skala mehrfach modifiziert. Einige Items wurden gestrichen, gleichzeitig wurden neue Inhaltsdimensionen einschließlich neuer Items aufgenommen. So rückte in der letzten Befragungswelle auch die Rolle des Mannes in das Interesse der ForscherInnen (vgl. Blohm, Harkness, Klein, & Scholz, 2003, S. 5 ff.). In den älteren Fassungen konzentrierten sich die Items überwiegend auf die Rolle der Frauen, was implizierte, dass lediglich ihr Verhalten als „problematisch“ angesehen wurde (vgl. Braun, 2008, S. 644 f.). In der neusten Version aus dem Jahr 2002 werden auch die Einstellungen hinsichtlich der Rolle des Mannes im privaten Bereich näher untersucht. Das Messinstrument besteht in dieser Fassung aus 10 Items und deckt inhaltlich vier Dimensionen ab: (1) Die „Konsequenzen der weiblichen Erwerbstätigkeit für die Familie“ (Bsp.: „Ein Kind, das noch nicht zur Schule geht, wird wahrscheinlich darunter leiden, wenn seine Mutter berufstätig ist.“), (2) die „Geschlechterrollen-Ideologie“ (Bsp.: „Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern.“), (3) die „Ökonomischen Konsequenzen weiblicher Erwerbstätigkeit“ (Bsp.: „Einen Beruf zu haben ist das beste Mittel für eine Frau, um unabhängig zu sein.“) (vgl. Braun, 2006) sowie (4) die „Rolle des Mannes hinsichtlich der Reproduktionsarbeiten“ (Bsp.: „Männer sollten einen größeren Anteil an der Kinderbetreuung übernehmen, als sie es jetzt tun.“). Gemäß einer traditionellen Geschlechterrollen-Ideologie sollte der Mann die außerhäusliche Erwerbstätigkeit übernehmen, während die Frau sich um die Familie und den Haushalt kümmern sollte. Das Ideal der familialen Kinderbetreuung durch die Mutter wird herausgestellt und folglich eine Erwerbstätigkeit von Frauen abgelehnt, da sie negativen Konsequenzen für das Wohl der Kinder/der Familie zur Folge hat (vgl. Braun, 2006, S. 8). In 5 der 10 Items werden Frauen in ihrer Rolle als Mutter angesprochen.

Vor den Modifikationen der Geschlechterrollen-Batterie im Jahr 2002 wurde neben einem quantitativen auch ein qualitativer Pretest in Form von kognitiven Interviews durchgeführt. Es sollte dabei herausgefunden werden, welche der bisherigen Items möglicherweise überarbeitet werden sollten und welche der neuen Items sich am besten für

einen internationalen Vergleich eignen (vgl. Blohm et al., 2003, S. 41; Scholz, Harkness, & Klein, 2003 S. 9).

#### **2.4.2.2 Moderner Sexismus**

##### *Theorie*

Im Zuge der verstärkten öffentlichen Debatte um die Gleichstellung von Frauen und Männern wurde eine offene Befürwortung traditioneller Geschlechterrollen immer weniger gesellschaftlich akzeptiert. Gleiches galt auch für Vorurteile sowie diskriminierendes Verhalten gegenüber Frauen. Folglich wurden derartige Einstellungen immer seltener offen geäußert, so dass der traditionelle Sexismus mit Hilfe von Befragungen immer schwerer erfasst werden konnte (vgl. Campbell, Schellenberg, & Senn, 1997, S. 89; Eckes & Six-Materna, 1998, S. 226 ff.; Swim, Becker, Lee, & Pruitt, 2010, S. 138). So sanken beispielsweise die Zustimmungsraten zur Attitudes Toward Women Scale (Spence & Hahn, 1997) sowohl für Frauen als auch für Männer im Laufe der Jahre immer weiter.

Angesichts dieser Entwicklungen gaben verschiedene TheoretikerInnen (Swim, Aiken, Hall, & Hunter, 1995; Tougas, Brown, Beaton, & Joly 1995) zu bedenken, dass von einem wirklichen Abklingen sexistischer Einstellungen vermutlich nicht die Rede sein könne, sondern lediglich eine Verschiebung in Richtung einer subtilen, verdeckten Form stattgefunden hätte (vgl. Goodwin & Fiske, 2001, S. 359). Die Notwendigkeit, neue Theorien und subtile Messinstrumente zu generieren wurde somit deutlich. In Anlehnung an den „modernen Rassismus“ (McConahay, 1986) bzw. den „symbolischen Rassismus“ (Kinder & Sears, 1981) wurden unabhängig voneinander die Konzepte des „modernen Sexismus“ und des „Neosexismus“ entwickelt (vgl. Swim, Becker, Lee, & Pruitt, 2010, S. 149 f.). Der moderne Sexismus zeichnet sich nach Swim et al. (1995) vor allem durch (a) die Leugnung der fortbestehenden Diskriminierung von Frauen, (b) den Unmut über Frauen, die sich über Diskriminierungen/Sexismus beklagen und (c) die Ablehnung von Frauenförder-Maßnahmen aus (vgl. Swim et al., 1995, S. 199). Eine Begriffsbestimmung des „Neosexismus“ liefern Tougas et al. (1995), sie definieren ihn als eine „manifestation of a conflict between egalitarian values and residual negative feelings towards women“ (S. 843).

Dem Grundgedanken des modernen Rassismus entsprechend nehmen moderne Sexisten eine nahezu vollständige Realisierung der Gleichberechtigung zwischen Frauen und Männern an und erachten weitere Forderungen diesbezüglich als überflüssig. Politische Maßnahmen zur Frauenförderung werden als eine Bedrohung egalitärer Werte eingestuft, da sie Frauen einen Vorteil verschaffen und Männer schlechter stellen (vgl. Eckes & Six-Materna, 1998, S. 226). Als modern sexistisch werden daher solche Personen bezeichnet, die Frauenfördermaßnahmen ablehnen, eine Benachteiligung von Frauen in Kauf nehmen bzw. leugnen und somit den Status quo der Geschlechterungleichheit aufrechterhalten wollen (vgl. Swim & Campbell, 2003, S. 221).

Aus der Definition des modernen Sexismus ergibt sich jedoch ein generelles Problem: Personen, die eine fortgesetzte Diskriminierung von Frauen leugnen, werden als modern sexistisch eingestuft. Dabei kann es jedoch leicht zu Fehleinschätzungen kommen, wenn die Befragten beispielsweise einfach nicht wissen, dass Frauen heutzutage immer noch in einigen Bereichen weniger Lohn erhalten als ihre männlichen Kollegen. Dies gilt insbesondere für jüngere Befragte, die bisher noch keine eigenen Erfahrungen im Berufsleben sammeln konnten und daher auch selbst noch nicht mit Diskriminierungen konfrontiert wurden. Ein weiteres Manko im Zusammenhang mit diesem Konzept besteht darin, dass fast alle Items als Meinungs- und nicht als Einstellungsitems formuliert sind (vgl. hierzu Schüßler, 2006, S. 78 ff.). Braun (2000) zufolge beziehen sich Einstellungsitems darauf, „wie die Befragten ein Einstellungsobjekt bewerten“ wohingegen Meinungsitems ansprechen, „(...) wie etwas ist bzw. von den Befragten gesehen wird, ohne dabei auszudrücken, wie die Befragten dies persönlich bewerten“ (Braun, 2000, S. 47, Hervorhebung im Original). (Beispielitem MSS: „Wenn Frauen tatsächlich einmal schlechter bezahlt werden als Männer, dann nur deshalb, weil sie einfachere Arbeit zu leisten haben.“). Meinungsitems erfassen somit lediglich eine Einschätzung einer befragten Person bezüglich bestimmter Sachverhalte oder wahrgenommener Eigenschaften. Sie verlangen keine Bewertung eines Einstellungsobjektes und müssen folglich auch nicht die Ansichten einer Person widerspiegeln (vgl. Braun, 2000, S. 52). Die eben genannten Probleme ergeben sich vor allem im Zusammenhang mit der Modern Sexism Scale, da diese sich vor allem auf die Leugnung der Diskriminierung bezieht, während die Neosexism Scale hauptsächlich die Ablehnung weiblicher Forderungen und den Widerstand gegen Frauenfördermaßnahmen abfragt (siehe nächster Abschnitt). Items, die Einstellungen zur positiven Diskriminierung (affirmative action) von Frauen erfassen, scheinen daher besser geeignet, da für sie die eben beschriebenen Probleme nicht bestehen.

#### *Messinstrumente*

Im englischsprachigen Raum wurden – den oben angeführten Definitionen entsprechend – die beiden folgenden Skalen entwickelt: Die „Modern Sexism Scale“ von Swim et al. (1995) mit acht Items sowie die „Neosexism Scale“ von Tougas et al. (1995) mit elf Items. Beide Skalen decken inhaltlich die drei genannten Dimensionen des Modernen Sexismus nach Swim et al. (1995) ab, dennoch werden unterschiedliche Schwerpunkte gelegt. Während sich die Items der Modern Sexism Scale überwiegend (fünf von acht Items) mit der Leugnung der fortbestehenden Diskriminierung von Frauen befassen, drehen sich die Items der Neosexism Scale hauptsächlich (sieben von elf Items) um die Ablehnung weiblicher Forderungen und den Widerstand gegen Frauenfördermaßnahmen (vgl. Swim et al., 2010, S. 150). Ähnlich der Attitudes Toward Women Scale (AWS) erfassen auch diese Skalen weniger direkte Einstellungen in Bezug auf Frauen, als Einstellungen hinsichtlich der Gleichstellung der Geschlechter (vgl. Rudman, 2005, S. 109). Beide Skalen wurden in verschiedenen europäischen Ländern adaptiert und auch validiert.

Eckes und Six-Materna (1998) generierten mit der „Skala zur Erfassung des modernen Sexismus“ (MSS) das deutschsprachige Instrument. Die Skala besteht insgesamt aus zehn Items, von denen fünf aus der „Modern Sexism Scale“ und zwei aus der „Neosexism Scale“ übernommen wurden. Die verbliebenen drei Items wurden von beiden AutorInnen selbst formuliert (vgl. Eckes & Six-Materna, 1998, S. 226 ff.). Kein einziges Item der deutschen Version bezieht sich auf die Subgruppe der Mütter.

Sowohl in ihren Ursprungsformen als auch in den adaptierten Versionen erwiesen sich die Skalen als reliabel und valide. Konfirmatorische Faktorenanalysen zeigten, dass die Modern Sexism Scale und die Neosexism Scale Konstrukte messen, die zwar mit der Befürwortung traditioneller Geschlechterrollen und Stereotype korreliert sind, sich aber dennoch davon unterscheiden (vgl. Swim et al., 2010, S. 150). In verschiedenen Studien konnte die konvergente Validität der beiden englischsprachigen Skalen sowie der deutschsprachigen Skala bestätigt werden: Sowohl die Modern Sexism Scale als auch die Neosexism Scale wiesen signifikante Korrelationen mit der AWS auf (vgl. Swim et al., 1995; Tougas et al., 1995) bzw. die MSS mit der GRO (vgl. Eckes & Six-Materna, 1998, S. 232 f.). Folglich besteht zwischen den offen negativen Einstellungen der traditionellen Skalen (AWS und GRO) und den modernen Messinstrumenten ein Zusammenhang: Beiden liegt eine Antipathie in Bezug auf Frauen zugrunde. Viele traditionelle Sexisten sind zugleich moderne Sexisten, nicht alle modernen Sexisten sind jedoch auch gleichzeitig traditionelle Sexisten (vgl. Goodwin & Fiske, 2001, S. 359).

Die Existenz modern sexistischer Einstellungen sowie die Verwendbarkeit der Messinstrumente müssen nach Ansicht von Swim et al. (2010) noch in anderen Ländern außer Nordamerika und Europa getestet werden, da sich der Status von Frauen in verschiedenen Ländern sehr unterscheidet. So ist es möglich, dass moderner Sexismus und somit die Leugnung von Diskriminierung lediglich in den Staaten vorherrscht, die (zumindest oberflächlich betrachtet) schon relativ große Fortschritte hinsichtlich der Gleichstellung von Frauen und Männern erreicht haben (vgl. Swim et al., 2010, S. 151).

Auch die Skalen des modernen Sexismus sind mit einem „Verfallsdatum“ versehen, denn wie beim traditionellen Sexismus können die gesellschaftlichen Entwicklungen die Inhalte der Skalen einholen, so dass sie keine Gültigkeit mehr besitzen (vgl. McHugh & Frieze, 1997, S. 7). Wie bereits ausgeführt sind Sexismus-Items und deren Interpretation stark vom Wandel der Zeit abhängig. So ist es möglich, dass das pragmatische Verständnis der Befragten aufgrund veränderter Einstellungen und Verhaltensweisen in der Gesellschaft nicht mehr im intendierten Sinne aufgefasst werden. Ebenso kann es vorkommen, dass Items zum modernen Sexismus auf Grund der fortschreitenden Gleichberechtigung der Geschlechter die Realität einfach nicht mehr treffend wiedergeben. Deutlich wird dies am folgenden Beispiel: In eigenen Untersuchungen (Schüßler, 2006; Schüßler & Schmidt, 2008) fiel auf, dass das Item „In der Schule werden Mädchen immer noch benachteiligt“ der MSS im Vergleich zu den anderen Items der Skala durchweg eine geringe Zustimmung erzielte. So lag der Mittelwert für die Gesamtstichprobe bei einer Skala von 1 „stimme voll und ganz zu“ bis 6 „stimme überhaupt nicht zu“ bei

$M = 4.88$  ( $SD = 1.15$ ). In der Subgruppe der Frauen betrug er  $M = 4.59$  ( $SD = .96$ ) und bei den Männern  $M = 5.15$  ( $SD = 1.25$ ). Insgesamt stimmten 89.4% der Befragten der Aussage nicht zu und stritten folglich eine Benachteiligung der Mädchen in der Schule ab. Dies trifft auch auf Personen zu, die bei anderen Items eine äußerst egalitäre, nicht-sexistische Einstellung aufwiesen. Folglich ist das Item auch nicht in der Lage zwischen den verschiedenen Einstellungen der Befragten zu differenzieren. Da auch viele Frauen die Diskriminierung im schulischen Bereich leugnen, scheint dieses Item modernen Sexismus in der heutigen Zeit nicht mehr valide erfassen zu können (vgl. Schüßler, 2006, S. 61). Die Lebenswelt der Frauen hat sich insbesondere im Bereich der Bildung in den letzten Jahrzehnten deutlich gewandelt:

*„In kaum einem gesellschaftlichen Bereich haben Frauen in den letzten 40 Jahren so aufgeholt können, wie in der Bildung. Mädchen und junge Frauen waren die Gewinner der Bildungsexpansion seit den 1970er Jahren. Heute erreichen sie im Durchschnitt höhere und bessere schulische Bildungsabschlüsse als junge Männer“* (Sachverständigenkommission, 2011, S. 212).

Es ist daher anzunehmen, dass im Falle dieses Items ein hoher Skalenwert nicht mit einer modern sexistische Einstellung einer Person gleichzusetzen ist.

#### **2.4.2.3 Ambivalenter Sexismus**

##### *Ritterlichkeit*

Wie eben angeführt, erfassen sowohl der traditionelle als auch der moderne Sexismus frauenfeindliche Einstellungen. Sie vernachlässigen somit Bewertungen von Frauen, die aus Sicht einer sexistisch eingestellten Person positiv wahrgenommen werden (siehe Unterkapitel 2.2.2). Darunter fallen vor allem die Zuschreibung von feminin-positiven Eigenschaften (wie z.B. warmherzig oder empathiefähig) sowie die Ansicht, Frauen seien zu behüten und zu verehren (vgl. Eckes & Six-Materna, 1998, S. 235).

Nadler und Morrow (1959) waren die ersten, die mit ihrem Konzept der „Ritterlichkeit“ („chivalry“) – neben offen unterordnenden Einstellungen gegenüber Frauen – auch subjektiv positiv empfundene, aber dennoch sexistische Einstellungen berücksichtigten. 14 Jahre später wurde dieses Konzept von Schmerl und Steinbach (1973) in Deutschland aufgegriffen. Ebenso wie der moderne Sexismus stellt auch die Ritterlichkeit eine subtile Form des Sexismus dar, was sie sogar noch viel wirksamer werden lässt. So dient Ritterlichkeit mitunter dazu, offensichtlich diskriminierende Handlungen zu rechtfertigen, indem die „(vermeintliche) Schutzbedürftigkeit von Frauen“ betont wird (vgl. Eckes & Six-Materna, 1999, S. 212).

Schmerl und Steinbach (1973) weisen darauf hin, dass Ritterlichkeit nicht eingesetzt wird, um Frauen etwa traditionell weibliche Haus- und Familienarbeiten abzunehmen (wie z.B. das Erledigen von Einkäufen oder das Reinigen der Wohnung). Vielmehr wird sie nur in solchen sozialen Situationen gezeigt, in denen die traditionelle Verteilung der

Geschlechterrollen offensichtlich ist und diese auch eingehalten wird. Nur dann werden Frauen für ihr „weibliches Wohlverhalten“ belohnt und Männer können sich als Ritter inszenieren und ihre Stärke unter Beweis stellen (vgl. Schmerl & Steinbach, 1973, S. 59 ff.).

#### *Theorie des Ambivalenten Sexismus*

Auch Glick und Fiske (1997) schlugen eine Erweiterung des Sexismus-Konzeptes um eine subjektiv positive Variante vor, da eine Konzeption von Sexismus als reine Feindseligkeit gegenüber Frauen zu kurz greife:

*“Definitions of sexism generally emphasize two components: hostility toward women (i.e., hostile affect and negative stereotypes) and the endorsement of traditional gender roles (i.e., restricting women’s conduct to fit societal prescriptions and confining women to roles accorded less status and power than those of men). (...) Although it has long been considered a truism that traditional attitudes about women’s roles go hand in hand with hostility toward women, recent research (...) (Glick & Fiske, 1996) suggests that this presumption is misleading because traditional attitudes are associated with highly positive as well as highly negative evaluations of women. We suggest a reconceptualization of the nature and measurement of sexism, which recognizes that sexism is fundamentally ambivalent, encompassing both subjectively benevolent and hostile feelings toward women” (Glick & Fiske, 1997, S. 119 f.).*

Dementsprechend differenzieren die beiden AutorInnen in ihrer Theorie des ambivalenten Sexismus (Glick & Fiske, 1996, 1997), zwischen feindselig (hostil) und wohlwollend (benevolent) sexistischen Einstellungen gegenüber Frauen. Die Theorie soll nun im Folgenden vorgestellt werden.

Ambivalenter Sexismus ist das Resultat zweier grundlegender Merkmale, die die Beziehungen zwischen Frauen und Männern kennzeichnen: (1) Einerseits die Macht- und Statusunterschiede zwischen beiden Geschlechtern (die sogenannte strukturelle Macht der Männer) und (2) andererseits ihre wechselseitige Abhängigkeit in heterosexuellen Paarbeziehungen (die sogenannte dyadische Macht der Frauen).

- (1) In den meisten Gesellschaften liegt die Leitung von bedeutenden wirtschaftlichen und politischen Einrichtungen nach wie vor überwiegend in den Händen von Männern (siehe hierzu auch Unterkapitel 1.1). Um die damit verbundenen Macht- und Statusunterschiede aufrechtzuerhalten und zu rechtfertigen, tendieren Männer dazu, feindselige Ideologien zu vertreten, die die Unterlegenheit von Frauen herausstellen. Dabei wird insbesondere die vermeintlich mangelnde Kompetenz von Frauen betont (vgl. Eckes & Six-Materna, 1999, S. 213; Glick & Fiske, 1996, S. 492; Glick & Fiske, 1997, S. 120 f.).



- (2) Trotz der männlichen Vorrangstellung im Berufsleben ist ein Mann im Privatleben auf seine Frau als (Ehe-)Partnerin, Hausfrau und Mutter angewiesen, die seine Bedürfnisse nach seelischer und körperlicher Nähe und Intimität befriedigt und ihm Kinder schenkt (vgl. Eckes & Six-Materna, 1999, S. 213; Glick & Fiske, 1996, S. 492 ff.). So besteht in Paarbeziehungen eine wechselseitige Abhängigkeit zwischen beiden Geschlechtern, was die Arbeitsteilung anbelangt. Wie in Unterkapitel 2.1.2 beschrieben, herrscht in Deutschland, aber auch in vielen anderen Gesellschaften, eine traditionelle Verteilung der Geschlechterrollen vor: Während der Mann vorwiegend als Ernährer der Familie fungiert, übernimmt die Frau den Hauptteil der häuslichen und familiären Pflichten (Kinderbetreuung und Hausarbeiten). Folglich wird allein durch diese Arbeitsteilung eine gegenseitige Abhängigkeit zwischen Frau und Mann begründet – beide sind aufeinander angewiesen. Dies führt dazu, dass Männer dem weiblichen Geschlecht auch wohlwollende, subjektiv positiv empfundene Einstellungen entgegenbringen und Frauen beispielsweise als rein und beschützenswert charakterisiert werden.

In der Theorie des ambivalenten Sexismus beziehen Glick und Fiske (1996, 1997) den hostile und den benevolenten Sexismus auf drei Bereiche, die die Beziehung zwischen Frauen und Männern kennzeichnen: (a) Das Patriarchat mit seiner Manifestierung in Form des Paternalismus, (b) die Geschlechterdifferenzierung und (c) die Heterosexualität. Für jede dieser drei Komponenten gibt es eine hostile und eine benevolente Variante. Gemeinsam dienen sie dazu, die sozialen Bedingungen, die das Verhältnis der Geschlechter kennzeichnen, zu erklären bzw. zu rechtfertigen.

Hostiler Sexismus umfasst die drei folgenden Komponenten:

- (a) Der *dominative Paternalismus* schließt die Ansicht ein, Männer sollten (aufgrund der vermeintlich geringeren Kompetenz des weiblichen Geschlechts) mehr strukturelle Macht als Frauen innehaben und sie zugleich kontrollieren. Verbunden damit ist die Angst, Frauen könnten versuchen, die männliche Dominanz/Vorrangstellung zu untergraben (indem sie z.B. ihre weiblichen Reize einsetzen (siehe *hostile Heterosexualität*)).
- (b) Die *kompetitive Geschlechterdifferenzierung* rechtfertigt ebenfalls die männliche strukturelle Macht und stärkt zugleich das Selbstwertgefühl der Männer, indem negative Stereotype über Frauen (z.B. ihre vermeintliche Unsicherheit oder Abhängigkeit) betont werden. Gleichzeitig werden den Frauen „typisch männliche“ Eigenschaften (wie z.B. Dominanz, Unabhängigkeit und Kompetenz) abgesprochen, die erforderlich sind, um soziale, politische oder wirtschaftliche Institutionen zu führen.
- (c) Die *hostile Heterosexualität* bezieht sich auf die Ansicht, die weibliche Sexualität stelle eine Bedrohung für die gesellschaftliche Vorrangstellung der Männer dar. So besteht die Befürchtung, Frauen könnten ihre sexuellen Reize gezielt einsetzen,

um Männer um den Finger zu wickeln und ihnen den Rang abzulaufen (Frauen als „femme fatale“ oder als „gate-keeper“) (vgl. Eckes & Six-Materna, 1999, S. 213 f.; Glick & Fiske, 1996, S. 493 f.; Glick und Fiske, 1997, S. 120 ff.).

Hostiler Sexismus zeichnet sich folglich durch ein rein negatives Frauenbild aus. Er spiegelt die Befürchtung wider, Frauen könnten mittels feministischer Ideologien oder mit Hilfe ihrer Sexualität versuchen, die Macht der Männer zu untergraben. Dementsprechend sagt hostiler Sexismus feindselige Einstellungen in Bezug auf Frauen voraus, die das traditionelle Rollenbild ablehnen, wie beispielsweise Karrierefrauen oder Feministinnen (vgl. Glick & Fiske, 2001a, S. 113).

Die benevolenten Gegenstücke zu den drei Komponenten des hostilen Sexismus sind diese:

- (a) Der *protektive Paternalismus* umfasst die Überzeugung, Männer sollten für Frauen (aufgrund der vermeintlichen Schwäche des weiblichen Geschlechts) die Rolle des Versorgers und Beschützers einnehmen.
- (b) Die *komplementäre Geschlechterdifferenzierung* weist Frauen als die „bessere Hälfte“ des Mannes aus, indem sie die positiven Stereotype über Frauen in den Vordergrund rückt, die die Wesenszüge des Mannes komplementieren. So werden Frauen beispielsweise feminine Eigenschaften wie Warmherzigkeit oder Einfühlungsvermögen zugeschrieben, die sie für die Rolle der Hausfrau und Mutter besonders geeignet erscheinen lassen. Männern hingegen werden eher maskuline Eigenschaften zugesprochen, die im Zusammenhang mit der Erwerbsrolle erforderlich sind (siehe hierzu auch Unterkapitel 2.2.3.2 und 2.2.3.4).
- (c) Die *intime Heterosexualität* rückt heterosexuelle Paarbeziehungen in ein romantisches, verklärendes Licht. Verbunden ist damit die Ansicht, ein Mann könne ohne die Liebe einer Frau im Leben keine Erfüllung finden, da nur sie seine Bedürfnisse nach Nähe und Intimität befriedige (vgl. Eckes & Six-Materna, 1999, S. 213 f.; Glick & Fiske, 1996, S. 493 f.; Glick und Fiske, 1997, S. 120 ff.; Glick & Fiske, 2001b, S. 120 ff.).

Benevolenter Sexismus ist demzufolge durch ein (scheinbar) positives Frauenbild gekennzeichnet, das sich jedoch nur auf eine bestimmte Subgruppe von Frauen bezieht: nämlich jene, die die traditionelle Geschlechterrolle erfüllen (wie beispielsweise Hausfrauen) (vgl. Glick & Fiske, 2001a, S. 113). Sie werden sozusagen für ihr „Wohlverhalten belohnt“, indem ihnen Schutz und Zuneigung angeboten werden (“women are wonderful when’ effect – when they are not in charge” (Rudman & Glick, 2008, S. 163)). Gleichzeitig werden Frauen allerdings als schwach und schutzbedürftig angesehen. Somit verhindert benevolenter Sexismus die Gleichstellung der Geschlechter, da er die traditionellen Stereotype und die männliche Dominanz untermauert. Hinzu kommt, dass

benevolenter Sexismus von vielen Menschen aufgrund seiner subtilen Form nicht als sexistisch wahrgenommen wird. Dies gilt auch für Frauen: Während sie hostilem Sexismus im Vergleich zu Männern stärker ablehnen, unterstützen sie häufig benevolenten Sexismus (vgl. Glick & Fiske, 2001a; Kilianski & Rudman, 1998).

Zusammenfassend kann festgehalten werden, dass Glick und Fiske (1996, 1997) zufolge die Koexistenz von Machtunterschieden und wechselseitiger Abhängigkeit zwischen Frauen und Männern zu ambivalent-sexistischen Einstellungen führen können, die einerseits negativ (feindselig) und andererseits scheinbar positiv (wohlwollend) sind. Sowohl hostiler als auch benevolenter Sexismus sehen Frauen als das schwache Geschlecht an, das auf häusliche und familiäre Pflichten reduziert werden sollte. Somit sind sie als „legitimierende Ideologien“ einzustufen, die dazu dienen, die gesellschaftliche Vorrangstellung der Männer zu begründen und aufrechtzuerhalten – auch wenn sie dies auf unterschiedliche Weise tun (vgl. Glick & Fiske, 2001a, S. 110 f.). Hostiler Sexismus rechtfertigt die Macht der Männer durch sexistische Antipathie, wohingegen benevolenter Sexismus die Macht eher auf eine subtile Art und Weise legitimiert, indem Frauen versprochen wird, dass sich Männer innerhalb intimer Paarbeziehungen um sie kümmern. HS und BS stellen ein Set ineinander greifender Einstellungen dar, das ein System aus Belohnungen (BS) und Bestrafung (HS) widerspiegelt und Frauen einen Anreiz liefert, die Machtunterschiede zwischen den Geschlechtern zu akzeptieren und sich nicht dagegen aufzulehnen (vgl. Glick & Fiske, 2001b, S. 117).

#### *Messung ritterlich-misogynen und ambivalent-sexistischer Einstellungen*

Die ersten beiden Skalen, die zwischen negativen (autoritären) und subjektiv positiven (ritterlichen) Einstellungen gegenüber Frauen unterscheiden, wurden bereits 1959 von Nadler und Morrow vorgelegt: die „Open Subordination of Women“-Skala und die „Chivalry“-Skala. Schmerl und Steinbach (1973) adaptierten die „Chivalry“-Skala für den deutschen Sprachraum und entwickelten die „Misogynie-Ritterlichkeits-Skala“ (MC-Skala). Sie besteht aus 30 Items, 3 davon erfassen Einstellungen der Befragten in Bezug auf die Subgruppe der Mütter (Item 1: „Durch die hohe Verantwortung für die Dinge des Haushalts, ihre zärtliche Zuneigung zu den Kindern und die Opferbereitschaft für den Erfolg der Söhne verdienen Frauen höchsten Respekt und Bewunderung.“; Item 2: „Als Familienmutter und Hüterin des heimischen Herdes kann eine Frau ihre wertvollsten Eigenschaften entfalten.“ und Item 3: „Eine Frau empfindet es als ihr vorrangiges Ziel, Leben zu schenken, zu pflegen und zu erhalten.“).

Bei allen drei Skalen wurde die faktorielle Struktur nicht untersucht. In Bezug auf die MC-Skala wurden darüber hinaus keine exakten Angaben zu den Item- und Skalenkennwerten gemacht. Sowohl inhaltlich als auch von ihrer Formulierung her sind die Items der MC-Skala mittlerweile veraltet (Beispielitem: „Es gehört sich nicht für einen Herrn, in der Gegenwart einer Dame zu fluchen oder zu schwören“). Es wird daher angenom-

men, dass die Mehrzahl der Befragten diese Items heutzutage verneinen und das Messinstrument dementsprechend nicht mehr zufriedenstellend zwischen den Einstellungen der einzelnen Befragten differenzieren kann.

Aufbauend auf ihrer Theorie des ambivalenten Sexismus generierten Glick und Fiske (1996) das „Ambivalent Sexism Inventory“ (ASI). Diese Skala besteht aus insgesamt 22 Items, wobei jeweils elf Items eine Subskala zum hostilem Sexismus (HS) bzw. zum benevolenten Sexismus (BS) bilden. Inhaltlich beziehen sich die Items auf die Gesamtgruppe der Frauen (teilweise auch auf Feministinnen) und decken sowohl für die hostile als auch für die benevolente Variante die drei oben beschriebenen Aspekte ab: (a) Paternalismus (dominant und protektiv), (b) Geschlechterdifferenzierung (kompetitiv und komplementär) und (c) Heterosexualität (feindselig und intim). Eine deutschsprachige Version des ASI legten Eckes und Six-Materna (1999) vor: Die „Skala zur Erfassung des ambivalenten Sexismus“ (ASS) (Beispielitem HS: „Frauen versuchen, Macht zu erlangen, indem sie Männer immer mehr beherrschen“; Beispielitem BS: „Eine Frau sollte von ihrem Mann auf Händen getragen werden.“).<sup>16</sup> Wie bereits eingangs erwähnt, bleiben Mütter im ASI bzw. in der ASS vollkommen unerwähnt.

In mehreren Untersuchungen erwies sich sowohl die englische als auch die deutsche Variante der ASS als reliabel und valide. Die hostile Sexismus-Skala (HS), die eine reine Antipathie gegenüber Frauen abbildet, korreliert mit offen sexistischen Einstellungen der Attitudes toward Women Scale (AWS, Spence & Helmreich, 1972), der Old-Fashioned Sexism Scale (Swim et al. 1995) sowie mit subtil sexistischen Einstellungen der Modern Sexism Scale (Swim et al. 1995). Die benevolente Sexismus-Skala (BS), die ein positives Bild von Frauen zeichnet, ist hingegen mit diesen Maßen nur schwach oder (im Falle der Partialkorrelationen) überhaupt nicht korreliert (vgl. Glick & Fiske, 1996, S. 502 ff.; Swim, Becker, Lee, & Pruitt, 2010, S. 152). Abrams, Viki, Masser und Bohner (2003) fanden allerdings deutliche signifikante positive Zusammenhänge zwischen wohlwollendem Sexismus und „victim blaming“. So neigten Personen, die hohe Werte beim benevolenten Sexismus aufweisen, eher dazu, weiblichen Vergewaltigungsopfern die Schuld zuzuweisen als Personen, die niedrige Werte beim benevolenten Sexismus besitzen. Gerger, Kley, Bohner und Siebler (2007) konnten in ihren Untersuchungen zudem signifikante positive Zusammenhänge zwischen benevolentem Sexismus und der Akzeptanz moderner Mythen über sexuelle Aggression (AMMSA) aufzeigen.

In verschiedenen Studien wurde die Vorhersagevalidität der HS- und der BS-Subskala bestätigt. Während Hostiler Sexismus negative Einstellungen in Bezug auf nicht-traditionelle Frauen voraussagt, konnte sich benevolenter Sexismus als zuverlässiger Prädiktor für positive Einstellungen in Bezug auf traditionelle Frauen erweisen (vgl. Glick & Fiske, 2001a, S. 113).

---

<sup>16</sup> Zur methodischen Vorgehensweise ist anzumerken, dass die Items der ASS von den AutorInnen eigenständig aus dem Englischen übersetzt und von einem „sachkundigen native speaker“ durchgesehen wurden. Eine Überprüfung des intendierten Itemverständnisses mit Hilfe von kognitiven Techniken wurde nicht berichtet (vgl. Eckes & Six-Materna, 1999).

Konfirmatorische Faktorenanalysen ergaben, dass ein hierarchisches Modell mit zwei Faktoren zweiter Ordnung (HS und BS) und drei BS-Subfaktoren erster Ordnung (Pater­nalismus, Geschlechterdifferenzierung und Heterosexualität) die beste Anpassung an die Daten erzielt. (Im Hinblick auf den hostile Sexismus ließ sich somit die theoretisch postulierte Dreigliedrigkeit empirisch nicht nachweisen.) Beide Faktoren (HS und BS) weisen durchgängig signifikante positive Korrelationen auf ( $.40 < r < .50$ ) (vgl. Eckes & Six-Materna, 1999, S. 221 f.; Glick & Fiske, 1996, S. 497 ff.).

Auch in kulturvergleichenden Untersuchungen konnte die Validität des Ambivalent Sexism Inventory bestätigt werden. Es zeigte sich, dass hostiler und benevolenter Sexismus ein kohärentes Set von Ideologien über verschiedene Kulturen hinweg darstellen. In insgesamt 19 verschiedenen Ländern konnte die oben beschriebene Faktorenstruktur repliziert werden. Während Männer in allen Nationen bezüglich des hostile Sexismus höhere Werte und somit sexistischere Einstellungen aufweisen, fallen die Geschlechterdifferenzen hinsichtlich des benevolent Sexismus nicht so eindeutig aus. In einigen Ländern besitzen Frauen sogar höhere Mittelwerte (vgl. Glick et al., 2000). Eine Ursache für dieses Ergebnis könnte sein, dass der benevolente Sexismus eine gewisse Verlockung für Frauen darstellt, der sie nicht widerstehen können: Männer erklären sich bereit, ihr eigenes Wohlergehen zu opfern, um Frauen zu versorgen und zu beschützen. Ebenso könnte es auch eine Art Selbstschutz für Frauen in Kulturen darstellen, in denen viele Männer hostile sexistisch eingestellt sind. Die Ironie dabei ist, dass Frauen dazu gezwungen werden gerade bei der Gruppe von Menschen Schutz zu suchen, von der die Bedrohung ausgeht und je größer die Bedrohung ist, desto größer ist auch die Versuchung, die beschützende Ideologie des benevolent Sexismus zu akzeptieren. Für diese Vermutung spricht die Beobachtung, dass gerade in den Gesellschaften, in denen die Prävalenz von Sexismus am stärksten ausgeprägt ist, Frauen höhere Werte hinsichtlich des benevolent Sexismus aufweisen als Männer (vgl. Glick et al., 2000; Glick & Fiske, 2001a, S. 109 ff.). Auf die Validität der Skala deuten zudem Korrelationen zwischen dem ambivalenten Sexismus und nationalen Indikatoren für Ungleichheit hin (vgl. Glick & Fiske, 2001a, S. 114).

Da die meisten Untersuchungen zur Validierung der Skala an Studierendenstichproben durchgeführt wurden, bedarf es nach Ansicht von Swim et al. (2010) weiterer Studien mit repräsentativen Stichproben (vgl. Swim et al., 2010, S. 152).

#### **2.4.2.4 Zusammenfassung**

Bisher werden in der einschlägigen Literatur drei Arten von Sexismus unterschieden: Der traditionelle, der moderne und der ambivalente Sexismus. In Deutschland etablierten sich vor allem sieben Skalen, die die entsprechenden Einstellungen erfassen sollen. Der traditionelle Sexismus wird durch die „Misogynie-Skalen“ von Krameyer und Schmidt (1973) abgebildet. Zur Erfassung von Geschlechterrolleneinstellungen sind drei Skalen erstellt worden: Die „Skala zur Messung der normativen Geschlechtsrollen-Orientierung“ von

Krampen (1979), der „Fragebogen zur Messung normativer Geschlechtsrollenorientierung“ von Athenstaedt (2000) und die „Geschlechterrollen-Batterie des ISSP“ aus dem Jahr 2002. Eckes und Six-Materna (1998) generierten mit der „Skala zur Erfassung des modernen Sexismus“ ein Messinstrument für den modernen Sexismus. Ambivalent sexistische Einstellungen werden mit Hilfe der „Misogynie-Ritterlichkeits-Skala“ von Schmerl und Steinbach (1973) sowie der „Skala zur Erfassung des ambivalenten Sexismus“ von Eckes und Six-Materna (1999) erfasst.

Hinsichtlich der Schwerpunktlegung lässt sich sagen, dass sich keine der angeführten Skalen ausschließlich auf sexistische Einstellungen in Bezug auf Mütter konzentriert. Sowohl die Modern Sexism Scale als auch die Neosexism Scale beziehen sich in ihren Items – abgesehen von einer Ausnahme – auf die Globalkategorie der Frauen. Lediglich in einem Item der Neosexism Scale werden Frauen in ihrer zukünftigen Rolle als Mutter und Hausfrau erwähnt (Item: „Universities are wrong to admit women in costly programs such as medicine, when in fact, a large number will leave their jobs after a few years to raise their children.“). In der deutschsprachigen Version bleibt die Subgruppe der Mütter vollständig unberücksichtigt. Gleiches gilt auch für die „Skala zur Erfassung des ambivalenten Sexismus“ von Eckes und Six-Materna (1999): Mütter bleiben hier unerwähnt. In der „Misogynie-Ritterlichkeits-Skala“ von Schmerl und Steinbach (1973) sprechen, wie bereits bemerkt, hingegen drei von 30 Items Frauen in ihrer Rolle als Mutter auf eine subtile, scheinbar positive Art und Weise an (Bsp.: „Als Familienmutter und Hüterin des heimischen Herdes kann eine Frau ihre wertvollsten Eigenschaften entfalten.“, übrige Items siehe Unterkapitel 2.4.2.1).

Die Skalen zur Erfassung des traditionellen Sexismus bzw. der Geschlechterrollen-Einstellungen beziehen sich – aufgrund ihrer Definition – noch am stärksten auf die Gruppe der Mütter. Dies gilt bei den genannten deutschsprachigen Messinstrumenten insbesondere für die Subskala „Misogynie durch spezielle Zuweisung von Aufgaben und Pflichten“ der Misogynie-Skalen von Krameyer und Schmidt (1973) mit 6 von 70 Items (Bsp.: „Sobald eine Frau Kinder hat, sollte sie ihre ganze Zeit und Kraft der Welt des Hauses widmen.“) sowie für die Geschlechterrollen-Batterie des ISSP (2002) mit 5 von 10 Items (Bsp. 1: „Ein Kind, das noch nicht zur Schule geht, wird wahrscheinlich darunter leiden, wenn seine Mutter berufstätig ist“; Bsp. 2: „Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern.“). Die Verantwortung der Mutter für die Kinderbetreuung, die die traditionelle Aufteilung der Geschlechterrollen festigt, wird in vielen Items in präskriptiver Weise betont: Frauen *sollten* sich um die Betreuung der Kinder kümmern. Eine positive Begründung, die die stereotypen, weiblichen Eigenschaften herausstellt und daraus die besondere Eignung der Frauen für die Mutterrolle unterstreicht, wird in keiner der beiden traditionellen Sexismus-Skalen vorgenommen.

Dies ist insofern von Bedeutung, als dass Items mit scheinbar positiven Zuschreibungen heutzutage subtile Vorurteile eher erfassen können, da ihnen der negative Anschein

genommen wird. Auch wenn das Resultat letztendlich gleich ist und Frauen von der öffentlichen Sphäre ausgeschlossen werden und die traditionelle Aufteilung der Arbeitsteilung zementiert wird, wirken solche Aussagen nicht negativ, da Frauen in ein positives Licht gerückt werden und ihre besondere Befähigung betont wird. Folglich erscheinen solche Items, die eine wohlwollende Rechtfertigung für eine Diskriminierung von Frauen vornehmen, auch eher sozial akzeptabel und Befragte scheuen sich weniger davor, offen ihre Einstellung zu äußern.

Mit der Entwicklung und Validierung einer Skala, die feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter (kurz FWSEM-Skala) erfasst, verfolgt diese Arbeit (wie bereits in Unterkapitel 1.2 dargelegt) zwei zentrale Anliegen:

- (1) *Inhaltliches Anliegen.* Das Hauptanliegen dieser Arbeit ist es eine Skala zur Erfassung sexistischer Einstellungen in Bezug auf Mütter zu entwickeln, da diese Subgruppe von Frauen in den bisherigen Messinstrumenten komplett unerwähnt blieb (MSS, ASS, NGRO) bzw. eine nachrangige Rolle spielte (MC, GRO, MS) bzw. nicht ausschließlich berücksichtigt wurde (GRB). Mit Hilfe der neuen Skala soll ein zeitgemäßes Messinstrument generiert werden, das gut zwischen den Einstellungen der Befragten differenzieren kann. Um dies zu erreichen, werden neben negativen, feindseligen Einstellungen auch scheinbar positive, wohlwollende Einstellungen berücksichtigt, die einen subtilen Charakter besitzen, nicht sofort als „sexistisch“ erkannt und daher auch eher in Umfragen offen geäußert werden. Als theoretische Grundlage für die Skalenentwicklung wird daher die Theorie des Ambivalenten Sexismus herangezogen, die zwischen hostilem und benevolentem Einstellungen differenziert und in 19 verschiedenen Ländern bestätigt werden konnte.
- (2) *Methodisches Anliegen.* Um neben den psychometrischen Eigenschaften (Analyse von Pretest-Daten) auch das Frageverständnis der einzelnen Items zu überprüfen, wird die Skala bereits in ihrer Entstehungsphase mit Hilfe kognitiver Techniken getestet. Fast alle der genannten Skalen wurden auf ihre Reliabilität sowie ihre konvergente und/oder diskriminante Validität hin untersucht, einem kognitiven Pretest wurden die einzelnen Items jedoch nicht unterzogen. (Einzige Ausnahmen bildet die Geschlechterrollen-Batterie des ISSP. Sie wurde im Jahre 2001 bei ihrer Neukonstruktion mit Hilfe kognitiver Techniken überprüft). Diese Tatsache ist umso schwerwiegender, da es in einer standardisierten Befragung weder den Befragten noch den InterviewerInnen möglich ist, sicherzustellen, dass die Itemformulierungen in der intendierten Weise aufgefasst werden (vgl. Schober, 1999, S. 83 f.; siehe das nächste Unterkapitel 2.5). Folglich sollten bereits bei der Fragebogenkonstruktion (neben anderen Pretest-Verfahren) auch kognitive Techniken eingesetzt werden, mit deren Hilfe das Frageverständnis der Befragten erfasst und mögliche Formulierungsdefizite aufgedeckt werden können. Clark und Schober

(1992) betonen, warum dies so wichtig ist: "Surveyors cannot possibly write perfect questions, self-evident to a respondent, that never need clarification. And because they cannot, the answers will often be surprising" (S. 29).

Dem methodischen Anliegen dieser Arbeit entsprechend werden nun im Folgenden die theoretischen Grundlagen zur Entwicklung und Testung von Messinstrumenten geliefert. Dabei wird zunächst auf die kognitiven und kommunikativen Prozesse in der Umfrageforschung eingegangen und erläutert, weshalb es zu Verständnisproblemen im Rahmen von standardisierten Befragungen kommen kann (Unterkapitel 2.5). Anschließend werden verschiedene Pretest-Verfahren zur Evaluierung einzelner Items bzw. gesamter Fragebögen vorgestellt (Unterkapitel 2.6).

## 2.5 Kognitive und kommunikative Prozesse in Befragungssituationen

### 2.5.1 Frage-Antwort-Prozess in standardisierten Befragungen

Um eine Frage im Rahmen einer standardisierten Umfrage beantworten zu können, müssen die TeilnehmerInnen verschiedene kognitive Aufgaben bewältigen. Die meisten theoretischen Modelle (siehe u.a. Tourangeau, 1984) nennen in diesem Zusammenhang die folgenden vier Prozesse:

- (1) Verstehen der Frage,
- (2) Abrufen fragerelevanter Informationen aus dem Gedächtnis,
- (3) Bildung eines Urteils und
- (4) Mitteilen der Antwort im vorgegebene Format (vgl. Groves et al., 2004, S. 202; Schaeffer & Presser, 2003, 67).

Sudman, Bradburn und Schwarz (1996) ergänzen die vierte Aufgabe um eine weitere und weisen darauf hin, dass die Antwort möglicherweise hinsichtlich sozialer Erwünschtheit editiert wird, bevor sie den InterviewerInnen mitgeteilt wird.

Der Frage-Antwort-Prozess wird im Folgenden gemäß der Einteilung von Sudman, Bradburn und Schwarz (1996) näher beschrieben (siehe Abbildung 3). Dabei wird die eben vorgestellte Reihenfolge beibehalten, da dies die Ausführungen vereinfacht. In einer realen Umfragesituation springen die TeilnehmerInnen jedoch häufig zwischen den einzelnen Aufgaben vor und zurück, mitunter lassen sie davon auch einige komplett aus oder führen sie nur nachlässig aus (vgl. Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2004, S. 203; Sudman, Bradburn, & Schwarz, 1996, S. 56).



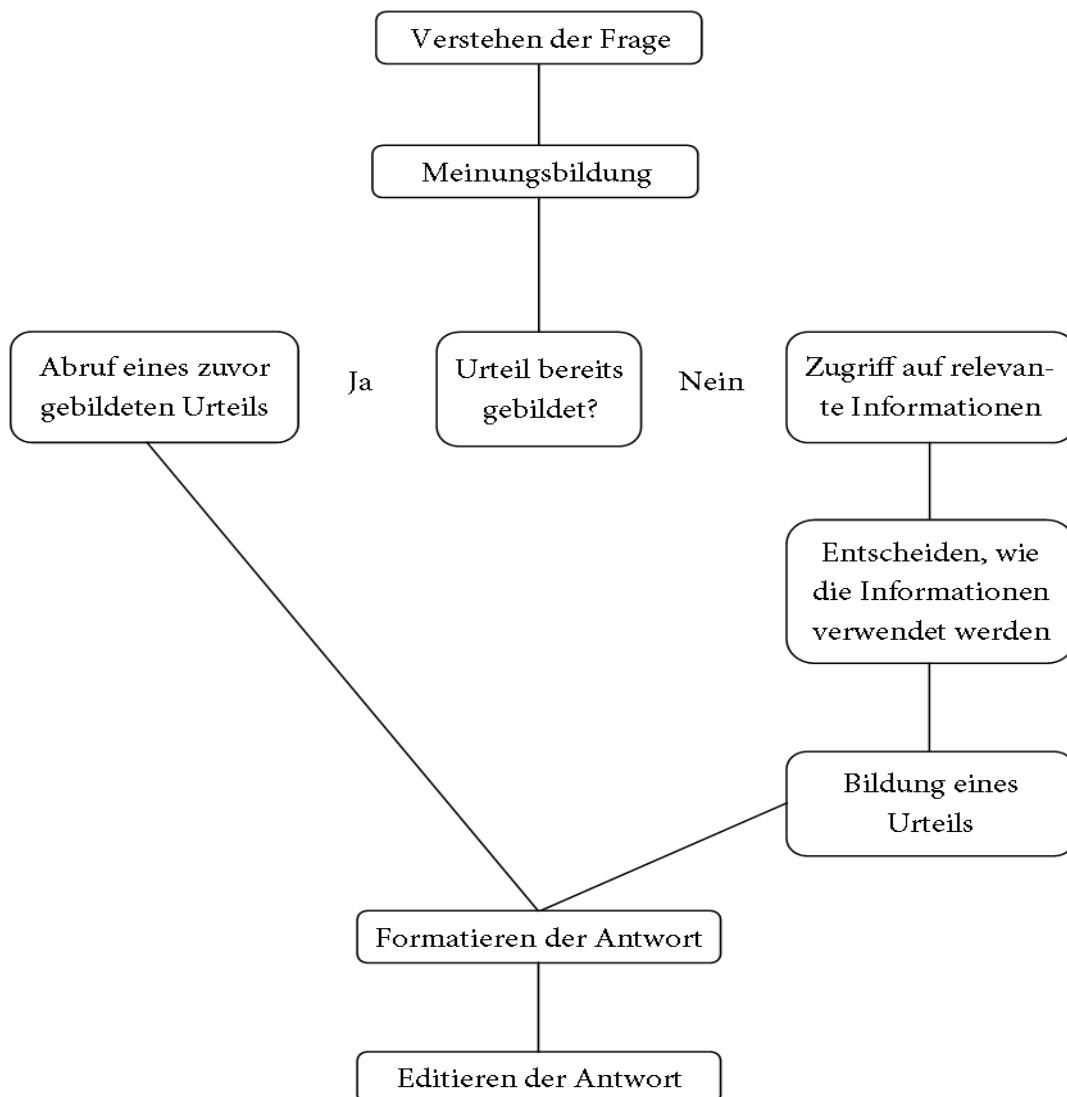


Abbildung 3. Modell des Frage-Antwort-Prozesses in einer standardisierten Umfragesituation. Quelle: Sudman, Bradburn, & Schwarz, 1996, S. 58. Eigene Übersetzung.

### 2.5.1.1 Verstehen der Frage

Sudman, Bradburn und Schwarz (1996) unterscheiden im Zusammenhang mit dem Frageverständnis zwei miteinander verknüpfte Prozesse: das semantische und das pragmatische Verständnis.

#### *Das semantische Verständnis*

Die erste Aufgabe der Befragten besteht darin, die *wörtliche Bedeutung* der Frage zu erschließen. Hierfür werden zunächst die Laute einer Äußerung bzw. die Symbole der

Schriftsprache in einzelne Wörter übertragen. Anschließend wird die Abfolge der einzelnen Wörter in mögliche syntaktische Einheiten untergliedert und die lexikalische Bedeutung aus dem semantischen Gedächtnis abgerufen.

Häufig hängen Verständnisprobleme im Rahmen von standardisierten Befragungen (aber auch in Alltagskonversationen) mit der lexikalischen oder strukturellen Ambiguität der Frage zusammen. Eine *strukturelle* Ambiguität liegt beispielsweise bei folgendem Satz vor: „Emily kaufte das Buch ihrer Schwester.“ In diesem Fall ist nicht klar, ob Emily ein Buch *für* ihre Schwester kaufte oder sie *ihr* das Buch beispielsweise auf einem Flohmarkt *abkaufte*. Der Satz „Emily ging zur Bank“ lässt ebenfalls zwei unterschiedliche Interpretationen zu. In diesem Fall handelt es sich um eine *lexikalische* Ambiguität vor, da das Wort „Bank“ sowohl für eine *Sitzgelegenheit*, als auch für das *Geldinstitut* stehen kann (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 59 f.).

Während strukturelle Ambiguität zumeist durch eine umsichtige Anordnung der einzelnen Satzglieder vermieden werden kann, ist lexikalische Ambiguität ein Bestandteil der Sprache und spiegelt die unterschiedlichen Auslegungsmöglichkeiten einiger Wörter wider. In den meisten Fällen wird der Interpretationsspielraum jedoch durch den Kontext, in dem die Wörter verwendet werden, eingegrenzt. So liegt es nahe, dass die Sitzgelegenheit gemeint ist, wenn der oben angeführte Satz folgendermaßen fortgeführt wird: „Emily ging zur Bank im Park und ließ sich erschöpft darauf nieder.“

Lässt der Kontext unterschiedliche Interpretationen zu, so wird zumeist auf die Auslegung des Wortes zurückgegriffen, die der befragten Person spontan in den Sinn kommt und diese Interpretation wird so lange beibehalten, bis sie sich möglicherweise als falsch erweist. Welche der verschiedenen Bedeutungen eines Wortes am leichtesten zugänglich sind, hängt wiederum von verschiedenen Variablen des Fragebogens (z.B. den vorhergehenden Fragen) und der befragten Person (z.B. dem persönlichen Bezug) ab (siehe Unterkapitel 2.5.1.2). Je höher beispielsweise der persönliche Bezug im Hinblick auf eine Interpretationsmöglichkeit ist bzw. je häufiger diese im alltäglichen Leben verwendet wird, desto eher ist sie auch salient (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 60 f.). So wird eine Befragte, die bei einer Bank angestellt ist, bei dem Wort „Blüte“ wahrscheinlich zuerst an eine falsche Geldnote denken, wohingegen einer Gärtnerin vermutlich spontan das Blatt einer Blume in den Sinn kommt. Für beide Personen liegt ihre Auslegung des Begriffes persönlich näher.

Dieses Beispiel verdeutlicht, wie wichtig es gerade im Bereich der Umfrageforschung ist, eindeutige Formulierungen zu verwenden. Wie in Unterkapitel 2.5.2 noch aufgezeigt wird, können die Befragten nämlich nicht, so wie dies bei einer normalen Konversation im Alltag der Fall ist, nach der Bedeutung eines Wortes fragen. Sollten sie dies dennoch tun, so dürfen sie keine zufriedenstellende Antwort erwarten. Die InterviewerInnen werden ihnen lediglich mitteilen, dass sie ihnen keine Hilfestellung diesbezüglich geben können und die Befragten sich die Interpretation selbst erschließen müssen. Dementsprechend weiß man in der Regel nicht, wie die UmfrageteilnehmerInnen einzelne Begriffe

oder auch die gesamte Formulierung einer Frage aufgefasst haben (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 61).

Mit dem semantischen Verständnis einer Frage ist folglich das *wörtliche Verständnis* gemeint. Dieses erfordert die Identifikation der einzelnen Begriffe, den Abruf der wörtlichen Bedeutung aus dem semantischen Gedächtnis sowie die Erschließung der Bedeutung der gesamten Frage. Dabei kommen zum einen der Fragekontext und zum anderen die Personenvariablen zum Tragen (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 62).

#### *Das pragmatische Verständnis*

Die reine Kenntnis der wörtlichen Bedeutung einer Frage genügt jedoch in der Regel nicht, um ein Item angemessen beantworten zu können. Damit die Befragten auch tatsächlich die gewünschten Informationen preisgeben, müssen sie herausfinden, worin der Kerngehalt der Frage besteht, weshalb die Frage gestellt wurde und was eine passende Antwort wäre (vgl. Schaeffer & Presser, 2003, S. 67). Das heißt sie müssen – als kooperative TeilnehmerInnen einer Umfrage – erschließen, worin die *intendierte Bedeutung* der Frage liegt. Hierfür greifen sie, obwohl sie sich eigentlich in einer standardisierten Umfragesituation befinden, zumeist unbewusst auf die impliziten Regeln der Alltagskonversation zurück (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 62 ff.).

Paul Grice (1975) hat diese Regeln als die vier „Maximen der Konversationslogik“ bezeichnet. Er nimmt an, dass die einzelnen GesprächsteilnehmerInnen sich daran halten und eine Konversation somit nach den folgenden kooperativen Prinzipien abläuft:

- (1) *Maxime der Quantität*: Es sollen nur die Informationen gegeben werden, auf die die FragestellerInnen angewiesen sind und keine darüber hinaus.
- (2) *Maxime der Qualität*: Es soll nichts gesagt werden, was als falsch erachtet wird oder wofür es an der nötigen Evidenz mangelt.
- (3) *Maxime der Relevanz*: Die einzelnen Redebeiträge sollen für das Ziel der Konversation bedeutsam sein.
- (4) *Maxime des Stils*: Die Gesprächsbeiträge sollen leicht zu verstehen sein: Mehrdeutigkeiten, komplizierte Formulierungen und unnötige Ausschweifungen gilt es zu vermeiden.

Grice zufolge sollen die Äußerungen der KonversationsteilnehmerInnen demzufolge *informativ, ehrlich, relevant* und *klar* sein (vgl. Grice, 1975, S. 45 ff.; Sudman, Bradburn, & Schwarz, 1996, S. 62 f.).

Übertragen die Befragten nun diese Maximen der Konversationslogik auf die Umfragesituation, gelangen sie zu den nachstehenden Annahmen:

- (a) Der Maxime der Relevanz folgend gehen die Befragten davon aus, dass die ForscherInnen bzw. InterviewerInnen keine sinnlosen Fragen stellen. Selbst wenn also Fragen vorkommen, deren Bedeutung sich ihnen nicht unmittelbar erschließt, so werden sie – als kooperative UmfrageteilnehmerInnen – versuchen,

die Intention dahinter zu erfassen. Als mögliche Hinweise ziehen sie verschiedene Kontextmerkmale der Frage, wie etwa den Einleitungstext, vorhergehende Items (oder in schriftlichen Befragungen auch nachfolgende Items) sowie die Abfolge und die Merkmale der vorgegebenen Antwortkategorien heran (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 63 f.).

- (b) Der Maxime der Quantität folgend werden die Befragten wiederum, nachdem sie sich eine Vorstellung darüber gebildet haben, was man von ihnen wissen möchte, nicht beliebig alles äußern, was ihnen zu dem Thema einfällt, sondern nur, was anscheinend von Belang ist. Auch bereits Gesagtes werden sie in ihrer Antwort nicht wiederholen. So werden sie sich überlegen, was die InterviewerInnen bereits wissen und welche zusätzlichen „neuen“ Informationen sie ihnen liefern können. Zu diesem Zweck werden die befragten Personen die Formulierung der Frage selbst berücksichtigen sowie den Kontext, in dem sie gestellt wurde. Ferner werden sie die vorgegebenen Antwortkategorien als Informationsquelle nutzen und ihre eigenen, vorangegangenen Äußerungen bedenken (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 63 f.).

Um das Vorgehen der Befragten zu verdeutlichen, stellen Sie sich vor, dass die folgende offene Frage im Rahmen einer standardisierten Befragung an Sie gerichtet wird: „Was haben Sie heute gemacht?“ Um eine Antwort liefern zu können, die den oben angeführten Maximen gerecht wird, müssen Sie zunächst bestimmen, worin die Intention dieser Frage liegt und welche Tätigkeiten die ForscherInnen interessieren könnten. Unerwähnt lassen Sie vermutlich solche Handlungen, die den InterviewerInnen bereits bekannt sind (wie beispielsweise Ihre Teilnahme an einer Befragung) oder die als selbstverständlich erachtet werden (wie etwa das morgendliche Aufstehen und Zähneputzen) – schließlich soll Ihre gegebene Antwort so informativ wie möglich sein (Maxime der Quantität) (vgl. ebd., S. 64).

Anders sähe dies hingegen aus, wenn die gleiche Frage in einem geschlossenen Format gestellt würde und die vorgegebenen Antwortkategorien genau diese Tätigkeiten enthielten („Ich habe an einer Befragung teilgenommen“ und „Ich habe mir die Zähne geputzt“). In diesem Falle nähmen Sie vermutlich an, dass die aufgelisteten Aktivitäten die intendierte Bedeutung der Frage und somit das Interesse der ForscherInnen widerspiegeln. Folglich kreuzten Sie die entsprechenden Kategorien auch an. Selbst bei einer halb-offenen Frage, die eigene Ergänzungen zulässt, würden Sie die Möglichkeit einer zusätzlichen Angabe vermutlich nur selten nutzen, da die vorgegebenen Antwortalternative scheinbar bereits die primären Interessen aufgreifen (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 64).

### 2.5.1.2 Bildung oder Abruf eines Urteils

#### *Abrufen fragerelevanter Informationen und Bildung eines Urteils*

Haben die Befragten die intendierte Bedeutung der Frage erschlossen, müssen sie in einem nächsten Schritt die relevanten Informationen zusammentragen, um die Frage beantworten zu können. In einigen Fällen, die eher die Ausnahme als die Regel darstellen, werden sie die Frage unmittelbar beantworten können, da sie direkt auf eine passende Antwort zurückgreifen können (siehe nächster Abschnitt). Zumeist trifft dies jedoch nicht zu. Damit sich die Befragten ein Urteil bilden können, müssen sie die fragerelevanten Informationen erst einmal aus dem Langzeitgedächtnis abrufen, welches neben autobiografischen Informationen auch generelles Wissen enthält (vgl. Groves et al., 2004, S. 205; Sudman, Bradburn, & Schwarz, 1996, S. 69).

Sollen sie beispielsweise bei einer Verhaltensfrage eine Aussage darüber treffen, wie häufig sie im letzten Monat eine Scheibe Brot gegessen oder Radio gehört haben, dann werden sie die Antwort vermutlich nicht sofort wissen, da dies alltägliche Handlungen sind, die sie häufig und zudem unregelmäßig ausführen. Selbst wenn ihnen mehrere konkrete Beispiele einfallen sollten, müssen die Befragten überprüfen, ob das erinnerte Verhalten auch in die angegebene Periode fällt. Um zu ermitteln, wie häufig das entsprechende Verhalten in dieser Zeit gezeigt wurde, nehmen die Befragten (je nach Motivation, verfügbarer Zeit und Erinnerungsvermögen) entweder eine genaue Zählung oder eine grobe Schätzung vor.

Ähnlich sieht dies im Falle von Einstellungsfragen aus. Selbst wenn die befragte Person bereits über das Thema nachgedacht haben sollte, spricht die Frage möglicherweise einen ganz speziellen Aspekt an, zu dem noch keine eindeutige Meinung besteht. Folglich ist es erforderlich, dass sich die Befragten zuerst verschiedene Informationen ins Gedächtnis rufen, um sich ein Urteil bilden zu können (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 70).

Suchen die Befragten ihr Gedächtnis ab, um ein Urteil fällen zu können, werden sie im Normalfall nicht alle erdenklichen Informationen heranziehen, die unter Umständen wichtig sein könnten. Stattdessen wird der Suchprozess beendet, sobald genügend Informationen gesammelt wurden, auf deren Basis ein Urteil gefällt werden kann. Dies hängt mit den speziellen Rahmenbedingungen zusammen, unter denen die meisten Befragungen stattfinden. Zunächst einmal wäre eine intensive Informationssuche nur dann wahrscheinlich, wenn eine hohe Motivation der Befragten vorliegt. Dies ist in der Regel dann der Fall, wenn die gegebene Antwort persönliche Konsequenzen für die TeilnehmerInnen hat. Auf die Umfrageforschung trifft dies allerdings äußerst selten zu. Des Weiteren besteht in den meisten standardisierten Befragungen Zeitdruck, was einer ausführlichen Informationssuche zuwiderläuft und die befragten Personen dazu veranlasst, den Suchprozess so schnell wie möglich zu Ende zu bringen.

In der Regel beruhen die Urteile, die in Umfragen gefällt werden, folglich in erster Linie auf den Informationen, die den Befragten direkt in den Sinn kommen („saliente Informationen“). Andere potentiell ebenso relevante Aspekte, die schwerer zugänglich

sind, bleiben zumeist unberücksichtigt (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 71). Ausschlaggebend dafür, welche Informationen schnell ins Gedächtnis gerufen werden können, sind nach Sudman, Bradburn und Schwarz (1996) mitunter folgende Faktoren:

- (a) Die *Zeit*, die *seit der letzten Nutzung* verstrichen ist: Ist erst wenig Zeit vergangen, seitdem die Informationen zuletzt abgerufen wurden (beispielsweise im Zusammenhang mit einer vorhergehenden Frage, die eine ähnliche Thematik behandelt), sind sie schnell wieder im Gedächtnis parat. Man bezeichnet dies als *temporär verfügbare Informationen*; sie werden maßgeblich von den Variablen des Fragebogens bestimmt und tragen zur Instabilität der Urteile zu verschiedenen Zeitpunkten bei, da sie je nach (Frage-)Kontext unterschiedlich ausfallen können.
- (b) Die *Häufigkeit*, mit der die *Informationen genutzt* werden: Werden Informationen sehr häufig verwendet, kommen sie Befragten auch eher in den Sinn. Mann spricht in diesem Fall von *chronisch verfügbaren Informationen*; sie hängen v.a. von personenbezogenen Variablen ab und führen selbst zu verschiedenen Zeitpunkten und auch unter verschiedenen Kontextbedingungen zu stabilen Urteilen. Eine schwangere Frau wird sich beispielsweise sehr häufig mit Themen wie der „Geburt“, dem „Stillen“, der „Säuglingspflege“ oder ähnlichem befassen, so dass die Informationen dazu ständig im Gedächtnis verfügbar sind. Folglich wird sie auch spontan daran denken, wenn ihr etwa eine Frage zum Thema „Kinder“ gestellt wird. Anderen Befragten kommen diese Aspekte vermutlich nur dann in den Sinn, wenn sie explizit in den vorhergehenden Fragen angesprochen wurden.
- (c) Die *Integration der Informationen in übergeordnete Wissensstrukturen*: Je stärker die Informationen in höhere Wissensstrukturen integriert sind, umso leichter sind sie auch verfügbar. Das menschliche Gedächtnis wird in der Kognitionspsychologie häufig als ein großes Netzwerk beschrieben. Darin bilden die persönlichen Informationen einzelne „Knotenpunkte“, welche wiederum durch diverse Pfade miteinander verknüpft sind. Je mehr Pfade zu einem Knotenpunkt führen, desto leichter kann man die dazugehörige Information abrufen (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 71 f.). So hat eine Befragte, die sich Kinder wünscht, jedoch keine bekommen kann, dieses Problem in Bezug auf viele verschiedene Aspekte ihres Privatlebens betrachtet – entsprechend viele Pfade wurden gebildet. Folglich taucht das Thema ungewollte Kinderlosigkeit sofort in ihrem Gedächtnis auf, sobald ein Aspekt angesprochen wird, der damit verbunden ist.

#### *Abruf eines Urteils*

Wie bereits angedeutet, können die Befragten in seltenen Fällen auch unmittelbar auf eine Frage antworten. Ob auf die passende Antwort unmittelbar zugegriffen werden kann, hängt laut Sudman, Bradburn und Schwarz (1996) von unterschiedlichen Faktoren ab, die sich wiederum je nach Frageart unterscheiden.

- (a) Für *Einstellungsfragen* gilt: Je größer die thematische *Bedeutung für das einzelne Individuum* ist, je höher der Grad an *persönlicher Erfahrung* in Bezug auf das Einstellungsobjekt ist und je weniger Zeit seit der letzten Beurteilung des Einstellungsobjektes zurückliegt, desto schneller ist die passende Antwort parat.
- (b) Für *Verhaltensfragen* gilt: Je größer die *persönliche Bedeutung* des Verhaltens und je seltener die *Auftretenshäufigkeit* ist, desto schneller kann die passende Antwort aus dem Gedächtnis abgerufen werden. Bei einem Verhalten, das sehr häufig gezeigt wird, ist ein schnelles Urteil nur möglich, wenn es gleichzeitig sehr *regelmäßig* auftritt (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 69 f.). Werden die Befragten beispielsweise gefragt, in welchem Jahr sie ihr Studium beendet haben oder ihr erstes Kind geboren wurde, so sind dies einmalige Ereignisse, die eine große persönliche Bedeutung für sie besitzen. Dementsprechend ist die Wahrscheinlichkeit auch sehr hoch, dass die passende Antwort sofort verfügbar ist.

Zusammenfassend lässt sich festhalten: Die meisten Antworten, die im Rahmen von standardisierten Befragungen gegeben werden, repräsentieren Urteile, die von den Befragten direkt an Ort und Stelle gebildet werden und die folglich stark unter dem Einfluss der jeweiligen Kontextbedingungen des Interviews stehen (vgl. Sudman, Bradburn, & Schwarz, 1996, S. 70).

### 2.5.1.3 Formatierung der Antwort

Nachdem die UmfrageteilnehmerInnen ein Urteil gefällt haben, muss die entsprechende Antwort den InterviewerInnen mitgeteilt werden. Handelt es sich nicht um eine offene Frage, die eine Formulierung in eigenen Worten erlaubt, sondern eine geschlossene Frage, so ist es erforderlich, die Antwort an die vorgegebenen Kategorien anzupassen.

Wie bereits im Zusammenhang mit dem pragmatischen Verständnis erläutert, übt die Auswahl der Antwortvorgaben einen maßgeblichen Einfluss auf das Frageverständnis aus – gleiches gilt auch für diese Stufe des Frage-Antwort-Prozesses. Da die Befragten den Maximen der Alltagskonversation folgend davon ausgehen, dass die Antwortkategorien nicht zufällig von den ForscherInnen ausgewählt wurden, ziehen sie verschiedene Rückschlüsse, die mit den Antwortoptionen zusammenhängen (vgl. Sudman, Bradburn & Schwarz, 1996, S. 73). Beinhalten die meisten Antwortoptionen eher niedrige Häufigkeitsangaben, so ist es wahrscheinlich, dass die Antworten der Befragten auch in diese Richtung verschoben werden (vgl. Groves et al., 2004, S. 207). Finden die Befragten beispielsweise keine geeignete Antwortvorgabe, die ihr Urteil treffend wiedergibt, so nehmen sie an, dass diese Angabe für die ForscherInnen nicht von Interesse ist und teilen sie daher i.d.R. auch nicht mit (vgl. Sudman, Bradburn & Schwarz, 1996, S. 73) (siehe Unterkapitel 2.5.1.1. Beispielfrage „Was haben Sie heute gemacht?“).

#### **2.5.1.4 Editieren der Antwort**

Bevor die Befragten ihre Antwort endgültig mitteilen, kommt es mitunter vor, dass sie sie im Hinblick auf die Konsistenz mit ihren zuvor gegebenen Antworten anpassen. Der Druck eine Konsistenz bezüglich berichteter Einstellungen oder auch Verhaltensweisen herzustellen, scheint allein schon durch den Prozess des Fragestellens hervorgerufen zu werden (vgl. Groves et al. 2004, S. 207; Tourangeau, 1984, S. 91). Ebenso ist es möglich, dass die Antworten hinsichtlich wahrgenommener sozialer Akzeptanz abgeändert werden. Ein sozial erwünschtes Antwortverhalten tritt insbesondere bei sensitiven Themen in face-to-face Befragungen auf, d.h. also in Situationen, in denen die TeilnehmerInnen den Fragebogen nicht selbst ausfüllen, sondern ihre Antwort einer anderen Person (dem/der InterviewerIn) mitteilen müssen. Bei Fragen, die eine potentielle Bedrohung des eigenen Ansehens darstellen können (z.B. bezüglich des Drogenkonsums oder des Sexualverhaltens) stehen Befragte vor einem Dilemma: Sie haben sich dazu bereit erklärt, an der Befragung teilzunehmen und Informationen preiszugeben, allerdings wollen sie ihr Gesicht wahren. Häufig tendieren die InterviewteilnehmerInnen dann dazu, eine Antwort zu verweigern oder sie geben vorsätzlich falsche Antworten. Da bei einer Frage nach dem Drogenkonsum eine Antwortverweigerung einem Eingeständnis gleichkäme, scheint es leichter, einfach zu verneinen, schon einmal einen Joint geraucht zu haben (vgl. Groves et al. 2004, S. 207 ff.).

#### **2.5.1.5 Zusammenfassung**

Im Rahmen des Frage-Antwort-Prozesses müssen die TeilnehmerInnen einer standardisierten Befragung eine Vielzahl verschiedener Aufgaben meistern. Grundvoraussetzung ist zunächst das (pragmatische) Verständnis. Um den ForscherInnen die gewünschten Informationen liefern zu können, ist es erforderlich, dass sich die Befragten den intendierten Sinn der Frage erschließen. Diese Aufgabe wird den befragten Personen jedoch – verglichen mit einer Alltagskonversation – deutlich erschwert, da die InterviewerInnen eine neutrale Haltung wahren sollen. (Auf die zentralen Unterschiede zwischen beiden Interaktionssituationen wird im folgenden Unterkapitel 2.5.2 näher eingegangen.) Um die intendierte Bedeutung der Frage zu erschließen, ziehen die Befragten (den Regeln der Alltagskonversation entsprechend) den Fragekontext heran. In einem nächsten Schritt müssen fragerrelevante Informationen aus dem Gedächtnis abgerufen und ein Urteil gebildet werden. Die meisten Antworten repräsentieren dabei Urteile, die von den Befragten unmittelbar während des Frage-Antwort-Prozesses generiert wurden und die dementsprechend stark von den jeweiligen Kontextbedingungen der Befragung beeinflusst werden. Zuletzt muss das eigene Urteil in die vorgegeben Antwortkategorien übertragen und den InterviewerInnen mitgeteilt werden. Dabei kann es mitunter vorkommen, dass die TeilnehmerInnen ihre Antwort zuvor hinsichtlich sozialer Erwünschtheit editieren oder sie sie so abändern, dass sie konsistent zu ihren bisherigen Auskünften ist.



## 2.5.2 Unterschiede zwischen einer Konversation im Alltag und in standardisierten Befragungen

### 2.5.2.1 Sinngebung bei Fragen in Alltagskonversationen

Genauso wie im Rahmen einer standardisierten Befragung setzt sich auch bei einer Alltagskonversation das Verständnis aus zwei miteinander verbundenen Prozessen zusammen: Dem semantischen und dem pragmatischen Verständnis. Clark und Schober (1992) zufolge nehmen es die GesprächspartnerInnen als selbstverständlich an, dass AdressatInnen herauszufinden versuchen, was die SprecherInnen mit ihrer Äußerung gemeint haben könnten (*principle of speaker's meaning*). Folglich reicht auch bei einer Alltagskonversation das reine Wortverständnis (das semantische Verständnis) nicht aus, sondern die Bedeutung, die die SprecherInnen dem Gesagten beigemessen haben (das pragmatische Verständnis), ist entscheidend (vgl. Clark & Schober, 1992, S.16). Damit die AdressatInnen den intendierten Sinn einer Äußerung erschließen können, greifen die KonversationsteilnehmerInnen laut Schober (1999) auf zwei interaktive Ressourcen zurück: auf das *Audience Design* und das *Grounding*.

#### *Audience Design*

Die SprecherInnen versuchen jede Äußerung so zu formulieren, dass ihre AdressatInnen herausfinden können, was sie gemeint haben (*principle of utterance design*, Clark & Schober, 1992). Damit dies gelingt, basieren ihre Äußerungen auf dem gemeinsamen Wissenshintergrund, dem *common ground*. Dieser schließt die Informationen ein, die beide GesprächsteilnehmerInnen teilen, ihr beidseitiges Wissen, ihre Meinungen (*beliefs*) sowie ihre geteilten Annahmen (vgl. Clark & Schober, 1992, S.17; Schober, 1999, S. 78). Das Gesagte wird folglich im Rahmen einer Alltagskonversation von den SprecherInnen explizit auf die AdressatInnen und deren Wissen zugeschnitten.

Der gemeinsame Wissenshintergrund der beiden GesprächsteilnehmerInnen lässt sich in zwei Arten unterteilen:

- (a) Das *kulturell geteilte Wissen* (*cultural common ground*): Gehören zwei Personen derselben kulturellen Gruppe an, so verfügen sie über ein gemeinsames Wissen, was diese Gemeinschaft angeht und welches typisch für diese ist. Die Zugehörigkeit zu einer Gruppe kann sich beispielsweise auf das Herkunftsland, die Sprache/den Dialekt, den Ausbildungsgrad, den Beruf oder auf Hobbys beziehen (vgl. Clark & Schober, 1992, S.18). Treffen sich beispielsweise Emily und Finn auf einer Party und stellen fest, dass sie beide Volleyball spielen, so können sie davon ausgehen, dass ihr Gegenüber über das spezielle gemeinsame Wissen derjenigen Personen verfügt, die diese Sportart ausüben.
- (b) Das *persönlich geteilte Wissen* (*personal common ground*): Die physikalische Umgebung, in der sich beide GesprächspartnerInnen befinden, kann als gemeinsames Wissen vorausgesetzt werden (z.B. Objekte innerhalb des Raumes, die von beiden

gesehen werden). Gleiches gilt auch in Bezug auf das Gesagte in aktuellen sowie in vorangegangenen Gesprächen oder gemeinsame Erlebnisse aus der Vergangenheit. Auch von ihnen kann angenommen werden, dass sie bekannt sind (vgl. Schober, 1999, S. 80).

### *Grounding*

Das persönlich geteilte Wissen erweitert sich im Verlauf einer Konversation durch die einzelnen Beiträge der TeilnehmerInnen (*principle of accumulation*). Damit das Gesagte zum Bestandteil des common ground werden kann, müssen die GesprächspartnerInnen jedoch zu verstehen geben, dass das Gesagte auch tatsächlich im intendierten Sinne verstanden wurde (vgl. Clark & Schober, 1992, S. 24). Dies geschieht im Rahmen eines interaktiven Prozesses, der als *Grounding* bezeichnet wird.

Um die intendierte Bedeutung einzelner Wörter oder auch der gesamten Äußerung, die Implikationen des Gesagten oder auch die Ziele, die hinter der Äußerung stehen, abzusichern, müssen SprecherInnen und AdressatInnen zusammenarbeiten (vgl. Clark & Schober, 1992, S. 25). Wenn Emily ihrem Freund Finn erzählt, was sie am vorigen Abend unternommen hat, muss er ihr verdeutlichen, dass er das Gesagte auch verstanden hat. Schließlich ist es durchaus möglich, dass er mit seinen Gedanken gerade abwesend war oder laute Nebengeräusche ihre Äußerung unhörbar werden ließen. Folglich muss er seiner Freundin einen Hinweis geben, dass der Konversationsbeitrag bei ihm auch tatsächlich angekommen ist, bevor sie sich dessen sicher sein kann. Hierfür stehen ihm verschiedene Möglichkeiten zur Verfügung: Er kann einfach mit dem Kopf nicken oder „Mhm“ sagen, um zu zeigen, dass er ihr zuhört oder die Konversation fortführen, indem er nachfragt: „In welchem Kinofilm warst Du denn?“ Vollständig abgeschlossen ist der Prozess des Groundings allerdings erst, wenn *beide* GesprächspartnerInnen zu erkennen gegeben haben, dass das Gesagte im intendierten Sinne aufgefasst wurde (vgl. Schober, 1999, S. 81). Folglich muss nun auch Emily im Gegenzug ähnliche Grounding-Techniken anwenden und beispielsweise fortführen, dass sie in dem neuen James Bond-Film war.

Auch dem Prozess des Groundings sind Grenzen gesetzt und er kann nicht immer ein korrektes beidseitiges Verständnis garantieren. So hängt es von den jeweiligen KonversationsteilnehmerInnen und deren Zielen ab, wie sehr ihnen daran gelegen ist, sich gegenseitig exakt zu verstehen: „For each contribution to discourse, the participants try to reach the mutual belief that the addressees have understood what the speaker meant to a criterion sufficient for current purposes“ (*principle of grounding*; Clark & Schober, 1992, S. 24). Trifft man beispielsweise auf einer Party einen flüchtigen Bekannten und unterhält sich kurz, mag es nicht so sehr von Bedeutung sein, jedes einzelne Wort im intendierten Sinne aufzufassen. Hat man hingegen ein technisches Problem mit dem Computer und ruft bei einer Service-Hotline an, besteht vermutlich ein sehr großes Interesse daran, die technischen Hinweise richtig zu begreifen. Unter diesen Umständen werden die zur Verfügung stehenden Grounding-Techniken vermutlich vollständig ausgeschöpft (vgl. Schober, 1999, S. 81).

Kehren wir noch einmal zu der Frage nach dem Tagesablauf aus Unterkapitel 2.5.1.1 zurück. Stellen Sie sich nun aber vor, dass Sie mit Ihrer Freundin telefonieren und Sie Ihnen die gleiche Frage stellt: „Was hast Du heute gemacht?“ Auf welche Art und Weise verleihen Sie dieser Frage einen Sinn?

Auch in Alltagskonversationen muss zunächst das semantische Verständnis der Frage gewährleistet sein. Der kontinuierliche Strom von Lauten muss in einzelne Wörter übertragen werden, hierzu sind phonologische Kenntnisse (Kenntnisse über die Organisation von Sprachlauten) nötig. Anschließend muss die lexikalische Bedeutung der einzelnen Wörter wie „du“, „heute“ und „gemacht“ aus dem semantischen Gedächtnis abgerufen werden. Die einzelnen Wörter werden wiederum in Satzglieder überführt (z.B. Subjekt, Prädikat, und Objekt) und so kombiniert, dass die gesamte Äußerung einen Sinn ergibt. Um zu erfassen, dass das Gesagte die syntaktische Form einer Frage besitzt, müssen Sie zudem über das erforderliche grammatikale Sprachwissen verfügen. Darüber hinaus sind Kenntnisse hinsichtlich der Abläufe sozialer Situationen erforderlich, wie beispielsweise die Konvention, dass auf eine gestellte Frage auch eine Antwort erwartet wird (vgl. Schober, 1999, S. 77).

Neben dem semantischen Verständnis gilt es, sich das pragmatische Verständnis der Frage zu erschließen. Um dies tun zu können, werden Sie den Kontext, in dem die Frage gestellt wurde, berücksichtigen. Dabei stellt der common ground zwischen SprecherInnen und AdressatInnen den Schlüssel dar, um die intendierte Bedeutung des Gesagten ermitteln zu können (vgl. Clark & Schober, 1992, S. 17).

Dass dieser Vorgang nicht immer leicht fällt, wurde bereits anhand verschiedener Beispiele aufgezeigt und gilt auch für diese Frage. So müssen Sie versuchen, sich die Intention zu erschließen, die hinter der Frage Ihrer Freundin steht. Was interessiert sie? Wie detailliert soll Ihre Schilderung sein? Auf welche Zeitspanne zielt die Frage ab? Ist auch die Zeit gemeint, die Sie im Büro verbracht haben oder nur die Zeit, seitdem Sie wieder daheim sind? Um die Intention der Frage zu ermitteln, werden Sie sich auf den gemeinsamen Wissenshintergrund beziehen und beispielsweise die Inhalte vorangegangener Gespräche (personal common ground) berücksichtigen.

Falls Sie sich unsicher sein sollten, worauf die Frage abzielt, werden Sie vermutlich direkt nachfragen („Interessiert es dich, ob es heute wieder viel im Büro zu tun gab?“ oder „Möchtest Du wissen, ob ich heute Sport gemacht habe?“), um das beidseitige Verständnis abzusichern. Ebenso ist es möglich, dass Sie überhaupt nicht darüber nachdenken, die Intention Ihnen deutlich erscheint und Sie die Frage folgendermaßen beantworten: „Nach dem Büro war ich noch schwimmen, dann habe ich mir etwas zu essen gekocht und danach noch ein Buch gelesen.“ Diese Antwort geben Sie in der Annahme, dass Ihre Freundin Ihnen einen Hinweis darauf gibt, ob Sie sie richtig verstanden haben oder eben nicht. Ist Ihre Freundin mit Ihrer Antwort zufrieden, da Sie die Intention ihrer Frage erfasst und die gewünschten Informationen geliefert haben, muss sie Ihnen dies zu verstehen geben, indem sie verschiedene Grounding-Techniken anwendet und beispielsweise

nickt, „Das hört sich gut an?!“ sagt oder Ihren Beitrag noch einmal aufgreift und die Konversation fortführt „Welches Buch liest Du denn gerade?“. Sollte sie hingegen nicht mit Ihrer Interpretation zufrieden sein, so muss sie Sie darauf hinweisen („Ich wollte eigentlich wissen, was Du heute zum Abendessen gekocht hast.“), denn nur dann kann das Gesagte Bestandteil des persönlich geteilten Wissens werden.

Dieses Beispiel zeigt auf, dass die Sinnzuweisung im Rahmen einer Alltagskonversation keineswegs nur von *individuellen* kognitiven Prozessen (wie z.B. dem Abrufen der lexikalischen Bedeutung aus dem Gedächtnis) abhängt, sondern auch ein *interaktiver* Austausch der GesprächspartnerInnen erforderlich ist. Um sich die Intention des Gesagten zu erschließen, stützen sich die KonversationsteilnehmerInnen auf ihre Annahmen über die Abläufe der sozialen Welt (wie z.B. die kooperativen Prinzipien/Regeln der Konversationslogik von Grice, 1975). Gleichzeitig verlassen sie sich bei ihrer Auslegung auf die Hilfe der FragestellerInnen im darauffolgenden Gespräch – insbesondere im Falle einer Fehlinterpretation (vgl. Schober, 1999, S. 77 f.).

### 2.5.2.2 Sinngebung bei Fragen in standardisierten Befragungen

Nehmen wir nun an, dass Ihnen dieselbe Frage „Was haben Sie heute gemacht?“ im Rahmen eines standardisierten Interviews gestellt wird. Wie gehen Sie in dieser Situation vor, um der Frage einen Sinne zu verleihen? Erschließen Sie sich die Bedeutung auf dieselbe Art und Weise, wie Sie es im Rahmen einer Alltagskonversation mit Ihrer Freundin getan hätten?

TeilnehmerInnen einer Umfrage wissen in der Regel, ohne explizit darüber nachgedacht zu haben, dass sich die Gesprächssituation, in der sie sich befinden, von der im Alltag in vielerlei Hinsicht abweicht (vgl. Schober, 1999, S. 81 f.). Clark und Schober (1992) nennen drei zentrale Unterschiede:

- (a) Umfragen sind um eine einzige Art des Austausches herum konzipiert: Fragen und Antworten.
- (b) Der Ablauf des Interviews ist durch die Struktur des Fragebogens und die entsprechenden Anweisungen und Instruktionen zur Durchführung vorherbestimmt.
- (c) Die Personen, die eigentlich die Fragen stellen (nämlich die ForscherInnen), sind nicht anwesend (vgl. Clark & Schober, 1992, S. 26).

Zudem unterliegen die Befragungen aufgrund ihrer Standardisierung einer Vielzahl von Einschränkungen, was die Interaktion der InterviewerInnen mit den Befragten anbelangt. Diese Vorgaben sollen Antwortverzerrungen vermeiden sowie eine Generalisierbarkeit der Antworten gewährleisten, haben jedoch auch zur Folge, dass die beiden interaktiven Prozesse des Audience Designs und des Groundings nicht genauso ablaufen, wie dies bei einem informellen Gespräch der Fall ist (vgl. Schober, 1999, S. 82).

### *Audience Design*

Wie bereits angedeutet, besteht ein zentraler Unterschied zwischen beiden Gesprächssituationen darin, dass im Alltag die FragestellerInnen in der Regel ihre eigenen Fragen formulieren, die ihrem Erkenntnisdrang entspringen. In einer standardisierten Befragung nehmen die InterviewerInnen hingegen nur eine vermittelnde Position ein und stellen nicht ihre eigenen Fragen, sondern die der ForscherInnen. Auch hinsichtlich der Formulierung bleibt den InterviewerInnen kein Gestaltungsspielraum. Um den Antwortprozess nicht in irgendeiner Art und Weise zu beeinflussen, wird ihnen aufgetragen, die Fragen wortgetreu vom Fragebogen abzulesen und keinerlei Änderungen vorzunehmen.

Ein weiterer Unterschied besteht im Hinblick auf die AdressatInnen, an die die ForscherInnen gedacht haben, als sie sich die Frage überlegten. Im Gegensatz zu einer Alltagskonversation, bei der die SprecherInnen ihre Äußerungen ganz speziell auf ihre AdressatInnen und den gemeinsamen Wissenshintergrund zuschneiden, werden Fragen im Rahmen einer standardisierten Befragung von den ForscherInnen für den „normalen Durchschnittsbefragten“ konstruiert (vgl. Schober, 1999, S. 82). Die Items sollen idealerweise für alle Personen, die an der Befragung teilnehmen, verständlich sein. Folglich basieren die Fragen auch nicht auf persönlich geteiltem, sondern auf kulturell geteiltem Wissen.

Die Befragten wiederum sind sich dieser Unterschiede zumeist bewusst. Folglich werden sie bei der Interpretation der Frage auch den ForscherInnen/FragebogendesignerInnen eine andere Intention unterstellen als beispielsweise einer Freundin. Auch was die Verwendung ihrer Antworten anbelangt, werden sie wahrscheinlich – je nachdem in welcher Gesprächssituation ihnen die Frage gestellt wurde – unterschiedliche Vermutungen anstellen (vgl. Schober, 1999, S. 82 f.).

### *Grounding*

Hinsichtlich des Groundings führt die Standardisierung der Befragungssituation ebenso zu starken Einschränkungen, die ein *beidseitiges* Absichern des pragmatischen Verständnisses unmöglich werden lassen. So werden die InterviewerInnen instruiert, eine neutrale Haltung zu wahren, wenn die Befragten um Erklärungen im Zusammenhang mit der gestellten Frage bitten. Durch diese Vorgehensweise soll eine Verzerrung des Antwortprozesses vermieden werden. So bleiben Nachfragen hinsichtlich einzelner Wortbedeutungen („Was ist mit dem Begriff „Familie“ in dieser Frage gemeint?) oder auch hinsichtlich der Intention der gesamten Aussage („Was ist mit dieser Frage gemeint?“) unbeantwortet, da die InterviewerInnen nicht dabei helfen dürfen, ein Grounding des Frageverständnisses sicherzustellen. Es ist ihnen lediglich gestattet, die Frage zu wiederholen („Die Frage lautet: Was haben Sie heute gemacht?“) oder explizit zu betonen, dass die Interpretation der Befragten gefragt sei („Es geht um *Ihre* Interpretation der Frage“) (vgl. Schober, 1999, S. 83, Willis, 2005, S. 19 f.).

Die InterviewerInnen selbst dürfen hingegen nachfragen, wie die Antwort der Befragten zu verstehen ist, die sie auf eine offene Frage hin gegeben haben. Allerdings ist es den

InterviewerInnen nur gestattet, auf bestimmte Nachfragen („Wie meinen Sie das?“ oder „Können Sie dies vielleicht noch näher ausführen/erläutern?“) zurückzugreifen. Dabei sollten sie weiterhin eine neutrale Haltung wahren.

Wie stark die InterviewerInnen in ihren Möglichkeiten bezüglich des Groundings eingeschränkt sind, sieht man auch im Falle von geschlossenen Fragen. Geben die Befragten eine Antwort ab, die nicht den vorgegebenen Antwortkategorien entspricht, so haben die InterviewerInnen nur die Möglichkeit, die Frage erneut vorzulesen und/oder die Antwortmöglichkeiten abermals zu nennen. Jede Abweichung von dieser Prozedur (z.B. die Wiederholung nur einiger Antwortmöglichkeiten) wird als „directive probe“ angesehen. In einer Alltagskonversation werden die FragestellerInnen hingegen anders agieren. So wird es nur selten vorkommen, dass sie bestimmte Antwortoptionen vorschreiben. Sollte dies doch einmal vorkommen („Hättest Du lieber Plätzchen oder Kuchen zum Kaffeetrinken?“), so würden sie sicherlich nicht auf diesen Vorgaben bestehen und sie beharrlich wiederholen, sollten sie von den AdressatInnen nicht explizit in der Antwort aufgegriffen worden sein („Ich hätte gerne Gebäck.“). Statt einer standardisierten Vorgehensweise zu folgen, würden die SprecherInnen wahrscheinlich anknüpfende Nachfragen stellen („Bevorzugst Du Cookies oder Spritzgebackenes?“), um ihrem Gegenüber zu verdeutlichen, dass ihre Antwort verstanden wurde (vgl. Schober, 1999, S. 83).

Geht es hingegen um das semantische Verständnis und um die Sicherstellung der akustischen Wahrnehmung, so ist der Prozess des Groundings weniger restringiert und die InterviewerInnen wenden gängige Grounding-Techniken an. So signalisieren die InterviewerInnen ihrem Gegenüber beispielsweise, dass sie ihre Antworten vernommen haben, indem sie „Okay“ oder „Mhm“ sagen. Sollten die Befragten eine Frage nicht gehört haben oder sollten sie gerade unaufmerksam gewesen sein, so werden die InterviewerInnen diese bei Anfrage (wenn auch nur wortwörtlich) wiederholen (vgl. Schober, 1999, S. 84).

Zusammenfassend lässt sich festhalten, dass im Rahmen einer standardisierten Befragung der Prozess des Groundings asymmetrisch abläuft. Wenn die Befragten sich nach dem intendierten Sinn einer Frage erkundigen, werden sie keine zufriedenstellende Antwort erhalten, da es den InterviewerInnen nicht gestattet ist, das pragmatische Verständnis der Befragten abzusichern. Die InterviewerInnen dürfen hingegen nachhaken, wenn ihnen nicht klar ist, was die Befragten mit ihrer Antwort meinen und diese sollten ihnen wiederum darüber Auskunft erteilen. Auch in einem solchen Fall dürfen die InterviewerInnen allerdings ihre Grounding-Techniken nicht frei wählen, sondern lediglich unverfängliche Nachfragen stellen, die die Befragten und somit den Antwortprozess nicht beeinflussen (vgl. Schober, 1999, S. 84).

Befragten ist es somit in einem standardisierten Interview nicht möglich, durch ein gemeinsames Grounding sicherzustellen, dass sie die Fragen im intendierten Sinne ausgelegt haben. Sie sind bei der Sinngebung auf sich alleine gestellt und können nur hoffen, dass ihre Interpretationen zutreffend sind. Ob es den Befragten jedoch immer gelingt, die intendierte Bedeutung zu erfassen, bleibt ungewiss:

„In standardized survey interviews, respondents are not licensed to ground understanding of words and phrases in the question. Respondents must answer questions with a presumption of interpretability (...) – they have to assume that their best guess as to an interpretation must be the one the survey designers intended. And their best guess can easily be wrong” (Schober, 1999, S. 89).

### 2.5.2.3 Mögliche negative Folgen der Standardisierung für die Datenqualität

Wie wir im vorangegangenen Abschnitt erfahren haben, ist es den TeilnehmerInnen einer standardisierten Befragung nicht möglich, ihr Frageverständnis durch ein Grounding abzusichern. Dies gilt sowohl für Selbstausfüller-Fragebögen, bei denen die Befragten auf sich alleine gestellt sind, als auch für persönlich geführte Interviews, bei denen die InterviewerInnen eine neutrale Haltung wahren sollen (vgl. Willis, 2005, S. 19).

Bereits auf der ersten Stufe des Frage-Antwort-Prozesses können daher zentrale Probleme auftreten, die mit dem pragmatischen Frageverständnis zusammenhängen. So werden häufig Fragen gestellt, die zu vage formuliert wurden. Lässt die Frage mehrere Interpretationen zu, fällt es den Befragten schwer, die Intentionen der ForscherInnen zu erschließen. Es ist nicht eindeutig für sie, nach welchen Informationen sie gefragt werden, und aufgrund des fehlenden beidseitigen Groundings können sie nur Vermutungen darüber anstellen. Folglich besteht die Möglichkeit, dass Befragte die Frage nicht in der intendierten Weise auffassen und dementsprechend eine Antwort geben, die nicht ihrem tatsächlich erfragten Verhalten oder ihrer tatsächlichen Einstellung entsprechen. Es kommt also zu einer Diskrepanz zwischen dem beobachteten Wert und dem theoretisch „wahren Wert“, die man als Antwortfehler (*response error*)<sup>17</sup> bezeichnet (vgl. Willis, 2005, S. 17 f.).

### 2.5.2.4 Implikationen für die Umfrageforschung

#### *Flexible, nicht-standardisierte Interviews*

Eine Option, um ein Grounding der Frage abzusichern und dementsprechend Antwortfehler zu reduzieren, ist ein flexibles, nicht-standardisiertes Interview. Hierbei ist es den InterviewerInnen erlaubt, auf Rückfragen der Befragten einzugehen, und die Intentionen der ForscherInnen zu verdeutlichen. Diese Form der Interviews konnte sich jedoch bislang im Rahmen von großen Bevölkerungsumfragen nicht durchsetzen. Dies hängt u.a. damit zusammen, dass ein flexibles Interview deutlich mehr Zeit in Anspruch nimmt als ein standardisiertes (vgl. Schober, 1999, S. 90, Willis, 2005, S. 19 f.).

Eine Standardisierung der Befragungssituation hat zudem noch weitere Vorteile. So müssen allen Befragten die Fragen in derselben Art und Weise präsentiert werden, um Fehler zu vermeiden, die auf unkontrollierte Variationen von einzelnen InterviewerInnen

---

<sup>17</sup> Willis (2005) beschreibt response errors folgendermaßen: “Characteristics of questions, and of respondent processing of those questions, may lead to incorrect answers” (S. 14).

zurückgehen. Viele FeldinterviewerInnen haben nur eine kurze Schulung durchlaufen und sind mit den Zielen und Intentionen der ForscherInnen überhaupt nicht vertraut. Jedes InterviewerInnen-Befragten-Paar würde daher sein eigenes beidseitiges Verständnis absichern, ohne dass dies zwangsläufig den Intentionen der ForscherInnen entsprechen muss. Denn schließlich stellen die InterviewerInnen nicht ihre eigenen Fragen, sondern sie nehmen nur eine vermittelnde Position zwischen ForscherInnen und Befragten ein. Eine Standardisierung erlaubt es daher, die Notwendigkeit des Groundings in Bezug auf zwei Konversationen – zum einen zwischen ForscherInnen und InterviewerInnen und zum anderen zwischen InterviewerInnen und Befragten – minimal zu halten (vgl. Willis, 2005, S. 20 f.).

#### *Leitlinien zur Itemformulierung und ihre Grenzen*

Folglich hat eine standardisierte Vorgehensweise durchaus verschiedene Vorteile, ihre Nachteile hinsichtlich des fehlenden Groundings lassen jedoch offensichtlich werden, wie bedeutsam der Prozess der Fragebogenentwicklung und die Formulierung von „guten“ Survey-Fragen ist, deren Sinn sich den Befragten auch von alleine erschließt. Wie gelingt dies am besten und wie sollten Fragen formuliert werden, um (Verständnis-) Probleme bereits im Vorfeld zu umgehen?

Es gibt eine Reihe von Methoden, die FragebogenentwicklerInnen dabei helfen sollen. Willis (2005) verweist insbesondere auf Leitlinien zur Fragebogenkonstruktion sowie auf Checklisten-Systeme (siehe Unterkapitel 2.6.2). Mit Hilfe von Leitlinien sollen den FragebogenentwicklerInnen Hinweise an die Hand gegeben werden, um gute Fragen konzipieren zu können. Da die UmfrageteilnehmerInnen hinsichtlich der Interpretation eines Items vollends auf sich allein gestellt sind und das gegenseitige Grounding nicht sicherstellen können, ist es von Bedeutung, ihre Gedankengänge gleich von Anfang an in die „richtigen Bahnen“ zu lenken. Damit dies gelingt und es zu keinen Fehlinterpretationen kommt, sollten die Items nach Porst (2000) u.a. möglichst klar und leicht verständlich formuliert sein. Dies gilt auch für die einzelnen Begriffe, die innerhalb einer Frage verwendet werden. Bei Fachbegriffen oder anderen Begriffen, die einem Teil der Befragten vermutlich nicht geläufig sind, ist es auf jeden Fall ratsam, eine Definition mitzuliefern. Die zeitlichen Bezüge, die in einer Frage verwendet werden, sollten ebenso eindeutig sein. Lange und komplexe Sätze sowie doppelte Stimuli<sup>18</sup> oder doppelte Verneinungen gilt es zu vermeiden. Ferner sollten hypothetische Fragen weggelassen werden. Dies trifft insbesondere auf solche Fragen zu, die Situationen ansprechen, in die sich die Befragten vermutlich nur schwer hineinversetzen können (vgl. Porst, 2000, S. 2).

Auch Leitlinien weisen jedoch gewisse Schwächen auf. Willis (2005) führt drei zentrale Gründe an, weshalb der Einsatz weiterer Pretest-Verfahren sinnvoll ist, selbst wenn

---

<sup>18</sup> Mit dem Begriff „doppelter Stimulus“ ist in diesem Fall die Bewertung zweier Sachverhalte in ein und derselben Frage gemeint, wie beispielsweise: „Mögen Sie Erdbeeren und Kiwis?“ Eine Befragte, die zwar Erdbeeren mag, Kiwis aber nicht, fällt vermutlich die Wahl einer Antwortkategorie schwer, wenn sie nur zwischen „ja“ und „nein“ wählen kann (vgl. Porst, 2008, S.102).



die Leitlinien zur Fragebogenkonstruktion berücksichtigt wurden. Erstens sind die Leitlinien nicht spezifisch genug. So kann man beispielsweise nicht genau sagen, ab wann eine Frage zu lang ist. Dies hängt u.a. von der Population ab, die befragt werden soll, oder auch von der Art der Frage (vgl. Willis, 2005, S. 27).

Zweitens gibt es mehrere Leitlinien, die nicht miteinander zu vereinbaren sind und einander sogar widersprechen. Fragen sollen beispielsweise so kurz und so einfach wie möglich gehalten werden, dennoch dürfen sie nicht zu vage formuliert sein. Snijkers (2004) beschreibt dieses Problem folgendermaßen: “However, a common dilemma in questionnaire design is: When to leave the interpretation of the question to the respondent (and have a simple and short question) and when to make it precise (and consequently have a long question containing a definition)?” (S. 196). Die verschiedenen Leitlinien zur Fragebogenkonstruktion können daher lediglich als Hilfestellung dienen, allgemein gültig sind sie jedoch nicht. Letztendlich muss jede Forscherin und jeder Forscher fallspezifisch entscheiden und abwägen, welches die beste Formulierung einer Frage ist – selbst wenn dadurch einige Leitlinien verletzt werden.

Drittens können Leitlinien dabei helfen, Fragen zu formulieren, die vermutlich weniger Verständnisprobleme verursachen, aber sie können nicht sicherstellen, dass die Items auch tatsächlich im intendierten Sinne aufgefasst werden. In diesem Zusammenhang betont Willis (2005) wie wichtig es ist, die Fragen etwa im Rahmen von kognitiven Interviews mit „realen“ Befragten zu testen. Es ist schwer, Probleme, die das pragmatische Verständnis betreffen, vorherzusehen. Bestimmte Interpretationen der Befragten werden erst ans Licht gebracht, wenn kognitive Interviews (oder auch andere Pretest-Verfahren) zum Einsatz kommen. So können die ForscherInnen zwar annehmen, dass die Intention des Items offensichtlich ist – schließlich wissen sie selbst, was mit der Frage gemeint ist –, trotzdem ist es möglich, dass kognitive Interviews eine völlig andere Interpretation seitens der Befragten offenbaren.

Zusammenfassend lässt sich festhalten, dass Leitlinien zur Fragebogenkonstruktion zwar nützlich sind, sie jedoch auch Schwächen aufweisen, da sie teilweise zu unspezifisch sind, sich mitunter widersprechen und das intendierte Frageverständnis nicht garantieren können. Folglich besteht die beste Vorgehensweise zur Entwicklung eines Fragebogens aus einer Kombination von Leitlinien und anschließendem Pretest.

*“If you don’t have the resources to pilot test your questionnaire, don’t do the study”*  
(Bradburn, Sudman, & Wansink, 2004, S. 317).

## 2.6 Pretest-Verfahren zur Evaluierung von Fragebögen

Jede Person, die bereits einmal selbst einen Fragebogen entwickelt hat, weiß um die Schwierigkeit, eine gute Frage zu formulieren. Bradburn, Sudman und Wansink (2004) sehen dies ähnlich: “Even after years of experience, no expert can write a perfect questionnaire” (S. 317). Einer der größten Fortschritte der Umfrageforschung in den letzten 20 Jahren ist daher die gestiegene Aufmerksamkeit hinsichtlich der systematischen Evaluierung von Fragen (vgl. Groves, Fowler, Couper, Lepkowski, Singer, & Tourangeau, 2009, S. 260). Doch was macht eine gute Frage aus und auf welche Aspekte hin gilt es einen Fragebogen zu testen?

Fowler (2001) zufolge sollte eine Frage acht Standards erfüllen:

- (1) *Inhaltliche Standards*. Einerseits müssen die Antworten auf die erhobenen Items dazu geeignet sein, um die Forschungsziele/-fragen beantworten zu können (Sprechen die Fragen also die richtigen Themen an?). Andererseits müssen die Fragen so gestellt sein, dass die Befragten willens und auch fähig sind, die erfragten Informationen zu liefern.
- (2) *Kosteneffektive Nutzung der Umfragezeit*. ForscherInnen sollten Prioritäten setzen und die zur Verfügung stehende Umfragezeit dazu nutzen, die Items zu erheben, die die wertvollsten Informationen für die eigenen Forschungsziele/-fragen liefern.
- (3) *Kognitive Standards*. Die Fragen sollten von allen Befragten in konsistenter Weise aufgefasst werden. Sie sollten so eindeutig formuliert sein, dass sie, wenn sie wortwörtlich von den InterviewerInnen vorgelesen werden, auch von den Befragten beantwortet werden können. Das pragmatische Verständnis sollte sich also den Befragten auch ohne wechselseitiges Grounding erschließen. Die Befragten müssen zu den Informationen, auf die die Frage abzielt, Zugang haben, und sie aus dem Gedächtnis abrufen können. Die Antworten, die auf die Fragen gegeben werden, sollten die Realität, nach der gefragt wurde, genau widerspiegeln.
- (4) *Interpersonelle Standards*. Wird der Fragebogen im Rahmen eines persönlichen Interviews durchgeführt, so gibt er (zumindest den InterviewerInnen) ein Skript für die Interaktion vor. Die Fragen sollten daher so konstruiert sein, dass sie wortwörtlich gestellt werden können und dabei so klar und eindeutig sind, dass die Befragten bereit sind, sie zu beantworten. Der Fragebogen muss zudem so konstruiert sein, dass sich bei dessen Befolgung durch die InterviewerInnen eine Interaktion ergibt, die den Befragten angemessen erscheint. Sich wiederholende, zu detaillierte, zu persönliche oder scheinbar irrelevante Fragen können zu einem

Nachlassen der Motivation der Befragten oder zu Antwortverweigerungen führen (vgl. Fowler, 2001, S. 49 ff.; Groves et al., 2009, S. 259).

- (5) *Benutzerfreundlichkeit (Usability)*. Fragebögen müssen, egal ob in schriftlicher oder in elektronischer Form, so einfach und selbst erklärend konstruiert sein, dass die InterviewerInnen bzw. die Befragten sich voll und ganz auf den Frage-Antwort-Prozess konzentrieren können, keine Probleme mit der Handhabung des Fragebogens haben und ihn wie vorgesehen ausfüllen können.
- (6) *Psychometrische Standards*. Eine gute Frage sollte so konstruiert sein, dass ihre Antworten das zugrunde liegende Konstrukt reliabel und valide messen (vgl. Fowler, 2001, S. 52; Prüfer & Rexroth, 2000, S. 3). Zudem sollten die Antwortverteilungen eine gewisse Variation aufweisen, da der Informationsgehalt der Frage ansonsten relativ gering ist (vgl. Fowler, 2001, S. 52).
- (7) *Verwendbarkeit bei verschiedenen Erhebungsmodi (Multi-Mode-Capability)*. Idealerweise sollten Fragen so konstruiert sein, dass sie sich für verschiedene Erhebungsformen (InterviewerIn vs. selbst-administrierte Fragebögen; Papier vs. Telefon vs. Online) eignen. Die Antworten auf die Fragen sollten dabei vergleichbar sein und somit nicht durch den Darbietungsmodus beeinflusst werden. Diese Aspekte sollten bereits bei der Formulierung der Items und auch bei der Wahl der Antwortkategorien berücksichtigt werden.
- (8) *Verwendbarkeit in verschiedenen Sprachen (Multi-Language-Capability)*. Bereits bei der Formulierung neuer Items sollte mitbedacht werden, wie die Fragen am besten gestellt sein sollten, damit sie sich gut in andere Sprachen übersetzen lassen (vgl. Fowler, 2001, S. 52 ff.).

Ähnlich wie in Bezug auf die Leitlinien besteht das Problem hinsichtlich der Standards darin, dass sie teilweise einander widersprechen. Wird eine Frage formuliert, die anhand eines Standards als gut einzustufen ist, kann sie gemäß eines anderen Standards als schlecht gelten (vgl. Fowler, 2001, S. 49). Liefert man beispielsweise detaillierte Definitionen und Erklärungen, um eine Frage leicht verständlich zu machen, lässt sie dies gleichzeitig kompliziert werden. Ein anderes Beispiel wäre der Konflikt zwischen psychometrischen Standards und der Einsatzmöglichkeit des Fragebogens mit verschiedenen Erhebungsmodi. So konnte gezeigt werden, dass eine größere Anzahl an Antwortalternativen sowie die Benennung (Labelung) der einzelnen Antwortkategorien zu besseren psychometrischen Kennwerten führt. Die größere Anzahl an Antwortvorgaben ist jedoch insbesondere bei Telefonbefragungen als nachteilig einzustufen, da es für die Befragten schwierig ist, sich an alle Kategorien zu erinnern (vgl. Fowler, 2001, S. 56). Eine Multi-Mode-Capability wäre somit vermutlich nicht gegeben.

Als grundlegendes Problem im Zusammenhang mit den Standards kann jedoch die Beobachtung von Fowler (2001) angeführt werden, dass viele Fragebögen erstellt werden,

ohne dass sich die FragebogenentwicklerInnen ernsthaft Gedanken über die acht aufgeführten Standards gemacht haben:

*„It is a common experience for those testing questions to learn that designers of questions have not even thought through what they want to measure and why (standard 1), much less begun to think through carefully the next and more complex step of which questions to ask to achieve their poorly specified objectives. Many questions are given birth without any real thought about standards at all”* (S. 54).

Ursache hierfür könnte die Tatsache sein, dass nicht alle Standards als gleichermaßen unabdingbar für die Fragebogenentwicklung angesehen werden. Lange Zeit lag der Fokus von Pretests lediglich auf der Überprüfung der Benutzerfreundlichkeit. In einigen Disziplinen werden routinemäßig Reliabilitäts- und Validitätstests durchgeführt, andere begnügen sich hingegen ohne einen empirischen Beleg mit der „face validity“. Neuere Forschungen zeigen jedoch die Bedeutsamkeit der übrigen Standards auf, die lange Zeit nicht berücksichtigt wurden und es teilweise bis heute noch nicht werden. Auch wenn es offensichtlich erscheint, dass eine Frage in konsistenter Weise von den Befragten erfasst werden sollte und das Verständnis auf Seiten der Befragten dem der ForscherInnen entsprechen sollte, werden beispielsweise kognitive Interviews erst in einer kleinen (wenn auch steigenden) Anzahl von Zentren routinemäßig durchgeführt (vgl. Fowler, 2001, S. 54).

Zur Evaluation von Fragebögen gibt es eine Reihe verschiedener Verfahren, die sich in zwei Kategorien unterteilen lassen. Zum einen die qualitativen Verfahren, die mit einer kleinen Anzahl von Befragten (zwischen 15 bis 30) arbeiten. Sie kommen in der Regel zu Beginn der Fragebogenentwicklung bzw. -testung in (Pretest-)Laboren zum Einsatz. Man bezeichnet sie daher auch als *Pre-Field-Techniken* (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 646). Dazu zählen u.a.:

- (1) Kognitive Interviews
- (2) Fokusgruppen
- (3) ExpertInnen-Review (Checklisten-Systeme)

Zum anderen stehen verschiedene quantitative Verfahren zur Verfügung, die mit deutlich größeren Stichproben (in der Regel über 100 Personen) durchgeführt werden und somit Aussagen über die Auftretenshäufigkeit einzelner Probleme erlauben. Aufgrund ihres Erhebungsortes im Feld werden sie auch *Field-Techniken* genannt. Sie werden zumeist in einer späteren Phase der Fragebogenentwicklung verwendet (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 646). Unter dieser zweiten Gruppe der Pretest-Verfahren lassen sich die folgenden Methoden subsumieren:

- (4) Beobachtungspretest (mit InterviewerInnen-Debriefing)
- (5) Behavior Coding
- (6) Befragten-Debriefing

- (7) Response Latency
- (8) Reinterview
- (9) Analyse von Pretest-Daten
- (10) Split-Ballot

Im Folgenden sollen nun die einzelnen Verfahren näher erläutert werden. Da die Methode der kognitiven Interviews zentral für die Evaluierung der FWSEM-Skala ist, wird sie als erste vorgestellt und erfährt besondere Aufmerksamkeit.

## 2.6.1 Kognitive Interviews

### 2.6.1.1 Historische Entwicklung

Das Problem, dass Fragen häufig nicht im intendierten Sinne aufgefasst werden, ist Kurz, Prüfer und Rexroth (1999) zufolge schon seit längerer Zeit bekannt: Bereits in den 1940er Jahren stellte Cantril gezielte Nachfragen („probes“), um zu erforschen, was die InterviewteilnehmerInnen während des Antwortfindungsprozesses gedacht haben. Nuckols arbeitete in den 1950er Jahren mit der Technik des Paraphrasierens und bat die Befragten den Iteminhalt in ihren eigenen Worten zu wiederholen (vgl. Groves, Fultz & Martin, 1992, S. 50; Kurz, Prüfer & Rexroth, 1999, S. 86). In den 1960er Jahren regte Schumann an, bei zufällig ausgewählten Items der Hauptstudie Nachfragen zum Verständnis einzusetzen („random probes“), um so die Validität der Items kontrollieren zu können (vgl. Kurz, Prüfer & Rexroth, 1999, S. 86).

Erst seit den 1980er Jahren befassen sich kognitive (Sozial-)PsychologInnen und UmfrageforscherInnen zielgerichtet mit den kognitiven Aspekten in der Umfragemethodologie („Cognitive Aspects of Survey Methodology“, kurz: CASM). Ihre interdisziplinäre Zusammenarbeit geht dabei vor allem auf zwei Treffen zurück: Im Herbst 1983 fand in den USA das „Advanced Research Seminar on Cognitive Aspects of Survey Methodology“ statt und im Sommer 1984 eine Tagung in Deutschland zu Thema „Social Information Processing and Survey Methodology“ (vgl. Schwarz, Strack, & Hippler, 1991, S. 175 f.; Willis, 2005, S. 34 f.). Ein Ergebnis dieser Workshops war, dass die UmfrageforscherInnen den Wert verschiedener Techniken, die ursprünglich von KognitionspsychologInnen entwickelt wurden, im Hinblick darauf untersuchten, inwieweit sie Auskunft über den Frage-Antwort-Prozess geben können. Ein Verfahren, das während dieser Treffen diskutiert wurde, war der Einsatz kognitiver Interviews zur Evaluierung von Survey-Fragen (vgl. Groves et al., 2009, S. 263 f.).

Nach und nach richteten mehrere Organisationen (z.B. the National Center for Health Statistics, Census Bureau, Bureau of Labor Statistics und das GESIS – Leibniz-Institut für Sozialwissenschaften, ehemals Zentrum für Umfragen, Methoden und Analysen) eigene kognitive Labore ein, in denen der Verlauf des Frage-Antwort-Prozesses

erforscht und Fragebögen auf mögliche Mängel hin getestet werden (vgl. Beatty & Willis, 2007, S. 288; Kurz, Prüfer & Rexroth, 1999, S. 86).

### 2.6.1.2 Ziele

Das primäre Ziel kognitiver Interviews ist es herauszufinden, wie gut ein Fragebogen bzw. seine einzelnen Items „funktionieren“ und ob potentielle Probleme bestehen. Um dies zu überprüfen, orientiert man sich an den vier Stufen des Frage-Antwort-Prozesses (siehe Unterkapitel 2.5.1) und testet folgende Aspekte:

- (1) *Verständnis der Frage*
  - (a) Intention der Frage: Was glauben die Befragten, was mit der Frage gemeint sein könnte? Wurde die Frage im intendierten Sinne aufgefasst?
  - (b) Bedeutung bestimmter Begrifflichkeiten: Wie interpretieren die Befragten die Schlüsselbegriffe und bestimmte Formulierungen des Items?
- (2) *Abruf fragerelevanter Informationen und Bildung eines Urteils*
  - (a) Abrufbarkeit von Informationen: Welche Art von Informationen müssen die Befragten aus dem Gedächtnis abrufen, um die Frage beantworten zu können? Sind die Befragten dazu in der Lage, fragerelevante Informationen aus dem Gedächtnis abzurufen?
  - (b) Motivation: Nehmen die Befragten auch größere mentale Anstrengungen in Kauf, um die Frage nach bestem Wissen und Gewissen möglichst korrekt zu beantworten?
  - (c) Erinnerungsstrategien: Welche Erinnerungsstrategien werden eingesetzt, um die Informationen abzurufen? Versuchen die Befragten sich beispielsweise an alle einzelnen Ereignisse ganz genau zu erinnern und summieren diese dann auf oder wird lediglich eine grobe Schätzung abgegeben?
- (3) *Formatierung der Antwort*

Verortung der Antwort: Sind die Befragten dazu in der Lage, ihre intern ermittelten Antworten in die vorgegebenen Antwortmöglichkeiten zu übertragen?
- (4) *Editierung der Antwort*

Sensitivität/soziale Erwünschtheit: Sind die Befragten bereit, eine ehrliche Antwort abzugeben oder versuchen sie sich in ein positives Licht zu rücken?  
(vgl. Prüfer & Rexroth, 2005, S. 4; Willis, 1999, S. 2)

Für Fragen, die nicht alltäglich sind, kann der Frage-Antwort-Prozess sehr komplex sein und mehrere kognitive Schritte beanspruchen. Einige dieser Prozesse laufen „bewusst“ ab, andere erfolgen hingegen automatisch und sind daher für den Befragten schwer zugänglich. UmfrageforscherInnen sind sich darüber im Klaren, dass sie trotz kognitiver Interview-Techniken keinen vollständigen Zugriff auf die Gedanken der Befragten erlangen können. Das Ziel kognitiver InterviewerInnen ist es daher, die Befragten

dazu zu bringen, ihnen Informationen mitzuteilen, die (zumindest) Hinweise bezüglich der oben beschriebenen Fragen liefern und somit hilfreich für die Evaluierung der Items sind (vgl. Willis, 2005, S. 2 f.).

### **2.6.1.3 Empfehlungen zur Durchführung**

Bislang existiert kein allgemein anerkanntes Regelwerk zur Durchführung kognitiver Interviews (vgl. Kurz, Prüfer & Rexroth, 1999, S. 86), gleichwohl liegen mittlerweile einige englisch- und auch deutschsprachige Empfehlungen hierzu vor, die die praktischen Erfahrungen der AutorInnen im Zusammenhang mit dieser Methode wiedergeben (siehe hierzu z.B. Beatty & Willis, 2007; Prüfer & Rexroth, 2005; Willis, 2005).

Willis (1999) empfiehlt, den kognitiven Interviews ein „Expert Appraisal“ vorzuschalten. Erfahrene kognitive InterviewerInnen sind oftmals dazu in der Lage, bestimmte Probleme, die im Zusammenhang mit einer Frage auftreten (könnten), zu antizipieren. Folglich ist es sinnvoll, zuerst ein ExpertInnen-Review (siehe Unterkapitel 2.6.3) anzusetzen und evtl. bereits einige Änderungen der Items vorzunehmen und erst anschließend die kognitiven Interviews durchzuführen (vgl. Willis, 1999, S. 29).

Im Zusammenhang mit den kognitiven Interviews selbst sind sich die ExpertInnen einig, dass es sinnvoll ist, mehrere Runden durchzuführen. Nachdem der Fragebogen bzw. die einzelnen Items in einer ersten kognitiven Runde getestet und überarbeitet wurden, empfehlen sie, diese Änderungen in weiteren Runden zu evaluieren (vgl. Beatty & Willis, 2007, S. 296; Prüfer & Rexroth, 2005, S. 17). Hinsichtlich der Frage, wie viele solcher Testrunden angesetzt werden sollten, macht Willis (1999, S. 30) folgende Anmerkung: „In one sense, a questionnaire could be tested forever, and still have problems (the perfect survey question may not exist).“ Im Idealfalle sollte seiner Ansicht nach ein Fragebogen so lange geprüft werden, bis die zentralen Probleme aufgespürt würden. Da die Entwicklung und Testung von Umfragen in der Regel jedoch einem strikten Zeitplan unterlägen, würden in der Praxis nur so viele Runden kognitiver Interviews durchgeführt, wie es die zeitlichen Ressourcen zuließen. Üblicherweise seien dies maximal 3 oder 4 Runden (vgl. Willis, 1999, S. 30).

Die Art der Interview-Runden ändert sich mit den verschiedenen Stadien der Fragebogentestung. Zu Beginn der Fragebogen-Überprüfung beziehen sich die Analysen neben spezifischen Itemformulierungen auch auf allgemeinere Aspekte, die das Konstrukt betreffen, welches mit Hilfe des Fragebogens erfasst werden soll. Beispielsweise wird überprüft, ob die erfragten Informationen das Wissen oder die Erinnerungsleistungen der Befragten übersteigen oder ob das Konstrukt zu kompliziert ist, um es mit einigen wenigen Items zu erfassen. Wenn die größten konzeptionellen Probleme behoben sind, konzentrieren sich spätere Runden kognitiver Interviews verstärkt auf die Formulierungen einzelner Items. Daneben wird auch der Fragebogen als Ganzes in den Blick genommen und auf mögliche Kontexteffekte (z.B. Reihenfolgeeffekte) geachtet (vgl. Willis, 1999, S. 30 f.).

Nicht exakt festgelegt ist, wie viele kognitive Interviews insgesamt durchgeführt werden sollten. Sowohl Prüfer und Rexroth (2005) als auch Willis (2005) geben an, dass sie innerhalb einer Test-Runde 5 bis 15 Personen befragen. Werden mehrere Runden kognitiver Interviews durchgeführt, so variiert i.d.R. auch hier die Anzahl der befragten Personen. Befinden sich beispielsweise Fragebögen noch in einer frühen Testphase (erste Runde der kognitiven Interviews), so ist es möglich, dass bereits nach wenigen Interviews schwerwiegende Probleme auftauchen, die zunächst behoben werden sollten, bevor noch weitere Personen befragt werden. Oftmals genügen dann bereits vier bis fünf Interviews, um die größten Schwachstellen oder auch Mängel hinsichtlich des zu messenden Konstruktes (konzeptionelle Probleme) ausfindig zu machen. In späteren Runden, die sich eher auf die Feinheiten der Frageformulierung beziehen, wird dann die Fallzahl erhöht (vgl. Willis, 1999, S. 29 ff.).

Blair, Conrad, Ackermann und Claxton (2006) weisen darauf hin, dass größere Stichproben generell empfehlenswert sind. Ihren Untersuchungen zufolge offenbarten kognitive Interviews, die nur mit wenigen Befragten ( $N = 5$ ) durchgeführt wurden, weniger Probleme als solche, in denen viele Personen ( $N = 50$ ) befragt wurden. Während mit steigender Fallzahl die Anzahl der identifizierten Fragemängel zunahm, konnte in kleineren Stichproben eine Vielzahl von Problemen nicht ausfindig gemacht werden. Darunter waren auch solche Probleme, die als schwerwiegend einzustufen sind. Stehen entsprechend Zeit und Geld zur Verfügung, kann die Durchführung einer größeren Anzahl Interviews als dies normalerweise üblich ist, durchaus hilfreich sein, so die Schlussfolgerung der AutorInnen.

Zur Rekrutierung der Befragten ist anzumerken, dass zumeist mit Quoten-Stichproben gearbeitet wird, die eine Variation des Geschlechts, des Alters, der Bildung sowie des sozioökonomischen Status gewährleisten soll. Sind weitere Merkmale im Zusammenhang mit der Thematik des Fragebogens relevant, so sollten auch diese bei der Quotenauswahl einbezogen werden (vgl. Rexroth & Prüfer, 2005, S. 17; Willis, 1999, S. 24). Handelt es sich beispielsweise um einen Fragebogen zum Einkaufsverhalten im Internet, so sollten alle Befragten das Internet auch nutzen. Als Aufwandsentschädigung empfehlen Rexroth und Prüfer (2005, S. 17), den InterviewteilnehmerInnen ein Geldbetrag in Höhe von 25 Euro pro einstündigem Interview (einschließlich Anfahrtszeiten und -kosten) zu zahlen.

Wenn möglich, sollten die Interviews in einem kognitiven Labor durchgeführt werden, in dem das Gespräch akustisch und visuell aufgezeichnet und zudem in einem Nebenraum beobachtet werden kann. Demgemäß wird diese Form der Befragung auch oftmals als „Laborinterview“ bezeichnet (vgl. Kurz, Prüfer & Rexroth, 1999, S. 87). Steht kein kognitives Labor zur Verfügung, kann als Alternative auch ein ruhiger (Büro-) Raum genutzt werden; eine akustische Aufzeichnung sollte jedoch auch hier erfolgen, da diese als Basis für die spätere Auswertung dient (vgl. Prüfer & Rexroth, 2005, S. 16; Willis, 1999, S. 26).



Für die Durchführung der Interviews werden speziell geschulte InterviewerInnen eingesetzt, zu denen auch MitarbeiterInnen des Forschungsteams oder auch KognitionspsychologInnen zählen können (vgl. Kurz, Prüfer & Rexroth, 1999, S. 87). Dabei verlangt die Durchführung der kognitiven Interviews eine völlig andere Vorgehensweise von den InterviewerInnen, als dies normalerweise bei einer Befragung im Feld der Fall ist. Während letztere etwa ein schnelles „Abarbeiten“ der Fragen erfordert (damit die Befragten nicht müde und unkooperativ werden), sollten sich die InterviewerInnen bei einem kognitiven Interview für jede einzelne Frage Zeit nehmen (vgl. Willis, 1999, S. 22). Zudem sollten FeldinterviewerInnen eine neutrale Haltung wahren, d.h. sie sollen nicht auf die Nachfragen der Befragten eingehen und nicht von standardisierten Vorgehensweisen abweichen, um so den Antwortprozess nicht zu beeinflussen. Ein kognitives Interview ähnelt hingegen viel eher einer Alltagskonversation – das Sicherstellen des beidseitigen Verständnisses, also das Grounding des Gesagten, ist von zentraler Bedeutung, um den Formulierungsdefiziten auf den Grund gehen zu können.

Willis (1999) zufolge sollten InterviewerInnen, die im Rahmen von kognitiven Interviews eingesetzt werden, idealerweise folgende Voraussetzungen mitbringen:

- (1) Vertrautheit mit den Methoden und Zielen von kognitiven Interviews;
- (2) Kenntnisse über die kognitiven und kommunikativen Prozesse der Umfrageforschung, dazu zählen u.a. der Frage-Antwort-Prozess, mögliche Kontexteffekte sowie Skaleneffekte;
- (3) Kenntnisse der Intentionen der einzelnen Items bzw. des Forschungszieles des Fragebogens, denn nur so können Diskrepanzen zwischen dem Frageverständnis der Befragten und der ursprünglichen Intention der ForscherInnen erkannt werden;
- (4) Erfahrung hinsichtlich der Itemformulierung und Fragebogenkonstruktion, da diese nötig sind, um die Items überarbeiten zu können;
- (5) Besitz von zwischenmenschlichen Fähigkeiten, das Vermögen den Befragten die Nervosität zu nehmen und urteilsfrei in den eigenen Äußerungen zu bleiben (vgl. Willis, 1999, S. 21).

Hinsichtlich des zeitlichen Rahmens der Interviews führt Willis (1999) eine Stunde als optimale Länge an. Kurz, Prüfer und Rexroth (1999) empfehlen eine Dauer von einhalb Stunden nicht zu überschreiten, da dies andersfalls zu Lasten der Motivation und Konzentration der Befragten gehen könnte. Folglich ist es daher in den meisten Fällen nicht möglich, alle Fragen der Haupterhebung zu testen, sondern lediglich eine Auswahl daraus (vgl. Kurz, Prüfer & Rexroth, 1999, S. 87; Willis, 1999, S. 21).

#### 2.6.1.4 Ablauf

Zu Beginn eines jeden Interviews sollten die Befragten von den InterviewerInnen über das Anliegen und die Ziele kognitiver Pretests informiert werden. So sollte ihnen mitgeteilt werden, dass es nicht um das Sammeln von Umfragedaten, sondern um die Testung eines Fragebogens geht. In diesem Zusammenhang bietet es sich an, die TeilnehmerInnen unmittelbar auf die Möglichkeit hinzuweisen, dass während des Interviews Fragen auftauchen können, die schwer zu verstehen oder zu beantworten sind und dass es gilt, genau diese Probleme zu identifizieren. Im Idealfall gelingt es den InterviewerInnen, den Befragten das Gefühl zu vermitteln, dass sie Kritik frei äußern dürfen, ja dass diese sogar erwünscht ist, da sie dabei hilft, den Fragebogen zu verbessern. Zu betonen ist dabei auch, dass alle Hinweise, selbst wenn sie noch so trivial oder irrelevant erscheinen, hilfreich sein können (vgl. Willis, 1999, S. 23).

Die Vorgehensweise des Pretest selbst sieht i.d.R. so aus, dass die Befragten das Item zunächst genauso wie bei einer normalen Befragung beantworten und sich für einen Skalenwert entscheiden.<sup>19</sup> Direkt im Anschluss daran (oder auch erst nach Abschluss des gesamten Fragebogens) erfolgt die Test-Phase. In dieser versuchen die InterviewerInnen unter Zuhilfenahme verschiedener kognitiver Methoden (siehe Unterkapitel 2.6.1.5), die Gedanken der Befragten während des Frage-Antwort-Prozesses aufzudecken (vgl. Kurz, Prüfer & Rexroth, 1999, S. 87).

Bei einer *flexiblen* (unstrukturierten) Vorgehensweise bleibt es den InterviewerInnen selbst überlassen, welche Methoden sie einsetzen und wie intensiv sie nachfragen, um den Frage-Antwort-Prozess zu untersuchen. Der Vorteil einer solchen Vorgehensweise liegt vor allem darin, dass den Ursachen für die Probleme seitens der Befragten (insbesondere durch erfahrene InterviewerInnen) viel genauer auf den Grund gegangen werden kann, als dies bei einem *standardisierten* Verfahren der Fall ist, bei dem bereits im Vorfeld festgelegt wird, welche Items mit welcher Technik untersucht werden. Der wesentliche Nachteil besteht in einer geringeren Vergleichbarkeit der gewonnenen Informationen, da in jedem Interview unterschiedliche Aspekte angesprochen werden.

Um eine Vergleichbarkeit zu erreichen, müsste ein standardisierter Fragebogen (ein sogenannter „Evaluationsfragebogen“) zum Einsatz kommen, der allerdings wiederum den InterviewerInnen keine flexible Vorgehensweise erlaubt. Zudem hängen die Menge und auch die Güte der gewonnenen Informationen dann stark von der Qualität des Evaluationsfragebogens ab, der neben den zu überprüfenden Items des Fragebogens die verschiedenen kognitiven Verfahren vorgibt, die zur Anwendung kommen sollen (vgl. Prüfer & Rexroth, 2000, S. 10 f.).

Als idealer Kompromiss ist daher eine halbstandardisierte Vorgehensweise anzusehen, welche die Vorteile beider Verfahren verbindet. Dabei kommen mehrere kognitive Techniken zum Einsatz, die bereits im Vorfeld vorbereitet werden. Gleichzeitig wird den InterviewerInnen die Möglichkeit gegeben, eigenständig weitere kognitive Techniken

---

<sup>19</sup> Einzige Ausnahme bildet die Concurrent-Think-Aloud-Methode (siehe Unterkapitel 2.6.1.5).

einzusetzen, wenn sie dies für angebracht halten (z.B. eine Kombination aus scripted und spontaneous probes; siehe Unterkapitel 2.6.1.5) (vgl. Willis, 1999, S. 8f.).

### **2.6.1.5 Kognitive Techniken**

Da die Befragten im Rahmen von kognitiven Interviews unabhängig von einem formal korrekten Antwortverhalten um Erläuterungen und Offenlegung ihrer Gedankengänge gebeten werden, bezeichnet man alle kognitiven Verfahren als „aktive“ Techniken (vgl. Prüfer & Rexroth, 2000, S. 7). Im Folgenden werden nun die wichtigsten kognitiven Techniken mit ihren jeweiligen Stärken und Schwächen vorgestellt.

#### *Card Sorting (Sortiertechnik)*

Die Sortiertechnik soll vor allem darüber informieren, wie eng oder weit gefasst relevante Begriffe einer Frage definiert werden bzw. welche Beispiele als einem Oberbegriff zugehörig wahrgenommen werden. Hierfür werden im Vorfeld Karten angefertigt, auf denen verschiedene Situationsbeschreibungen, Aussagen oder Unterbegriffe abgetragen sind. Die Befragten sollen dann für jede Karte entscheiden, ob sie sie in die Oberkategorie einordnen würden oder nicht. Geht es beispielsweise um den Begriff „Familie“, könnte untersucht werden, ob auch „Patenkinder“, „Cousinen zweites Grades“ o.ä. nach Ansicht der Befragten dazu zählen.

Im Zusammenhang mit dieser Technik ist positiv anzumerken, dass sie sehr verlässlich funktioniert und für die Befragten eine willkommene Abwechslung darstellt. Negativ zu erwähnen ist hingegen der Aufwand, der im Vorfeld nötig ist, um die Karten zu erstellen, die alle relevanten Aspekte des zu prüfenden Begriffs abdecken (vgl. Prüfer & Rexroth, 2005, S. 13 f.).

#### *Confidence Rating<sup>20</sup> (Methode zur Bewertung der Verlässlichkeit der Antwort)*

Bei dieser Technik werden die Befragten im Anschluss an die Beantwortung der Frage gebeten, eine Bewertung über die Verlässlichkeit ihrer Antwort vorzunehmen. Dies kann in Form einer offenen Frage („Was denken Sie, wie exakt Ihre Antwort ist?“) oder in Form einer geschlossenen Frage erfolgen („Wie sicher waren Sie sich, als Sie Ihre Antwort abgegeben haben: sehr sicher, eher sicher, eher unsicher, sehr unsicher?“) (vgl. Kurz, Prüfer & Rexroth, 1999, S. 88; Prüfer & Rexroth, 2005, S. 10). Der Einsatz des Confidence Ratings eignet sich nach Einschätzung von Rexroth und Prüfer (2005) v.a. zur Überprüfung von Rückerinnerungsfragen (z.B. „Seit wie vielen Jahren fahren Sie Ihr Auto?“) (vgl. Prüfer & Rexroth, 2005, S. 10).

Der Vorteil dieser Technik liegt in der einfachen Anwendbarkeit und Auswertung. Der Nachteil besteht in dem geringen Informationsgehalt. Gibt die befragte Person an,

---

<sup>20</sup> Sowohl die Technik des Confidence Ratings als auch die des Paraphrasings ordnet Willis (1999) der Methode des Probing zu, da es sich hierbei ebenfalls um verschiedene Formen von Nachfragen handelt.

sie sei sich bei ihrer Antwort sehr unsicher gewesen und sie schätze sie als wenig exakt ein, so kann das eigentliche Frage-Problem – wenn überhaupt – nur dann ausfindig gemacht werden, indem andere Methoden zum Einsatz kommen. So müsste man beispielsweise eine zusätzliche Nachfrage (siehe *Probing*) anschließen, die sich nach dem Grund für die geringe Verlässlichkeit erkundigt (vgl. Kurz, Prüfer & Rexroth, 1999, S. 90; Prüfer & Rexroth, 2005, S. 11).

#### *Paraphrasing (Methode des Paraphrasierens)*

Ziel des Paraphrasierens ist es herauszufinden, welche inhaltlichen Aspekte mit dem Item assoziiert werden und ob alle wichtigen Details der Frage im Zuge des Antwortprozesses berücksichtigt wurden (vgl. Prüfer & Rexroth, 1996, S. 108). Um diesbezüglich die gewünschten Informationen zu erhalten, werden die Befragten gebeten, den Inhalt der Frage in ihren eigenen Worten wiederzugeben. Dabei ist es von zentraler Bedeutung, dass sie den Fragetext sinngemäß paraphrasieren, da eine wortwörtliche Wiederholung nur einen Rückschluss über die Erinnerungsleistung, nicht jedoch hinsichtlich des Frageverständnisses zulässt (vgl. Kurz, Prüfer & Rexroth, 1999, S. 89 f.; Prüfer & Rexroth, 2005, S. 12).

Auch bei der Methode des Paraphrasierens liegt der klare Vorteil in der einfachen Durchführbarkeit für InterviewerInnen und Befragte. Der erste Nachteil dieser Technik liegt in der eingeschränkten Einsatzmöglichkeit. So ist ihre Anwendung bei kurzen Faktfragen wenig sinnvoll und bietet sich eigentlich nur bei langen, komplizierten Fragetexten an. Zudem eignet sich das Paraphrasieren nur begrenzt zur Aufdeckung von Verständnisproblemen; konkrete Interpretationen eines Begriffes oder einer Formulierung können nicht zufriedenstellend untersucht werden (vgl. Kurz, Prüfer & Rexroth, 1999, S. 89 f.). Prüfer und Rexroth (2004) stießen in ihren Untersuchungen auf eine weitere Schwachstelle hinsichtlich der Zuverlässigkeit dieser Technik. Gezielte Nachfragen, die im Anschluss an die Technik des Paraphrasierens gestellt wurden, zeigten Folgendes auf: Nicht in allen Fällen griffen die befragten Personen alle wichtigen Aspekte eines Items in ihren Paraphrasierungen auf – selbst nicht unbedingt dann, wenn sie diese bei der Beantwortung der Frage miteinbezogen hatten. Auch für den umgekehrten Fall gilt dies: Wird der Aspekt beim Paraphrasieren genannt, bedeutet dies nicht zwangsläufig, dass er auch beim Antwortprozess berücksichtigt wurde.

Als Fazit kann somit festgehalten werden, dass die Methode des Paraphrasierens nur einen sehr groben Aufschluss über das Frageverständnis liefern kann und sie daher immer nur in Kombination mit weiteren kognitive Verfahren zum Einsatz kommen sollte (vgl. Prüfer & Rexroth, 2005, S. 13).

#### *Think aloud (Methode des „lauten Denkens“)*

Die Methode des lauten Denkens wird von Willis (1999) neben dem *Probing* als eine der beiden zentralen Techniken beschrieben, die im Rahmen von kognitiven Interviews ein-

gesetzt werden. Bei diesem Verfahren werden die Befragten instruiert, alle Gedankengänge, die ihnen *während der Antwortfindung* durch den Kopf gehen, laut auszusprechen. Die InterviewerInnen unterbrechen die befragten Personen dabei so wenig wie möglich. Lediglich wenn diese längere Pausen einlegen und vergessen, weiterhin „laut zu denken“, sollen die InterviewerInnen sie daran erinnern, ihre Gedanken wieder auszusprechen („Erzählen Sie mir bitte, was Sie gerade denken.“) (vgl. Willis, 1999, S. 3).

Kurz, Prüfer und Rexroth (1999) unterscheiden von dieser Vorgehensweise, die sie als *Concurrent-Think-Aloud* bezeichnen, die Methode des *Retrospective-Think-Alouds*, bei der die Befragten erst *im Nachhinein* rekonstruieren sollen, wie sie zur Wahl ihres Skalenwertes gelangt sind (vgl. Kurz, Prüfer, & Rexroth, 1999, S. 88).<sup>21</sup>

Möchte man die Methode des lauten Denkens einsetzen, so muss diese zu Beginn des Interviews mit den Befragten geprobt werden. Als Übung schlägt Willis (1999) vor, die Befragten die Anzahl der Fenster, die ihre Wohnung besitzt, zählen zu lassen. Dabei sollen sie sich ihre Wohnung vorstellen, beschreiben, was sie sehen und laut aussprechen, was sie denken. Sollte die Bewältigung dieser Aufgabe den InterviewteilnehmerInnen noch schwer fallen (haben sie z.B. nicht kontinuierlich laut gedacht oder sind sie nicht auf problematische Grenzfälle, wie eine Balkontür o.ä. eingegangen), ist ein weiteres Training nötig, bevor mit dem eigentlichen kognitiven Interview begonnen werden kann (vgl. Willis, 1999, S. 4).

Den Erfahrungen von Prüfer und Rexroth (2005) zufolge fällt die Technik des lauten Denkens vielen befragten Personen schwer und einige sind überhaupt nicht dazu in der Lage, sie adäquat auszuführen. Nur etwa jede zweite Person schafft es, ihre Gedanken kontinuierlich in Worte zu fassen. Am besten funktioniert die Methode noch bei retrospektiven Fragen oder Fragen, die eine Problemlösung erfordern. So fällt es vielen Personen leichter, einen Rückerinnerungsprozess zu beschreiben (z.B. wie häufig sie in den letzten 12 Monaten im Kino waren) als eine Einstellung zu einem bestimmten Sachverhalt laut kund zu tun (z.B. ob sie für oder gegen die Legalisierung von Abtreibung sind) (vgl. Prüfer & Rexroth, 2005, S. 16).

Dennoch bietet diese Technik auch verschiedene Vorteile: Zunächst einmal können auch sehr unerfahrene bzw. ungeübte InterviewerInnen bei dieser Methode eingesetzt werden, da sie lediglich das Item vorlesen und den Befragten bei längeren Pausen daran erinnern müssen, laut zu denken. Diese Vorgehensweise bringt mit sich, dass die Gefahr einer Antwortverzerrung des Befragten durch Einwürfe der InterviewerInnen minimiert wird. Da die Äußerungen der Befragten in keine bestimmte Richtung gelenkt werden, besteht zudem die Möglichkeit, dass auf Probleme hingewiesen wird, die von den ForscherInnen/ExpertInnen nicht antizipiert wurden (vgl. Willis, 1999, S. 4).

---

<sup>21</sup> Diese Variante lässt sich allerdings nur schwer von der Methode des *Follow-Up-Probing* (siehe nächster Abschnitt) unterscheiden und könnte auch in die Kategorie *Hinterfragen von Antworten* eingeordnet werden (z.B. „Was haben Sie gedacht, als Sie die Frage eben beantwortet haben?“).

Die Nachteile dieses Verfahrens lassen sich so zusammenfassen: Wie bereits erwähnt, ist es nötig, diese Technik vor dem Interview mit den TeilnehmerInnen zu üben, was wiederum von der Zeit für das eigentliche kognitive Interview abgeht (vgl. Prüfer & Rexroth, 2005, S. 16; Willis, 1999, S. 5). Selbst nach dem Training schaffen es einige Personen nicht, ihre Gedanken laut auszusprechen. Sie neigen dazu, lediglich die Antwort auf die Fragebogenfrage zu geben, ohne weitere Erläuterungen vorzunehmen, da ihnen mitunter die Fähigkeit zur Introspektion und zur Verbalisierung fehlt (vgl. Kurz, Prüfer & Rexroth, 1999, S. 89; Willis, 1999, S. 5). Nach einer Einschätzung von Willis (1999) liegt die Hauptlast bei dieser Methode somit auf den Schultern der Befragten und stellt extrem hohe Anforderungen an diese.

Ein weiterer Nachteil dieser Technik liegt darin, dass viele Befragte in ihren Ausführungen von der eigentlichen Fragestellung abschweifen. Dies stellt die InterviewerInnen vor das Problem, dass sie die Befragten wieder zum Kern des Themas zurückführen müssen (vgl. Willis, 1999, S. 5). Zudem werden auch die Auswertungen sehr zeitintensiv, da einfach mehr Interviewmaterial produziert wird, das bearbeitet werden muss (vgl. Prüfer & Rexroth, 2005, S. 16). Da die Befragten sich bei der Technik des lauten Denkens unter Umständen relativ lange zu einem Item äußern, können innerhalb eines Interviews auch – im Vergleich zu anderen Methoden (wie z.B. dem Probing) – nur sehr wenige Items eines Fragebogens getestet werden (vgl. Willis, 1999, S. 5).

Befragte haben häufig nur einen Zugang zu dem Resultat ihrer Denkanstrengungen und nicht zu den kognitiven Prozessen selbst (d.h. die Antwort auf eine Frage ist sofort im Gedächtnis parat, ohne dass sie wissen, wie sie zu dieser Antwort gekommen sind) (siehe hierzu Nisbett & Wilson, 1977). Oftmals sind Befragte daher nicht in der Lage, ihre Gedanken während der kognitiven Interviews laut auszusprechen. Selbst wenn ihnen dies gelingen sollte, ist es möglich, dass der Prozess des lauten Denkens den „normalen“ Frage-Antwort-Prozess beeinflusst (vgl. Schaeffer & Presser, 2003, S. 82). So liegt es nahe, dass die Befragten aufgrund der Aufforderung laut zu denken weitaus mehr Anstrengungen unternehmen und ihre Gedanken weiter kreisen lassen, als sie dies vermutlich sonst bei der Beantwortung einer Frage getan hätten (vgl. Willis, 1999, S. 5). Dies ist mitunter auch der Grund, weshalb im Rahmen von kognitiven Interviews von vielen ForscherInnen zu meist die Technik des Probing eingesetzt wird

#### *Probing (Hinterfragen von Antworten)*

Bei der Methode des *Probing* handelt es sich um ein bewährtes Verfahren, mit dessen Hilfe das Frageverständnis überprüft wird (vgl. Kurz, Prüfer & Rexroth, 1999, S. 88). Bei dieser Technik stellen die InterviewerInnen eine oder mehrere Zusatzfragen (sogenannte *Probes*), um Hinweise bezüglich des pragmatischen Verständnisses und/oder des gesamten Frage-Antwort-Prozesses zu erhalten (vgl. Kurz, Prüfer & Rexroth, 1999, S. 90). Je nachdem, zu welchem Zeitpunkt die Methode eingesetzt wird, bezeichnet man das Verfahren folgendermaßen:

- (a) *Follow-Up-Probing* (auch *Concurrent Probing* genannt). Hierbei erfolgt das Probing direkt nach der Beantwortung jedes einzelnen Items (also direkt nach der Wahl einer Antwortkategorie) und somit noch während des Interviews;
- (b) *Post-Interview-Probing* (auch *Retrospective Probing* genannt). Bei dieser Variante werden die Nachfragen erst im Anschluss an das gesamte Interview gestellt, also nachdem für alle zu testenden Items eine Antwortoption ausgewählt wurde (vgl. Prüfer & Rexroth, 1996, S. 107; Willis, 1999, S. 7).

In der Praxis wird zumeist das Follow-Up-Probing eingesetzt, da die Informationen, die man erfragen möchte, sich dann noch unmittelbar im Gedächtnis des Befragten befinden. Stellt man die Nachfrage hingegen erst, nachdem alle Items dargeboten wurden, dann besteht die Gefahr, dass sich die TeilnehmerInnen nicht mehr daran erinnern können, was sie gedacht haben, als sie die Frage beantwortet haben.

Der Einsatz des Post-Interview-Probing kann dennoch sinnvoll sein, da die Fragephase erst nach der Beantwortung aller Items erfolgt und somit die Antwortsituation eher einer standardisierten Befragung im Feld entspricht. Wenn man einen Fragebogen in seiner Endphase der Entwicklung unter den (annähernd) gleichen Bedingungen wie im Feld testen möchte, kann es somit sinnvoll sein, ein Post-Interview-Probing einzusetzen. Auch im Zusammenhang mit Selbstausfüller-Fragebögen kann diese Vorgehensweise vorteilhaft sein, da so überprüft werden kann, ob die Befragten alleine mit den Instruktionen, der Filterführung, den Antwortskalen usw. zurechtkommen (vgl. Willis, 1999, S. 7).

Neben dem Zeitpunkt des Einsatzes gibt es noch ein weiteres Unterscheidungsmerkmal: den Zeitpunkt der Entwicklung. Willis (1999) differenziert hier zwei Nachfrage-Formen:

- (a) *Scripted Probes*: Diese Nachfragen werden *vor* dem Interview entwickelt und von allen InterviewerInnen erhoben.
- (b) *Spontaneous Probes*: Diese Nachfragen werden *während* des Interviews generiert und dementsprechend auch nur von einzelnen InterviewerInnen gestellt.

Der Einsatz von *Scripted Probes* eignet sich vor allem bei unerfahrenen InterviewerInnen, die von einem strukturierten „Leitfaden“ (siehe standardisierter Evaluationsfragebogen in Unterkapitel 2.6.1.4) profitieren (vgl. Willis, 1999, S. 8). Zudem gewährleisten sie einen Vergleich des erhobenen Materials, da in allen Interviews die gleichen Nachfragen gestellt wurden. Der große Vorteil von *Spontaneous Probes* ist, dass die (erfahrenen) InterviewerInnen direkt auf das Gesagte der Befragten eingehen können. Zeichnet sich also ein mögliches Problem im Verlauf des Gespräches ab, das im Vorfeld nicht antizipiert wurde, so können die InterviewerInnen mit Hilfe spontaner Nachfragen diesem auf den Grund gehen. Da beide Varianten somit unterschiedliche Stärken aufweisen, eignet sich am besten eine Kombination beider Nachfrage-Formen (vgl. Willis, 1999, S. 8 f.).

Unabhängig von dem Zeitpunkt, zu dem die Probing-Frage gestellt oder entwickelt wird, gibt es auf inhaltlicher Ebene eine Reihe verschiedener Nachfragen:

- (a) *Comprehension Probe* bzw. *Interpretation Probe* (Nachfrage zum Frageverständnis): „Was bedeutet der Begriff ... für Sie?“
- (b) *Information Retrieval Probe* bzw. *Recall Probe* (Nachfrage zu den Strategien der Informationsbeschaffung / Nachfrage zum Besitz relevanter Informationen): „Wie haben Sie sich daran erinnert, dass ...?“ oder „Wie sind Sie vorgegangen, um zu dieser Antwort zu gelangen?“
- (c) *Category Selection Probe* (Nachfrage zur Wahl der Antwortkategorie): „Warum haben Sie den Skalenwert ... gewählt?“
- (d) *General Probe* (Unspezifische Nachfrage): „Gibt es weitere Unklarheiten im Zusammenhang mit dieser Frage, die wir bisher noch nicht besprochen haben?“ oder „Hatten Sie bei dieser Frage noch mit anderen Aspekten Probleme?“ (vgl. Prüfer & Rexroth, 2005, S. 6 ff.; Willis, 1999, S. 6).<sup>22</sup>

Prüfer und Rexroth (2005) weisen im Zusammenhang mit den unspezifischen Nachfragen darauf hin, dass diese Technik nur effektiv ist, wenn sich die Befragten bewusst sind, dass ihnen ein bestimmter Aspekt der Frage Probleme bereitet. Dies ist jedoch nicht immer der Fall. Ergibt die Interpretation eines Items für die Befragten Sinn und fällt ihnen die Wahl der Antwortkategorie leicht, so werden sie vermutlich die oben angeführten unspezifischen Nachfragen verneinen. Trotzdem ist es möglich, dass sie – ohne es zu wissen – einen Begriff innerhalb des Items anders ausgelegt haben, als es intendiert war und somit ein Verständnisproblem vorlag („silent misunderstanding“, Willis, 2005, S. 210). Dies kann jedoch nur durch eine spezifische Nachfrage zur Interpretation des Begriffes aufgedeckt werden. Unspezifische Nachfragen eignen sich daher eher als abschließende Frage nach einer Reihe spezifischer Nachfragen. Auf diese Weise soll sichergestellt werden, dass kein weiteres Problem übersehen wurde, dessen sich die Befragten bewusst sind, welches jedoch im Vorfeld nicht von den ForscherInnen antizipiert oder von den InterviewerInnen während der Erhebung bemerkt wurde (vgl. Prüfer & Rexroth, 2005, S. 9 f.).

In der Gesamtbetrachtung liegt ein Vorteil der Methode des Probings in der einfachen Anwendbarkeit, die im Gegensatz zum lauten Denken auch keines Trainings bedarf. Die Befragten müssen lediglich auf die gestellten Nachfragen antworten. Willis (1999) zufolge ist es oftmals der Fall, dass die InterviewteilnehmerInnen nach einiger Zeit von selbst beginnen, spontan ihre Gedanken und ihre Kritik bezüglich der zu testenden Items zu äußern. Sie tun dies häufig unmittelbar nach der Darbietung des Items, also noch bevor die InterviewerInnen spezifische Nachfragen stellen. Die Interviewsituation ähnelt dann stark einem Think-Aloud (vgl. Willis, 1999, S. 6 f.). Auch in eigenen Interviews konnte diese Beobachtung gemacht werden (vgl. Schüßler, 2006).

---

<sup>22</sup> Willis (1999) klassifiziert die Methode des Paraphrasierens und die Technik zur Bewertung der Verlässlichkeit der Antwort ebenfalls als spezielle Nachfrage-Varianten.



Ein weiterer Vorzug dieser Technik ist, dass die InterviewerInnen (im Normalfall) durch das zielgerichtete Nachfragen den Verlauf des Gespräches steuern können. Irrelevante oder unproduktive Diskussionen können durch diese Vorgehensweise zumeist vermieden werden. Gleichzeitig besteht die Möglichkeit, sich auf diejenigen Aspekte zu konzentrieren, die als potentielle Ursache von Antwortfehlern in Frage kommen (vgl. Willis, 1999, S. 6). Je präziser dabei die Nachfragen auf den zu überprüfenden Aspekt zugeschnitten sind, desto wahrscheinlicher ist es auch, dass man tatsächliche Schwachstellen der Frage aufdeckt (vgl. Kurz, Prüfer & Rexroth, 1999, S. 90).

Bei der Formulierung der einzelnen Nachfragen muss jedoch aufgepasst werden, da sie zu einer Verzerrung der Antworten führen kann. Durch eine vorsichtige, nicht suggestive Wortwahl kann diese Gefahr minimiert werden (vgl. Kurz, Prüfer & Rexroth, 1999, S. 90; Willis, 1999, S. 7).

Gelegentlich werden zudem Bedenken hinsichtlich der Validität des (Follow-Up-) Probings angeführt, da durch die Nachfragen der Ablauf des Interviews verändert wird und er nicht mehr mit einer standardisierten Befragung im Feld identisch ist. Willis (1999) hält dem entgegen, dass die Technik des Verbal Probings nicht weniger unrealistisch sei als die alternative Methode des Concurrent-Think-Alouds. Zudem bestehe die Möglichkeit, die Nachfragen erst nach Durchführung des gesamten Interviews zu stellen (Post-Interview-Probing bzw. Retrospective Probing). Seines Erachtens ist die Kritik auch insofern zu vernachlässigen, als beide Erhebungen ein anderes Ziel hätten: Während kognitive Interviews Fragen analysierten, würden mit standardisierten Befragungen im Feld Daten erhoben (vgl. Willis, 1999, S. 7).

### **2.6.1.5 Auswertung**

Nach Abschluss der Durchführung der kognitiven Interviews liegt eine Fülle an Datenmaterial vor, das es auszuwerten gilt: So gibt es Tonbandaufzeichnungen und handgeschriebene Notizen zu den Ausführungen der Befragten während des Think-Aloud-Prozesses bzw. zu ihren Antworten, die sie auf die Nachfragen oder andere kognitive Techniken hin gegeben haben. Gleiches gilt auch für die spontanen Äußerungen der Befragten, die unmittelbar angebracht wurden, nachdem das zu testende Item präsentiert wurde. Des Weiteren liegen schriftliche Kommentare und Anmerkungen der InterviewerInnen vor, die sie während und/oder nach dem Interview angefertigt haben.

Bislang existiert kein allgemein akzeptiertes Regelwerk darüber, wie dieses Datenmaterial ausgewertet werden sollte. Den Ausführungen von Willis (1999, 2005) konnte entnommen werden, dass es vor allem in Bezug auf drei Aspekte unterschiedliche Vorgehensweisen gegeneinander abzuwägen gilt, die jeweils ihre eigenen Vorteile, aber auch Nachteile aufweisen:

- (a) Basis der Analyse: handgeschriebene Kommentare der InterviewerInnen vs. Tonbandaufzeichnungen und Transkripte;

- (b) Art der Analyse: informell vs. formell;
- (c) Fokus der Analyse und der Dokumentation der Ergebnisse: einzelnes Interview vs. einzelne Frage.

(a) *Basis der Analyse*

Zunächst muss eine Entscheidung getroffen werden, auf welcher Basis die Analyse der Interviews stattfinden soll: anhand der mitgeschriebenen Kommentare der InterviewerInnen oder der Tonbandaufzeichnungen und/oder der entsprechenden Transkriptionen.

Willis (2005) zufolge stellt die Verwendung der *handgeschriebenen Kommentare* die einfachste und zugleich kostengünstigste Form der Analyse dar (vgl. Willis, 2005, S. 156). Hierbei werden die Mitschriften, die die InterviewerInnen während (und auch nach) der Befragung anfertigen, in eine elektronische Version des Fragebogens unterhalb jedes einzelnen Items eingetragen und anschließend auf mögliche Formulierungsdefizite der Items oder andere Fehlerquellen hin untersucht (vgl. Willis, 2005, S. 161). An die InterviewerInnen stellt diese Vorgehensweise jedoch hohe (kognitive) Anforderungen, da sie gleichzeitig zuhören, das Gesagte gedanklich aufnehmen, es mitschreiben und ggf. spontane Nachfragen stellen müssen. Eine Alternative stellt daher eine Live-Evaluierung des Interviews durch eine zweite Person dar. Diese verfolgt das Gespräch und notiert die eigenen Eindrücke darüber, wie gut die Items funktionieren (vgl. Willis, 2005, S. 156).

Ebenso ist es möglich, das Interview aufzunehmen und im Nachhinein eine *Transkription* anzufertigen. Diese Vorgehensweise ist wünschenswert, da sie die detaillierteste Analyse erlaubt, allerdings ist sie auch sehr zeitintensiv und kostspielig. Zudem besteht die Gefahr, dass aufgrund der großen Datenmenge der Blick für das Wesentliche verlorengeht (vgl. Willis, 2005, S. 161).

Die Entscheidung, welches Material man für die Auswertungen verwenden möchte, hängt letztendlich von der Art der Fragebogentestung ab. Willis (1999) empfiehlt eine Auswertung der handgeschriebenen Kommentare, wenn es sich um eine Auftragsarbeit („production work“) handelt, bei der die Überarbeitungen der Fragen innerhalb einer kurzen Zeitspanne vorgenommen werden (müssen). Die Analyse der Interviewmitschnitte und/oder der Transkriptionen kann hingegen für Forschungszwecke sowie für die genaue Evaluation von Fragebögen bzw. deren Entwicklung nützlich sein (vgl. Willis, 1999, S. 28).

(b) *Art der Analyse*

Willis (2005) führt zwei verschiedene Formen der Analyse an: die informelle und die formelle. Bei der *informellen Analyse* werden die Ergebnisse der Befragung qualitativ ausgewertet. Dabei werden die Antworten der Befragten sowie die Kommentare der InterviewerInnen einzeln analysiert und im Hinblick auf mögliche Probleme, die mit der

Frageformulierung zusammenhängen, durchgesehen. Zentrale Fragemängel, die ausfindig gemacht werden, gilt es festzuhalten und zu beschreiben. Je präziser dabei die Defizite geschildert und mit den Merkmalen der Frage in Verbindung gebracht werden, desto hilfreicher ist dies für die spätere Überarbeitung des Items (vgl. Willis, 2005, S. 157). So liefert beispielsweise der Hinweis, dass eine Frage offensichtlich aufgrund einer bestimmten Formulierung falsch interpretiert wurde, genaue Anhaltspunkte für die Umformulierung.

Bei der *formellen Analyse* wird hingegen ein Kodierungsschema verwendet (vgl. Willis, 2005, S. 164). Conrad und Blair (1996) haben beispielsweise vorgeschlagen die verbalen Inhalte der kognitiven Interviews in eine Tabelle zu übertragen, die verschiedene Arten von „Problemtypen“ (lexikalische, zeitliche, logische, etc.) auf der einen Achse und die verschiedenen Stufen des Frage-Antwort-Prozesses (z.B. Verstehen der Frage, Formulieren der Antwort) auf der anderen Achse abbildet. Alle Item-Probleme, die im Rahmen der kognitiven Interviews ausfindig gemacht werden, müssen bei dieser Vorgehensweise dem entsprechenden Tabellenfeld zugeordnet und so als ein bestimmtes Problem gekennzeichnet werden (vgl. Conrad & Blair, 1996, S. 2 ff.).

Verglichen mit der informellen Analyse ist die formelle Analyse systematischer und auch objektiver, allerdings bringt sie auch erheblichen Zeitaufwand mit sich. Ein weiterer Nachteil ist, dass durch die Kodierung Informationen verloren gehen. So wird in die Tabelle lediglich eingetragen, dass im Zusammenhang mit dem Item ein Verständnisproblem aufgetreten ist. Im ursprünglichen Kommentar der InterviewerInnen wird hingegen genau beschrieben, welcher Begriff oder welcher Frageaspekt die Ursache des Problems war. Um das Item angemessen überarbeiten zu können, müsste folglich nach der formellen Analyse wieder der qualitative Kommentar herangezogen werden, was die Kodierung mehr oder weniger überflüssig werden lässt.

Daher sind qualitative Kommentare zumeist besser geeignet, um Probleme aufzudecken und beheben zu können, so dass in der Regel auch auf die informelle Vorgehensweise zurückgegriffen wird. Wenn Kodierungsschemata verwendet werden, geschieht dies eher zum Zweck einer Methodenevaluation (vgl. Willis, 2005, S. 167).

### (c) *Fokus der Analyse und Dokumentation der Ergebnisse*

Des Weiteren muss eine Entscheidung getroffen werden, worauf der Fokus bei der Analyse und der Dokumentation der Ergebnisse gelegt werden soll: auf die einzelnen Interviews oder die einzelnen Items. Im ersten Fall wird ein Bericht für jedes *einzelne Interview* angefertigt (vgl. Willis, 2005, S. 161). Da auf diese Weise die gesamte Verlaufsgeschichte der Befragung festgehalten wird, offenbaren sich mögliche Kontexteffekte eher.

Im zweiten Fall wird ein gemeinsamer Bericht zu jeweils einem *einzelnen Item* über alle Interviews hinweg erstellt. Dieser liefert eine zusammengefasste fragenspezifische Version der Ergebnisse. Dabei kann einerseits qualitativ festgehalten werden, welche Arten von Problemen im Zusammenhang mit jedem einzelnen Item aufgetreten sind, und andererseits quantitativ beschrieben werden, ob die identifizierten Mängel sich über die

verschiedenen Befragungen hinweg ähnelten und wie häufig sie insgesamt auftraten. Auf diese Weise erhält man einen Überblick über die zentralen Frage-Probleme, die sich während der kognitiven Interviews gezeigt haben (vgl. Willis, 2005, S. 161).

Prüfer und Rexroth (2005) geben für die Auswertung der kognitiven Interviews folgende Empfehlung ab: Für jede/n InterviewteilnehmerIn sollte eine itemspezifische Auflistung der Äußerungen angefertigt werden, die sich aus folgenden Informationen zusammensetzt:

- (1) der gewählten Antwortkategorie des bzw. der Befragten;
- (2) den spontanen Äußerungen des bzw. der Befragten, die noch vor dem Einsatz der kognitiven Techniken abgegeben wurden;
- (3) den Antworten, die auf die kognitiven Techniken hin gegeben wurden sowie
- (4) den Kommentaren der InterviewerInnen (vgl. Prüfer & Rexroth, 2005, S. 18).

Das Manko der itemspezifischen Analyse hinsichtlich der Kontexteffekte kann abgeschwächt werden, indem die InterviewerInnen nicht nur zu den einzelnen Items Kommentare anfertigen, sondern auch bezüglich des gesamten Interviews generelle Anmerkungen machen, die beispielsweise mögliche Reihenfolgeeffekte der Fragen ansprechen.

#### **2.6.1.6 Stärken und Schwächen des Verfahrens**

Da es sich bei der Methode der kognitiven Interviews um eine aktive Technik handelt, können dabei auch solche Probleme aufgedeckt werden, die den Befragten selbst nicht bewusst sind. Im Gegensatz dazu basiert beispielsweise der Beobachtungspretest (siehe Unterkapitel 2.6.4) auf einer passiven Vorgehensweise, so dass Frageprobleme bei dieser Methode auch nur eher zufällig gefunden werden können (vgl. Prüfer & Rexroth, 2000, S. 7). Der Vorteil kognitiver Interviews gegenüber der Methode des ExpertInnen-Reviews (siehe Unterkapitel 2.6.3) besteht darin, dass „reale“ Personen befragt werden.

Eine Grenze kognitiver Interviews liegt darin, dass sie in erster Linie dazu dienen, mögliche Probleme im Zusammenhang mit einzelnen Fragen zu evaluieren. Schwachstellen, die den gesamten Fragebogen betreffen (wie z.B. potentielle Probleme der InterviewerInnen mit den Instruktionen) können nicht bzw. nur schwer ausfindig gemacht werden, da der Fragebogen nicht in seiner eigentlichen Form unter Feldbedingungen erhoben wird (vgl. Prüfer & Rexroth, 1996, S. 109; Prüfer & Rexroth, 2000, S. 18).

Des Weiteren ist anzumerken, dass kognitive Interviews in der Regel nur mit kleinen Stichproben durchgeführt werden, die nicht zwangsläufig repräsentativ für die Zielpopulation sein müssen. Folglich lassen sich auch die identifizierten Frage-Probleme nicht einfach generalisieren (vgl. Groves et al. 2009, S. 270; Prüfer & Rexroth, 1996, S. 109). Dieser Kritik, dass kognitive Interviews im Vergleich zum Beobachtungspretest unzureichend sind, da ihre Fallzahlen zu klein sind, um begründete Schlussfolgerungen ziehen zu können, hält Willis (1999) Folgendes entgegen:

- (1) Der Zweck kognitiver Interviews bestehe nicht in statistischen Schätzungen. Folglich würden auch bei der Auswahl der Befragten keine großen Fallzahlen, sondern eine Mannigfaltigkeit einzelner Individuen angestrebt.
- (2) Die Natur kognitiver Interviews sei qualitativer und nicht quantitativer Art. So gehe es nicht darum, Frage-Probleme zu evaluieren, indem die Anzahl der Interviews, in denen Frage-Mängel vorgekommen sind, ausgezählt würden. Natürlich sei ein Formulierungsdefizit, das in jedem einzelnen Interview in gleicher Weise vorkomme, signifikant, dies könne jedoch auch für ein Problem gelten, das nur in einem einzigen Fall auftrete. Sei es einleuchtend und schwerwiegend, müsse es nicht erst mit Hilfe einer großen Fallzahl verifiziert werden, da es offensichtlich vorhanden sei und daher behoben werden sollte (vgl. Willis, 1999, S. 33).

Eine weitere Grenze kognitiver Methoden liegt darin begründet, dass sie auf mündlichen Äußerungen der Befragten beruhen, die die Gedanken nicht immer vollständig wiedergeben. So können nicht alle kognitiven Prozesse in Worte gefasst werden, da manche so schnell ablaufen, dass sie keine Spuren im Arbeitsgedächtnis hinterlassen. Erkundigt man sich z.B. bei den Befragten nach ihrem Geschlecht, so werden sie die Antwort sofort im Gedächtnis parat haben, ohne dass sie groß darüber nachdenken müssen. Folglich können die Befragten auch nicht angeben, wie sie zu dieser Antwort gelangt sind (vgl. Collins, 2003, S. 236). Es stellt sich daher die Frage, inwieweit kognitive Interviews dazu in der Lage sind, einen Zugang zu den Gedankengängen der Befragten zu schaffen, und inwieweit die verbalen Ausführungen der Befragten dann auch tatsächlich den Gedanken in einer „normalen“ Befragungssituation entsprechen (vgl. Willis, 2005, S. 208 f.).

Der Erfolg kognitiver Interviews basiert auf der Annahme, dass sie auch in der Tat dazu imstande sind, zentrale Defizite einer Frage ans Licht zu bringen. Damit dies gelingt, müssen die AuswerterInnen auch in der Lage sein, den Äußerungen der Befragten Hinweise auf problematische Aspekte zu entnehmen und diese dann mit den Merkmalen der Frage in Verbindung zu bringen (vgl. Beatty, 2004, S. 18 f.). Dabei muss die Möglichkeit in Erwägung gezogen werden, dass die Schlussfolgerungen der kognitiven Interviews lückenhaft oder nicht zutreffend sind – schließlich basieren sie auf menschlichen Einschätzungen.

Beatty (2004) zufolge ist es nicht möglich, mit Hilfe kognitiver Interviews *alle* potentiellen Defizite einer Frage aufzufindig zu machen (selbst wenn dies als ihr Zweck angesehen werden kann). So besteht stets die Möglichkeit, dass ein zusätzliches Interview zu einer neuen Erkenntnis führt oder andere InterviewerInnen oder AuswerterInnen ein weiteres Frage-Problem identifizieren. Demnach ist es unwahrscheinlich, dass kognitive Interviews immer dieselben Problemen aufdecken und somit zu reliablen Ergebnissen führen (vgl. Beatty, 2004, S. 20 f.). Auch Willis (2005) räumt ein, dass die Evaluation kognitiver Interviews diesbezüglich gemischt ausfällt. So verweist er mitunter auf eine eigene Studie, bei der es zwar gelungen sei aufzuzeigen, dass kognitive Interviews Probleme identifizieren, die auch bei standardisierten Befragungen im Feld auftauchen. Allerdings sei nicht

klar, ob die Aufdeckung der Ergebnisse auch in einer reliablen Art und Weise geschehe, da das grundlegende Prinzip der Wiederholbarkeit über verschiedene InterviewerInnen und Labore hinweg bislang noch nicht belegt wurde (vgl. Willis, 2005, S. 228).

So informativ die Methode der kognitiven Interviews folglich auch ist, sie kann zu Schlussfolgerungen führen, die unvollständig, irreführend oder auch nicht korrekt sind. Beatty und Willis (2007) nennen verschiedene potentielle Fehlerquellen:

- (1) Kognitive Interviews können Probleme aufzeigen, die im Rahmen einer „realen“ Befragung im Feld nicht vorkommen („falscher Alarm“).
- (2) Auch der umgekehrte Fall ist möglich: Probleme, die im Feld auftreten, werden nicht ausfindig gemacht.
- (3) Es ist möglich, dass unabhängige Gruppen von kognitiven InterviewerInnen zu unterschiedlichen Ergebnissen gelangen, wenn sie ein und denselben Fragebogen überprüfen (siehe hierzu z.B. DeMaio & Landreth 2004; Forsyth, Rothgeb, & Willis 2004). Beatty und Willis (2007) führen in diesem Zusammenhang an, dass dies mehrere Ursachen haben kann. Wird beispielsweise mit einer flexiblen Vorgehensweise gearbeitet, die es den InterviewerInnen erlaubt bei potentiellen Problemen selbst aktiv zu werden und spontan verschiedene kognitive Techniken einzusetzen, so ist es nicht unwahrscheinlich, dass verschiedene Probleme identifiziert werden. Ebenso ist es möglich, dass die Differenzen auf einen unterschiedlichen Hintergrund der beiden Gruppen sowie eine unterschiedliche Sensitivität in Bezug auf bestimmte Arten von Problemen zurückgeführt werden können (vgl. Beatty & Willis, 2007, S. 303).

Die ersten beiden Punkte sprechen Schwächen hinsichtlich der Validität kognitiver Techniken an, der letzte Punkt macht hingegen auf ein Reliabilitätsproblem des Verfahrens aufmerksam (vgl. Beatty & Willis, 2007, S. 302). Beatty und Willis (2007) merken hierzu an:

*“It seems unlikely that cognitive interviewing would generate reliable findings in the sense that survey researchers might use the term (i.e., each set of interviews identifies the same set of problems with questions). When findings are different, yet not necessarily contradictory, this may indicate that no one set of findings is complete, and they need to be examined more closely. Findings that are difficult to reconcile might indicate either faulty reasoning by analysts, or that interviewing has not yet yielded an adequate understanding of responses associated with a question. The former case calls for a closer look at the data, while the latter indicates a need for continued data collection”* (S. 303).

Kognitive Interviews können zur Evaluation von Items und auch zur Überprüfung der vorgenommenen Itemüberarbeitungen eingesetzt werden, allerdings können sie keine quantitativen Beweise dafür liefern, dass die neu formulierten Itemversionen besser sind als die Originale (vgl. Collins, 2003, S. 236). Um die Effektivität kognitiver Interviews

zu beurteilen, schlägt Willis (2005) für die zukünftige Forschung folgende mögliche „Gütemaße“ vor:

- (a) Untersuchung der Reliabilität unterschiedlicher Fragebogenversionen mit Hilfe von Reinterviews.
- (b) Record check studies, die objektive Angaben mit den berichteten Antworten auf die Fragen vergleichen, um auf diese Weise Unterschiede im Antwortverhalten aufzudecken.
- (c) Untersuchung von offensichtlichen Defiziten oder problematischen Daten in Feld-Studien (z.B. wie eine hohe Anzahl an „Weiß ich nicht“-Angaben oder anderer Missings) (vgl. Willis, 2005, S. 217).

Allerdings sind auch die Studien zur Effektivität kognitiver Interviews mit Vorsicht zu interpretieren: Zeigt sich die Überarbeitung eines Fragebogens nicht in einer Verbesserung der untersuchten Gütekriterien, kann dies an den mangelnden Fähigkeiten der Fragebogen-DesignerInnen gelegen haben. So sind kognitive Interviews zwar in der Lage spezifische Frageprobleme aufzudecken („*finding* a problem“), aber sie müssen nicht unbedingt auch einen Hinweis darauf liefern, welche Änderungen vorgenommen werden müssen, um die Frage zu verbessern („*fixing* a problem“). Letztendlich liegt es an den FragebogenentwicklerInnen eine Lösung zu finden, die zu einer besseren Fragevariante führt. Auch wenn die Quelle des Problems mit Hilfe der kognitiven Interviews aufgedeckt werden konnte, sind am Ende die Fähigkeiten der FragebogenentwicklerInnen ausschlaggebend dafür, ob eine Frage verbessert werden kann oder nicht (vgl. Willis, 2005, S. 215).

Nachdem nun die Methode der kognitiven Interviews ausführlich beschrieben wurde, wird im Folgenden auf die übrigen Pretest-Verfahren mit ihren individuellen Stärken und Schwächen eingegangen.

### 2.6.2 Fokusgruppen

Fokusgruppen werden üblicherweise in einer frühen Phase der Fragebogenentwicklung eingesetzt (noch bevor ein Instrument vorliegt, das es zu testen gilt). In der Regel werden sechs bis zehn TeilnehmerInnen aus der Zielpopulation rekrutiert, die gemeinsam über das Befragungsthema diskutieren sollen – dabei übernimmt eine Moderatorin bzw. ein Moderator eine leitende Funktion (vgl. Groves et al., 2009, S. 261). Auf diese Weise soll herausgefunden werden, was „normale Personen“ über die Materie, Schlüsselkonzepte und zentrale Sachverhalte des Fragebogens denken und welche Begriffe sie in diesem Zusammenhang verwenden. Fokusgruppen-Diskussionen dienen somit eher der Erforschung potentieller Themen und gängiger Begrifflichkeiten für einen Fragebogen als der detaillierten Evaluation bereits existierender Fragen (vgl. Willis, 2005, S. 233 f.; Statsitics Sweden, 2004, S. 26).

Da mehrere Personen zur gleichen Zeit befragt werden, sind Fokusgruppen hinsichtlich der Durchführung äußerst effizient (vgl. Prüfer & Rexroth, 1996, S. 110). Darin liegt jedoch auch ein Nachteil des Verfahrens begründet: Aufgrund des Gruppen-Settings ist die Methode nicht dazu geeignet, eine reale Befragungssituation zu simulieren und herauszufinden, wie einzelne Individuen konkrete Fragen verstehen und wie sie vorgehen würden, um diese zu beantworten (vgl. Groves et al., 2009, S. 263). Kognitive Interviews sind hingegen dazu in der Lage.

### 2.6.3 ExpertInnen-Review (Checklisten-Systeme)

Im Rahmen eines ExpertInnen-Reviews werden (mehrere) ExpertInnen zu Rate gezogen, um die Qualität einzelner Items und des Fragebogenentwurfs zu untersuchen. Einerseits kann es sich dabei um ExpertInnen bezüglich des Themenbereiches handeln, die die Items dahingehend überprüfen, ob ihre Inhalte dafür geeignet sind, die intendierten Konzepte zu erfassen. Andererseits können in Bezug auf die Fragebogenentwicklung ExpertInnen befragt werden, die die Items hinsichtlich der unter 2.6 angeführten Standards untersuchen. Im Zentrum des Interesses stehen dabei u.a. die Formulierungen der einzelnen Items, die Fragestruktur, die Antwortvorgaben, die Reihenfolge der Items sowie die Instruktionen für die InterviewerInnen bzw. die Befragten (vgl. Groves et al., 2009, S. 260). Eine Vergleichbarkeit der Einschätzungen kann etwa dadurch gewährleistet werden, dass ein Checklisten-System verwendet wird (vgl. Prüfer & Rexroth, 1996, S. 110).

Das Checklisten-System dient vor allem dazu, schlechte Fragen zu vermeiden, indem potentielle Probleme identifiziert und behoben werden, noch bevor andere Pretesting-Verfahren (wie etwa kognitive Interviews) zum Einsatz kommen. Willis und Lessler (1999) haben ein eigenes System entwickelt, das *Question Appraisal System* (QAS), welches auf Probleme im Zusammenhang mit persönlichen Befragungen (insbesondere am Telefon) spezialisiert ist.<sup>23</sup> Zum Großteil lässt es sich jedoch auch auf Selbstausfüller-Fragebögen übertragen. Im Folgenden wird es näher erläutert:

Der Fragebogen wird beim QAS auf mögliche Schwächen überprüft, indem er verschiedene Stufen der Analyse durchläuft. Dabei wird eine Vielzahl von bekannten Problemen berücksichtigt, die bei der Frageformulierung vermieden werden sollten. Jede einzelne Stufe widmet sich einer potentiellen Fehlerquelle. So zielt die dritte Stufe etwa auf die Klarheit und Verständlichkeit der Frage ab. Dabei gilt es mitunter folgende Punkte zu überprüfen: Sind einzelne Begriffe zu schwierig, unklar oder vage? Wird ein umständlicher, langer Satzbau verwendet? Gibt es mehrere Auslegungsmöglichkeiten in Bezug auf die Frage? Werden den Befragten die Intentionen der ForscherInnen deutlich? Können die Befragten verstehen, was mit der Frage gemeint ist (pragmatisches Verständnis)? Die

---

<sup>23</sup> In Anlehnung an das Checklisten-System von Willis und Lessler (1999) entwickelten Faulbaum, Prüfer und Rexroth (2009) mit dem *Bewertungssystem zur Qualitätsüberprüfung einzelner Fragen eines Fragebogens* (FBS) eine deutschsprachige Variante.



anderen Stufen des QAS beziehen sich beispielsweise auf die Lesefreundlichkeit, die Instruktionen oder die vorgegebenen Antwortkategorien (vgl. Willis, 2005, S. 23 ff.).

Zu den Vorteilen des ExpertInnen-Reviews zählen der kostengünstige Einsatz sowie die einfache Durchführbarkeit. Allerdings gewährleistet selbst der Einsatz von Checklisten-Systemen keine Übereinstimmung in den Urteilen der ExpertInnen. So können etwa zwei ExpertInnen, die ein und dieselbe Frage bewerten, zu unterschiedlichen Ansichten darüber gelangen, ob die Frage eindeutig formuliert ist oder ob die Aufgaben, die die Befragten zu bewältigen haben, zu schwierig sind. Ein weiteres Defizit dieses Verfahrens besteht darin, dass – im Gegensatz zu kognitiven Interviews – keine „realen Befragten“ eingesetzt werden und folglich auch keine Aussage darüber getroffen werden kann, welche Items etwa von den Befragten am exaktesten beantwortet werden können und welche somit die gewünschten Informationen liefern (vgl. Groves et al., 2009, S. 269 f.). Willis (2005) führt zudem als Schwäche von Checklisten-Systemen an, dass dabei die einzelnen Fragen im Vordergrund der Analyse stehen und nicht der Fragebogen als Ganzes. Das Verfahren kann somit keine Gesamteinschätzung des Instrumentes liefern (vgl. Willis, 2005, S. 27).

#### 2.6.4 Beobachtungspretest mit InterviewerInnen-Debriefing

Ein Beobachtungspretest (auch Standard-Pretest, klassischer oder konventioneller Pretest genannt) ist eine – zumeist einmalige – quantitative Testung des Fragebogens im Feld, wobei die Rahmenbedingungen denen der Hauptstudie so ähnlich wie möglich sein sollten (vgl. Prüfer & Rexroth, 2000, S. 14). Wie die Bezeichnungen bereits nahelegen, ist diese Form des Pretests die gängigste und älteste. Gleichwohl gibt es Prüfer und Rexroth (1996) zufolge bis heute keine einheitliche Konvention hinsichtlich der Vorgehensweise: Die Empfehlungen hinsichtlich der Fallzahl schwanken zwischen 10 und 200 Personen (Quoten- oder Zufallsstichprobe), die entweder über die Testcharakter aufgeklärt („participating pretest“) oder darüber im Unklaren gelassen werden („undeclared pretest“).

Die InterviewerInnen sollen während der Erhebung auf mögliche Probleme im Zusammenhang mit dem Fragebogen achten (z.B. Äußerungen oder Verhaltensweisen der Befragten, die eine Unsicherheit zum Ausdruck bringen wie etwa ein Zögern, die Bitte um Erklärungen oder eine Wahl der Antwortkategorie „weiß ich nicht“) und diese im Anschluss an das Interview berichten (vgl. Presser et al., 2004, S. 110; Prüfer & Rexroth, 1996, S. 98). Dies kann entweder schriftlich in einem *Erfahrungsbericht* bzw. *Pretest-Report* oder mündlich im Rahmen des sogenannten *InterviewerInnen-Debriefings* geschehen. Das Debriefing kann mit den InterviewerInnen einzeln oder mit allen gemeinsam in der Gruppe stattfinden (vgl. Prüfer & Rexroth, 1996, S. 98).

Diese Vorgehensweise basiert auf der Annahme, dass Probleme der Befragten bei der Antwortfindung auf ein mangelndes Frageverständnis und somit eine „schlecht konstruierte“ Frage hindeuten, wohingegen ein „formal korrektes“ Antwortverhalten als Indikator für eine „gut konstruierte“ Frage angesehen wird. Häufig trifft diese Schlussfolgerung

jedoch nicht zu. Selbst wenn ein formal korrektes Antwortverhalten gezeigt wird, kann es vorkommen, dass ein nicht intendiertes Frageverständnis seitens der Befragten vorliegt.<sup>24</sup> Folglich können die Fragen in einem solchen Fall auch nicht als gut konstruiert gelten, da sie offensichtlich nicht eindeutig genug formuliert sind (vgl. Prüfer & Rexroth, 1996, S. 98).

Als größte Schwäche des Beobachtungspretests kann daher seine „passive“ Vorgehensweise genannt werden. Mängel einer Frage können nur dann identifiziert werden, wenn sie von den Befragten selbst zur Sprache gebracht werden (vgl. Prüfer & Rexroth, 2000, S. 14). Dies ist jedoch nur selten der Fall, da Befragte problematische Fragen für sich häufig so abändern, dass sie für sie leichter zu beantworten sind. Handelt es sich beispielsweise um eine Verhaltensfrage, die sich auf einen Zeitraum bezieht, an den sich die Befragten nicht komplett erinnern können, ist es durchaus möglich, dass sie sich in ihren Antworten nur auf die Periode beziehen, über die sie eine Aussage treffen können (vgl. Prüfer & Rexroth, 2005, S. 3). Bei der Methode der kognitiven Interviews wird daher, wie bereits geschildert, „aktiv“ vorgegangen, um Frage-Probleme aufdecken zu können, die den Befragten selbst nicht bewusst sind.

Ein weiteres Problem im Zusammenhang mit dem Beobachtungspretest besteht in den überwiegend unspezifischen InterviewerInnen-Anweisungen hinsichtlich der Vorkommnisse, die während der Befragung beobachtet und später berichtet werden sollen. Selbst wenn eine intensive Schulung stattgefunden hat, sind die InterviewerInnen zudem nicht in der Lage, alle Auffälligkeiten wiederzugeben, die sich während des Beobachtungspretests ereignet haben (vgl. Prüfer & Rexroth, 1996, S. 99). Informationen, die mit Hilfe dieser Methode gewonnen werden, sind somit eher zufälliger Art (vgl. Prüfer & Rexroth, 2000, S. 14).

Beobachtungspretests eignen sich folglich weniger zur Identifizierung von Verständnisproblemen einzelner Fragen, sondern eher zum Aufdecken von Problemen, die mit dem Einsatz des gesamten Fragebogens unter realen Feldbedingungen verbunden sind. Zum einen sind damit Schwierigkeiten gemeint, die InterviewerInnen beispielsweise mit der Filterführung, den Instruktionen o.ä. haben, und zum anderen Probleme, die sich für die Befragten ergeben, wie etwa eine nachlassende Konzentration oder Motivation.

Positiv zu erwähnen ist im Zusammenhang mit diesem Verfahren die Tatsache, dass es eine relativ genaue Schätzung ermöglicht, welchen zeitlichen Rahmen der Fragebogen im Feld in Anspruch nehmen wird, da der Beobachtungspretest unter den gleichen Bedingungen wie die Haupterhebung stattfindet (vgl. Prüfer & Rexroth, 2000, S. 14). Des Weiteren erlauben die gegebenen Antworten der Befragten, die während des Beobachtungspretests miterhoben werden, eine weiterführende quantitative Analyse der Pretest-Daten (vgl. Groves et al., 2009, S. 265 f.). Für einen Teil der Analysen ist allerdings eine ausreichend große Stichprobe Voraussetzung (siehe Unterkapitel 2.6.7).

---

<sup>24</sup> Derartige Situationen konnten im Rahmen der kognitiven Interviews (Pretest-Phase 2), die zur Entwicklung und Validierung der FWSEM-Skala durchgeführt wurden, mehrfach beobachtet werden.

Ein weiterer Vorteil dieses Verfahrens besteht darin, dass die Kosten und der organisatorische Aufwand vergleichsweise gering ausfallen – dies trifft vor allem dann zu, wenn keine randomisierte Auswahl der Befragten vorgenommen, sondern stattdessen mit einer Quotenstichprobe gearbeitet wird (vgl. Prüfer & Rexroth, 1996, S. 99).

### 2.6.5 Behavior Coding

Das Behavior Coding kann als Ergänzung zum Beobachtungspretest dienen (vgl. Schaeffer & Presser, 2003, S. 82), bei dem die Interaktionen zwischen InterviewerInnen und Befragten systematisch analysiert werden. (Dies kann entweder während des Interviews oder anhand einer Aufzeichnung im Anschluss daran geschehen.) Dabei wird das Frage- und Antwortverhalten mit Hilfe eines bestimmten Codesystems klassifiziert. Die vergebenen Codes sollen dabei Auskunft über die Qualität und mögliche Probleme im Zusammenhang mit einzelnen Fragen bzw. dem gesamten Messinstrument geben (vgl. Prüfer & Rexroth, 1996, S. 99).

Beispielhaft soll das Codesystem von Cannell, Oksenberg, Kalton, Bishoping und Fowler (1989) angeführt werden, demzufolge das InterviewerInnenverhalten folgendermaßen eingruppiert wird: (a) als exaktes Vorlesen des Fragetextes, (b) als leichte Abwandlung, durch die die intendierte Bedeutung der Frage nicht tangiert wird, (c) als starke Abweichung vom Fragetext, die die Bedeutung verändert oder (d) als Abbruch (der Interviewer konnte die Frage nicht zu Ende vorlesen, da ihn die/der Befragte unterbrochen hat). Das Antwortverhalten der Befragten wird hingegen u.a. folgenden Kategorien zugeordnet: (a) adäquat, (b) adäquat, aber eine Unsicherheit wird angedeutet (c) inadäquat (Befragte/r gibt eine Antwort, die nicht dem Frageziel entspricht), (d) Wiederholung oder Klärung der Frage wird erbeten oder (e) Wahl der Antwortkategorie „weiß nicht“ oder Antwort verweigert (vgl. Cannell, Oksenberg, Kalton, Bishoping & Fowler, 1989, S. 56).

Neben dem verwendeten Codesystem muss im Rahmen dieser Technik geklärt werden, welches Verhalten analysiert werden soll: nur die erste Reaktion der Befragten, nachdem ihnen das Item vorgelesen wurde oder aber alle Reaktionen. Wählt man die zweite Variante, muss zudem entschieden werden, ob die gleichen Codes mehrfach oder nur einmal verwendet werden dürfen – also auch in solchen Fällen, in denen das entsprechende Verhalten häufiger gezeigt wurde (vgl. Prüfer & Rexroth, 1996, S. 100 f.)

Der Vorteil dieser Technik liegt eindeutig in ihrem objektiven, systematischen Vorgehen, welches reliable Ergebnisse liefert (vgl. Schaeffer & Presser, 2003, S. 82). Durch die quantitative Vorgehensweise erlaubt das Verfahren zudem objektive Aussagen über das Ausmaß der Probleme (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 645).

Allerdings weist das Behavior Coding auch verschiedene Nachteile auf: Zum einen erhält man – anders als im Falle kognitiver Techniken – keine Informationen über die Ursachen der Probleme im Zusammenhang mit den Fragen und kann sie dementsprechend auch nicht beheben. Zum anderen setzt auch diese Technik voraus, dass sich die

Befragten der Verständnisprobleme bewusst sind. Trifft dies nicht zu, so sind diese auch nicht aus ihrem Verhalten abzulesen. Selbst wenn also formal korrekte Antworten gegeben werden, können (unbewusste) Probleme im Zusammenhang mit der Frage bestehen, die möglicherweise zu Antwortverzerrungen (Abweichungen vom „wahren Wert“) führen (vgl. Prüfer & Rexroth, 1996, S. 101). Zudem ist der Aufwand für diese Methode relativ groß, da entweder eine zeitgleiche Evaluierung des Interviews durch eine weitere Person vorgenommen wird oder ein nachträgliches Anhören der Aufzeichnungen zur Vercodung erforderlich ist (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 645).

### 2.6.6 Befragten-Debriefing

Wie bereits erwähnt, liegen dem Beobachtungspretest und dem Behavior Coding eine passive Vorgehensweise zugrunde, die nicht alle Fragen-Probleme offenbart, da trotz eines formal korrekten Antwortverhaltens eine nicht intendierte Interpretation des Items vorliegen kann. Zudem bleiben die Ursachen derjenigen Probleme, die identifiziert wurden, in der Regel ungewiss. Um diese Defizite auszugleichen, kann als ergänzendes Verfahren ein Befragten-Debriefing (respondent debriefing) durchgeführt werden. Dabei werden den Befragten gezielte Nachfragen gestellt, um herauszufinden, ob die Items im intendierten Sinne aufgefasst wurden. So wird beispielsweise gefragt, wie die befragten Personen einzelne Begriffe oder auch das gesamte Item aufgefasst haben, wie sie sich für ihre Antwortkategorie entschieden haben, ob die gewählte Antwortvorgabe auch die eigene Sichtweise treffend wiedergibt und wie bedeutsam es für sie grundsätzlich ist, eine exakte Antwort zu liefern. Das Debriefing kann dabei entweder in Form von strukturierten Follow-up-Einzelinterviews, in Form von Gruppeninterviews oder auch mit Hilfe eines Feedback-Fragebogens, der von den Befragten selbständig auszufüllen ist, erfolgen. Die Debriefings werden in der Regel erst am Ende des Beobachtungspretests oder sogar einige Tage danach durchgeführt (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 645 f.), damit der normale Ablauf des Interviews und die Antworten der Befragten nicht durch die Inhalte der Nachfragen verzerrt werden. Schließt die Debriefing-Phase unmittelbar an den Beobachtungspretest an, so müssen die Befragten (beispielsweise durch eine Instruktion) darauf aufmerksam gemacht werden, dass es bei den folgenden Nachfragen darum geht, den Fragebogen zu verbessern und sich ihre Befragten-Rolle nun zu einer InformantInnen-Rolle wandelt (vgl. DeMaio, Mathiowetz, Rothgeb, Beach, & Durant, 1993, S. 41).

Im Gegensatz zu den Pre-Field-Techniken, die in der Regel mit einer kleinen Fallzahl arbeiten und qualitative Informationen hervorbringen, stellt das Befragten-Debriefing eine quantitative Methode dar, um Probleme im Zusammenhang mit dem Fragebogen aufzudecken. Dies ermöglicht eine objektive Dokumentation der Existenz und des Ausmaßes eines Problems.

Um die Belastung der Befragten zu minimieren und einer möglichen Ermüdung vorzubeugen, ist es nötig, das Debriefing auf nur wenige problematische Items zu beschränken. (Folglich ist keine so breite Testung wie etwa bei der Methode der kognitiven Interviews möglich). Neben der Entscheidung, welche Items getestet werden sollen, ist es für den Erfolg des Debriefings ebenso wichtig, die richtigen Zusatzfragen zu stellen. Fragebogen-DesignerInnen und ForscherInnen müssen somit eine klare Vorstellung davon haben, welche Items möglicherweise Probleme bereiten und worin potentielle Schwierigkeiten bestehen könnten, damit sie eine geeignete Auswahl der zu testenden Items treffen und entsprechende Nachfragen entwickeln können. Hinweise darauf können die sorgfältige Analyse des Fragebogens, die Ergebnisse von vorhergehenden Pre-Field-Tests (wie z.B. kognitive Interviews), Beobachtungen aus vorangegangenen Interviews oder auch bereits durchgeführte quantitative Datenanalysen der Umfrage geben (vgl. DeMaio et al., 1993, S. 35 ff.).

### 2.6.7 Analyse von Pretest-Daten

Die Daten, die während eines Feld-Tests gewonnen werden, können ebenfalls (vorausgesetzt, die Stichproben sind groß genug) genutzt werden, um die Qualität eines Fragebogens zu evaluieren.<sup>25</sup> Verschiedene Arten der Analyse sind aufschlussreich:

- (a) *Item-Nonresponse-Raten.* Die Analyse von partiellen Antwortausfällen kann in zweierlei Hinsicht informativ sein: Erstens können hohe „Weiß ich nicht“-Raten ein Indikator dafür sein, dass ein Item für die Befragten zu schwierig zu beantworten ist. Zweitens können Antwortverweigerungs-Raten aufzeigen, ob bestimmte Fragen oder bestimmte Versionen einer Frage als zu vertraulich empfunden werden, um sie im Rahmen einer Befragung zu beantworten (siehe Unterkapitel 2.5.1.4) (vgl. DeMaio et al., 1993, S. 57).
- (b) *Antwortverteilungen.* Antwortverteilungen geben die Häufigkeiten wieder, mit denen die einzelnen Antwortkategorien eines Items von den Befragten gewählt wurden. Sie können u.a. dazu genutzt werden, um zu bestimmen, ob unterschiedliche Formulierungen einer Frage oder unterschiedliche Fragereihenfolgen zu unterschiedlichen Antwortverteilungen führen (siehe Split-Ballot-Test in Unterkapitel 2.6.10). Antwortverteilungen sollten jedoch niemals alleine verwendet werden, um die Modifikationen einer Frage zu testen, da sie keine Auskunft darüber geben, welche Frageversion im intendierten Sinne verstanden wird. Diese Form der Analyse ist nur in Verbindung mit anderen Evaluations-Methoden (wie

---

<sup>25</sup> „Analysis of pretest data is generally done using raw data (unedited and unimputed), because the focus is on how well the respondent understands and responds to the questions. Editing and imputation improve the quality of the data and thereby have the potential to mask questionnaire problems” (DeMaio et al., 1993, S. 57).

z.B. dem Befragten-Debriefing) sinnvoll (vgl. DeMaio et al., 1993, S. 57 ff.). Auch Prüfer und Rexroth (1996) sind der Ansicht, dass die Häufigkeitsverteilungen nur einen groben Hinweis auf die Qualität einer Frage geben können. Sind einzelne Antwortkategorien überhaupt nicht oder nur in sehr geringen Maße besetzt oder handelt es sich um eine extreme Antwortverteilung, so kann dies ein erstes Anzeichen für etwaige Fragemängel sein (vgl. Prüfer & Rexroth, 1996, S. 104).

- (c) *Reliabilität*. Mit der Reliabilität ist die Zuverlässigkeit gemeint, mit der ein Messinstrument ein Merkmal misst. Um diese zu ermitteln, kommen vor allem zwei Verfahren zum Einsatz: (1) Entweder wird den Befragten die gleiche Frage zweimal zur Beantwortung vorgelegt und die Antworten werden miteinander korreliert (Retest-Reliabilität) oder (2) es werden innerhalb eines Interviews mehrere Items eingesetzt, um das Konstrukt zu erfassen. In diesem Fall wird die interne Konsistenz der Antworten (in der Regel ermittelt über Cronbachs Alpha) über die verschiedenen Items hinweg bestimmt (vgl. Groves et al. 2009, S. 281 ff.; Rammstedt, 2004, S. 5 ff.).
- (d) *Validität*. Die Validität (Gültigkeit) gibt den Grad der Genauigkeit an, mit der ein Messinstrument das misst, was es zu messen beabsichtigt. Die Bestimmung der Validität kann zum einen mit Hilfe eines Außenkriteriums erfolgen. Dabei wird ermittelt, inwieweit die Antworten, die im Rahmen des Fragebogens gegeben werden, mit den Ergebnissen eines Außenkriteriums übereinstimmen (Kriteriumsvalidität). Zum anderen kann untersucht werden, inwieweit das Messinstrument das zugrunde liegende Konstrukt (Merkmal) erfasst, in dem überprüft wird, ob die durch die Umfrage erhobenen Daten verschiedenen theoretischen Annahmen folgen (Konstruktvalidität). So können beispielsweise Korrelationen des zu testenden Messinstrumentes mit anderen Instrumenten, die das gleiche oder ein verwandtes Konstrukt erfassen, untersucht oder eine Dimensionalitätsüberprüfung mit Hilfe einer Faktorenanalyse durchgeführt werden (vgl. Groves et al. 2009, S. 274 ff.; Rammstedt, 2004, S. 17 ff.).

### 2.6.9 Response Latency

Bei der Response-Latency-Methode wird die Zeit, die zwischen der Darbietung des Items und der Wahl einer Antwortkategorie vergeht, erfasst. Dies kann entweder durch eine exakte Messung oder durch eine subjektive Schätzung der InterviewerInnen („Qualitative Timing“ genannt) geschehen. Die Annahme, die diesem Verfahren zugrunde liegt, ist, dass lange „Reaktionszeiten“ einen Hinweis auf Fragedefizite liefern (vgl. Prüfer & Rexroth, 1996, S. 109).

Die einfache und kostengünstige Anwendung dieser Methode (insbesondere im Rahmen von computerbasierten Befragungen) kann als Vorteil angesehen werden (vgl. Presser et al., 2004, S. 117). Der eindeutige Nachteil dieses Verfahrens liegt darin, dass lange

Reaktionszeiten nicht zwangsläufig auf ein Frage-Problem hindeuten müssen. So kann es verschiedene Gründe dafür geben, weshalb die Wahl einer Antwortkategorie bei einem Item etwas mehr Zeit in Anspruch nimmt, als dies bei anderen Items der Fall ist. So ist es möglich, dass lange Reaktionszeiten nicht auf Schwierigkeiten im Zusammenhang mit einer Frage zurückzuführen sind, sondern auf einen sorgfältigen, gewissenhaften Prozess der Antwortfindung auf Seiten der Befragten (vgl. Presser et al., 2004, S. 117). So kann es etwas länger dauern (a) sich eine Meinung zu bilden, wenn es sich um eine Thematik handelt, mit der sich die Befragten zuvor noch nicht auseinandergesetzt haben, oder wenn (b) ein erfragtes Verhalten schwer zu rekonstruieren ist, weil es z.B. sehr häufig, aber unregelmäßig vorkommt (siehe Unterkapitel 2.5.1) (vgl. Schüßler, 2006, S. 25). Allein sollte dieses Verfahren daher nicht angewendet werden, sondern nur in Kombination mit anderen Methoden.

### 2.6.8 Reinterviews

Für Fowler (2001) liefern Reinterviews einen Beitrag zur Evaluation von Fragen, der schwer mit Hilfe eines anderen Verfahrens abgedeckt werden kann. So werden beim Reinterview den Befragten die gleichen Items zu zwei Messzeitpunkten vorgelegt, was einen Vergleich der Antworten ermöglicht. Auf diese Weise kann die Stabilität des Antwortverhaltens und somit die Reliabilität der Antworten überprüft werden. Items, von denen man annimmt, dass sie etwas Beständiges messen, die allerdings eine geringe Retest-Reliabilität besitzen, sollten entweder überarbeitet oder aus der Skala entfernt werden.

Wird das Reinterview im Rahmen eines Computergestützten Interviews (computer-assisted interviewing) eingesetzt, eröffnet es eine weitere Möglichkeit: Die InterviewerInnen können (sofort, nachdem die zweite Antwort abgegeben wurde, oder erst am Ende des Reinterviews) über ein nicht konstantes Antwortverhalten informiert werden. Bei den entsprechenden Items können die InterviewerInnen dann nachhaken und mit Hilfe verschiedener Nachfragen den Frage-Antwort-Prozess näher untersuchen und möglichen Ursachen für die Unterschiede im Antwortverhalten auf den Grund gehen. Diese Methode stellt somit eine wertvolle Ergänzung zu kognitiven Interviews dar, da sie dazu in der Lage ist, mögliche Probleme und Unsicherheiten seitens der Befragten, die zu einem inkonsistenten Antwortverhalten führen, unter realistischen Feldbedingungen aufzudecken (vgl. Fowler, 2001, S. 61 f.).

### 2.6.10 Split-Ballot

Bei Split-Ballot-Experimenten werden verschiedene Varianten eines oder mehrerer Items (z.B. mit unterschiedlichen Formulierungen oder Antwortmöglichkeiten) bzw. eines Fragebogens (z.B. mit einer unterschiedlichen Reihenfolge der Items) oder verschiedene Erhebungsmodi (z.B. Papier vs. Online) in unterschiedlichen Teilstichproben getestet (vgl.

Blanke, Gauckler, & Sattelberger, 2008, S. 646). Hinsichtlich der Anzahl der Personen, die im Rahmen eines Split-Ballot-Experimentes befragt werden sollten, treffen Blanke, Gauckler und Sattelberger (2008) folgende Aussage: „Der Stichprobenumfang sollte je Subgruppe ausreichend groß sein, sodass statistisch gesicherte Rückschlüsse auf die Grundgesamtheit gezogen werden können“ (Blanke, Gauckler, & Sattelberger, 2008, S. 646). Prüfer und Rexroth (1996) geben an, dass die Fallzahl für dieses Verfahren daher mindestens bei 100 liegen sollte (vgl. Prüfer & Rexroth, 1996, S. 104). Groves et al. weisen darauf hin, wenn es wichtig ist, bereits relativ kleine Unterschiede zwischen den Teilstichproben nachzuweisen, so seien große Stichproben nötig (vgl. Groves et al., 2009, S. 268). Zudem ist eine zufällige Zuweisung der Befragten auf die Teilstichproben unabdingbar, damit die Unterschiede in den Ergebnissen auch auf die verschiedenen Fragebogen-Varianten bzw. Erhebungsmodi zurückgeführt und andere Ursachen (differential sampling error or incomparable samples) ausgeschlossen werden können (vgl. Blanke, Gauckler, & Sattelberger, 2008, S. 646).

Mit Hilfe eines Split-Ballot-Experimentes können unterschiedliche Ziele verfolgt werden. Es kann getestet werden,

- (a) welche Variante einer Frage bzw. eines Fragebogens jeweils die beste ist;
- (b) ob die Überarbeitung eines Items auf Basis vorangegangener Pretest-Ergebnisse zu einer Verbesserung psychometrischer Kennwerte (survey estimates) geführt hat;
- (c) ob es Skaleneffekte, Mode-Effekte, Kontexteffekte o.ä. gibt (vgl. DeMaio et al., 1993, S. 55; Presser et al., 2004, S. 118).

Zur Evaluierung der Ergebnisse des Experimentes können beispielsweise die partiellen Antwortausfälle (item nonresponse) untersucht oder die Antwortverteilungen herangezogen werden. Hinweise auf Fragemängel sind u.a. eine hohe Missing-Rate (z.B. Verweigerung der Antwort oder die Wahl der Antwortkategorie „weiß nicht“) oder schiefe Verteilungen (vgl. Prüfer & Rexroth, 1996, S. 104). Des Weiteren können auch die Informationen der bereits vorgestellten Pretest-Verfahren wie etwa des InterviewerInnen- und Befragten-Debriefings oder des Behavior Codings genutzt und die Ergebnisse in beiden Teilstichproben miteinander verglichen werden (vgl. DeMaio et al., 1993, S. 56 f.).

### 2.6.11 Fazit Pretest-Verfahren

Die verschiedenen Methoden, die vorgestellt wurden, variieren in der Art der Frage-Probleme, für deren Identifizierung sie sich am besten eignen. Keines der Verfahren kann alle Standards, die eine gute Frage erfüllen sollte, überprüfen. Im Folgenden sind diejenigen



Verfahren aufgelistet, die am besten dazu geeignet sind, eine Frage im Hinblick auf die einzelnen Standards zu evaluieren (siehe Tabelle 3 für einen Gesamtüberblick):<sup>26</sup>

- (1) *Inhaltliche Standards: Zielen die Fragen auf die richtigen Informationen ab? Werden die richtigen Fragen gestellt?* Der inhaltliche Standard bezieht sich auf zwei Aspekte: Ersten muss die Frage aus der Perspektive der ForscherInnen die Informationen liefern, die die Forschungsziele abdecken. Der einzige Weg, um dies herauszufinden ist es, inhaltliche ExpertInnen zu befragen (ExpertInnen-Review). Zweitens geht es darum, zu erfassen, wie gut die Befragten auch in der Lage sind, die Fragen zu beantworten. Um dem nachzugehen, eignen sich vor allem Fokusgruppen-Diskussionen und die Methode der kognitiven Interviews. Mit Hilfe von Fokusgruppen kann herausgefunden werden, was die Befragten über die angesprochenen Themen wissen. Kognitive Interviews können Auskunft darüber geben, ob ein bestimmtes Set von Fragen in konsistenter Weise beantwortet werden kann und ob die Antworten auf die Fragen auch die Informationen liefern, nach denen die ForscherInnen suchen (vgl. Groves et al., 2009, S. 268 f.).
- (2) *Kognitive Standards: Können die Befragten die Frage verstehen und sie auch beantworten?* Um dies zu untersuchen, eignen sich kognitive Interviews am besten, da sie genau zu diesem Zweck konstruiert werden. Nach Groves et al. (2009) können drei weitere Verfahren zumindest Hinweise auf mögliche kognitive Probleme (z.B. mehrdeutige Wörter und Konzepte sowie Aufgaben, die für Befragte schwer zu bewältigen sind) im Zusammenhang mit einer Frage liefern: (a) Fokusgruppen-Diskussionen, (b) ExpertInnen-Reviews, (c) Behavior Codings (vgl. Groves et al., 2009, S. 269). Fowler (2001) ist der Meinung, dass darüber hinaus auch (d) das InterviewerInnen-Debriefing, (e) das Befragten-Debriefing sowie (f) Reinterviews dazu beitragen können, uneindeutige und schwer zu beantwortende Items ausfindig zu machen (vgl. Fowler, 2001, S. 60). Darüber hinaus ist die Response-Latency-Methode zu nennen.

Im Vergleich zu kognitiven Interviews weisen diese Verfahren folgende Schwächen auf: Aufgrund des Gruppen-Settings sind Fokusgruppen weniger gut dazu geeignet, um eine reale Befragungssituation zu simulieren und herauszufinden, wie einzelne Befragte die Items verstehen und wie sie vorgehen, um sie zu beantworten. Beim ExpertInnen-Review werden wiederum überhaupt keine „realen Befragten“ eingesetzt. Dem InterviewerInnen-Debriefing und dem Behavior Coding liegt eine „passive“ Vorgehensweise zugrunde, die Frageprobleme nur aufdecken kann, wenn diese den Befragten bewusst sind. Selbst in diesen Fällen

---

<sup>26</sup> Groves et al. (2009) beziehen sich in ihrem Text nur auf drei Standards (Inhaltliche und Kognitive Standards sowie die Usability) und Fowler (2001) gibt lediglich in einer Tabelle einen Überblick über die Eignung einzelner Verfahren ohne jedoch nähere Erläuterungen hierzu abzugeben. Folglich wurden in Bezug auf die übrigen vier Standards (Interpersonell, Psychometrisch, Multi-Language, Multi-Mode) eigene Schlussfolgerungen angestellt.

werden die Ursachen der Probleme nicht immer ausfindig gemacht. Um die Belastung der Befragten zu reduzieren, ist es nötig, sich beim Befragten-Debriefing auf einige wenige problematische Items zu beschränken. Dies ist insofern von Nachteil, als dass es nicht immer möglich ist, potentielle Schwierigkeiten im Vorfeld zu antizipieren und dementsprechend die „richtige“ Auswahl der zu testenden Fragen zu treffen. Am wenigsten hilfreich ist vermutlich die Response-Latency-Methode, da lange Reaktionszeiten nicht zwangsläufig auf ein Frage-Problem hindeuten müssen, sondern auch andere Gründe haben können. Zudem liefert sie keinerlei Hinweise darauf, worin mögliche Frageprobleme bestehen und wie diese behoben werden könnten.

- (3) *Interpersonelle Standards: Sind die Fragen so formuliert, dass sie einerseits von den InterviewerInnen genauso gestellt werden können, wie sie niedergeschrieben sind und sie andererseits auch nicht zu größeren Nachfragen auf Seiten der Befragten führen? Ergibt sich durch die vorgegebene Struktur und die Inhalte des Fragebogens eine Interaktion, die den Befragten angemessen erscheint?* Das Behavior Coding ist nach Ansicht von Fowler (2001) die beste Methode, wenn es darum geht, die Interaktionen zwischen InterviewerInnen und Befragten und ihr jeweiliges Verhalten systematisch zu analysieren. Allerdings erhält man keine Informationen hinsichtlich der Ursachen möglicher Probleme, so dass InterviewerInnen-Debriefing und Befragten-Debriefing als ergänzende Methoden hilfreich sein können.
- (4) *Usability: Ist der Fragebogen für die AnwenderInnen (die InterviewerInnen oder auch die Befragten) einfach und selbsterklärend konstruiert?* Herauszufinden, wie gut ein Fragebogen sich im Feld einsetzen lässt, ist das primäre Ziel des Beobachtungspretests (mit anschließendem InterviewerInnen-Debriefing). Zusätzlich kann im Vorfeld ein ExpertInnen-Review die Fragen oder Instruktionen identifizieren, die den InterviewerInnen und oder den Befragten Probleme bereiten (vgl. Groves et al., 2009, S. 269). Fowler (2001) attestiert zudem der Methode des Befragten-Debriefings eine Brauchbarkeit in diesem Zusammenhang (vgl. Fowler, 2001, S. 60).
- (5) *Psychometrische Standards: Weisen die Antwortverteilungen eine gewisse Variation auf? Führen die Fragen zu Antworten, die das, was beschrieben und erklärt werden soll, valide und reliabel messen?* Hinweise bezüglich psychometrischer Standards liefern die drei folgenden Verfahren: Die Analyse von Pretest-Daten, Split-Ballot-Experimente sowie Reinterviews. Um eine Auskunft über Häufigkeitsverteilungen zu erhalten, ist vor allem eine Analyse der Pretest-Daten sinnvoll. Auch hinsichtlich der Bestimmung der Reliabilität sollten Pretest-Daten analysiert werden, um die interne Konsistenz von Multi-Item-Skalen zu bestimmen. Darüber hinaus können Reinterviews sinnvoll sein, wenn es sich um Skalen handelt, die ein stabiles Merkmal erfassen sollen und/oder die nur aus ein oder zwei Items bestehen (Retest-Reliabilität) (vgl. Rammstedt, 2004, S. 24). Des Weiteren

kann mit Hilfe der Pretest-Daten die Kriteriums- und Konstruktvalidität überprüft werden. Daneben sind Split-Ballot-Experimente empfehlenswert, wenn nicht nur eine Version eines Items/Fragebogens getestet werden soll, sondern verschiedene Versionen im Hinblick auf ihre psychometrischen Kennwerte verglichen werden sollen.

- (6) *Multi-Mode-Capability: Eignen sich die Fragen für unterschiedliche Erhebungsmodi? Sind die Daten über verschiedene Erhebungsmodi hinweg vergleichbar?* Fowler (2001) attestiert den Split-Ballot-Experimenten in diesem Zusammenhang den größten Nutzen. Schließlich ist es ein Anliegen dieses Verfahrens herauszufinden, ob sich ein Fragenbogen beispielweise im Rahmen eines Telefoninterviews ebenso gut einsetzen lässt wie bei einer Online-Erhebung oder ob sich ein Darbietungsmodus negativ auf die Qualität der Daten auswirkt. Neben den Split-Ballot-Experimenten sieht er das ExpertInnen-Review als hilfreiche Methode an. So können ExpertInnen beispielsweise antizipieren, ob eine mehrstufige Antwortskala, für einen Selbstausfüller-Fragebogen akzeptabel ist, sie im Falle eines Telefoninterviews jedoch die Befragten möglicherweise überfordert. Gleiches gilt etwa für Filterführungen, die für erfahrene InterviewerInnen ohne Probleme zu bewältigen sind, „normale“ Befragte hingegen verwirren können (vgl. Fowler, 2001, S. 55 ff.). Wie bereits zuvor erwähnt, können ExpertInnen-Reviews jedoch keine Informationen bezüglich der Vergleichbarkeit der Daten liefern. Gleichzeitig werden keine realen Personen befragt, was jedoch wiederum bei der Methode der kognitiven Interviews der Fall ist. Auch wenn kognitive Interviews üblicherweise face-to-face erhoben werden, besteht durchaus die Möglichkeit, andere Erhebungsformen (wie z.B. eine telefonische Befragung oder eine Online-Befragung) zu simulieren und die Eignung des Erhebungsinstrumentes für verschiedene Erhebungsmodi zu testen (siehe hierzu u.a. Beatty & Schechter, 1994; Gray, Blake, & Campanelli, 2014).
- (7) *Multi-Language-Capability: Lassen sich die Fragen gut in andere Sprachen übersetzen? Sind die Daten über verschiedene Sprachen hinweg vergleichbar?* Im Hinblick auf die Einsatzmöglichkeit des Fragebogens in verschiedenen Ländern kann ähnlich argumentiert werden wie im Zusammenhang mit der Multi-Mode-Capability. Auch hier sind Split-Ballot-Experimente die Methode der ersten Wahl, während das ExpertInnen-Review als hilfreiches Verfahren eingestuft werden kann. Gleiches gilt hier auch für die kognitiven Interview-Techniken. Sie geben Aufschluss darüber, ob bestimmte Begriffe bzw. das gesamte Item immer im intendierten Sinne verstanden werden und sie somit erfolgreich in eine andere Sprache übersetzt werden konnten. Denn nur, wenn auch das pragmatische Verständnis über verschiedene Länder hinweg sichergestellt werden kann, ist eine interkulturelle Validität der Messinstrumente gegeben, welche wiederum Voraussetzung für

Vergleiche ist (siehe hierzu u.a. Fitzgerald, Collins, Widdop, & Gray, 2009; Levin et al., 2009).

*Tabelle 3: Nutzen verschiedener Pretest-Verfahren zur Evaluierung der Fragen im Hinblick auf die zentralen Standards*

Verfahren	Standards zur Evaluation von Fragen						
	Inhalt	Kognitiv	Interpersonell	Usability	Psychometrisch	Multi-Mode-Capability	Multi-Language-Capability
<b>Qualitative Verfahren (Pre-Field Testing)</b>							
Kognitive Interviews	Ja (G) H (F)	Ja			H	Ja (S)	Ja (S) H (F)
Fokusgruppen	Ja	H					
ExpertInnen-Review	Ja	H		H (G)		H	H
<b>Quantitative Verfahren (Testverfahren im Feld)</b>							
Beobachtungspretest mit InterviewerInnen-Debriefing		H	H	Ja			
Behavior Coding		H	Ja				
Befragten-Debriefing		H	H	H	H		
Analyse von Pretest-Daten					Ja		
Response Latency		H					
Reinterviews		H			Ja		
Split-Ballot					Ja	Ja	Ja

*Anmerkungen.* Ja = Verfahren liefert die entsprechenden Informationen; H = Verfahren ist hilfreich, aber nicht ausreichend; F = Einschätzung von Fowler 2001; G= Einschätzung von Groves et al., 2009; S = Schüsler, eigene Einschätzung). Quelle: Tabelle übernommen aus Fowler, 2001, S. 60. Eigene Übersetzung, eigene Ergänzung in Bezug auf die Response-Latency-Methode. Zudem wurden Ergänzungen hinsichtlich unterschiedlicher Ansichten in Groves et al. (2009), Fowler (2001) und eigene Einschätzung in Bezug auf die Eignung des ExpertInnen-Reviews sowie in Bezug auf die Methode der kognitiven Interviews vorgenommen.

Fowler (2001) fasst die Eignung der verschiedenen Pretest-Verfahren für die Evaluation von einzelnen Items bzw. des kompletten Fragebogens im Hinblick auf die oben angeführten Standards wie folgt zusammen:

*“[T]he most common testing strategies – expert review, focus groups, cognitive testing, and behaviour coding of field interviews – address the first three standards plus usability. However, they do very little to help with the other four. Particularly lacking are measures of data quality, validity and reliability, and comparability of data across modes and across languages. (...) [T]here are three underutilized kinds of evaluation steps that would add considerably to knowledge about key elements of survey instruments: analysis of pilot or pretest data, reinterviews, and split-sample testing” (Fowler, 2001, S. 61).*

Auch verschiedene Studien zur Effektivität von Pretest-Verfahren (vgl. z.B. DeMaio & Landreth, 2004; Presser & Blair, 1994) kommen zu dem Ergebnis, dass die einzelnen Methoden nicht alle gleichermaßen effektiv sind (einige deuten auf Probleme hin, geben jedoch keine Auskunft über die jeweiligen Ursachen), da sie verschiedene Stärken und Schwächen aufweisen und hinsichtlich unterschiedlicher Problembereiche Auskunft über die Qualität einer Frage bzw. eines Messinstrumentes liefern können. Keine der Methoden ist demnach dazu geeignet, alle acht aufgeführten Standards zu kontrollieren. Dementsprechend wird eine Kombination verschiedener Verfahren empfohlen (vgl. Prüfer & Rexroth, 1996, S. 112). Willis (1999) schließt sich dem an:

*“[I]t is fruitless to determine which particular pretesting technique is superior, as they are ‘stacked up’ against one another. Rather, they may be best applied for different purposes, and at different points in the developmental sequence (...). The challenge to pretesting is to utilize a cohesive developmental plan that takes advantage of the strengths of each method” (Willis, 1999, S. 34).*

## 3 Empirischer Teil

### 3.1 Überblick Entwicklung und Testung der FWSEM-Skala

Ziel dieser Arbeit ist es, eine reliable und valide Skala zu entwickeln, die feindselig- und scheinbar wohlwollend-sexistische Einstellungen in Bezug auf Mütter erfasst. Die Basis für die Skalenkonstruktion bildete dabei, wie bereits in der Einleitung geschildert, die Theorie des Ambivalenten Sexismus von Glick und Fiske (1996, 1997). Der Fokus des Pretests liegt neben der Überprüfung der psychometrischen Eigenschaften auf der Kontrolle des Itemverständnisses mit Hilfe kognitiver Interviews. Es wurde folgendermaßen vorgegangen:

- (1) Anknüpfend an die theoretischen Überlegungen von Glick und Fiske (1996, 1997) wurde in einem ersten Schritt ein neuer Itempool generiert, der 59 Fragen umfasste. In dieser *Entwicklungs-Phase* wurden insbesondere inhaltliche Standards in den Blick genommen. So erfolgte die Formulierung der einzelnen Items unter Zuhilfenahme verschiedener Leitlinien zur Fragebogenkonstruktion (siehe Unterkapitel 2.5.2.4) und in enger Anlehnung an die Definition des Konstruktes.
- (2) Für die anschließende Itemselektion und Validierung der Skala wurde ein dreistufiges Verfahren eingesetzt, welches qualitative und quantitative Methoden kombiniert. Insgesamt wurden sieben unabhängige Studien ( $N_{\text{Gesamt}} = 923$ ) durchgeführt (siehe Tabelle 6 für die exakte Zusammensetzung der Stichproben). Ziel der ersten beiden quantitativen Studien (*Pretest-Phase 1*) war es, die Itemanzahl des ursprünglichen Itempools, der als Basis für die Generierung der Skala diente, zu reduzieren. Hierfür wurde eine psychometrische Analyse der Pretest-Daten durchgeführt. Neben verschiedenen Itemkennwerten wurde v.a. die faktorielle Struktur untersucht. Die Selektion der Items erfolgte sowohl theorie- als auch empiriegeleitet.
- (3) Anschließend wurde in drei qualitativen Studien (*Pretest-Phase 2*) das Frageverständnis der verbliebenen Items mit Hilfe kognitiver Techniken überprüft. Der Empfehlung von Willis (2005) zufolge wurde den kognitiven Interviews ein ExpertInnen-Review<sup>27</sup> vorgeschaltet, um mögliche Fragen-Probleme bereits *vor* der Durchführung der kognitiven Interviews zu antizipieren und alternative Formulierungen zu erarbeiten. Neben den inhaltlichen und kognitiven Standards sollte in dieser Pretest-Phase die Multi-Mode-Capability (Einsetzbarkeit der Skala bei verschiedenen Erhebungsmodi) der Items überprüft werden.
- (4) In den letzten beiden Studien (*Pretest-Phase 3*), die erneut quantitativ angelegt waren, wurden schließlich die besten Items für die endgültige Skala ausgewählt. Primäres

---

<sup>27</sup> Bei den ExpertInnen handelte es sich zum einen um inhaltliche ExpertInnen (SexismusforscherInnen) und zum anderen um methodische Expertinnen (die vier externen Interviewerinnen, die in Runde 1 der kognitiven Interviews mitwirkten). Sie nahmen die Items bezüglich ihrer Inhalte sowie hinsichtlich ihres Sprachflusses und ihre Verständlichkeit hin in Augenschein.

Anliegen dieser Pretest-Phase war es, die psychometrischen Standards der einzelnen Items sowie der neu entwickelten Skala zu überprüfen. Zudem wurde die Multi-Mode-Capability der Skala getestet, da neben einer Online-Umfrage auch eine schriftliche Befragung durchgeführt wurde.<sup>28</sup>

Im Rahmen einer repräsentativen Bevölkerungsumfrage wurde die neu entwickelte FWSEM-Skala schließlich erstmals eingesetzt und geprüft, ob sich die Skala auch im „realen“ Feldeinsatz bewährt. Die einzelnen Schritte der Skalenentwicklung und -testung werden nun im Folgenden Phase für Phase beschrieben (siehe Tabelle 4 für einen Gesamtüberblick über die einzelnen Pretest-Phasen und Tabelle 5 für einen Überblick über die vorgenommenen Veränderungen des Itempools).

---

<sup>28</sup> Im Rahmen des Drei-Phasen-Pretestes sollte somit die Güte und Einsetzbarkeit der FWSEM-Skala für sich alleine genommen untersucht werden. Fragebogenspezifische Probleme hinsichtlich der BenutzerInnenfreundlichkeit, welche beispielsweise die Instruktionen, die Filterführung, das Layout und Design oder auch die Zeitdauer einer Befragung betreffen, waren nicht von Interesse. Diese Aspekte müssten bei jeder weiteren Erhebung, bei der die FWSEM-Skala (neben weiteren Messinstrumenten) in einem endgültigen Fragebogen eingesetzt wird, extra überprüft werden.

Tabelle 4: Überblick über die einzelnen Phasen der Entwicklung und Testung der FWSEM-Skala

Phase	Studie	Primäres Ziel	Überprüfte Standards	Verwendete Verfahren
Entwicklungs-Phase		Generierung des Itempools	Inhaltliche Standards	Leitlinien
Pretest-Phase 1	Quantitative Studien 1-2: Papier-Fragebogen ( $N_1 = 276$ ) und Online-Fragebogen ( $N_2 = 310$ )	Reduzierung des Itempools	Psychometrische Standards (Konstruktvalidität)	Analyse von Pretest-Daten
Pretest-Phase 2	Qualitative Studien 3-5: Kognitive Interviews Runde 1 ( $N_3 = 26$ ), Runde 2 ( $N_4 = 9$ ) und Runde 3 ( $N_5 = 5$ )	Überprüfung des intendierten Frageverständnisses	Inhaltliche Standards, Kognitive Standards, Multi-Mode-Capability	Kognitive Interviews mit vorgeschaltetem ExpertInnen Review
Pretest-Phase 3	Quantitative Studien 6-7: Online-Fragebogen ( $N_6 = 207$ ) und Papier-Fragebogen ( $N_7 = 90$ )	Endgültige Auswahl der Items für die FWSEM-Skala und Überprüfung der psychometrischen Eigenschaften	Psychometrische Standards (Reliabilität, Konstrukt- und Übereinstimmungsvalidität), Multi-Mode-Capability	Analyse von Pretest-Daten
⇩				
Einsatz der Skala im Feld (Quantitative Studie 8: telefonische Befragung ( $N_8 = 397$ ))				



*Tabelle 5: Überblick über die vorgenommenen Veränderungen am Itempool während der einzelnen Phasen der Entwicklung und Testung der FWSEM-Skala*

Phase	Primäres Ziel	Bezeichnung der Itemsammlung	Vorgenommene Änderungen am Itempool	Anzahl der Items nach Abschluss der Phase
Entwicklungs-Phase	Generierung des Itempools	Ursprünglicher Itempool		59
Pretest-Phase 1	Reduzierung des ursprünglichen Itempools	Reduzierter Itempool	Entfernung von 23 Items aus dem Itempool auf Basis der empirischen Ergebnisse (auch theoretische Überlegungen kamen hierbei zum Tragen, siehe Unterkapitel 3.3.3)	36
Pretest-Phase 2	Überprüfung des intendierten Frageverständnisses	Überarbeiteter Itempool		
ExpertInnen-Review	Antizipierung möglicher Frage-Probleme vor den kognitiven Interviews,		kleinere Itemüberarbeitungen (z.B. Ersetzen eines Begriffs), Itemstreichungen und Entwicklung möglicher Alternativ-Formulierungen im Falle von 5 Items	41
Kognitive Interviews Runde 1	Aufdeckung und Behebung zentraler Formulierungsdefizite		Erneute Itemüberarbeitungen und Entfernung von 6 ähnlichen Items. Den Empfehlungen von Willis (2005) zufolge wurden bei massiven Formulierungsdefiziten auch schon im Rahmen der ersten drei Interviews Änderungen an den Items vorgenommen, die in Interview 4 bis 26 getestet wurden.	35

Phase	Primäres Ziel	Bezeichnung der Itemsammlung	Vorgenommene Änderungen am Itempool	Anzahl der Items nach Abschluss der Phase
Kognitive Interviews Runde 2	Überprüfung der Itemüberarbeitung		Erneute Itemüberarbeitungen	35
Kognitive Interviews Runde 3	Überprüfung der Itemüberarbeitung, Testung des Itemverständnisses im Rahmen einer simulierten telefonischen Befragung		Erneute Itemüberarbeitungen, Entwicklung einer neuen Item-Formulierungen und Entfernung von 14 Items, die (a) eine ähnliche Formulierung besaßen (hier wurde jeweils die Variante ausgewählt, die sich im Rahmen der kognitiven Interviews bewährt hatte) oder (b) nicht zufriedenstellend umformuliert werden konnten	28
Pretest-Phase 3	Endgültige Auswahl der Items für die FWSEM-Skala und Überprüfung der psychometrischen Eigenschaften	FWSEM-Skala	Entfernung von 16 Items aus dem Itempool auf Basis der empirischen Ergebnisse und Auswahl der besten 12 Items für die FWSEM-Skala	12



Phase	Studie							
	1	2	3	4	5	6	7	8
	Pretest-Phase 1 (quantitativ)			Pretest-Phase 2 (qualitativ)		Pretest-Phase 3 (quantitativ)		Einsatz der Skala (quantitativ)
Bundesgebiet								
West	N. E.	N. E.	N. E.	N. E.	N. E.	N. E.	76 (84.4%)	256 (64.5%)
Ost	N. E.	N. E.	N. E.	N. E.	N. E.	N. E.	7 (7.8%)	141 (35.5%)
Erwerbstätigkeit								
Erwerbstätige	N. E.	134 (43.2%)	42.31%	5 (55.56%)	5 (100%)	54 (26.1%)	N. E.	217 (54.7%)
Nicht-Erwerbstätige	N. E.	175 (56.5%)	57.69%	4 (44.44%)	0 (0%)	149 (72%)	N. E.	179 (45.1%)
im Altersruhestand	N. E.	N. E.	N. E.	N. E.	N. E.	2 (1%)	N. E.	129 (32.5%)
Hausfrau oder - mann	N. E.	N. E.	N. E.	N. E.	N. E.	3 (1.4%)	N. E.	15 (3.8%)
arbeitslos	N. E.	N. E.	N. E.	N. E.	N. E.	2 (1%)	N. E.	10 (2.5%)
StudentIn	N. E.	N. E.	N. E.	N. E.	N. E.	136 (65.7%)	N. E.	8 (2.0%)
SchülerIn	N. E.	N. E.	N. E.	N. E.	N. E.	N. E.	N. E.	5 (1.3%)
Sonstiges	N. E.	N. E.	N. E.	N. E.	N. E.	5 (2.4%)	N. E.	12 (3.0%)
Elternschaft								
Mit Kinder	10 (3.6%)	47 (15.2%)	14 (53.85%)	4 (44.44%)	3 (60%)	28 (13.5%)	3 (3.3%)	320 (80.6%)
Ohne Kinder	263 (95.3%)	262 (84.5%)	12 (46.15%)	5 (55.56%)	2 (40%)	178 (86%)	85 (94.4%)	77 (19.4%)

Anmerkungen. N. E. = Nicht erhoben. In den Spalten, in denen sich die jeweiligen Prozentzahlen nicht auf 100 aufaddieren lassen, sind fehlende Werte enthalten. D.h. die befragten Personen haben für diese entsprechenden Kategorien keine Angaben gemacht.

### 3.2 Entwicklungs-Phase: Generierung des ursprünglichen Itempools

Um die „Skala zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter“ (kurz FWSEM-Skala) zu entwickeln, wurden in einem ersten Schritt auf der Basis der Theorie des ambivalenten Sexismus (Glick & Fiske, 1996, 1997) verschiedene Inhaltskategorien ausgearbeitet. Diese Inhaltskategorien waren nicht als zusätzliche Subdimensionen von feindseligem oder wohlwollendem Sexismus angelegt, wie dies bei der ASS von Glick und Fiske (1996) der Fall war, sondern bildeten lediglich die Basis für die Ausformulierung der einzelnen Items. Sowohl für die Ausarbeitung der Inhaltskategorien als auch für die Itemformulierungen dienten neben den eigenen Überlegungen im Zusammenhang mit den in Kapitel 2 vorgestellten Theorien bereits etablierte Sexismus-Skalen als Ausgangspunkt. Das gleiche gilt für Beiträge und Meinungsäußerungen in Internetforen zu den Themen „Rabenmütter“, „Hausfrauen“ und „Karrierefrauen“ sowie für das Buch „Das Eva-Prinzip“ von Herman (2007). Insgesamt wurden anhand von sechs Inhaltskategorien 59 Items generiert. Sie formten den ursprünglichen Itempool, der in den nachfolgenden Studien reduziert und getestet werden sollte und aus dem letztendlich die spätere Skala hervorging (siehe Tabelle 5).<sup>29</sup>

Im Folgenden wird nun für jede Facette des ambivalenten Sexismus getrennt beschrieben, welche Einstellungen erfasst werden sollen, wie die zentralen Inhaltskategorien lauten und welche Überlegungen im Zuge der Itemformulierungen angestellt wurden. Zudem werden erste Itembeispiele genannt.

*„Frauen [werden, d. Verf.] mit dem Vorurteil konfrontiert,  
dass es dem Kind schade, wenn die Mutter arbeite. (...)  
Obwohl mein Mann ganztätig für ihn [den Sohn, d. Verf.] sorgte,  
wurde mir mehr als einmal signalisiert, dass meine Entscheidung,  
frühzeitig an den Arbeitsplatz zurückzukehren, äußerst egoistisch sei.  
Mir wurde – wenn auch unterschwellig – signalisiert ich sei eine ‚Rabenmutter‘.“*  
(Stachelhaus, 2010, o.S.)

#### 3.2.1 Feindseliger Sexismus

Wie bereits in Unterkapitel 2.4.2.3 erläutert, stellt der feindselige Sexismus (FS) eine offene, negative Form von Sexismus dar. Er erfüllt die klassische Definition von Vorurteilen, da Frauen, die aus der traditionellen Rolle ausbrechen, eine reine Antipathie

---

<sup>29</sup> Anzumerken ist in diesem Zusammenhang, dass die ursprüngliche Idee der Skalenentwicklung weiter gefasst war und sich nicht ausschließlich auf die Subgruppe der Mütter fokussierte. Im Laufe der weiteren Auswertung der Literatur kristallisierte sich jedoch zunehmend die Diskriminierung von Müttern als relevant heraus und auch die empirischen Analysen der Pretest-Daten zeigten, dass gerade die Mütter-Items inhaltlich etwas Neues abbildeten, wohingegen die anderen Items teilweise mit den Inhalten bereits bestehender Skalen redundant waren. Folglich wurde der Schwerpunkt im Laufe der Skalenentwicklung und somit auch beim Verfassen dieser Arbeit auf die Diskriminierung von Müttern und die Erfassung sexistischer Einstellungen in Bezug auf Mütter gelegt.

entgegengebracht wird. Dementsprechend sollten auch die neuen Items der FWSEM-Skala ein negatives Bild von Karrierefrauen bzw. berufstätigen Müttern zeichnen.

Feindseliger Sexismus ist nach Glick und Fiske (1996, 1997) das Resultat von Macht- und Statusunterschieden, die die Männer mittels feindseliger Ideologien zu rechtfertigen und aufrechtzuerhalten versuchen. Gerade erfolgreiche berufstätige Mütter stellen m. E. eine Bedrohung für die gesellschaftliche Vorrangstellung der Männer dar, da durch ihre Erwerbstätigkeit das traditionelle Hausfrauen-Ernährer-Modell in Frage gestellt wird. So müssen Männer einerseits befürchten, ihre Rolle als Versorger und Beschützer der Familie zu verlieren, was für sie zu einer Identitätskrise führen könnte (siehe hierzu auch „Rollenkonflikt“ Sachverständigengutachten, 2011, S. 154). Andererseits möchten sie vermutlich ihre strukturelle Macht und ihren Status nicht einbüßen, die mit einem hohem Ansehen und (finanzieller) Unabhängigkeit verbunden sind. Folglich werden Männer, die feindselig sexistisch eingestellt sind, berufstätige Mütter bzw. Karrierefrauen abwerten, indem sie sie beispielsweise als „Rabenmütter“ charakterisieren.

Der Definition von Sexismus zufolge (siehe Unterkapitel 2.4.1) müssen sexistische Einstellungen nicht zwangsläufig nur von Männern vertreten werden. Es ist ebenso möglich, dass auch Frauen anderen Frauen gegenüber eine Antipathie hegen (siehe z.B. Glick & Fiske, 2001a). Dies wird vermutlich v.a. dann der Fall sein, wenn sie einer anderen Subgruppe von Frauen angehören.<sup>30</sup> So werden Frauen, die aus Überzeugung Hausfrau und Mutter sind, vermutlich solche Frauen negativ beurteilen, die ihren Beruf und ihre Karriere weiter verfolgen möchten und daher ihre Kinder zur Betreuung weggeben. Auf diese Weise können Hausfrauen ihr eigenes Lebensmodell rechtfertigen und zugleich das Verhalten und die Lebensweise der Fremdgruppe, also der berufstätigen Mütter, abwerten.

Zur Ausformulierung der einzelnen Items dienten v.a. die drei folgenden Inhaltskategorien<sup>31</sup> als Orientierung, die einen Verstoß gegen den Muttermythos (siehe Unterkapitel 2.1.1.2) darstellen:

(a) **Karrierefrauen werden feminin-positive Eigenschaften (Warmherzigkeit) abgesprochen.**

Da Karrierefrauen nicht (nur) ihrer „natürlichen“ Rolle als Hausfrau und Mutter nachkommen, sie eine Mutterschaft vielleicht nicht als ultimative Erfüllung ansehen (Mutterschaft als Berufung) oder diese Aufgabe sie nicht alleine vollkommen erfüllt (natürliche Prädisposition), können sie aus Sicht einer feindselig-sexistisch

---

<sup>30</sup> Siehe hierzu „Social Identity Theory“ (Tajfel & Turner, 1986) und „in-group favouritism“ und „out-group derogation“ (Brewer, 2001).

<sup>31</sup> Da die ursprüngliche Idee der Skalenentwicklung nicht ausschließlich auf Mütter beschränkt war, gab es neben den drei aufgeführten Inhaltskategorien noch eine weitere Kategorie, die sich auf die Abwertung berufstätiger Frauen im Allgemeinen bezog (z.B. wurden ihnen Kompetenzen in Bezug auf Führungspositionen abgesprochen oder ihnen wurde unterstellt, ihre beruflichen Erfolge seien lediglich auf Frauenfördermaßnahmen zurückzuführen (Abwertung von sogenannten „Quotenfrauen“).

eingestellten Person keine „richtigen“ Frauen sein – sie werden als „vermännlicht“ angesehen. Mit dieser Charakterisierung nimmt die erste Inhaltskategorie neben dem Muttermythos auch auf das Stereotype Content Model Bezug, demzufolge Karrierefrauen als kompetent, aber emotional kalt eingestuft werden (siehe Unterkapitel 2.2.3.4). Da der feindselige Sexismus ein negatives Bild von Frauen zeichnet, wurde der inhaltliche Schwerpunkt der Items auf die fehlende Warmherzigkeit gelegt.

Itembeispiele:

*„Viele erfolgreich berufstätige Frauen verlieren ihre Weiblichkeit.“*

*„Erfolgreich berufstätige Frauen sind häufig ‚Mannweiber‘.“*

*„Viele erfolgreich berufstätige Frauen sind gefühlskalt.“*

**(b) Die Motive berufstätiger Frauen bzw. Mütter werden abgewertet.**

Gute Mütter zeigen dem Muttermythos zufolge eine grenzenlose Bereitschaft, sich für ihre Kinder aufzuopfern und ihre eigenen Bedürfnisse hinten anzustellen (mütterliche Hingabe). Frauen, die auch als Mütter weiterhin an ihrer Erwerbstätigkeit festhalten wollen und nicht die Anerkennung, die sozialen Kontakte oder auch die damit verbundene (finanzielle) Unabhängigkeit missen möchten, wird folglich Unverständnis entgegengebracht und ihre Motive werden abgewertet.

Itembeispiele:

*„Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine versorgen könnte.“*

*„Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.“*

**(c) Karrierefrauen werden als Rabenmütter und schlechte Ehefrauen angesehen.<sup>32</sup>**

Der Muttermythos postuliert eine Unabkömmlichkeit der Mutter – ihre ganztägige Zuwendung wird als unerlässlich für das kindliche Wohlergehen angesehen, was insbesondere für die ersten Lebensjahre gilt (siehe auch Mutterleitbild in Westdeutschland in Unterkapitel 2.1.1.3). Frauen, die sich aufgrund ihrer Erwerbstätigkeit nicht selbst um die Betreuung des Kindes kümmern können, werden abgewertet und in ein schlechtes Licht gerückt.

---

<sup>32</sup> Hinsichtlich dieser letzten Kategorie ähnelt die Feindselige-Subskala den Items der Geschlechterrollen-Ideologie der Geschlechterrollen-Batterie (GRB) des ISSP 2002 (siehe Unterkapitel 2.4.2.1). Auch dort werden die negativen Konsequenzen einer Berufstätigkeit der Mutter für die Familie angesprochen. Allerdings liegt bei diesen Items der Fokus stärker auf den Kindern – die Mutter wird nicht in ein schlechtes Licht gerückt und verurteilt, wie dies bei mehreren Items der FS-Subskala der Fall ist. (Beispiel-Item GRB: „Ein Kind, das noch nicht zur Schule geht, wird wahrscheinlich darunter leiden, wenn seine Mutter berufstätig ist.“)

Itembeispiele:

*„Ich finde Frauen, die ihre Karriere für die Betreuung ihrer Kinder nicht zurückstellen wollen, egoistisch.“*

*„Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.“*

*„Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.“*

Ziel der Feindseligen-Subskala (FS) ist es somit, die Einstellungen der Befragten in Bezug auf Mütter zu erfassen, die aus der traditionellen Rolle ausbrechen und ihre Kinder nicht während der ersten drei Jahre zu Hause betreuen, sondern zeitnah in den Beruf zurückkehren und evtl. sogar ihre Karriere vorantreiben möchten. Bewusst werden dabei Frauen beschrieben, die gegen den Muttermythos verstoßen, indem sie Mutterschaft nicht als „Berufung“ ansehen oder nicht komplett ihre eigenen Bedürfnisse zurückstellen wollen. So werden in den Items Frauen beschrieben, die *gerne* berufstätig sein *möchten*. Eine finanzielle Notwendigkeit ist dementsprechend nicht der vorherrschende Beweggrund. Vielmehr geht es darum, sich eine eigene Existenz aufzubauen, unabhängig zu bleiben, soziale Kontakte zu haben, Anerkennung für die erbrachte Leistung zu erhalten und/oder sich selbst zu verwirklichen.

Wie an den aufgeführten Beispielen zu erkennen ist, sind die Items der FS-Subskala aus Sicht einer feindselig-sexistisch eingestellten Person geschrieben, die die berufstätigen Mütter als „schlechte“ Mütter – eben als Rabenmütter – ansieht. Ihrer Ansicht nach erfüllen diese Frauen ihre „natürliche Aufgabe“ nicht, kommen ihrer Mutterrolle nicht nach, sind egoistisch und sollten dementsprechend ein schlechtes Gewissen haben. Feindselig-sexistisch eingestellte Personen verurteilen das oben beschriebene (aus ihrer Sicht normabweichende) Verhalten und versuchen auf diese Weise Frauen auf die traditionelle Rolle einzuschränken. Eine hohe Zustimmung zu den Items bedeutet dementsprechend eine hohe Ausprägung an feindseligem Sexismus.

### 3.2.2 Wohlwollender Sexismus

Der wohlwollende Sexismus (WS) stellt mit seiner subtilen, subjektiv positiv empfundenen Form des Sexismus das Gegenstück zum feindseligen Sexismus dar. Frauen, die die traditionelle Rolle erfüllen, werden durchweg positiv beschrieben, ihnen wird Unterstützung, Schutz und Bewunderung entgegengebracht. Dies hängt nach der Theorie des Ambivalenten Sexismus (Glick & Fiske, 1996, 1997) damit zusammen, dass Männer in Paarbeziehungen auf Frauen angewiesen sind, was die sexuelle Reproduktion sowie die Bedürfnisse nach seelischer und körperlicher Intimität anbelangt. Wie bereits in Unterkapitel 2.4.2.3 beschrieben, gilt diese Abhängigkeit auch in Bezug auf die familiäre Arbeitsteilung. So sind Männer häufig, um ihren Beruf erfolgreich ausüben zu können, im Privaten auf ihre Partnerinnen angewiesen. Würden die Frauen ihnen den Rücken von



häuslichen Pflichten nicht frei halten und die Betreuung der Kinder übernehmen, könnten sie den Anforderungen eines „idealen Arbeitnehmers“ nicht gerecht werden bzw. diese nur schwer erfüllen. Männliche Befragte, die wohlwollend-sexistisch eingestellt sind, werden Frauen folglich insbesondere solche positiven Eigenschaften attestieren, die sie für die Rolle als Mutter und Hausfrau prädestinieren (siehe Unterkapitel 2.1.1.2 und „women are wonderful“ effect von Eagly und Mladinic (1994) in Unterkapitel 2.2.2). Darunter fallen sogenannte feminine Merkmale, die auf die Interessen und Bedürfnisse anderer bezogen sind und u.a. Selbstlosigkeit, Empathievermögen oder auch Warmherzigkeit umfassen (siehe Unterkapitel 2.2.3).

Auch wenn derartige Zuschreibungen (deskriptive Stereotype) zunächst positiv erscheinen mögen, können sie zum Fortbestehen der Geschlechterhierarchien beitragen. Dies ist der Fall, wenn sie in Verhaltenserwartungen (präskriptive Stereotype) münden und beispielsweise Frauen nahegelegt wird, ihre Erwerbstätigkeit zu unterbrechen oder gar vollständig aufzugeben, um die Kinderbetreuung übernehmen zu können, da sie (und eben nicht die Männer) aufgrund ihres natürlichen Wesens besonders für diese Aufgabe geeignet sind. Auf diese Weise werden traditionelle Arbeitsteilungsmuster zementiert und soziale Ungleichheiten zwischen Frauen und Männern fortgeschrieben (siehe Unterkapitel 2.2.3.5 und Unterkapitel 2.4.1).

Auch im Hinblick auf wohlwollend-sexistische Einstellung gilt, dass sie nicht an das Geschlecht gekoppelt sind: Folglich sind es nicht nur Männer, die Frauen eine besondere Befähigung für die familiären Sorge- und Betreuungsaufgaben zusprechen, auch Frauen können derartige benevolent-sexistische Ansichten vertreten (siehe Unterkapitel 2.4.1 und 2.4.2.3). Koppetsch und Burkart (1999) gelangten in ihren Untersuchungen zu dem Ergebnis, dass insbesondere Frauen aus dem traditionellen Milieu versuchen, „typisch weibliche“ Zuständigkeitsbereiche vor „männlichen Übergriffen zu schützen“. Folglich lassen sie Männern mitunter überhaupt keine Chance, sich im häuslich-familiären Bereich einzubringen. Die AutorInnen erklären das gezeigte Verhalten dadurch, dass dies der Bereich sei, in dem Frauen nach eigenem Empfinden glänzten – im Haushalt und bei der Kinderbetreuung könne ihnen niemand etwas vormachen, und daher wollten sie diese Domäne nicht aufgeben. Zu einer ähnlichen Erkenntnis gelangten auch Fthenakis, Kalicki und Peitz (2002): „Frauen, bei denen die Mutterrolle den zentralen Aspekt ihrer Identität ausmacht, legen (...) keinen großen Wert auf eine aktive und gleichberechtigte Beteiligung des Vaters an der Sorge um das Kind. Für sie stellt ihre exklusive Expertise als Mutter eine wichtige Quelle für die Bestätigung des eigenen Werts dar“ (S.171). Es ist daher anzunehmen, dass insbesondere Frauen, die aus Überzeugung Hausfrau und Mutter sind, den Items der WS-Subskala zustimmen werden.

Für die Skalenentwicklung wurden die drei folgenden Inhaltskategorien berücksichtigt, die neben dem Muttermythos (siehe Unterkapitel 2.1.1.2) insbesondere auf die beiden Subkategorien des protektiven Paternalismus sowie die komplementäre Geschlechterdifferenzierung in der Theorie des ambivalenten Sexismus von Glick und Fiske (1996, 1997) (siehe Unterkapitel 2.4.2.3) Bezug nehmen:

**(a) Frauen bzw. Mütter sollten von Männern beschützt und umsorgt werden.**

Die Idee, die hinter der Formulierung der ersten Inhaltskategorie steht, ist die des edlen Ritters in weißer Rüstung. In diesem Kontext werden die Hilfsbedürftigkeit sowie die passive Rolle von Frauen (feminin-negative Geschlechterstereotype) thematisiert. Damit greift diese Kategorie den protektiven Paternalismus auf, der die Überzeugung umfasst, Männer sollten aufgrund der vermeintlichen Schwäche der Frauen die Rolle des Beschützers einnehmen und für das weibliche Wohlergehen sorgen.<sup>33</sup>

Itembeispiele:

*„Eine Frau braucht die starke Schulter eines Mannes, an die sie sich anlehnen kann.“*

*„Viele Frauen wünschen sich einen Mann, der ihr Herz im Sturm erobert.“*

*„Viele Frauen sehnen sich nach einem Mann, in dessen Armen sie Schutz und Geborgenheit finden.“*

**(b) Frauen bzw. Mütter sollten nur wenig (Teilzeit) oder überhaupt nicht erwerbstätig sein, damit die Doppelbelastung aus Beruf und Familie reduziert wird.**

Mit dieser Inhaltskategorie werden die Vereinbarkeitsproblematik und die damit verbundene Doppelbelastung der Frauen angesprochen. Aus der Sicht einer benevolent-sexistisch eingestellten Person ist es für Mütter besser, wenn sie sich auf den privaten Bereich konzentrieren und sich aus der öffentlichen Sphäre (d.h. der Erwerbstätigkeit) zurückziehen (siehe indirekte, scheinbar wohlwollende Formen der Diskriminierung sowie benevolente präskriptive Stereotype in Unterkapitel 2.3.2). An der traditionellen Arbeitsteilung der Geschlechter wird auf diese Weise nicht gerüttelt, allerdings wird sie den Frauen durch eine scheinbar wohlwollende Argumentationsweise „schmackhaft“ gemacht (siehe legitimierende Mythen in Unterkapitel 2.2.2). Damit stellt diese Kategorie eine Erweiterung zu den traditionellen Geschlechterrollen-Items dar, wie sie etwa im ISSP zu finden sind. Dort wird lediglich die Rollenverteilung zwischen den Geschlechtern festgeschrieben, ohne eine positive Begründung zu liefern. (Beispielitem GRB: „Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern.“).

Aufgrund der scheinbar positiven Absichten, die dieser Argumentationsweise zugrunde liegen (die Frau wird von ihrer Erwerbsrolle entlastet und der Mann in seiner Rolle als Versorger in die Pflicht genommen), könnte auch diese Inhaltskategorie dem protektiven Paternalismus zugerechnet werden, obgleich

---

<sup>33</sup> Die nachfolgenden Itembeispiele beziehen sich nicht auf die Subgruppe der Mütter, sondern auf Frauen im Allgemeinen. Dies hängt mit der ursprünglich thematisch weiter gefassten Idee der Skalenentwicklung zusammen (siehe Fußnote 29).

eine ähnliche thematische Schwerpunktsetzung (Mütter und Arbeitsteilung) in den entsprechenden Items von Glick und Fiske (1996) nicht zu finden ist.

Itembeispiele:

*„Wenn es die finanzielle Lage der Familie erlaubt, sollte eine Mutter kleiner Kinder lieber keinen Job haben, da ansonsten die Doppelbelastung aus Beruf und Familie für sie vielleicht zu groß wird.“*

*„Berufstätige Mütter sind heutzutage einer so enormen Doppelbelastung ausgesetzt, dass es eigentlich besser für sie wäre, wenn sie sich ausschließlich der Betreuung ihrer Kinder widmen würden.“*

(c) **Frauen sind für die Rolle als Mutter und Hausfrau prädestiniert.**

Diese Inhaltskategorie bezieht sich auf die natürliche Prädisposition des Muttermythos, der Frauen eine natürliche Veranlagung zur Erfüllung der reproduktiven Tätigkeiten attestiert. Dabei ähneln die Items der FWSEM-Skala denen der komplementären Geschlechterdifferenzierung von Glick und Fiske (1996) (siehe Unterkapitel 2.4.2.3) insofern, als dass sie Frauen als die „bessere Hälfte“ des Mannes erscheinen lässt: Es werden positive Stereotype über Frauen in den Vordergrund gerückt, die die Wesenszüge des Mannes ergänzen. Beispielsweise werden ihnen feminine Eigenschaften wie Warmherzigkeit (siehe Stereotype Content Model in Unterkapitel 2.2.3.4) oder Einfühlungsvermögen zugeschrieben. Dabei wird immer wieder die „naturgemäße“ bzw. „wesensgemäße Veranlagung“ der Frau für die Haus- und Familienarbeiten betont, die dem Mann nicht (bzw. nicht in gleichem Maße) gegeben sei. Auf diese Weise soll die Sinnhaftigkeit und Unumstößlichkeit der Rollenaufteilung untermauert werden.

Gerade dieser wesentliche inhaltliche Aspekt, der Frauen explizit auf ihre traditionelle Rolle als Hausfrau und Mutter festlegt, blieb bei der Skalenkonstruktion der ASS von Glick und Fiske (1996) unberücksichtigt (siehe Unterkapitel 2.4.2.3). So greifen die AutorInnen in ihren Items das ausgeprägte moralische Empfinden von Frauen, ihre Ehrlichkeit sowie ihren feineren Sinn für Kultur und ihren vermeintlich besseren Geschmack auf. Die besonderen mütterlichen Fähigkeiten bleiben hingegen unberücksichtigt.

Anders als dies bei Glick und Fiske (1996) der Fall ist, wurde bei der Itemformulierung zudem darauf geachtet, dass nicht nur feminin positive Geschlechterstereotype benannt werden, sondern gleichzeitig eine direkte Rollenzuschreibung vorgenommen wird, die Frauen auf die (unbezahlten) Reproduktionsarbeiten verweist. So wurden Frauen nicht nur als besonders empathiefähig beschrieben, sondern zugleich ihre sich daraus ergebende besondere Eignung für die Kinderbetreuung betont. Folglich wurden aus deskriptiven Stereotypen, präskriptive Stereotype abgeleitet, die wiederum durch die traditionelle Rollenaufteilung die Geschlechterungleichheit festigen und daher als sexistisch angesehen werden können. Verzichtet man auf diesen Zusatz wäre es meiner Ansicht nach möglich,

dass die Befragten fälschlicherweise als benevolent-sexistisch eingestuft werden. So könnte aus der Annahme, dass Frauen ein besseres moralisches Empfinden besitzen, auch die Ansicht abgeleitet werden, dass sie sich beispielsweise besonders gut als Richterinnen eignen. (Schließlich sind mitunter auch zunehmend typisch weibliche Eigenschaften („soft skills“) im Berufsleben gefragt.) Um diese Unsicherheiten zu vermeiden, werden die feminin-positiven Geschlechterstereotype, die nicht ohnehin im Zusammenhang mit reproduktiven Aufgaben stehen, explizit in diesem Kontext benannt. So taucht beispielweise die Multitasking-Fähigkeit, die Frauen häufig attestiert wird, bei den neu konstruierten Items der FWSEM-Skala im Rahmen von Familienaufgaben (Haushalt und Kinderbetreuung) auf. Ohne diese Spezifizierung hätte man die Merkmalszuschreibung durchaus auch im Zusammenhang mit einer verantwortungsvollen Position im Erwerbsleben sehen können.

Einen Nachteil hat diese Vorgehensweise dennoch: Durch die Verknüpfung von Geschlechterstereotypen und Rollenzuschreibungen erhalten die Items teilweise einen doppelten Stimulus. Dieser sollte eigentlich bei der Itemformulierung vermieden werden, da die Befragten ansonsten möglicherweise nicht wissen, wie sie auf das Item antworten sollen, falls sie einem Teil der Aussage zustimmen würden, dem anderen hingegen nicht. Auch bei der Auswertung können sich die ForscherInnen nicht sicher sein, auf welchen Teil des Items sich die Meinungsäußerung der Befragten bezieht (vgl. Porst, 2000, S. 7). Dennoch sollte daran festgehalten werden, um einer möglichen „Fehleinstufung“, wie sie oben beschrieben wurde, zu entgehen. Dementsprechend sollte auch im Rahmen der kognitiven Interviews überprüft werden, ob die zweifachen Stimuli den befragten Personen bei Beantwortung des Items Probleme bereiten oder beide Aspekte der Aussage für sie im Einklang stehen und daher die Wahl ihrer Antwortkategorie unkritisch ist.

Itembeispiele:

*„Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.“*

*„Frauen verfügen über ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.“*

*„Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater oder eine andere Bezugsperson.“*

Mit Hilfe der Wohlwollenden-Subskala (WS) sollen somit die Einstellungen der Befragten hinsichtlich von Frauen erfasst werden, die die traditionelle Rolle als Hausfrau und Mutter erfüllen. Solche Frauen werden durchweg positiv beschrieben („women are wonderful“ effect, Eagly & Mladinic, 1994), dabei werden ihnen feminine Merkmale zugesprochen, die sie für die traditionell weibliche Geschlechterrolle prädestinieren (siehe Social Role Theory und Stereotype Content Model in Unterkapitel 2.2.3.4). Gleichzeitig

wird indirekt die Schwäche des weiblichen Geschlechts in den Items thematisiert: Frauen werden als passiv und hilfsbedürftig angesehen. Der Mann wird hingegen in romantisch verklärender Sichtweise, als ihr „Retter“ beschrieben, der eine starke Schulter bietet und die Rolle des Versorgers übernimmt.

Den Formulierungen der feindselig-sexistischen Items entsprechend sind auch die Items der WS-Subskala aus der Sicht einer sexistisch eingestellten Person geschrieben. Eine hohe Zustimmung bedeutet dementsprechend eine hohe Ausprägung an benevolenten Sexismus.

### 3.3 Pretest-Phase 1: Reduzierung des ursprünglichen Itempools

#### 3.3.1 Stichprobe 1 und 2

Die ersten beiden quantitativen Studien in Pretest-Phase 1 waren als schriftliche Befragung (Studie 1) bzw. als Online-Befragung (Studie 2) angelegt. Sie dienten dazu, die Anzahl der Items des ursprünglichen Itempools vor der Phase der kognitiven Interviews zu reduzieren.

An der schriftlichen Befragung nahmen 276 Personen (159 Frauen und 115 Männer; eine Person machte bezüglich des Geschlechts keine Angabe) im Alter zwischen 19 und 59 Jahren ( $M = 23$ ,  $SD = 3.69$ ) teil. Die Gelegenheitsstichprobe setzte sich aus 11 BerufsschülerInnen aus Hessen sowie 262 Studierenden der Justus-Liebig-Universität Gießen zusammen. Sie wurden im Sommersemester 2008 auf freiwilliger Basis aus mehreren Lehrveranstaltungen rekrutiert. Im Einzelnen verteilten sie sich wie folgt prozentual auf die verschiedenen Studienfächer: BA Social Sciences (55.7%), Dipl. Sozialwissenschaften (17.6%), Dipl. Pädagogik (8.1%), MA Soziologie (5.5%) und Sonstige (9.1%). Für ihre Teilnahme erhielten die Befragten einen Schokoladenriegel.

Studie 2 war als Online-Befragung konzipiert. Der dazugehörige Link wurde über das Adressbuch der Autorin sowie über verschiedene Verteiler-Listen der Philipps-Universität Marburg weitergeleitet. Insgesamt beteiligten sich 310 Personen, davon 248 Frauen und 62 Männer. Das Alter der Befragten lag zwischen 16 und 75 Jahren ( $M = 27.33$ ,  $SD = 8.3$ ). Alle Befragten hatten die Möglichkeit, an einer Verlosung teilzunehmen, bei der es drei Gutscheine im Wert von je 40 Euro zu gewinnen gab (siehe Tabelle 6 für eine exakte Zusammensetzung der Stichprobe).

#### 3.3.2 Erhebungsinstrumente in Stichprobe 1 und 2

Die TeilnehmerInnen beantworteten einen Fragebogen zum Thema „Frauen und Männer und deren Verhältnis zueinander“, der zuerst die „Skala zur Erfassung des ambivalenten Sexismus“ (ASS) (Eckes & Six-Materna, 1999) mit 22 Items erhob und anschließend die Einstellungen der Befragten zu den 59 Items des ursprünglichen Itempools

erfasste (siehe Anhang 2). Die Items wurden allen Personen in dieser Reihenfolge dargeboten. Zur Bewertung der Aussagen stand ihnen eine sechsfach abgestufte Ratingskala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“ zur Verfügung. Im letzten Teil des Fragebogens wurden die demographischen Angaben der TeilnehmerInnen abgefragt.

### 3.3.3 Ergebnisse aus Studie 1 und 2

Um eine Entscheidung darüber treffen zu können, welche Items auch weiterhin im Itempool verbleiben sollten, wurde jeweils getrennt für beide Stichproben eine psychometrische Analyse der Pretest-Daten durchgeführt.<sup>34</sup> Die Items sollten in Anlehnung an Eckes (1998) den folgenden Kriterien genügen:

- (a) *mittlere Itemschwierigkeiten* (definiert als Itemmittelwerte größer 2.0 und kleiner 5.0),
- (b) *hohe Trennschärfen*<sup>35</sup> (definiert als eine korrigierte Itemwert-Skalenwert-Korrelation größer .40),
- (c) möglichst *reine Faktorladungen* (definiert als Hauptladungen größer .55 und Nebenladungen kleiner .30.) auf dem theoretisch zu verortenden Faktor (d.h. Items, die inhaltlich dem hostilem Sexismus zuzuordnen sind, sollten nicht bzw. nur in sehr geringem Maße auf dem Faktor des benevolenten Sexismus laden oder umgekehrt).

Darüber hinaus wurden die *Item-Nonresponse-Raten* untersucht, um einen Hinweis darauf zu erhalten, welche Items möglicherweise schwer zu verstehen sind, allerdings gab es hier keine großen Ausreißer.

Nach Anwendung dieser Kriterien wurden 23 Items aus dem ursprünglichen Itempool aussortiert. Die Selektion der Items erfolgte dabei allerdings nicht ausschließlich empirie-, sondern auch theoriegeleitet. So wurden einige Items, die zwar nicht alle oben genannten Kriterien erfüllten, aber einen relevanten, inhaltlichen Aspekt abdeckten, erst einmal in dem Itempool belassen. Im Rahmen der kognitiven Interviews sollte dann überprüft werden, ob die unzufriedenstellenden Itemkennwerte möglicherweise mit einer unklaren Itemformulierung zusammenhängen. Andere Items wurden hingegen aus inhaltlichen Gründen aus dem Itempool entfernt (z.B. aufgrund thematischer Doppelungen). Letztendlich verblieben somit 36 Items in dem reduzierten Itempool (siehe Tabelle 5 und Anhang 1 für einen Überblick über die vorgenommenen Veränderungen und Reduzierungen des Itempools im Anschluss an die einzelnen Pretest-Phasen).

---

<sup>34</sup> Alle nachfolgenden quantitativen Analysen wurden mittels SPSS 16.0 bzw. AMOS 20.0 für Windows durchgeführt.

<sup>35</sup> Die Trennschärfen wurden jeweils getrennt für den feindseligen und den benevolenten Sexismus berechnet.

## 3.4 Pretest-Phase 2: Überprüfung des Itemverständnisses mit Hilfe kognitiver Interviews

### 3.4.1 Überblick

Um das intendierte Frageverständnis zu überprüfen, folgte im Anschluss an Pretest-Phase 1 die Testung der Items mit Hilfe der Methode der kognitiven Interviews. Wie bereits erwähnt, wurde den Empfehlung von Willis (2005) zufolge ein ExpertInnen-Review vorgeschaltet, um mögliche Frage-Probleme bereits *vor* der Durchführung der kognitiven Interviews zu antizipieren. In gemeinsamer Beratung wurden an mehreren Items kleine Änderungen (v.a. Spezifizierungen) vorgenommen und fünf zusätzliche Items entwickelt, deren Formulierungen als Alternative zu den jeweiligen Ausgangsitems geprüft werden sollten. Ziel dabei war es, die Frage-Version zu identifizieren, die am leichtesten verständlich ist. Somit wurden insgesamt 41 Items (überarbeiteter Itempool) in Pretest-Phase 2 mit Hilfe von kognitiven Techniken analysiert.

Insgesamt gab es drei Runden mit kognitiven Interviews. In Runde 1 galt es, zentrale Formierungsdefizite aufzudecken und zu beheben, in Runde 2 und 3 wurde die jeweilige Wirksamkeit der Überarbeitungen evaluiert. Diejenigen Items, die auch nach der ersten und zweiten Umformulierung Probleme bereiteten, wurden abermals revidiert. Mit der dritten Runde der kognitiven Interviews verfolgte ich darüber hinaus ein weiteres Anliegen: Das Frageverständnis sollte im Rahmen eines simulierten Telefoninterviews überprüft werden, um sicherzustellen, dass sich die neu entwickelte Skala auch für verschiedene Erhebungsmodalitäten eignet (Multi-Mode-Capability). Dementsprechend wurden den Befragten die Items auch vorgelesen und nicht wie in Runde 1 und 2 schriftlich vorgelegt.

### 3.4.2 Stichprobe 3, 4 und 5

In der Zeit vom 28.11.2008 bis zum 16.04.2009 wurden drei aufeinander folgende Runden von kognitiven Interviews mit insgesamt 40 Befragten von der Autorin sowie von vier weiteren Interviewerinnen durchgeführt.<sup>36</sup> Bei der Rekrutierung der TeilnehmerInnen wurden folgende Merkmale berücksichtigt: Geschlecht, Alter und Bildungsabschluss. Darüber hinaus sollte ein Teil der Befragten bereits Kinder haben, da Mutterschaft das zentrale Thema der zu testenden Skala ist. Bei der Auswahl der Befragten wurde allerdings kein repräsentativer Anspruch verfolgt. Die Zusammensetzung der Stichprobe sollte le-

---

<sup>36</sup> Vor der Feldphase hatten diese vier Interviewerinnen an einer dreitägigen Schulung teilgenommen, die sie über die kognitiven und kommunikativen Prozesse von Befragungssituationen, die Ziele, die Durchführung, die einzelnen Techniken und die Problempunkte kognitiver Interviews sowie über die Theorie des ambivalenten Sexismus informierte. Die Interviewerinnen führten in der ersten Runde mit Hilfe eines halbstandardisierten Evaluationsfragebogens insgesamt 16 Interviews (Id.11 bis Id.26) durch – jede von ihnen vier.

diglich heterogen sein, so dass das Itemverständnis anhand unterschiedlicher Bevölkerungsgruppen untersucht werden konnte. Für die erste Runde wurden 26, für die zweite Runde 9 und für die dritte Runde 5 InterviewpartnerInnen ausgewählt. (Für die genaue Zusammensetzung der Stichprobe siehe Tabelle 6).<sup>37</sup>

Die Erhebungen, die von den vier Interviewerinnen aus Bielefeld durchgeführt wurden, fanden in verschiedenen privaten Räumlichkeiten statt und dauerten zwischen 33 und 77 Minuten ( $M = 52.03$ ,  $SD = 12.83$ ). Die Befragten, die allesamt aus dem Familien-, Freundes- und Bekanntenkreis der Interviewerinnen stammten, erhielten z. T. für ihre Teilnahme als kleine Aufmerksamkeit eine Tafel Schokolade.

Als Erhebungsort für die übrigen Interviews, die allesamt von der Autorin durchgeführt wurden, diente ein Büroraum an der Philipps-Universität Marburg. Die Interviews dauerten zwischen 59 und 101 Minuten ( $M = 83.46$ ,  $SD = 9.97$ ). Als Aufwandsentschädigung bekamen die TeilnehmerInnen, die per Kleinanzeige im Raum Gießen und Marburg rekrutiert worden waren, jeweils 30 Euro.

### 3.4.3 Erhebungsinstrumente und Durchführung in Stichprobe 3, 4 und 5

Der Ablauf der kognitiven Interviews sah folgendermaßen aus: Zu Beginn wurde den Befragten eine standardisierte Instruktion vorgelesen, in der sie über die Vorgehensweise und die Ziele kognitiver Interviews informiert wurden. Zur Thematik des Fragebogens wurde ihnen – ebenso wie bei den quantitativen Studien in Pretest-Phase 1 – gesagt, dass es um „Frauen und Männer und deren Verhältnis zueinander“ gehe. Alle Befragten erhielten eine Liste (siehe Anhang 3) mit den zu testenden Items und der sechsfach abgestuften Skala (von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“), die von der jeweiligen Interviewerin erläutert wurde. Abschließend wurde das Einverständnis zur Aufzeichnung der Interviews eingeholt, die vertrauliche Behandlung der Daten versichert sowie evtl. noch offen gebliebenen Fragen zur Vorgehensweise geklärt. Danach begann das eigentliche Interview mit der Untersuchung der einzelnen Items.

Um mögliche Verständnisschwierigkeiten seitens der Befragten aufdecken zu können, wurde von den Interviewerinnen in allen drei Stichproben ein halbstandardisierter „Evaluationsfragebogen“ eingesetzt (siehe Anhang 4). Dieser enthielt einerseits die Items, die überprüft werden sollten, und andererseits die kognitiven Methoden, mit deren Hilfe dies geschehen sollte (siehe Unterkapitel 2.6.1.4) (vgl. Prüfer & Rexroth, 2000, S. 10 f.). Überwiegend wurde mit der kognitiven Technik des Probing (Hinterfragen von Antworten) gearbeitet, da diese Methode sich in der Vergangenheit als besonders effektiv erwiesen hat (vgl. Kurz, Prüfer & Rexroth, 1999; Willis, 2005). In der Regel wurde den Befragten zunächst eine allgemeine Nachfrage gestellt: Entweder erkundigten sich die

---

<sup>37</sup> Da in Runde 1 Personen mit höherem Bildungshintergrund überrepräsentiert waren, wurden in Runde 2 gezielt Personen ausgesucht, die als höchsten schulischen Abschluss einen Hauptschulabschluss bzw. eine Mittlere Reife besaßen.



Interviewerinnen bei den Befragten, was sie bei der Beantwortung der Frage gedacht hätten<sup>38</sup> oder sie wurden gebeten anzugeben, weshalb sie sich einen bestimmten Skalenwert gewählt hatten („category selection probing“). Lagen Vermutungen hinsichtlich der Verständnisschwierigkeit etwa von Fachbegriffen oder vermeintlich vagen Formulierungen vor, so wurden anschließend eine oder mehrere Nachfragen zum Begriffsverständnis („comprehension probing“) gestellt.

Der überwiegende Teil der Probingfragen wurde bereits im Vorfeld der kognitiven Interviews formuliert („scripted probes“, Willis, 1999). Diese wurden allen Befragten gestellt, so dass die Ergebnisse aus den einzelnen Interviews über die verschiedenen TeilnehmerInnen hinweg besser miteinander verglichen werden konnten. Des Weiteren wurden, je nach Verlauf des Gespräches, spontan weitere kognitive Methoden (Paraphrasieren oder „spontaneous probes“, Willis, 1999) eingesetzt, um den eigentlichen Ursachen für die Verständnisschwierigkeiten noch genauer auf den Grund gehen zu können. Nachdem alle Items auf diese Weise hinterfragt worden waren, wurde zuletzt die folgende unspezifische Zusatzfrage („general probing“) gestellt: „Hatten Sie eben bei Beantwortung der Fragen ansonsten irgendwelche Probleme (mit einem Begriff, einer bestimmten Formulierung usw.)? [Wenn ja:] Welche waren das?“ Auf diese Weise sollten den Befragten die Möglichkeit gegeben werden, weitere Schwierigkeiten zu schildern, die zuvor durch die Interviewerinnen nicht angesprochen worden waren.

Der Evaluationsfragebogen war so konzipiert, dass die Dauer der Befragung insgesamt einen Zeitrahmen von eineinhalb Stunden nicht überschreitet, da dies andernfalls zu Lasten der Motivation und Konzentration der Befragten gehen könnte (vgl. Kurz, Prüfer & Rexroth, 1999, S. 87). Folglich konnten nicht alle 41 Items des überarbeiteten Itempools gleichermaßen intensiv überprüft werden, sondern es musste eine Auswahl getroffen werden. Hierfür wurden einerseits die Ergebnisse der Analyse der Pretest-Daten aus Phase 1, wie z.B. eine geringe Itemwert-Skalenwert-Korrelation oder uneindeutige Faktorladungen, miteinebezogen. Andererseits wurden die Ergebnisse des ExpertInnen-Reviews berücksichtigt, die Hinweis auf mögliche Verständnisschwierigkeiten im Zusammenhang mit einzelnen Items gaben. Des Weiteren flossen eigene Vermutungen bezüglich etwaiger Formierungsdefizite in die Entscheidung mit ein, welche Items im Rahmen der kognitiven Interviews eingehender überprüft werden sollten.

Die 41 Items des überarbeiteten Itempools wurden folglich in zwei Kategorien eingeteilt: (a) Die vermeintlich „problematischen“ Items, von denen angenommen wurde, dass sie den Befragten Verständnisschwierigkeiten bereiten könnten und die daher intensiv getestet werden sollte (Runde 1: 17 Items; Runde 2: 23 Items; Runde 3: 20 Items) und (b) solche, die auf den ersten Blick unproblematisch erschienen (Runde 1: 24 Items; Runde 2: 12 Items; Runde 3: 20 Items).

---

<sup>38</sup> Im Verlauf der einzelnen Interviews wurde immer wieder die folgende Beobachtung gemacht: Einige Befragten begannen mit der Zeit von sich aus sofort, nachdem sie das Item gelesen hatten, laut auszusprechen, was ihnen gerade durch den Kopf geht. (Man könnte dies als *spontaneous*-think-aloud bezeichnen). Wenn dies der Fall war, wurde die oben aufgeführte Nachfrage nicht mehr gestellt.

Die kognitiven Interviews wurden dementsprechend in zwei Teile untergliedert. Im ersten Teil des Interviews wurde das Verständnis der „problematischen“ Items (wie oben beschrieben) ausführlich hinterfragt. Dabei lasen sich die Befragten die einzelnen Items zuerst selbst auf ihrer Itemliste durch und beantworteten sie dann anhand der vorgegebenen Skala.<sup>39</sup> Direkt nachdem sich die Befragten für eine Antwortkategorie entschieden hatten und diese von der Interviewerin auf dem Evaluationsfragebogen vermerkt worden war, wurden die vorbereiteten Nachfragen zum Verständnis gestellt, um die Items zu testen. Kurz, Prüfer und Rexroth (1999) zufolge hat diese Vorgehensweise den Vorteil, dass den Befragten ihre Überlegungen und Gedankengänge zu den jeweiligen Items noch leichter zugänglich sein sollten, als nach Beantwortung des gesamten Fragebogens. Allerdings ist es bei dieser Vorgehensweise nicht (bzw. nur schwer) möglich, Kontexteffekte zu prüfen, da die „normale“ Itemabfolge durch die jeweiligen Nachfragen unterbrochen wird (vgl. Kurz, Prüfer, & Rexroth, 1999, S. 91). Die Testung von Kontexteffekten war jedoch in dieser Untersuchung nachrangig, da kein endgültiger Fragebogen, sondern lediglich eine Itemsammlung geprüft werden sollte.

Im zweiten Teil des Interviews wurden hingegen die übrigen – vermeintlich unproblematischen Items – mit Hilfe einer „Mischform“ aus Beobachtungspretest und kognitivem Interview (vgl. Faulbaum, Prüfer, & Rexroth, 2009, S. 99) folgendermaßen überprüft: Die TeilnehmerInnen wurden gebeten, die verbliebenen Items selbständig wie bei einer „normalen“ schriftlichen Befragung auf ihrer Itemliste zu beantworten und dabei jedes Item, das ihnen Verständnisprobleme bereitete, durch einen Kreis an Rande des Fragebogens zu markieren. Nur diese Items wurden dann mit Hilfe spontaner Probingfragen und/oder der Technik des Paraphrasierens intensiver untersucht.

Am Ende des Interviews wurden die demographischen Angaben der befragten Personen (Geschlecht, Alter, höchster Bildungsabschluss, Familienstand, Anzahl und Alter der Kinder sowie die eigene Erwerbstätigkeit als auch die der Partnerin / des Partners) erhoben.

---

<sup>39</sup> In Runde 3 wurde ein anderes Vorgehen gewählt, da annähernd die gleichen Rahmenbedingungen wie bei einer Telefonbefragung herrschen sollten. Um die Behaltensleistung und das Verständnis der Befragten angemessen untersuchen zu können, wurden die einzelnen Items daher nicht visuell, sondern lediglich akustisch dargeboten. Auf diese Weise sollte überprüft werden, ob die Items beispielweise zu lang oder zu kompliziert formuliert sind und sich daraus Probleme ergeben. Aus diesem Grund wurde den Befragten in dieser Runde auch keine Itemliste vorgelegt, sondern die Items von der Interviewerin vorgelesen. Lediglich die sechsfach abgestufte Skala erhielten die Befragten auf einem extra Blatt Papier. Dies sollte als Hilfestellung dienen, da es normalerweise unüblich ist, eine so große Zahl an Antwortmöglichkeiten bei einem Telefoninterview einzusetzen. Da jedoch in den beiden vorhergehenden Runden die sechstufige Skala verwendet wurde, wurde auch in der dritten Runde der kognitiven Interviews der Einheitlichkeit halber daran festgehalten.

### 3.4.4 Auswertung der kognitiven Interviews

Da bisher kein einheitlicher Standard für die Auswertung kognitiver Interviews existiert, wurde ein eigenes Verfahren entwickelt (zu den unterschiedlichen Vorgehensweisen siehe Unterkapitel 2.6.1.5). So wurde eine *informelle Form der Analyse* (qualitative Vorgehensweise) auf der Basis von *Transkriptionen* gewählt, da die Verschriftlichung der Interviews die detaillierteste Form der Auswertung erlaubt. Die informelle Analyse wurde der formellen vorgezogen, da bei der quantitativen Vorgehensweise durch die Kodierung Informationen verloren gehen und sie nur eine Bezifferung der aufgetreten Probleme, nicht jedoch deren exakte Beschreibung erlaubt. Für eine angemessene Umformulierung der Items wäre daher ohnehin ein Rückgriff auf die exakteren Problembeschreibungen der formellen Analyse nötig gewesen. Des Weiteren lag der *Fokus der Analyse auf den einzelnen Items* und nicht auf der Untersuchung der Verlaufsgeschichte der Interviews, weil vor allem mögliche Formierungsdefizite einzelner Items und weniger mögliche Kontextprobleme des Fragebogens im Zentrum des Interesses standen.

Die Vorgehensweise der Analyse sah folgendermaßen aus: Die Antworten der einzelnen Befragten wurden stets sofort stichpunktartig mitgeschrieben und zusätzlich mit einem Laptop oder Diktiergerät aufgezeichnet. Darüber hinaus wurden bereits während des Interviews die Items markiert, die den Befragten anscheinend Verständnisschwierigkeiten bereiteten. Im Anschluss an die einzelnen Interviews notierten die Interviewerinnen – getrennt für jedes Item – Hinweise auf besonders problematisch erscheinende Formulierungen sowie allgemeine Kommentare und Beobachtungen im Zusammenhang mit den einzelnen Fragen in einer elektronischen Version des Fragebogens („Problemtabelle“) (siehe Anhang 5). Auf diese Weise sollten erste Eindrücke bzgl. der Fragemängel festgehalten werden.

In einem zweiten Schritt wurde für die Analyse weiter in die Tiefe gegangen. Zuerst wurden die Interviewaufnahmen transkribiert und hinterher itemweise gruppiert, so dass sich die Äußerungen aller InterviewteilnehmerInnen zu ein- und derselben Frage in einer Word-Datei befanden. Auf diese Weise konnten die aufgetretenen, itemspezifischen Probleme über die verschiedenen Befragten hinweg besser miteinander verglichen werden.

Die Interviewsequenzen wurden in einem dritten Schritt auf Verständnisschwierigkeiten, die im Zusammenhang mit den einzelnen Items aufgetreten waren, hin analysiert. Dabei wurden die jeweiligen Kernaussagen der Befragten sowie Hinweise auf potentielle Frage-Probleme farbig markiert, um die Datenmenge übersichtlicher zu gestalten. Besonders aussagekräftige Teilsequenzen wurden in die Problemtabelle, die bereits die Anmerkungen der Interviewerinnen enthielt, kopiert.

Gemäß den Empfehlungen von Prüfer und Rexroth (2005) enthielt die Problemtabelle somit für jedes untersuchte Item die folgenden fallspezifischen Informationen:

- (1) Die gewählten Antwortkategorien der Befragten,
- (2) die spontanen Reaktionen der Befragten,

- (3) die Kernaussagen der Befragten auf die Nachfragen,
- (4) die Kommentare der InterviewerInnen (vgl. Prüfer & Rexroth, 2005, S. 18).

Auf Basis dieser Informationen wurden die zentralen Formulierungsdefizite identifiziert und die Items letztendlich überarbeitet.

### 3.4.5 Ergebnisse der kognitiven Interviews aus Runde 1 bis 3

Im Folgenden werden die Ergebnisse der kognitiven Interviews auf Grund der Fülle des Materials nur beispielhaft für einige Items vorgestellt. Ausgewählt wurden hierfür diejenigen Items, die letztendlich auch in die endgültige FWSEM-Skala aufgenommen wurden (siehe hierzu Unterkapitel 3.5.7.1) sowie ein weiteres Items, bei dem keine zufriedenstellende Umformulierung gefunden werden konnte.

Jeweils getrennt für den feindseligen und den wohlwollenden Sexismus wird itemweise beschrieben, wie die ursprüngliche Formulierung der Items lautete, was die zugrunde liegende Intention war und welche zentralen Probleme im Rahmen der kognitiven Interviews ausfindig gemacht wurden. Anschließend werden die Überlegungen erläutert, die im Rahmen der Überarbeitungen angestellt wurden, und die endgültige Itemversion benannt.

In den Interviewausschnitten, die jeweils zur Veranschaulichung angeführt werden, sind den Antworten der Befragten die jeweilige laufende Nummer des Interviews (z.B. „Id. 01“ für Interview Nummer 01) vorrangestellt. Die Nachfragen der Interviewerinnen werden hingegen durch das Kürzel „Int.“ eingeleitet. Interviewpassagen, während denen die InterviewteilnehmerInnen kurze Pausen einlegten oder sie einen Gedankengang nicht zu Ende brachten, sind durch drei Punkte ... gekennzeichnet. Sequenzen, die hingegen aufgrund ihrer Länge, Unbedeutsamkeit oder Doppelung weggekürzt wurden, sind durch drei Punkte innerhalb einer runden Klammer (...) markiert. Eigene Anmerkungen oder Ergänzungen wurden in eckigen Klammern [ ] eingefügt.

#### 3.4.5.1 Feindseliger Sexismus

Um die Ergebnisse der kognitiven Interviews bezüglich des feindseligen Sexismus (FS) auszugsweise vorzustellen, wird auf insgesamt sieben Items näher eingegangen. Die ersten sechs Items wurden in die endgültige Version der FS-Subskala aufgenommen, das letzte Item wurde hingegen aus dem Itempool entfernt, da es nicht gelang, es in einer zufriedenstellenden Weise zu überarbeiten. Während die ersten vier Items in Anlehnung an die dritte Inhaltskategorie (c) „Karrierefrauen werden als Rabenmütter und schlechte

Ehefrauen angesehen“ formuliert wurden, basieren die folgenden drei Items auf der zweiten Kategorie (b) „Die Motive berufstätiger Frauen bzw. Mütter werden abgewertet“ (für die exakte Formulierung der Ursprungsitems siehe Tabelle 7).<sup>40</sup>

*Tabelle 7: Itemformulierungen der FS-Items, die in Runde 1 der kognitiven Interviews getestet wurden*

Ursprungsitems	Itemkürzel
Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.	<i>Rabenmütter</i>
Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.	<i>Schlechtes Gewissen</i>
In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass sie ihren Kindern zu wenig Aufmerksamkeit schenken.	<i>Zu wenig Aufmerksamkeit</i>
Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen dabei die Bedürfnisse ihrer Familie.	<i>Bedürfnisse vernachlässigen</i>
Ich finde Frauen egoistisch, die bereits wenige Wochen nach der Geburt ihres Kindes wieder in ihren Beruf zurückkehren, weil sie ihre Karriere weiter verfolgen möchten.	<i>Selbstverwirklichung</i>
Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.	<i>Berufliche Anerkennung</i>
Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.	<i>Verständnis</i>

Bei den ersten beiden Items *Rabenmütter* und *Schlechtes Gewissen* werden berufstätige Mütter als „schlechte“ Mütter hingestellt, da sie die Betreuung der Kinder aufgrund ihrer Erwerbstätigkeit nicht selbst rund um die Uhr übernehmen können und daher ihr Kind

<sup>40</sup> Es ist v.a. auf die Ergebnisse der Faktorenanalysen in Pretest-Phase 3 (siehe Unterkapitel 3.5) zurückzuführen, dass die Inhaltskategorie (a) „Karrierefrauen werden feminin-positive Eigenschaften abgesprochen“ nicht mehr in der endgültigen FS-Subskala vertreten ist. So luden die dazugehörigen Items u.a. mit bereits bestehenden Sexismus-Skalen auf einem gemeinsamen Faktor und brachten somit inhaltlich nichts Neues.

außerhäuslich betreuen lassen. Die Items *Zu wenig Aufmerksamkeit* und *Bedürfnisse vernachlässigen* betonen schließlich die Unvereinbarkeit der Mutterrolle und der Berufsrolle sowie die scheinbar negativen Folgen für die Kinder, die sich aus der außerhäuslichen Erwerbstätigkeit der Mütter ergeben. Den Müttern wird dabei unterstellt, sie würden ihrer Mutterrolle nicht gerecht und vernachlässigten die Kinder.

Die drei Items *Selbstverwirklichung*, *Berufliche Anerkennung* und *Verständnis* rücken hingegen die Ursache für die Berufstätigkeit der Mutter in den Vordergrund. Es sollte hierbei keine finanzielle Notwendigkeit für ihre Berufstätigkeit bestehen, da sich in einem solchen Fall die Frauen schließlich nicht freiwillig für ihren Beruf entschieden hätten und dies vermutlich anders bewertet worden wäre, da ihnen dann keine wirklich Wahl bliebe (vgl. hierzu Bridges & Etaugh, 1995). Vielmehr sollten persönliche Gründe ausschlaggebend sein, wie zum Beispiel der Wunsch nach beruflicher Anerkennung. Diese Motive stoßen jedoch auf Unverständlichkeit, da eine „gute“ Mutter die eigenen Bedürfnisse hinter denen des Kindes zurückstellen sollte.

### **Ursprungitem *Rabenmütter***

*„Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.“*

Das Item *Rabenmütter* bezieht sich auf die „Unabkömmlichkeit der Mutter“ (siehe Muttermythos in Unterkapitel 2.1.1.2), nach der eine intensive, ganztägige Betreuung durch die eigene Mutter das Beste für das Kind ist. So werden in dem Item Frauen beschrieben, die ihre Kinder verhältnismäßig früh (was durch den Begriff „kleine Kinder“ in Kombination mit der Betreuungsform, nämlich einer „Kinderkrippe“, signalisiert werden sollte) in eine außerhäusliche, staatliche Betreuungseinrichtung geben. Diese Form der Betreuung wurde bewusst gewählt, da sie in Westdeutschland weniger gerne gesehen wird als die mütterliche Betreuung – insbesondere dann nicht, wenn die Kinder sich noch nicht im Kindergartenalter befinden (siehe Unterkapitel 2.1.1.3; vgl. Schneider & Dorbritz, 2011, S. 32 f.).

Des Weiteren wird in dieser Aussage der Grund, der eine Fremdbetreuung des Kindes erforderlich macht, betont: Die Mütter entscheiden sich nicht aus einer finanziellen Notwendigkeit heraus zu diesem Schritt, sondern haben scheinbar ein ganz egoistisches Motiv: ihre Karriere. Dieses Verhalten lässt die Frauen aus der Sicht einer feindselig-sexistisch eingestellten Person zu schlechten Müttern, eben „Rabenmüttern“ werden, da sie ihre Mutterrolle – vermeintlich – nicht angemessen erfüllen.<sup>41</sup>

---

<sup>41</sup> Durch die Wertung der Mütter als „Rabenmütter“, was eine sehr starke Verurteilung darstellt und eine Missbilligung des Verhaltens deutlich ausdrückt, stellt dieses Item – trotz eines ähnlichen inhaltlichen Schwerpunktes – eine extremere Variante dar als das nachfolgende Item *Schlechtes Gewissen*. Folglich ist zu erwarten, dass auch nur Personen, die eine äußerst feindselige Einstellung besitzen, dieser Aussage voll

Als Ergebnis der kognitiven Interviews bleibt festzuhalten, dass das Ursprungsitem *Rabenmütter* noch nicht in zufriedenstellender Weise verstanden wurde. So wussten nicht alle Befragten, was genau unter dem Begriff „Kinderkrippe“ zu verstehen ist und in welchem Alter Kinder üblicherweise dorthin gehen. Dementsprechend schien auch die Formulierung „kleine Kinder“ noch zu ungenau zu sein, da die Befragten auf die entsprechende Nachfrage hin Altersangaben zwischen null und sechs Jahren machten. Die Hälfte (13) der befragten Personen dachte an Kinder im Alter zwischen 0 und 3 Jahren, welches auch die intendierte Altersspanne war, wobei 8 von ihnen Säuglinge (d.h. Kinder unter einem Jahr) in ihre Überlegungen mit berücksichtigten. Allerdings gab es ebenfalls sieben Befragte die die Aussage auf deutlich ältere Kinder bezogen, nämlich auf Kindergartenkinder (zwischen 3 und 6 Jahren). Die übrigen Befragten bewegten sich zwischen diesen beiden Altersspannen und dachten gleichermaßen an Kinder unter sowie über drei Jahren. Damit bezogen sie zum Teil deutlich ältere Kinder in ihren Überlegungen mit ein, als dies beabsichtigt war. Wie erläutert, sollte es in der Aussage gerade darum gehen, Kinder frühzeitig außerhäuslich betreuen zu lassen und eben nicht erst – wie in den alten Bundesländern lange üblich – im Kindergartenalter.

Folglich wurde das Item hinsichtlich folgender Aspekte überarbeitet: Zum einen wurde der Begriff „Kinderkrippe“ durch das Wort „Kindertagesstätte“ ersetzt, da dieses den Befragten geläufiger zu sein schien. Zum anderen wurde der Begriff „kleine Kinder“ zunächst durch eine exakte Altersangabe in Klammern ( $\frac{1}{2}$  -  $2\frac{1}{2}$  Jahre) spezifiziert. Diese Lösung erwies sich jedoch noch nicht als zufriedenstellend, da diese Ergänzung in nachfolgenden Interviews teilweise überlesen wurde. Ausschlaggebend für die erneute Überarbeitung war zudem die Überlegung, dass die neue Skala auch bei einem Telefoninterview einsetzbar sein sollte, bei der die Items lediglich von den InterviewerInnen vorgelesen werden und die visuelle Darbietung fehlt. Unter diesen Umständen wäre die Angabe in Klammern vermutlich zu verwirrend gewesen. Daher wurde der Begriff „kleine Kinder ( $\frac{1}{2}$  -  $2\frac{1}{2}$  Jahre)“ in einem zweiten Schritt durch die Formulierung „Kinder mit einem halbem Jahr“ ersetzt. Die Untergrenze von einem halben Jahr wurde daher gewählt, da dies eine kritische Altersgrenze darzustellen scheint. Gerade das erste Lebensjahr wurde von mehreren Befragten als besonders wichtig für die Entwicklung des Kindes angesehen, so dass sie der Anwesenheit der Mutter in dieser Zeit eine große Bedeutung beimaßen. Demzufolge wird das Item für Kinder unter einem Jahr wahrscheinlich noch besser geeignet sein, um extreme Einstellungstendenzen der Befragten aufzeigen zu können.

Im Zuge dieser Überarbeitung wurde der Satzbaustein „die aus Karrieregründen“ an den Anfang des Items gestellt, um so den Lesefluss zu erleichtern („Frauen, die aus Karrieregründen ihre Kinder mit einem halbem Jahr ...“ statt „Frauen, die ihre Kinder mit einem halbem Jahr aus Karrieregründen ...“). Ferner rückte dieses Element somit an eine

---

und ganz zustimmen werden. Dieses Item sollte demnach dazu beitragen, besser zwischen den verschiedenen Abstufungen feindselig-sexistischer Einstellungen differenzieren zu können.

prominente Stelle, nämlich den Satzanfang, was die erhöhte Aufmerksamkeit der Befragten gewährleisten sollte („primacy effect“). Das Item lautete nun: „Frauen, die aus Karrieregründen ihr Kind mit einem halben Jahr in eine Kindertagesstätte geben, sind für mich Rabenmütter.“

Insgesamt lassen die Äußerungen der befragten Personen in der dritten Runde der kognitiven Interviews darauf schließen, dass das Item nach der erneuten Umformulierung so interpretiert wurde, wie es beabsichtigt war. So war ein Befragter beispielsweise der Ansicht, dass ein halbes Jahr noch sehr früh sei, um ein Kind außerhäuslich betreuen zu lassen. Folglich stimmte er der Aussage auch eher zu:

**Interview Nr. 03 (männlich): „stimme eher zu“**

**Id. 03:** (spontane Äußerung vor Nachfrage) Ein bisschen schwierig. Da sag ich dann mal ... Ein halbes Jahr ist natürlich noch sehr, sehr früh ... Ich sag mal „stimme ich eher zu“.

**Int.:** Warum haben Sie jetzt eben am Anfang gesagt, das ist schwierig?

**Id. 03:** Ja, also ich hab jetzt den Deal zwischen „stimme ich eher nicht zu“ oder „stimme ich eher zu“. Aber ich finde, das ist schon zu früh, ein halbes Jahr, aber jetzt natürlich vom Beruf her, behaupte ich mal, dass das geht.

Eine andere Befragte wiederum verneinte die Aussage, da sie den Begriff „Rabenmutter“ ablehne und eine Mutter, nur weil sie einer Berufstätigkeit nachgehe, auch nicht so bezeichnen würde:

**Interview Nr. 04 (weiblich): „stimme überhaupt nicht zu“**

**Id. 04:** (spontane Äußerung vor Nachfrage) Auf gar keinen Fall. Also: „stimme ich überhaupt nicht zu“.

**Int.:** Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen, bis Sie die Antwort gegeben haben?

**Id. 04:** Dass ich selbst zweieinhalb Jahre zu Hause geblieben bin, aber mir das auch nur leisten konnte, weil der Vater meine Tochter und mich unterstützt hat und im ungünstigsten Fall hätte ich auch sofort arbeiten müssen. Und schlimm finde ich das dann nicht. Obwohl ich der Ansicht bin, dass es sehr wichtig ist, dass in den ersten Jahren die Mutter bzw. der Vater als Bezugsperson für das Kind da ist. Aber das kann natürlich auch 'ne gute Erzieherin leisten oder 'ne Tagesmutter leisten, oder, oder, ...

**Int.:** Und warum haben Sie dann eben geantwortet „stimme überhaupt nicht zu“? Können Sie das noch mal so zusammenfassen, was den Ausschlag gegeben hat, dass Sie gesagt haben, der Aussage würde ich „überhaupt nicht zustimmen“?

**Id. 04:** Dass das 'ne Rabenmutter ist?



**Int.:** Mhm.

**Id. 04:** Nee, das ist ja keine „Rabenmutter“. Ich find den Begriff immer so schlimm. Nur, weil man aus entweder finanziellen Gründen oder aus persönlichen Gründen arbeiten will, ist man noch lange keine „Rabenmutter“. Also ich find den Begriff ganz schlimm. Kein Mensch sagt „Rabenvater“, nur, weil der sagt, ich muss das Geld verdienen. Ich find das immer so suggestiv ... Ja, das suggeriert den Menschen dann: Hier, du bist nicht in Ordnung.

Die Aussagen verdeutlichen, dass das Item nun in beabsichtigter Weise interpretiert wird, keine Unklarheiten mehr offen lässt und daher nun so beibehalten werden kann.

Endgültige Itemversion *Rabenmütter*

*„Frauen, die aus Karrieregründen ihr Kind mit einem halbem Jahr in eine Kindertagesstätte geben, sind für mich Rabenmütter.“*

**Ursprungitem *Schlechtes Gewissen***

*„Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.“*

Das Item *Schlechtes Gewissen* ähnelt von der Intention her stark dem vorhergehenden Item *Rabenmutter*. Erneut werden Mütter beschrieben, die aufgrund ihrer Berufstätigkeit nicht selbst rund-um-die-Uhr für ihre Kinder da sein können und sie daher zur Betreuung weggeben. Folglich sollten die Frauen auch – aus der Sicht einer feindselig-sexistisch eingestellten Person – ein schlechtes Gewissen haben, da sie ihrer Rolle als Mutter nicht in angemessener Weise nachkommen.

Hinsichtlich des Alters wurde angegeben, dass die Kinder sich noch *vor* dem Kindergartenalter befinden, da dem westdeutschen Mutterbild zufolge die Mutter idealerweise die ersten drei Jahre zu Hause bei ihrem Kind bleiben sollte (siehe Unterkapitel 2.1.1.3). Bezüglich der Betreuungsform wurde die Formulierung „fremde Hände“ verwendet, um auszudrücken, dass es sich nicht um eine Betreuung durch Familienangehörige handelt. Diese würde vermutlich (z.B. im Falle der Großmutter) in den alten Bundesländern weniger kritisch gesehen, als beispielsweise die Betreuung in einer Kindertagesstätte.

Auch der Begriff „erfolgreich berufstätige Frau“ wurde bewusst gewählt. So sollte er eine Alternative zu dem Wort „Karrierefrau“ darstellen, von dem angenommen wurde, dass es vermutlich eher negativ besetzt ist. Gleichzeitig sollte durch den Zusatz „erfolgreich“ deutlich werden, dass es sich bei der beschriebenen Frau um eine Person handelt, die voll im Berufsleben steht und berufliche Ambitionen hat.

Die Ergebnisse der kognitiven Interviews zeigten auf, dass auch das Item *Schlechtes Gewissen* in seiner ersten Variante zu unspezifisch formuliert war. So schien auch die Formulierung „Kinder schon vor dem Kindergartenalter“ nicht eindeutig genug zu sein. Auf die Nachfrage hin, an welches Alter die Befragten gedacht hätten, schwankten ihre Angaben zwischen sechs Monaten als Untergrenze und vier Jahren als Obergrenze. Die Mehrheit der Befragten hatte die Aussage auf Kinder bezogen, die ganz kurz vor dem Kindergartenalter standen, also knapp drei Jahre alt waren. Kinder unter einem Jahr wurden dabei von einigen Befragten explizit ausgeklammert, da sie die Anwesenheit der Mutter in dieser Zeit als besonders wichtig ansahen.

Des Weiteren erwies sich der Begriff „fremde Hände“ als nicht gut gewählt. Einerseits war er z. T. negativ besetzt und rief die Assoziation hervor, die Mütter würden ihre Kinder wahllos einer fremden Person anvertrauen. Andererseits schien ein Großteil der InterviewteilnehmerInnen „fremde Hände“ mit „anderen Händen als die der Mutter“ gleichzusetzen. Dementsprechend nannten sie im Rahmen des comprehensive probings neben Tagesmüttern, der Kinderkrippe oder Au-pair-Mädchen, auch Familienangehörige wie die Großeltern oder die Tante als Betreuungspersonen. Letztere waren jedoch bei der Formulierung des Items nicht intendiert gewesen.

Zudem ließen die Äußerungen mehrerer befragter Personen erkennen, dass ihnen der Grund fehlte, weshalb die Mutter einer Erwerbstätigkeit nachgeht und ihre Kinder wegibt. Die Formulierung „erfolgreich berufstätige Frau“ wurde zwar, wie beabsichtigt, positiver ausgelegt als der Begriff „Karrierefrau“, aber er erweckte nicht die gleichen Assoziationen hinsichtlich des Umfangs der Erwerbstätigkeit und der Karriereambitionen. So gaben mehrere InterviewteilnehmerInnen an, für sie könne eine „erfolgreich berufstätige Frau“ auch eine Frau sein, die nicht zwangsläufig vollzeiterwerbstätig sein müsse oder eine gehobene Stellung anstrebe. (Sogar Hausfrauen wurden mitunter dazugezählt.) Es gehe vielmehr darum, dass die Frau erfolgreich sei, in dem, was sie tue und dass sie Anerkennung dafür bekomme.

Beispiel 1:

**Int.:** Was bedeutet für Sie der Begriff „erfolgreiche berufstätige Frau“? Was verstehen Sie darunter?

**Id. 03 (weiblich):** Der Unterschied vielleicht zwischen den zweien? [Anmerkung d. Verf.: Der Unterschied zwischen den beiden Begriffen „erfolgreiche berufstätige Frau“ und „Karrierefrau“]

**Int.:** Ja, können Sie auch sagen.

**Id. 03:** „Erfolgreich“ ... Also ich denke, die „Karrierefrau“, die, die ist halt bestrebt danach immer höher zu kommen und die „erfolgreiche Frau“ kann auch jemand sein, die in ihrem Beruf auf ihrer Position einfach erfolgreich ist. [Pause] Die „erfolgreiche

Frau“ kann ja auch ’ne Mutter sein. ’Ne Hausfrau und Mutter kann ja auch ’ne erfolgreiche ...

**Int.:** Aber wenn da steht: „erfolgreich berufstätige Frau“.

**Id. 03:** Okay. Ja, dann kann sie aber, egal auf welcher Position, sagen wir’s so, kann sie ja erfolgreich sein. Aber „Karrierefrau“ würde ich eher halt sagen, dass das Bestreben ist, immer höher zu kommen. Und „erfolgreich“ hat auch nichts unbedingt damit zu tun, sag ich mal... dann muss ich nicht Überstunden machen. Dass hat auch nichts mit der Arbeitszeit zu tun, auch das nicht. „Erfolgreich“ kann ich ja auch so sein.

Beispiel 2:

**Int.:** Was bedeutet für Sie der Begriff „erfolgreiche berufstätige Frau“? Was verstehen Sie darunter?

**Id. 07 (weiblich):** Das ist noch mal, dadrunter [Anmerkung d. Verf.: unter Karrierefrau]. So eine Stufe drunter würde ich die ein ... [Pause] Ich bin auch erfolgreich berufstätig [lacht]. Also das ist, wenn man, ähm, gute Rückmeldungen bekommt von Mitarbeitern, Vorgesetzten. Dann ist man schon erfolgreich. Und das kann man auch, wenn man, weiß ich nicht, wenn man jetzt irgendwie nur dreimal in der Woche arbeiten gehen würde, so.

Aufgrund dieser Erkenntnisse wurde das Item im Hinblick auf drei Aspekte umformuliert. Erstens wurde eine exakte Altersangabe vorgegeben. Da die Formulierung „Kinder schon vor dem Kindergartenalter“ scheinbar noch zu ungenau ist, hängt deren Auslegung von der Interpretation der jeweiligen Befragten ab. Dementsprechend ist es durchaus möglich, dass einige Personen davon ausgehen, es seien Kinder kurz vor dem Kindergartenalter gemeint, andere Personen könnten hingegen annehmen, auch Kinder unter einem Jahr seien intendiert. Letztere Auslegung könnte wiederum dazu führen, dass die Befragten eher geneigt sind, der Aussage zuzustimmen, da Kinder noch sehr klein sind. Eine Vergleichbarkeit der Antworten bzw. eine Einordnung der Extremheit der Einstellungen ist somit unter diesen Bedingungen nicht möglich, da man nicht weiß, von welchen Umständen die Befragten ausgegangen sind. Folglich wurde das Alter explizit festgelegt. Parallel zu den Überlegungen, die im Rahmen der Itemüberarbeitung *Rabenmütter* angestellt wurden, erfolgte in einem ersten Schritt eine detaillierte Altersangabe in Klammern („Kinder schon vor dem Kindergartenalter (½ - 2½ Jahre)“), die jedoch aufgrund der bereits geschilderten Probleme in einem zweiten Schritt durch die Formulierung „Kind mit einem halben Jahr“ ausgewechselt wurde.

Zweitens wurde der Begriff „fremde Hände“ ersetzt, da mehrere Befragte – entgegen der Intention – an Verwandte dachten, die die Betreuung der Kinder übernehmen. Stattdessen war eine nicht-familiäre, „externe“ Betreuungsform gemeint gewesen. Um Missverständnissen vorzubeugen, wurde daher auch die Art der Betreuung explizit festgelegt.

Da in dem Item *Rabenmutter* bereits die Kinderkrippe bzw. Kindertagesstätte vorkam, fiel die Wahl für dieses Item auf die Tagesmutter.

Drittens erzielte die Bezeichnung „erfolgreiche berufstätige Frau“ nicht die intendierte Interpretation einer Frau, die voll im Erwerbsleben steht und für die ihr berufliches Vorankommen bedeutsam ist. Mehreren Befragten fehlte folglich auch ein Beweggrund, weshalb die Mutter ihr Kind zur Betreuung weggibt. Um dieses Problem zu beheben, wurde auf eine Formulierung zurückgegriffen, die bereits bei dem Item *Rabenmütter* zum Einsatz kam und die die Karriere ausdrücklich als Anstoß für die Erwerbstätigkeit der Mutter benennt. Die endgültige Version des Items lautet daher folgendermaßen:

Endgültige Itemversion *Schlechtes Gewissen*

„Frauen sollten ein schlechtes Gewissen haben, wenn sie aus Karrieregründen ihr Kind mit einem halben Jahr tagsüber zur Betreuung zu einer Tagesmutter geben.“

**Ursprungsitem *Zu wenig Aufmerksamkeit***

„In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass sie ihren Kindern zu wenig Aufmerksamkeit schenken.“

Das Item *Zu wenig Aufmerksamkeit* thematisiert die Unvereinbarkeit der Mutter- und der Berufsrolle (siehe Unterkapitel 2.3.2.2). Der Fokus liegt dabei auf den (vermeintlich) negativen Konsequenzen, die sich aus der außerhäuslichen Erwerbstätigkeit der Mütter für die Familie ergeben. So wird unterstellt, die Mütter hätten nicht mehr genügend Zeit für ihre Kinder und könnten ihnen nicht mehr genügend Aufmerksamkeit zuteil werden lassen. Auch in diesem Fall werden die Frauen somit, wenn auch eher indirekt, als schlechte Mütter hingestellt, die ihrer Rolle nicht gerecht werden.

Als zentrales Ergebnis der kognitiven Interviews bleibt festzuhalten, dass das Item *Zu wenig Aufmerksamkeit* unerwarteterweise Probleme bereitete. So wies das Item ein Formulierungsdefizit auf, welches weder im Rahmen der Fragenkonstruktion noch im ExpertInnen-Review antizipiert wurde: Es enthielt zwei Stimuli („In der heutigen Zeit sind viele Frauen auf ihre Karriere fixiert“ und „Diese Frauen schenken ihren Kindern zu wenig Aufmerksamkeit“), was jedoch bei der Itemformulierung vermieden werden sollte. Für Befragte kann es nämlich schwierig werden, sich für eine Antwortkategorie zu entscheiden, wenn sie einem Teil der Aussage zustimmen, dem anderen hingegen nicht (vgl. Porst, 2000, S. 7). Darüber hinaus ist der erste Teil als „Unterstellung“ (vgl. Porst, 2000, S. 7) formuliert („In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert“), was einigen Befragten ebenfalls Schwierigkeiten bereitete. Die folgenden Interviewsequenzen verdeutlichen die Probleme:

Beispiel1:

**Interview Nr. 03 (weiblich): „stimme eher zu“**

**Id. 03:** (spontane Äußerung vor Nachfrage) Da denke ich halt jetzt drüber nach ... Ich denke, dass es gar nicht so „viele“ Frauen gibt, die da drauf fixiert sind.

**Int.:** Mhm, ah, okay.

**Id. 03:** Ja, das muss ich ja dann jetzt aber auf die beziehen, die das machen, oder? Weil ich denke von vorneherein erst mal, dass es gar nicht so „viele“ Frauen sind, die da drauf aus sind, nur ... Jetzt müsste ich wissen, inwieweit ... auf was beantworte ich die Frage? Soll ich das auf die Karrierefrauen beziehen?

Beispiel2:

**Interview Nr. 15 (männlich): „stimme nicht zu“**

**Id. 15:** Also, z.B. Frage 28. Man hat oben immer wieder gesagt, „ja, die Frau, dass die Frau sich zu wenig der Familie oder grade kleinen Kindern, zu wenig widmen kann, wenn sie Karriere macht, oder, wenn sie im Beruf überhaupt tätig ist und erfolgreich ist. Und hier wird jetzt gesagt: „In der heutigen Zeit sind viele Frauen ...“ Also, man geht nicht nur von wenigen aus, sondern von „vielen“, die derart „auf ihre Karriere fixiert, dass sie ihren Kindern zu wenig Aufmerksamkeit schenken“.

**Int.:** Ja.

**Id. 15:** Also, da sind so viele Aussagen drin! Erstmal sind's viele Frauen, das ist die erste Aussage, also im Vergleich zur Gesamtheit der Frauen. Dann geht man davon aus, dass diese vielen Frauen Karriere machen wollen, was bei den meisten Frauen gar nicht der Fall ist. Die wollen einfach nur Geld verdienen, weil der Mann nicht genug verdient. „Dass sie ihren Kindern zu wenig Aufmerksamkeit schenken“. Also die Frage an sich ist komplex. Da werden mehrere Aspekte ... wie soll ich sagen ... eines Sachverhalts angesprochen. Und darauf eindeutig zu antworten, das ist schwierig. Jetzt hab ich aber hier mal gesagt: „Viele Frauen auf ihre Karriere fixiert“... und da geh ich schon mal nicht von aus, das ist schon Grund genug für mich zu sagen: „stimme nicht zu“. „Dass sie ihren Kindern zu wenig Aufmerksamkeit schenken“... d.h., in diesen Fragen sind eigentlich zwei Fragen drin, ne. Das ist sehr komplex muss ich sagen, also diese Frage wäre mir zu kompliziert.

Auch wenn nur zwei Befragte dieses Problem ansprachen, schien es so nachvollziehbar, dass es eine Überarbeitung erforderlich werden ließ. Bei der Entscheidung, welcher der beiden Satzteile beibehalten werden sollte, fiel die Wahl auf den zweiten Part, da dieser den zentralen Aspekt enthält: Frauen, die aus ihrer traditionellen Rolle ausbrechen und eine Karriere verfolgen, kommen aus der Sicht einer feindselig-sexistisch eingestellten Person ihren mütterlichen Aufgaben nicht angemessen nach, worunter die Kinder zu

leiden haben. Um den ersten Satzteil nicht komplett wegfallen zu lassen und die beruflichen Ambitionen der Frauen zu betonen, wurde die Formulierung „erfolgreich berufstätige Frau“ verwendet („Erfolgreiche berufstätige Frauen schenken ihren Kindern häufig zu wenig Aufmerksamkeit“). Da dieser Begriff jedoch, wie bereits erwähnt, nicht die gewünschten Assoziationen bei den Befragten hervorrief, wurde er nach der zweiten Runde der kognitiven Interviews durch das Wort „Karrierefrau“ ersetzt.

Endgültige Itemversion *Zu wenig Aufmerksamkeit*

„Ich denke, dass Karrierefrauen ihren Kindern häufig zu wenig Aufmerksamkeit schenken.“

### **Ursprungitem *Bedürfnisse vernachlässigen***

„Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen dabei die Bedürfnisse ihrer Familie.“

Das Item *Bedürfnisse vernachlässigen* thematisiert ebenfalls die negativen Konsequenzen der Berufstätigkeit einer Mutter für die Familie. Erneut ist in diesem Fall die Formulierung „erfolgreiche berufstätige Frauen“ gewählt worden, um auszudrücken, dass Mütter gemeint sind, die nicht – wie häufig der Fall in Deutschland – einen Teilzeit- oder Minijob haben (siehe Unterkapitel 2.1.3), sondern eine Vollzeitstelle. Zudem wird betont, dass sie Karrierebestrebungen besitzen und nicht etwa aufgrund einer finanziellen Notwendigkeit heraus so viel Zeit in ihren Beruf investieren, die dann den Kindern abhanden kommt. Die negativen Konsequenzen, die sich aus der Sicht einer feindselig-sexistisch eingestellten Person aus der Erwerbstätigkeit der Mütter ergeben, betreffen die emotionalen Bedürfnisse der Familie und insbesondere der Kinder.

Auch das Item *Bedürfnisse vernachlässigen* wurde nicht sofort von allen Befragten in der intendierten Weise aufgefasst. So deuteten die Interviewsequenzen zweier Personen darauf hin, dass der Begriff „vergessen“ anders ausgelegt wurde, als dies beabsichtigt war, nämlich im Sinne von: „in Vergessenheit geraten“, „nicht mehr daran denken“, sich die Konsequenzen des eigenen Handelns „nicht bewusst überlegen“. Ein Befragter merkte beispielsweise an, dass Frauen die Bedürfnisse nicht vergessen, sondern, dass sich dies einfach aufgrund ihres beruflichen Erfolges ergebe, also mehr oder weniger automatisch passiere. Diese Interpretation war jedoch nicht gemeint. Vielmehr ging es um ein „Vernachlässigen“ der Bedürfnisse des Ehepartners und vor allem der Bedürfnisse der Kinder, einem „Nicht-Gerecht-Werden“ der Mutterrolle, einem Agieren auf Kosten der Familie. Um dies zu verdeutlichen, wurde im Zuge der Itemüberarbeitung das Wort „vergessen“ durch das Wort „vernachlässigen“ ersetzt, welches mehrere InterviewteilnehmerInnen in ihren Ausführungen im Zusammenhang mit diesem Items verwendeten und welches eindeutiger zu sein schien.

Die Ergebnisse hinsichtlich des comprehensive probings zur Formulierung „Bedürfnisse ihrer Familie“ waren nicht ganz eindeutig. Überwiegend wurden sie im intendierten Sinne aufgefasst. So implizierte der Begriff „Familie“ für die meisten Befragten die Anwesenheit von Kindern und sie bezogen die Bedürfnisse vor allem auf den Nachwuchs.<sup>42</sup> Als Beispiele hierfür wurden u.a. folgende Aspekte genannt: Zeit mit den Kindern zu verbringen, mit ihnen zu spielen, sich mit ihnen zu beschäftigen, ihnen bei Hausaufgaben zu helfen oder bei anderen Problemen für sie da zu sein. Allerdings schlossen einige Personen in ihren Gedankengängen auch die Hausarbeit oder das Zubereiten der Mahlzeiten ein, was nicht der Intention des Items entsprach. Eine Interviewteilnehmerin gab zudem an, sie wisse nicht, ob in dieser Aussage Frauen gemeint seien, die bereits Kinder hätten oder solche, die kinderlos seien. Folglich sei es auch schwer für sie, sich für eine Antwortkategorie zu entscheiden. Im Zuge der Itemformulierung wurde angenommen, dass das Wort „Familie“ automatisch mit Kindern in Verbindung gebracht würde, was offenbar jedoch nicht der Fall war. Folglich wurde im Zuge der Itemüberarbeitung das Wort „Familie“ durch das Wort „Kinder“ ersetzt. Durch diese Änderung fielen zwar die Bedürfnisse des Partners weg, dies schien jedoch im Sinne der Eindeutigkeit des Items und der Schwerpunktlegung der Skala auf berufstätige Mütter, die bessere Wahl zu sein.

Streng genommen enthält auch das Item *Bedürfnisse vernachlässigen* einen doppelten Stimulus. Allerdings schien dies den Befragten (im Gegensatz zu dem vorhergehenden Item *Zu wenig Aufmerksamkeit*) keine Probleme zu bereiten. Dies mag damit zusammenhängen, dass in diesem Fall nur eine bestimmte Gruppe von Frauen angesprochen wurde (nämlich die der erfolgreich berufstätigen) und nicht Frauen im Allgemeinen, was den ersten Teil der Aussage („Viele erfolgreiche berufstätige Frauen denken nur an berufliches Vorankommen“) vermutlich weniger absolut erscheinen lässt und folglich auch seltener in Frage gestellt wird. Eine Änderung wurde aber dennoch an diesem Satzelement vorgenommen: Aufgrund der bereits beschriebenen Probleme bezüglich der Formulierung „erfolgreiche berufstätige Frauen“, wurde abermals auf das Wort „Karrierefrauen“ zurückgegriffen. Die Verwendung dieses Begriffs bringt noch einen weiteren Vorteil bezüglich des eben beschriebenen Problems mit sich: Die Befragten zweifeln wahrscheinlich noch weniger an, dass eine Karrierefrau daran interessiert ist, die Karriereleiter weiter hochzusteigen. Es ist daher anzunehmen, dass der Zusatz („denken nur an berufliches Vorankommen“) nicht als Unterstellung aufgefasst wird und daher auch keine Probleme bereitet, wie dies hingegen bei der Formulierung „heutzutage sind viele Frauen derart auf ihre Karriere fixiert“ in dem Ursprungstext *Zu wenig Aufmerksamkeit* der Fall war.

---

<sup>42</sup> Dies dürfte vermutlich auch auf einen Kontexteffekt mit den vorgehenden Fragen zurückzuführen sein, in denen überwiegend Kinder angesprochen wurden.

Endgültige Itemversion *Bedürfnisse vernachlässigen*

„Viele Karrierefrauen denken meiner Meinung nach nur an ihr berufliches Vorankommen und vernachlässigen die Bedürfnisse ihrer Kinder.“

### **Ursprungsitem *Selbstverwirklichung***

„Ich finde Frauen egoistisch, die bereits wenige Wochen nach der Geburt ihres Kindes wieder in ihren Beruf zurückkehren, weil sie ihre Karriere weiter verfolgen möchten.“

Das Item *Selbstverwirklichung* geht im Gegensatz zu den beiden vorhergehenden Items weniger auf die negativen Konsequenzen ein, die sich aufgrund der mütterlichen Erwerbstätigkeit für die Kinder ergeben, sondern es rückt die Motive der Frauen, wieder in ihren Beruf zurückzukehren, stärker in den Vordergrund. In der Ursprungsvariante des Items wurden die Beweggründe der Frau noch relativ offen gelassen und lediglich gesagt, sie wolle ihre Karriere weiter verfolgen. Durch das Wort „möchte“ sollte deutlich gemacht werden, dass keine finanzielle Notwendigkeit besteht, sondern dies dem Wunsch der Frau entspricht. Hinsichtlich des Zeitpunktes für die Rückkehr ins Erwerbsleben wurde angegeben, dass sie bereits wenige Wochen nach der Geburt erfolgt, was deutlich früher ist, als dies in der Regel in Deutschland praktiziert wird.<sup>43</sup> Aus Sicht einer feindselig-sexistisch eingestellten Person erfüllt die Mutter somit ihre Rolle nicht in angemessener Weise. Sie stellt ihre eigenen Bedürfnisse nicht hinter denen des Kindes zurück und wird folglich als egoistisch angesehen (siehe mütterliche Hingabe in Unterkapitel 2.1.1.2).

Als Ergebnis der kognitiven Interviews bleibt festzuhalten, dass das Item *Selbstverwirklichung* in seiner Ursprungsvariante noch zu viele Aspekte offen ließ. So erwies sich die Formulierung „wenige Wochen nach der Geburt“ als zu unpräzise. Eine Befragte sprach dieses Problem spontan an und erkundigte sich, was diese Formulierung bedeute.

Noch in einem weiteren Punkt war die erste Variante des Items zu vage, nämlich hinsichtlich der Erwerbstätigkeit der Mutter. So wurde lediglich gesagt, dass sie in ihren Beruf zurückkehre, allerdings nicht in welchem Umfang. Folgende Interviewsequenz verdeutlicht die Problematik:

#### **Interview Nr. 19 (männlich): „stimme eher nicht zu“**

**Int.:** Warum hast Du Dich für den Skalenwert „stimme eher nicht zu“ entschieden?

**Id. 19:** Hm, also die Frage beantworte ich genauso wie die vorhergehende Frage quasi, ja, von meinem normativen Verständnis, von meinem gesellschaftspolitischen Bild her. Und die Frage ist natürlich noch einmal ein bisschen verschärft, da es um „wenige

<sup>43</sup> Während im Jahr 2000 in Westdeutschland nur 25% und in Ostdeutschland nur 15% der Frauen *direkt* nach dem Mutterschutz in ihren Beruf zurückkehrten, unterbrachen 57% der westdeutschen und 70% der ostdeutschen Frauen ihre Erwerbstätigkeit für längere Zeit und nahmen Elternzeit in Anspruch (vgl. Dresel, Cornelißen & Wolf, 2005, S. 331 f.).



Wochen nach der Geburt“ geht. Ähm, sicher glaube ich, dass das Kind in den ersten Wochen und Monaten nach seiner Geburt die Nähe der Mutter braucht und auch eine Versorgung durch die Mutter. Im Einzelfall muss man das dann konkret sehen wie, ja, auch was das bedeutet, ja. [Pause] Wenn die Frau wieder in den Beruf einsteigt, ja. Ich meine, dass ist ja auch ein Unterschied, ob das ein 20-Stunden-Job ist oder ein 80-Wochenstunden-Job. (...) „Eher weniger“ habe ich gesagt, oder „stimme eher nicht zu“, ähm, weil auch hier im Unklaren bleibt, wie die konkreten Umstände sind, ne.

Des Weiteren legten die Äußerungen mehrerer befragter Personen nahe, dass die Begründung für die Rückkehr der Mutter in die Erwerbstätigkeit, nämlich der Wunsch die eigene Karriere weiter zu verfolgen, ebenfalls zu ungenau ist. Eine Interviewteilnehmerin merkte an, dass dahinter noch andere Gründe stehen müssten, wie z.B. ein höherer Verdienst der Frau im Vergleich zu dem des Mannes. Auch andere Befragte gingen davon aus, dass bestimmte Lebensumstände oder Sachzwänge (wie etwa ein drohender Arbeitsplatzverlust) zu dieser Entscheidung geführt haben müssten.

Folglich war es nötig, diese Unklarheiten zu beseitigen und das Item weiter zu spezifizieren. Hinsichtlich des Umfangs der Erwerbstätigkeit wurde eine Vollzeitstelle gewählt, da es bei der feindseligen Sexismus-Skala um Frauen geht, die aus der traditionellen Rolle ausbrechen, was sie hingegen bei einem Minijob oder einer Teilzeittätigkeit nicht wirklich tun. Um das Item jedoch nicht zu extrem zu formulieren, wurde die Formulierung „wenige Wochen nach der Geburt“ in ein „halbes Jahr nach der Geburt“ abgeändert, was gleichzeitig eine präzisere Angabe darstellt. Bezüglich des Beweggrundes der Frau, ihre Erwerbstätigkeit wieder aufzunehmen, wurden in der dritten Runde der kognitiven Interviews verschiedene Alternativen getestet. Letztendlich fiel die Wahl darauf, die Selbstverwirklichung der Frau als Begründung anzuführen. Dies hing v.a. damit zusammen, dass das übernächste Item *Verständnis* aufgrund seiner Komplexität um diesen inhaltlichen Aspekt gekürzt wurde. Um ihn trotzdem weiter beizubehalten, fand er daher in diesem Item Berücksichtigung.

#### Endgültige Itemversion *Selbstverwirklichung*

*„Ich finde Frauen egoistisch, die bereits ein halbes Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um sich selbst zu verwirklichen.“*

### Ursprungitem *Berufliche Anerkennung*

„Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.“

Auch bei dem Item *Berufliche Anerkennung* liegt der Fokus auf dem Wunsch der Mutter, einer Erwerbstätigkeit nachzugehen, obwohl das Kind noch nicht im Kindergartenalter ist. Erneut wird die mütterliche Hingabe angezweifelt und die Frau als selbstsüchtig erachtet.

Die kognitiven Interviews offenbarten ähnliche Schwächen wie im Zusammenhang mit dem vorhergehenden Item *Selbstverwirklichung*. Dabei erwies sich insbesondere der erste Teil des Satzes „Frauen, die gerne weiter erwerbstätig sein möchten“ als problematisch. So war er einerseits nicht exakt genug, was den Umfang der Erwerbstätigkeit betraf, andererseits ließ er offen, welcher Grund hinter dem Wunsch der Frau steht, einer Erwerbstätigkeit nachzugehen.

Ein Interviewteilnehmer wies auf ein weiteres Formulierungsdefizit im Zusammenhang mit diesem Satzelement hin. So machte er darauf aufmerksam, dass es sich hierbei lediglich um eine hypothetische Formulierung handle, da ein Konjunktiv verwendet werde. Schließlich werde nicht gesagt, dass der Wunsch auch in eine Handlung münde und die Frau tatsächlich ihren Beruf weiter verfolge:

**Interview Nr.15 (männlich):** von „stimme eher zu“ zu „stimme eher nicht zu“ zu „stimme überhaupt nicht zu“

**Id. 15:** [liest Frage noch einmal, betont „nur an sich“] ... NUR AN SICH! ... Da hätte ich eigentlich sagen müssen..., ähm..., mal überlegen, auf wen trifft das zu [liest; Pause]? Z.B. sagt das ja gar nicht, dass die Frau wirklich erwerbstätig ist. Frauen, die weiter erwerbstätig SIND, nicht sein MÖCHTEN. Das ist... Das ist schon ganz schön blöd, ne, die Frage... sondern, die weiter erwerbstätig sind... nicht mehr „gerne“... Das ist ja ein Konjunktiv, ne. [liest Frage] Ist Quatsch. Also, die, genau genommen muss das dahin [Anmerkung d. Verf.: von „stimme eher zu“ zu „stimme eher nicht zu“]. Also, diese Frage ist ein wenig unklar. (...)

**Int.:** Warum war das „nur“ jetzt noch mal ein Problem?

**Id. 15:** Also, hier ist nicht eindeutig gesagt, dass die Frau wirklich erwerbstätig ist. Das ist ja nur der Wunsch, erwerbstätig sein zu können, auch wenn sie Kinder unter drei Jahren haben. Dann denken sie ja nicht nur an sich. Ich denke, die meisten möchten gerne... Hier müsste ich also genau genommen sagen „stimme überhaupt nicht zu“, ne. Das wäre die richtige Antwort. Weil das ist ja 'ne Frage, da testet man wahrscheinlich nur, ob der Proband noch wach ist, ne. Aber das ist doch klar [lacht].

Folglich war es nötig, das Item hinsichtlich mehrerer Aspekte zu überarbeiten. Im Hinblick auf den Umfang der Erwerbstätigkeit wurde abermals eine Vollzeitstelle benannt. Des Weiteren wurde nun ein konkreter Beweggrund für die Entscheidung der Frau angeführt, weiterhin ihrem Beruf nachzugehen: der Wunsch nach Anerkennung.<sup>44</sup>

Durch die Verwendung eines Konjunktivs, handelte es sich bei dem Item um eine hypothetische Aussage. Dementsprechend könnte eine sexistisch eingestellte Person die Überlegung anstellen, dass eine solche Handlungsweise generell zu verurteilen wäre, solange eine Frau jedoch nur den Wunsch hege, ihn aber nicht in die Tat umsetze, sie auch nicht egoistisch sei. Folglich würde sie die Aussage verneinen und ihre „wahre“ Einstellung käme nicht ans Licht. Aus diesem Grund wurde dieses Satzelement komplett entfernt. Stattdessen wurde die Formulierung des Items *Selbstverwirklichung* benutzt („wenn sie wieder Vollzeit in ihren Beruf zurückkehren“), da diese sich als eindeutig erwiesen hatte. So wird in diesem Fall nicht von einem Wunsch, sondern von einer konkreten Handlung (nämlich der Rückkehr ins Erwerbsleben) gesprochen.

Hinsichtlich des Alters des Kindes war die Angabe „Kinder unter drei Jahren“ nicht zu ungenau, aber für mein Empfinden im Nachhinein doch zu wenig extrem. Deshalb entschloss ich mich auch, für dieses Item die Altergrenze auf ein Jahr festzulegen. In den vorherigen Ausführungen war bereits erwähnt worden, dass einige Befragte andernfalls dazu neigen, diese für sie kritische Altersphase auszuklammern und eher an ältere Kinder denken. Die überarbeitete Version des Items lautete daher:

#### Endgültige Itemversion *Berufliche Anerkennung*

*„Frauen denken meiner Meinung nach nur an sich selbst, wenn sie ein Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um berufliche Anerkennung zu bekommen.“*

---

<sup>44</sup> Einer Studie des BMFSFJ (2008b) zufolge gaben Frauen neben finanziellen Aspekten u.a. folgende Gründe für eine Rückkehr ins Erwerbsleben an: Die außerhäusliche Erwerbstätigkeit verleihe ihnen Unabhängigkeit, sei wichtig für ihr Selbstwertgefühl und eine Tätigkeit als Hausfrau und Mutter fülle sie alleine nicht aus (vgl. BMFSFJ, 2008b, S. 12; Sachverständigenkommission, 2011, S. 169). Da Haus- und Familienarbeiten zudem nur eine geringe (gesellschaftliche) Anerkennung erfahren, scheint es somit nachvollziehbar, dass sich die Frauen die fehlende Bestätigung in einem anderen Bereich, nämlich in diesem Fall dem Erwerbsleben, holen möchte.

### Ursprungitem *Verständnis*

„Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.“

In dem Item *Verständnis* wird abermals eine Mutter beschrieben, die gegen den Muttermythos verstößt, da sie aus scheinbar egoistischen Motiven (nämlich der Selbstverwirklichung) heraus einer Berufstätigkeit nachgeht, auch wenn die Kinder noch klein sind und keine finanzielle Notwendigkeit hierfür besteht. Dieses Item ist somit spezifischer als die beiden vorhergehenden Items in ihren Ursprungsvarianten waren. Dort wurde zuerst nicht explizit gesagt, weshalb die Mutter den Wunsch hegt, eine Erwerbstätigkeit aufzunehmen. (Dieser Aspekt wurde erst im Zuge der Überarbeitungen ergänzt). Zudem wurde auch keine Aussage hinsichtlich der finanziellen Situation der Familie getroffen. In dem Item *Verständnis* werden beide Punkte hingegen bereits in der ursprünglichen Itemvariante angesprochen, allerdings ist es folglich aufgrund dieser detaillierten Angaben relativ komplex (siehe Widersprüchlichkeit einiger Leitlinien zur Frageformulierung in Unterkapitel 2.5.2.4).

Obwohl das Item *Verständnis* somit teilweise sehr genaue Angaben enthielt, zeigte sich im Lauf der kognitiven Interviews, dass es hinsichtlich des Umfangs der Erwerbstätigkeit sowie bezüglich des Alters des Kindes noch zu vage formuliert war. Des Weiteren wurde deutlich, dass einige Befragte mit der Verneinung zu Beginn der Aussage („Ich habe *kein* Verständnis“) leichte Schwierigkeiten hatten. So mussten sie kurz überlegen, welche Antwortrichtung (eine Verneinung oder eine Zustimmung) ihre persönliche Meinung treffend wiedergibt. Folgender Ausschnitt aus Runde 2 verdeutlicht dies:

**Id. 01 (weiblich): „stimme nicht zu“**

**Id. 01:** (spontane Äußerung vor Nachfrage) (...) „habe kein Verständnis dafür“ hm. Dann „stimme ich nicht zu“.

**Int.:** Was haben Sie jetzt überlegt eben?

**Id. 01:** Ich hab gedreht quasi, die Frage gedreht. Ist das ein stimme „stimme eher zu“, ja. Ähm, hab gedacht, ich hab ja ähm, ich kann das sehr gut nachvollziehen, also stimme ich der Aussage so wie sie da steht nicht zu. Also das war jetzt ein bisschen schwierig für mich.

**Int.:** Wegen dieser Verneinung, wegen dem „kein Verständnis“?

**Id. 01:** Ja, genau.

Ein weiteres Frage-Problem, das sich im Laufe der kognitiven Interviews herauskristallisierte, war die Komplexität des Items. So enthielt die Formulierung so viele bedeutungsvolle Details, dass beispielsweise eine Befragte in Runde zwei den Umfang der Erwerbstätigkeit (der nach Runde 1 präzisiert worden war) überlas. Als sie dies bemerkte, wechselte sie die Wahl ihrer Antwortkategorie von „stimme eher nicht“ zu in „stimme

eher zu“, da sie für eine Vollzeitwerbstätigkeit einer Mutter unter diesen Umständen kein Verständnis habe:

**Id. 07 (weibliche): Wechsel von „stimme eher nicht zu“ zu „stimme zu“**

**Id. 07:** „Stimme ich eher nicht zu“, weil ich denke, die Mutter möchte einfach nicht komplett auf ihren Partner angewiesen sein. Sie möchte sich nicht nur um die Kinder kümmern. Wenn sie vielleicht einen 400 € Job macht oder so was, ja.

**Int.:** Also d.h., Sie hätten Verständnis dafür, oder... hätten Sie kein Verständnis, wenn sie das macht?

**Id. 07:** Ja, he, ...

**Int.:** Was überlegen Sie denn jetzt?

**Id. 07:** Ja, wenn es nur darum geht, sich selbst zu verwirklichen, dann würde ich sagen, dafür habe ich kein Verständnis. Ja.

**Int.:** Ja, aber Sie haben ja eben gesagt, dass sie vielleicht nicht nur auf ihren Partner angewiesen sein möchte ...

**Id. 07:** Ja, da habe ich dieses „Vollzeit“ nicht richtig... hatte ich das irgendwie nicht so ganz, ja, so ein bisschen überlesen.

**Int.:** Also, das „Vollzeit“ hatten Sie ein bisschen überlesen?

**Id. 07:** Ja.

(...)

**Int.:** Wenn Sie die Aussage jetzt noch mal lesen, so wie sie hier steht, würden Sie dann genauso antworten oder würden Sie anders antworten?

**Id. 07:** „Stimme ich zu“, ja.

**Int.:** Warum würden Sie dann jetzt sagen „stimme zu“?

**Id. 07:** Ja, wenn sie Vollzeit erwerbstätig sein möchte, nur um sich selbst zu verwirklichen, ist das doof, finde ich jetzt. (...)

**Int.:** Heißt das, Sie haben das jetzt von dem Vollzeit erwerbstätig abhängig gemacht?

**Id. 07:** Ja.

Im Rahmen des simulierten Telefoninterviews in Runde drei setzte sich das Problem bezüglich der Komplexität fort. So musste die Aussage zweimal vorgelesen werden, da der Befragte die Formulierung nicht mehr wusste.

**Id. 02 (männlich): „stimme eher nicht zu“**

**Id. 02:** (spontane Äußerung vor Nachfrage) Die Frage muss ich jetzt noch mal... Wenn die Frau nach einem Jahr wieder arbeiten geht,...

**Int.:** Ich lese es noch mal vor. [Item wird wiederholt].

**Id. 02:** Ich hab leider den Anfang schon wieder vergessen.

**Int.:** „Ich habe kein Verständnis dafür...“

**Id. 02:** So genau. „Stimme ich eher nicht zu“, weil ich hab da doch Verständnis für. So, genau. Soll ich wieder weit ausholen? (...)

Um die geschilderten Defizite zu beheben, wurden folgende Itemveränderungen vorgenommen. Es wurde eine Spezifizierungen hinsichtlich des Alters des Kindes (nach Runde 1: „eine Mutter kleiner Kinder (1 - 3 Jahre)“ bzw. nach Runde 2 „eine Mutter eines 1-jährigen Kindes“) vorgenommen. Zudem wurde angegeben, die Mutter gehe einer Vollzeitberufstätigkeit nach. Gleichzeitig musste die Formulierung jedoch auch vereinfacht und verkürzt werden, da sie (vermutlich insbesondere im Rahmen von telefonischen Befragungen) zu kompliziert gewesen wäre. So wurde zum einen die Negation am Satzanfang entfernt und das Item folglich in einer egalitären, nicht-sexistischen Weise formuliert („Ich habe Verständnis dafür“). Zum anderen wurde ein Element aus dem Item entfernt, um die Komplexität zu reduzieren. Da in den vorhergehenden Items nie explizit angegeben wurde, dass keine finanzielle Notwendigkeit für die Erwerbstätigkeit der Mutter besteht, wurde auf das Motiv der Mutter verzichtet, einer außerhäuslichen Erwerbstätigkeit nachzugehen. Folglich wurde das Satzelement „nur um sich selbst zu verwirklichen“ gestrichen. Stattdessen wurde gesagt, dass die Mutter erwerbstätig sein „möchte“, um trotzdem weiterhin zu verdeutlichen, dass es dem *Wunsch* der Mutter entspricht. Die endgültige Formulierung des Items nach Runde drei der kognitiven Interviews lautete folgendermaßen:

Endgültige Itemversion *Verständnis*

*„Ich habe Verständnis dafür, wenn die Mutter eines 1-jährigen Kindes Vollzeit erwerbstätig sein möchte, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.“*

Das Item wurde in dieser Form zwar auch in Pretest-Phase 3 mitgenommen (siehe Anhang 1), anschließend wurde es jedoch aus dem Itempool entfernt, da die Umformulierung des Items nicht als zufriedenstellend angesehen wurde. Der besondere Reiz dieses Items bestand ursprünglich darin, möglichst viele relevante Aspekte anzusprechen, die einen Einfluss auf die Wahl der Antwortkategorie nehmen könnten: (1) das Alter des Kindes, (2) das exakte Motiv für die Erwerbstätigkeit, (3) die Angabe, dass die Familie auch ohne das Einkommen der Frau finanziell abgesichert ist und (4) nach Runde 1 der kognitiven Interviews auch der Umfang der Erwerbstätigkeit der Mutter. In der endgültigen Variante des Items war auf das Motiv für die Erwerbstätigkeit auf Grund der Komplexität des Items verzichtet worden. Stattdessen war eine Formulierung gewählt worden, die vermutlich neue Schwierigkeiten mit sich brachte. Die Aussage, dass die Mutter gerne „erwerbstätig sein möchte“ könnte nämlich wieder als hypothetisch angesehen werden, was mitunter wieder zu Fehleinstufungen führen könnte, wenn einige sexistisch eingestellte Personen der Aussage zustimmen, da in dem Item nur ein Wunsch geäußert werde, dem

nicht unbedingt auch Taten folgen müssten (siehe Ausführungen zum Item *Berufliche Anerkennung*). Zusammenfassend kann im Zusammenhang mit dem Item *Verständnis* festgehalten werden, dass die Formulierung auch nach mehrmaliger Überarbeitung nicht vollends zufriedenstellend war. Die zentralen Frage-Defizite konnten zwar mit Hilfe der kognitiven Interviews aufgedeckt, jedoch nicht in einer zufriedenstellenden Weise behoben werden.

#### 3.4.5.2 Wohlwollender Sexismus

Um die Ergebnisse der kognitiven Interviews hinsichtlich des wohlwollenden Sexismus (WS) vorzustellen, wurden die folgenden sechs Fragen ausgewählt, die auch in die endgültige WS-Subskala aufgenommen wurden. Sie stammen alle aus der dritten Inhaltskategorie (c) Frauen sind für die Rolle als Mutter und Hausfrau prädestiniert.<sup>45</sup> (Siehe Tabelle 8 für die ursprünglichen Itemformulierungen, die in Runde 1 der kognitiven Interviews getestet wurden.)

*Tabelle 8: Itemformulierungen der WS-Items, die in Runde 1 der kognitiven Interviews getestet wurden*

Ursprüngliche Itemformulierung	Itemkürzel
Eine Mutter spürt viel besser als ein Vater, was die eigenen Kinder brauchen.	<i>Spüren, was Kinder brauchen</i>
Frauen besitzen einfach eine viel größere Fürsorglichkeit als Männer.	<i>Größere Fürsorglichkeit</i>
Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater.	<i>Auf Bedürfnisse eingehen</i>
Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.	<i>Sanft und Warmherzig</i>
Frauen besitzen ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.	<i>Hausarbeit und Kinderbetreuung</i>
Frauen haben einfach einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.	<i>Blick Hausarbeiten</i>

<sup>45</sup> Wie bereits im Zusammenhang mit dem feindseligen Sexismus erläutert wurde, ist es v.a. auf die Ergebnisse der Faktorenanalysen in Pretest-Phase 3 (siehe Unterkapitel 3.5) zurückzuführen, dass ein Teil der ursprünglich formulierten Inhaltskategorien nicht mehr in der endgültigen Skala vertreten ist. Dies gilt auch für die Kategorien (a) „Frauen bzw. Mütter sollten von Männern beschützt und umsorgt werden“ und (b) „Frauen bzw. Mütter sollten nur wenig (Teilzeit) oder überhaupt nicht erwerbstätig sein, damit die Doppelbelastung aus Beruf und Familie reduziert wird“ der WS-Subskala.

Die ersten vier Items *Spüren, was Kinder brauchen*, *Größere Fürsorglichkeit*, *Auf Bedürfnisse eingehen* und *Sanft und Warmherzig* beziehen sich auf die natürlichen Mutterinstinkte einer Frau und ihr mütterliches Wesen, welches sie für die Betreuung und Versorgung eines Kindes auszeichnet, sie unersetzlich werden lässt. Die beiden letzten Items *Hausarbeit und Kinderbetreuung* sowie *Blick Hausarbeiten* fokussieren auf allgemeine Haus- und Familienarbeiten, die neben der Kinderbetreuung anfallen und sprechen die Multitasking-Fähigkeiten von Frauen an.

### **Ursprungsitem *Spüren, was Kinder brauchen***

„Eine Mutter spürt viel besser als ein Vater, was die eigenen Kinder brauchen.“

Wie in Unterkapitel 3.2.2 erläutert, werden Frauen im Rahmen der WS-Subskala solche positiven Eigenschaften zugeschrieben, die sie für die traditionell weibliche Rolle auszeichnen. Das Item *Spüren, was Kinder brauchen* bezieht sich auf die vermeintlich natürlichen Mutterinstinkte einer Frau, ihre weibliche Intuition und ihr Einfühlungsvermögen. Es hebt hervor, dass eine Mutter – aufgrund ihres Geschlechts – im Gegensatz zu einem Vater einfach merkt, wie die emotionale Befindlichkeit des eigenen Kindes ist, worin seine Bedürfnisse bestehen und wie man am besten auf diese eingeht. Auf diese Weise wird die besondere Eignung der Mutter für die Kinderbetreuung betont und ihre Festlegung auf den familiären Bereich plausibel gemacht.

Die Antworten der Befragten im Rahmen der kognitiven Interviews deuteten darauf hin, dass das Ursprungsitem *Spüren, was Kinder brauchen* in der beabsichtigten Weise verstanden wurde und keine größeren Verständnisschwierigkeiten bestanden. So gingen die Interpretationen der Befragten im Zusammenhang mit dem Wort „spüren“ in die intendierte Richtung. Auf die entsprechende Nachfrage hin führten die InterviewteilnehmerInnen an, sie hätten den Begriff als Intuition aufgefasst, ihn als Einfühlungsvermögen, als Muttergefühl, als „besonderen Draht“ zum Kind, als Gefühl für die Bedürfnisse des Kindes verstanden.

Beispiel 1:

**Interview Nr. 06 (männlich): „stimme eher zu“**

**Int.:** Als Sie das Wort „spüren“ gelesen haben, an was haben Sie dabei gedacht? Können Sie ein paar Beispiele geben?

**Id. 06:** Ähm, ja gut, da hab ich einfach so an Intuition gedacht, das irgendwo, was man halt so als Muttergefühle umschreibt, dass man... dass die Mutter es halt eher merkt, spürt, ja. War ja... Ja, Intuition einfach. Auf 'ne emotionale Art und Weise wahrnehmen, fühlen ob's den Kindern gut oder schlecht geht oder dass..., dass vielleicht ... ja, fühlen, wenn was nicht in Ordnung ist.



Beispiel 2:

**Interview Nr. 12 (weiblich): „stimme nicht zu“**

**Int.:** Als Du das Wort „spüren“ gelesen hast, an was hast Du dabei gedacht? Kannst Du da ein paar Beispiele geben?

**Id. 12:** Mhm, also, so ein Gefühl dafür zu haben, wie es dem Kind geht. Vielleicht auch etwas zu sehen, ohne..., oder zu merken, ohne, dass das Kind groß was sagen muss. Ich glaube, das sind Eigenschaften, die sowohl Frauen als auch Männer besitzen.

Auch die allgemeinen Answerläuterungen der Befragten lassen darauf schließen, dass das Item in der beabsichtigten Weise interpretiert wurde. So gaben mehrere Personen an, dass sie der Aussage nicht zustimmen, da sie die Fähigkeit, die Bedürfnisse der Kinder zu spüren und darauf einzugehen, nicht am Geschlecht festmachten. Ausschlaggebender Faktor sei die Beziehung zum Kind, die gemeinsame Bindung. Väter könnten per se genauso gut ein Gespür für die Bedürfnisse des Kindes haben wie Mütter, bzw. es gäbe auch Mütter, die nicht darüber verfügten. In seiner Allgemeingültigkeit wollten sie dem Item daher nicht zustimmen:

Beispiel 1:

**Interview Nr. 12 (Id. 12) (weiblich): „stimme nicht zu“**

**Id. 12:** (spontane Äußerung vor Nachfrage) „Stimme ich nicht zu.“ Ich glaube, da haben heutzutage Frauen und Männer ein gleiches Gespür, wenn sie eins haben. Und es gibt da sicherlich genauso Frauen wie Männer, die da wenig oder kein Gespür haben. Und da kann man nicht pauschal sagen, 'ne Mutter kann das besser als ein Vater. Das hängt ganz eindeutig von der Person ab.

Beispiel 2:

**Interview Nr. 21 (männlich): „stimme nicht zu“**

**Int.:** Als Du eben die Frage beantwortet hast, an was hast Du da gedacht? Was ist Dir durch den Kopf gegangen, bis Du die Antwort gegeben hast?

**Id. 21:** Ich habe versucht mir vorzustellen, welche grundlegenden Bedürfnisse, ähm, Kinder haben in Bezug auf verschiedene Lebensaspekte und bin zu dem Schluss gekommen, dass, äh, Eltern, ob Mutter oder Vater, dies gleich gut erspüren können. Sofern sie mit, äh, ja, sofern sie eine ähnliche, ähnlich starke Bindung und Beziehung zu ihren Kindern bereits aufgebaut haben. Also es hängt nicht am, ähm, daran, ob man Frau oder Mann, Mutter oder Vater ist. Sondern es hängt, die Frage, ob man erspüren kann, was die eigenen Kinder brauchen, macht sich daran fest, wie gut die Beziehung zu den Kindern aufgebaut ist.

Daneben gab es eine Reihe von Befragten, die dem Item zustimmten, da sie annahmen, Mütter hätten in der Regel einen „besseren Draht“, eine engere Verbindung zum Kind und könnten daher auch „besser spüren, was ein Kind braucht“.

Beispiel 1:

**Interview Nr. 4 (männlich): „stimme zu“**

**Int.:** Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?

**Id. 04:** Ja, dass die Mutter eigentlich zu ihren Kindern, sag ich mal, einen besseren Draht noch hat, als der Vater. Oder Frauen... oder eine Mutter das früher erkennt oder spürt, dass irgendwas mit den Kinder nicht in Ordnung ist, dass ihnen irgendwas fehlt oder, dass sie irgendwas bedrückt oder... Wie auch immer, so diese Richtung.

Beispiel 2:

**Interview Nr.18 (männlich): „stimme eher zu“**

**Id. 18:** (spontane Äußerung vor Nachfrage) Generell ist das schon so, weil, wie gesagt, die Mütter oftmals eine viel stärkere Bindung haben. Da geht's nicht um's wo. Ich denke, das ist überall so. Ob man jetzt in irgendeinen Wildstamm geht oder nach Ostdeutschland oder in die USA, ist ja das... Ich mein, die Mütter haben die Kinder auch als erstes im Arm und fangen dann an zu heulen. Denen bedeutet das richtig was. Und, ähm, ein Vater kann das sicherlich auch spüren, wenn er sich mit dem Kind auseinandersetzt, aber oftmals sind es die Frauen, die eine engere Bindung haben und eher darauf achten.

Folglich wurde die ursprüngliche Formulierung des Items beibehalten.

Endgültige Itemversion *Spüren, was Kinder brauchen*

„Eine Mutter spürt besser als ein Vater, was die eigenen Kinder brauchen.“

**Ursprungsitem *Größere Fürsorglichkeit***

„Frauen besitzen einfach eine viel größere Fürsorglichkeit als Männer.“

In diesem Item wird Frauen mit der Fürsorglichkeit eine feminine Eigenschaft und somit ein klassisches weibliches Stereotyp zugeschrieben. Abermals wurde bewusst ein Vergleich zwischen den Geschlechtern eingebaut und betont, dass Männer nicht in derselben Weise über dieses Merkmal verfügen. Folglich schwingt in der Aussage mit, dass Frauen eher dafür geeignet sind, sich im familiären Bereich um den Partner und insbesondere die Kinder zu kümmern, sie zu versorgen, ihre Bedürfnisse zu erkennen und auf diese einzugehen.

Die Ausführungen der InterviewteilnehmerInnen verdeutlichen, dass das Item in seiner Ursprungsvariante noch zu allgemein formuliert war. Während der Begriff Fürsorglichkeit an sich keine Schwierigkeiten bereitete und von den Befragten im intendierten Sinne definiert wurde, war es für sie jedoch nicht ganz eindeutig auf wen sich die Fürsorge bezieht. Dementsprechend unterschiedlich waren auch die Assoziationen der InterviewteilnehmerInnen, als sie die Aussage beantworteten. Während ein kleiner Teil der befragten Personen (vermutlich aufgrund der thematischen Ausrichtung der vorhergehenden Fragen) ausschließlich an Kinder dachte, nahm eine Befragte sogar an, dass bei diesem Item Kinder ausdrücklich nicht gemeint sein könnten, da sie nicht erwähnt würden (siehe hierzu Maxime der Konversationslogik in Unterkapitel 2.5.1.1.).

**Interview Nr.16 (weiblich): „stimme eher zu“**

**Int.:** Warum hast Du Dich für den Skalenwert „stimme eher zu“ entschieden?

**Id. 16:** Weil Männer... Vielleicht ist das anerzogen, aber ich glaube, dass Männer mehr Fürsorge brauchen, als dass sie die geben können. Bei kleinen Kindern und bei eigenen Kindern vielleicht ausgenommen (...) ... ja gut, aber das ist ja jetzt nicht die Frage gewesen...

**Int.:** Also geht die Frage für dich nicht um Kinder.

**Id. 16:** Nee, um generelle Fürsorglichkeit. Hier steht nichts von Kindern.

Andere TeilnehmerInnen bezogen die Aussage wiederum entgegen der Intention mitunter auf ältere Familienmitglieder, Bekannte oder den Freundeskreis und somit eher ganz allgemein auf Mitmenschen aus dem persönlichen Umfeld.

Um das Item zu spezifizieren und den ursprünglichen Kerngedanken hervorzuheben, wurde im Zuge der Itemüberarbeitung der Fokus ausschließlich auf Mütter und Väter gelegt. Auf diese Weise sollte deutlich gemacht werden, dass die Fürsorglichkeit in Bezug auf die eigenen Kinder gemeint ist. Zugleich wurde durch diese Umformulierung auch der familiäre Bereich als der relevante hervorgehoben und somit die besondere Eignung der Frauen für ihre traditionelle Rolle als Mutter betont. Wäre hingegen die mehrdeutige Formulierung des Ursprungsitems beibehalten worden, hätte die besondere Fürsorglichkeit der Frauen durchaus auch im Kontext einer außerhäuslichen Erwerbstätigkeit (z.B. im Pflegebereich) als vorteilhaft angesehen werden können. Dies würde dann allerdings keine Festlegung der Frau auf die private Sphäre und somit auch keine Zementierung der geschlechtsspezifischen Arbeitsteilung bedeuten (siehe Unterkapitel 3.2.2). Die überarbeitete Formulierung des Items lautete daher folgendermaßen:

Endgültige Itemversion *Größere Fürsorglichkeit*

„Mütter besitzen eine größere Fürsorglichkeit als Väter.“

### **Ursprungitem *Auf Bedürfnisse eingehen***

*„Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater.“*

Das Item *Auf Bedürfnisse eingehen* spricht die besondere Befähigung der Mutter an, sich in der ersten Zeit (d.h. direkt nach der Geburt des Kindes, aber auch noch in den ersten Lebensjahren) um den Nachwuchs zu kümmern. Dabei greift es das deskriptive Stereotyp auf, Frauen (bzw. Mütter) seien besonders einfühlsam, empathisch, auf das Wohlergehen anderer Personen bedacht und könnten daher besonders gut auf die Bedürfnisse anderer Personen (bzw. des Kindes) eingehen. Auch die vermeintlich natürlichen Mutterinstinkte einer Frau sowie ihre besondere Verbindung zum Kind schloss die Intention dieses Items ein. Gleichzeitig wird ein präskriptives Stereotyp angedeutet: Der erste Teil des Satzes, in dem geäußert wird, insbesondere kleine Kinder würden ihre Mutter brauchen, impliziert die Verhaltenserwartung an die Frauen, sich in der ersten Zeit selbst um die Betreuung ihrer Kinder zu kümmern und dies nicht dem Vater zu überlassen.

Wie bereits im Zusammenhang mit den Items der feindseligen Subskala offenbarten die kognitiven Interviews bei diesem Item ein Problem mit der Formulierung „kleine Kinder“. Sie erwies sich als zu unspezifisch. So gab eine Befragte an, sie hätte überlegen müssen, auf welches Alter sich das Item bezieht:

#### **Interview Nr. 09 (weiblich): „stimme eher zu“**

**Int.:** Was haben Sie jetzt eben überlegt?

**Id. 09:** Ähm, also ich hab erst mal überlegt, was „kleine Kinder“ heißt. Ähm, „ganz kleine Kinder“, also gerade frisch Geborene und bis zu einem Jahr, würde ich jetzt mal sagen, brauchen natürlich noch ihre Mutter, einfach aus biologischen Gründen. Ähm... Also, natürlich kann der Vater auch das Fläschchen geben und so, ähm, aber einfach, äh, weil die ja auch noch gestillt werden, zum Großteil. Und, ähm ja..., ist einfach eigene Erfahrung, weil ich einfach sehe, dass mein Jungspund noch mal ein anderes Verhältnis zu mir hat, als zum Vater. Obwohl mein Mann..., obwohl wir uns eigentlich die Betreuung 50:50 teilen, so... ähm, aber ja... Also, ich das einfach erfahren habe, dass ganz kleine Kinder die Mutter schon noch ein bisschen mehr brauchen, als den Vater.

Um die Unklarheit bezüglich des Alters zu beseitigen, wurde äquivalent zu den Überarbeitungen der FS-Subskala zunächst eine Altersangabe in Klammern gemacht ( $\frac{1}{2}$  - 2 Jahre), die allerdings nach Runde 2 (aufgrund möglicher Schwierigkeiten im Rahmen von Telefonbefragungen) durch die Formulierung „Kinder unter zwei Jahren“ ersetzt wurde. Während bei der Umformulierung der feindseligen Sexismus-Items das Alter für Kinder auf unter einem halben Jahr bzw. unter einem Jahr festgelegt wurde, ist hier eine Altersobergrenze von zwei Jahren angesetzt worden. Diese Entscheidung hing damit zusammen, dass mehrere Personen im Laufe der kognitiven Interviews darauf hinwiesen, kleine

Kinder würden ihre Mütter in der ersten Zeit vor allem aufgrund des Stillens brauchen. Folglich stimmten sie der Aussage auch zu. Um diesem biologischen Unterschied zwischen Frauen und Männern als Hauptargument das Gewicht zu nehmen und eine Debatte um Stillen vs. Milch abpumpen vs. Flaschenmilch geben zu vermeiden, sollten auch ältere Kinder in die Überlegungen miteinbezogen werden. Einerseits sollten diese zwar noch vor dem Kindergartenalter stehen (da dies eine der kritischen Altergrenzen zu sein scheint, in der der Anwesenheit der Mutter eine besondere Bedeutung beigemessen wird), andererseits sollten sie aber auch nicht mehr so klein sein, dass sie unbedingt noch gestillt werden müssen. Eine Altersobergrenze von zwei Jahren schien folglich angemessen.

Bezüglich der Fragestruktur offenbarten die kognitiven Interviews ebenfalls eine Schwachstelle des Items *Auf Bedürfnisse eingehen*. Wie bei der Beschreibung der Intention der Wohlwollenden Sexismus-Skala bereits angeführt wurde, besitzt ein Teil der Items durch die Verknüpfung von deskriptiven Geschlechterstereotypen und präskriptiven Rollenzuschreibungen einen doppelten Stimulus (siehe Unterkapitel 3.2.2). Bei diesem Item stellte sich nun heraus, dass zwei InterviewteilnehmerInnen tatsächlich Schwierigkeiten hatten, sich für eine Antwortkategorie zu entscheiden. So hätten beide Personen dem ersten Teil der Aussage („Insbesondere kleine Kinder brauchen ihre Mutter“) zugestimmt, dem zweiten („Eine Mutter kann auf die Bedürfnisse des Kindes viel besser eingehen als der Vater“) hingegen nicht. Während der erste Befragte dem einleitenden Part der Aussage ein größere Bedeutung beimaß und sich folglich für die Antwortkategorie „stimme eher zu“ entschied, wählte die zweite Person aufgrund des Nachsatzes ein „stimme nicht zu“.

Beispiel 1:

**Interview Nr. 06 (männlich): „stimme eher zu“**

**Id. 06:** (spontane Äußerung vor Nachfrage) Na ja, wenn man die Frage, ähm, halt splitten könnte... Ich würde auf den ersten Teil zustimmen, auf den zweiten nicht [lacht].

**Int.:** Was würden Sie dann antworten, wenn Sie die Aussage jetzt in einem Fragebogen beantworten müssten?

**Id. 06:** Ähm... [Pause] das ist jetzt schwer zu sagen. Es ist natürlich... Vielleicht wenn man den zweiten Teil der Frage, den hat man ja als letztes, man tendiert ja dann vielleicht eher dazu, den zu beantworten. Dem würde ich „eher nicht zustimmen“. Aber wenn man jetzt die Frage so durchliest und wirklich noch mal drüber überlegt und es getrennt sieht und dann überlegt, was ist die wichtigere Komponente, dann würde ich „eher zustimmen“, weil ich glaube, dass der erste Teil halt auch sehr wichtig ist. Dass halt... Das ist die wichtigere Aussage, dass „kleine Kinder ihre Mutter brauchen“. (...)

Beispiel 2:

**Interview Nr. 07 (weiblich): „stimme nicht zu“**

**Int.:** Warum würden Sie hier jetzt sagen „stimme nicht zu“?

**Id. 07:** Ähm, weil ich allgemein sagen kann, „kleine Kinder brauchen ihrer Mutter“. Das ist mir klar. Aber dass sie auf die Bedürfnisse des Kindes besser eingehen kann, das kann ich nicht, äh, sagen. Ähm, wenn jetzt nur stehen würde: „Insbesondere kleine Kinder brauchen ihre Mutter“, dann würde ich sagen „stimme voll und ganz zu“. Bin aber dann mit diesem Zusatz da hinten dran noch, ähm, kann ich ja nicht zustimmen. Geht nicht [lacht].

Offenbar schienen für diese InterviewteilnehmerInnen die beiden Satzteile nicht im Einklang miteinander zu stehen, so dass ihnen eine Antwortfindung schwer fiel. Um dieses Problem zu vermeiden und auch bei der Auswertung der Daten sicher sein zu können, worauf sich die gegebene Antwort der Befragten bezieht, wurde der zweite Part des Items (das positive deskriptive Stereotyp Mütter könnten auf die Bedürfnisse kleiner Kinder besser eingehen als Väter) herausgegriffen und auf den ersten Teil verzichtet (das implizite präskriptive Stereotyp, Mütter müssten sich um die Betreuung des Kindes kümmern, da diese sie brauchten). Diese Schwerpunktlegung wird auch schon allein aufgrund der Theorie des ambivalenten Sexismus nötig, da hier vermeintliche positive Zuschreibungen vorgenommen werden und ein positives Bild von Frauen gezeichnet wird. Dies ist nur beim zweiten Satzteil der Fall. Hinzu kommt, dass die erste Aussage („*Insbesondere kleine Kinder brauchen ihre Mutter.*“) allein genommen vermutlich eine sehr breite Zustimmung erfährt und nicht unbedingt sexistische Einstellungen erfassen muss. Wird die Betonung auf die Wörter **INSBESONDERE** und **KLEIN** gelegt, so könnte man schließlich auch so argumentieren, dass kleine Kinder selbstverständlich die Mutter in stärkerem Maße brauchen als ältere Kinder. Zudem müsse dies nicht auf die Mutter beschränkt sein, sondern könne auch für den Vater oder eine andere feste Bezugsperson gelten.

Die „Gefahr“ einer möglichen Fehleinstufung der Befragten besteht in diesem Fall auch nach der Kürzung nicht, da hier die besonderen Fähigkeiten der Frauen (nämlich auf andere Personen einzugehen) durch den Begriff „Mutter“ und die Spezifizierung der Bedürfnisse als die des eigenen Kindes bereits auf den familiären Bereich festgelegt sind und nicht etwa in einen Zusammenhang mit dem öffentlichen Bereich gebracht werden können (siehe hierzu die Ausführungen zur Inhaltskategorie (c) Frauen sind für die Rolle als Mutter und Hausfrau prädestiniert in Unterkapitel 3.2.2). Zudem ist ein Vergleich mit den Fähigkeiten des Vaters eingebaut, der meines Erachtens die Konsequenzen für die Arbeitsteilung offensichtlich werden lässt: Wenn also die Auffassung besteht, die Mutter könne besser auf die Bedürfnisse des Kindes eingehen als der Vater, dann sollte sie sich vermutlich, aus Sicht einer benevolent-sexistisch eingestellten Person, im Idealfall auch um die Betreuung des Kindes kümmern.

Endgültige Itemformulierung *Auf Bedürfnisse eingehen*

„Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.“

#### **Ursprungitem *Sanft und Warmherzig***

„Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.“

Wie bereits bei der vorhergehenden Frage, handelt es sich auch bei dem Item *Sanft und Warmherzig* um eine Kombination aus einem deskriptiven Stereotyp (Eine Mutter ist sanft und warmherzig) sowie einem präskriptiven Stereotyp (Die Mutter sollte sich um die Betreuung des Kindes kümmern, da sie besser dafür geeignet ist). Inhaltlich wird in dieser Aussage auf die beiden Kerndimensionen des Stereotype Content Models (SCM) (siehe Unterkapitel 2.2.3.4) zurückgegriffen. Einerseits wird die besondere Warmherzigkeit von Frauen angesprochen, gleichzeitig wird daraus ihre Kompetenz abgeleitet. Allerdings ist hier, im Gegensatz zum SCM, nicht die Kompetenz in Bezug auf die öffentliche Sphäre gemeint, sondern die besondere Befähigung der Mutter hinsichtlich eines familiären Bereiches – in diesem Fall der Kinderbetreuung. Somit wird eine aus Sicht des Befragten vermutlich positiv empfundene Zuschreibung vorgenommen, die jedoch letztendlich Frauen auf den häuslichen Bereich verweist und damit die bestehenden Geschlechterhierarchien festigt.

Die Antworten der Befragten, die im Rahmen des kognitiven Interviews gegeben wurden, legen die Vermutung nahe, dass das Item *Sanft und Warmherzig* in der beabsichtigten Weise verstanden wurde und auch der doppelte Stimulus keine Probleme bereitete. Im Gegensatz zu dem Item *Auf Bedürfnisse eingehen*, bei dem im Prinzip zwei gleichrangige Aussagen miteinander verbunden werden, stellt in diesem Fall der zweite Teil des Satzes eine Schlussfolgerung aus dem ersten Satzteil dar, der wiederum eine Vorbedingung ist. Stimmt man also dem ersten Satzteil nicht zu, so muss man folglich die gesamte Aussage verneinen. Ist man hingegen der Ansicht, dass Mütter tatsächlich sanfter und warmherziger sind als Väter, diese Eigenschaften sie jedoch nicht für die Betreuung eines Kindes prädestinieren, so muss man die Aussage ebenfalls verneinen, da man die Schlussfolgerung als falsch ansieht. Die letzte Antwortmöglichkeit besteht darin, dass beide Elemente des Items für richtig erachtet werden und dementsprechend der Aussage im Gesamten zugestimmt wird. Dies traf auch auf mehrere Befragte zu, sie stimmten der Aussage ausnahmslos zu.

**Interview Nr. 18 (männlich): „Stimme zu“.**

**Id. 18:** Ähm, „Eine Mutter ist...“. Ja doch, es geht ja nicht um eine Frau, sondern eine Mutter. Also, eine Mutter ist dem Kind gegenüber mit Sicherheit „sanfter und warmherziger“.

**Int.:** Als der Vater?

**Id. 18:** Ja, glaub ich schon. Wegen der engeren Bindung oftmals. Soll aber nicht bedeuten, dass Väter unsensibel sind. Es geht einfach nur um sanfter und warmherziger und auch wenn es nur ein bisschen ist, ist es meistens da. (...) Ja, es ist definitiv größer, aber nicht viel größer – also, das soll kein Vergleich sein, so richtig, sondern einfach nur, ja, was ich grad schon gesagt hab – und daher auch besser für die Betreuung des Kindes in einem jüngeren Alter geeignet. (...)

Andere InterviewteilnehmerInnen hielten hingegen nur den ersten Teil des Items für zutreffend, den zweiten hingegen nicht und entschieden sich aufgrund des konsekutiven Konjunkionaladverbs (daher) folgerichtig für eine verneinende Antwortkategorie. So räumte beispielsweise eine Befragte ein, es könne schon der Fall sein, dass eine Mutter das Kind sanfter behandle, dies bedeute jedoch nicht, dass sie auch besser für dessen Betreuung geeignet sei, als der Vater. Folglich stimmte sie dem Item nicht zu.

**Interview Nr. 08 (weiblich): „stimme nicht zu“**

**Int.:** Darf ich da gleich nachfragen? Warum haben Sie hier geantwortet „stimme nicht zu“? Bei der 39.

**Id. 08:** Also, dass, dass 'ne Mutter vielleicht „sanfter“ mit 'nem Kind umgeht als'n Vater, das kann durchaus sein, aber sie ist deswegen nicht „besser für die Betreuung des Kindes geeignet“. Also, die Mutter nimmt das Kind sicherlich, was weiß ich, sanfter vom Boden auf, der Vater macht dann auch bald mal 'nen Salto mit ihm und dann quiekt es vor Vergnügen, vor Vergnügen. Aber die Wertigkeit: Vater, Mutter, gut für die Kinderbetreuung, wird dadurch nicht betroffen. (...)

Auch der Fall, dass beide Elemente des Satzes als falsch erachtet wurden, kam vor. Die InterviewteilnehmerInnen stimmten dementsprechend der Aussage im Gesamten nicht zu.

Beispiel 1:

**Interview Nr. 09 (weiblich): „stimme überhaupt nicht zu“**

**Int.:** Und dann die letzte Frage zu der Aussage 39, was haben Sie da geantwortet?

**Id. 09:** Ich hab geantwortet, „stimme überhaupt nicht zu“.

**Int.:** Warum?

**Id. 09:** Weil „eine Mutter vom Wesen her“, also, weil das auch wieder eine... biologische Aussage ist. Also, äh... das, also, es hat so was, das Wesen der Frau ist, äh..., „sanft und warmherzig“... So. [Pause] Und auch Väter können genauso warmherzig und sanft sein. Ja. [Pause] Und ich find es einfach ganz schwierig zu... Also, weil das impliziert ja, eine Mutter kann ja eigentlich nur eine Frau sein, und das impliziert ja, dass Frauen einfach, ähm, dass es ihre natürliche Aufgabe ist, ein Kind zu betreuen.



**Int.:** Mhm. Und dem würden Sie nicht zustimmen?

**Id. 09:** Nee, dem würde ich überhaupt nicht zustimmen.

**Int.:** Haben Sie auch beide Teile der Aussage mitberücksichtigt oder jetzt nur den ersten Teil, als Sie das verneint haben? Also, haben Sie nur verneint, dass sie vom Wesen her sanfter sind oder auch den zweiten Teil, dass sie...?

**Id. 09:** Nee, ich hab da beide Teile gleichwertig verneint.

Beispiel 2:

**Interview Nr. 21 (männlich): „stimme nicht zu“**

**Int.:** Als Du eben die Frage beantwortet hast, an was hast Du da gedacht? Was ist Dir durch den Kopf gegangen, bis Du die Antwort gegeben hast?

**Id. 21:** Ich habe überlegt, erstens sind Frauen, sprich Mütter, sanfter und warmherziger zu Kindern und wäre da vielleicht noch zu dem Ergebnis gekommen: Ja, es kann sein, dass Mütter zärtlicher mit ihren Kindern umgehen. Aber die Folge, sie sind daher besser für die Betreuung ihres Kindes geeignet, sehe ich so nicht. Selbst wenn man unterstellt, dass Frauen und Mütter sanfter und warmherziger sind, befähigt sie das nicht besser zur Betreuung von Kindern oder ihres eigenen Kindes, weil es bei qualifizierter Kinderbetreuung nicht nur auf Sanftheit und Warmherzigkeit ankommt. Im Gegenteil ist es, ähm, vielleicht auch für die Entwicklung des Kindes besser, etwas weniger sanft und..., oder sagen wir mal, nicht nur sanft und warmherzig zu sein.

**Int.:** Warum hast Du Dich für den Skalenwert „stimme nicht zu“ entschieden?

**Id. 21:** Weil ich glaube, dass erstens Väter ähnlich sanft und warmherzig sein können und zweitens, wie bereits gesagt, ähm, aus Sanftheit und Warmherzigkeit sich keine Qualifikation für gute Betreuung des Kindes ergibt.

Der Befragte (Id. 06), der im Zusammenhang mit dem Item *Auf Bedürfnisse eingehen* aufgrund des zweifachen Stimulus Probleme bei der Antwortfindung hatte, besaß dieses Mal keine Schwierigkeiten, da er beide Elemente der Aussage verneinte. Er räumte zwar ein, dass er vermutlich bei seiner Entscheidungsfindung in einen Konflikt geraten wäre, wenn es sich bei dem Kind um ein Neugeborenes gehandelt hätte. Unter dieser Bedingung brauche das Kind, seiner Ansicht nach, die Mutter mehr als den Vater. Letzten Endes hätte er jedoch auch in diesem Fall das Item verneint, da er sich als Mann bzw. Vater persönlich diskreditiert gefühlt hätte.<sup>46</sup>

---

<sup>46</sup> Durch die positive Formulierung in Bezug auf die Mütter, die dem benevolenten Sexismus zugrunde liegt, werden indirekt auch Aussagen über die Väter getroffen. So wird das weibliche mit dem männlichen Geschlecht verglichen und dem weiblichen eine besondere Befähigung im häuslich-familiären Bereich attestiert und gleichzeitig dem männlichen abgesprochen. Es ist daher nicht verwunderlich, dass Männer und insbesondere Väter sich durch die Items der wohlwollenden Subskala herabgesetzt sehen. So hätte ich auch von allen Männern, die nicht benevolent-sexistisch eingestellt sind, eine derartige Reaktion erwartet.

Insgesamt betrachtet schien das Item im intendierten Sinne aufgefasst worden zu sein. Auch der doppelte Stimulus führte zu keinen Problemen bei der Antwortfindung. Auch hinsichtlich des Alters wollte ich keine Änderung vornehmen. Die Andeutung des Befragten Id. 06 in Runde 1, das in der ursprünglichen Version des Items das Alter des Kindes nicht spezifiziert sei, zog aus dreierlei Gründen keine Umformulierung der ursprünglichen Aussage nach sich: Erstens schien das Alter für dieses Item nicht so ausschlaggebend zu sein, als dass man es auch ohne eine genauere Angabe nicht hätte genauso gut/sinnvoll beantworten können. Dies wurde auch daran deutlich, dass im Gegensatz zu den Items der feindseligen Subskala in diesem Fall nicht mehrere Befragte angaben, ihnen fehle diese Information. Zweitens wäre die Aussage durch eine präzisere Altersangabe noch länger und komplexer geworden. Dies sollte jedoch vermieden werden (vgl. Porst, 2008, S. 99 f.). Drittens stellen nicht alle Probleme, die man während eines kognitiven Interviews aufspürt, auch tatsächlich ein Problem dar. So ist es möglich, dass eine vermeintliche Schwierigkeit, die identifiziert wird, in Wirklichkeit lediglich ein „falscher Alarm“ ist (vgl. Conrad & Blair, 2009, S. 52 f.) Ich gehe davon aus, dass dem in diesem Fall so war. Erst auf meine Nachfrage hin, welchen Teil des Satzes er verneint hätte, kam der Befragte auf die Idee, dass man die Antwort je nach Alter des Kindes auch hätte weiter ausdifferenzieren können. Folglich wurde die ursprüngliche Formulierung des Items *Sanft und Warmherzig* beibehalten.

Endgültige Itemversion *Sanft und Warmherzig*

„Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung eines Kindes geeignet.“

#### **Ursprungsitem *Hausarbeit und Kinderbetreuung***

„Frauen besitzen ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.“

Das Item *Hausarbeit und Kinderbetreuung* greift das deskriptive Stereotyp auf, Frauen besäßen Multitasking-Fähigkeiten und könnten folglich mehrere Dinge gleichzeitig „unter einen Hut bringen“. Wie bereits erwähnt, werden Frauen im Rahmen der WS-Subskala solche positiven Eigenschaften attestiert, die sie für die traditionell weibliche Rolle prädestinieren (siehe Unterkapitel 3.2.2). Da die Multitasking-Fähigkeit nicht automatisch mit reproduktiven Aufgaben in Verbindung steht, wird sie explizit im Kontext von Haushaltstätigkeiten und Kinderbetreuung benannt. Um die besondere Eignung der Frauen für die Familienarbeiten zu unterstreichen, wird ihnen eine natürliche, wesensgemäße Veranlagung für diese Aufgaben attestiert, die dem Mann nicht bzw. nicht in gleichem Maße gegeben sei („Frauen besitzen ein größeres Talent als Männer“). Auf diese Weise sollte die Sinnhaftigkeit der traditionellen Rollenaufteilung herausgestellt und zugleich deren Unumstößlichkeit untermauert werden.

Im Laufe der kognitiven Interviews erwies sich jedoch gerade diese Formulierung als problematisch. So war sich eine Befragte zunächst nicht sicher, ob das Wort „Talent“ etwas sozial Erlerntes oder etwas Angeborenes impliziere. Da dieser Punkt jedoch ausschlaggebend für die Wahl ihrer Antwortkategorie war, bereitete ihr diese Unklarheit anfänglich Schwierigkeiten. Letztendlich interpretierte sie das Item wie beabsichtigt. Sie entschied sich aufgrund des Wortes „besitzen“ dafür, dass eine natürliche Veranlagung der Frauen angesprochen werde und verneinte daher das Item.

**Interview Nr. 09 (weiblich): „stimme eher nicht zu“**

**Id. 09:** (spontane Äußerung vor Nachfrage) Hm... „stimme eher nicht zu“, wobei Moment... schwierig. Ich schwanke zwischen „stimme eher nicht zu“ und „stimme eher zu“, ähm, weil ich denke, also, das ist vielleicht ein „Talent“, was Frauen entwickeln mit der Zeit, weil sie einfach da drauf angewiesen sind, das so unter einen Hut zu bekommen, ähm, insofern, also, stimmt das schon. Wenn das was ist, also ein „Talent“, was sich entwickelt, aber nichts, was ihr angeboren ist.

**Int.:** Wenn ich jetzt nicht hier wär, was würden Sie denn dann antworten oder was würden Sie ankreuzen?

**Id. 09:** Da würde ich, glaub ich, auch..., obwohl... Das „besitzen“ heißt ja eigentlich, dass das was ist, was sie von Natur aus haben. Insofern würde ich eher sagen, „stimme eher nicht zu“.

**Int.:** Würden Sie bitte die Frage noch einmal mit Ihren eigenen Worten wiederholen. Was denken Sie, worum es bei dieser Frage geht?

**Id. 09:** Frauen sind von Natur aus, äh, oder haben von Natur aus dazu die Veranlagung irgendwie, ja, Haushalt und Kinder, ähm, gleichermaßen zu bewältigen. Ja. Dahinter steht für mich sowas, wie, ähm, eben auch bei der Frage vorher schon, dass Frauen, wenn sie dann doch das größere Talent dazu haben, das beides zu machen, dann sollen sie das auch. So. Also, sollen sie es auch machen. (...)

Andere InterviewteilnehmerInnen interpretierten das Item hingegen nicht in der intendierten Weise und gingen davon aus, dass von einer sozial erlernten Fähigkeit die Rede sei. Eine derartige Auslegung ist problematisch. Wie bereits in Unterkapitel 2.2.3.4 erläutert, tendieren Personen, die die Ansicht vertreten, Geschlechterdifferenzen hinsichtlich bestimmter Merkmale (in diesem Falle bezüglich der Multitasking-Fähigkeit) seien in erster Linie auf biologische Faktoren zurückzuführen (und somit naturgegeben) in stärkerem Maße dazu, normative Verhaltenserwartungen zu hegen, als solche Personen, die die gleichen Geschlechtsunterschiede auf Sozialisationsprozesse zurückführen (vgl. Swim & Campbell, 2003, S. 223).

Wird also im Zusammenhang mit diesem Item von einer erlernten Fähigkeit ausgegangen, sind Fehleinstufungen der InterviewteilnehmerInnen eher wahrscheinlich. So ist

es vorstellbar, dass eine befragte Person dem Item nur zustimmt, weil es ihrer Beobachtung entspricht, dass Frauen derzeit eher die Familienarbeiten übernehmen und daher auch geübter darin sind, diese zu vereinbaren. Gleichzeitig könnte sie jedoch die Ansicht vertreten, dass auch Männer durchaus lernen könnten, die reproduktiven Aufgaben zu übernehmen und dass eine Aufteilung der Arbeiten daher nicht automatisch am Geschlecht festgemacht werden sollte. Folglich würde die befragte Person auch vermutlich keine normativen Verhaltenserwartungen an Frauen stellen, die Hausarbeit und die Kinderbetreuung zu übernehmen. Aufgrund ihrer Zustimmung zu dem Item wäre die Person unter diesen Umständen fälschlicherweise als benevolent-sexistisch eingestuft worden.

Um das identifizierte Formulierungsdefizit zu beheben, wurde das Wort „Talent“ gestrichen und stattdessen die Formulierung „natürliche Begabung“ in das Item aufgenommen.

Endgültige Itemversion *Hausarbeit und Kinderbetreuung*

*„Frauen besitzen eine größere natürliche Begabung als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.“*

**Ursprungitem *Blick Hausarbeiten***

*„Frauen haben einfach einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.“*

Wie auch bei dem vorangegangenen Item *Hausarbeit und Kinderbetreuung* wurde in der Frage *Blick Hausarbeiten* die besondere, wesensgemäße Eignung der Frauen („Frauen haben einfach“) für die Reproduktionsarbeiten aufgegriffen, über die Männer aus stereotyper Sicht nicht gleichermaßen verfügen. Auf diese Weise sollte die traditionelle Arbeitsteilung, bei der die Frau sich um den Haushalt (und die Kinderbetreuung) kümmert, als ideale Lösung herausgestellt werden. Die Folge wäre somit eine Zementierung der Geschlechterhierarchien in der Gesellschaft.

Als zentrales Ergebnis der kognitiven Interviews bleibt festzuhalten, dass auch bei diesem Item die natürliche Veranlagung nicht als solche aufgefasst wurde. So legten mehrere Befragte die Formulierung „einfach einen besseren Blick haben“ als etwas Erlerntes aus. Die folgende Interviewsequenz verdeutlicht zugleich das oben beschriebene Problem einer möglichen Fehleinstufung – so stimmt die Befragte der Aussage eher zu, da sie davon ausgeht, Frauen hätten aufgrund der geschlechtsspezifischen Sozialisation eher einen Blick dafür, welche Hausarbeiten es zu erledigen gälte. Sollte die Aussage allerdings auf eine biologische, wesensgemäße Veranlagung der Frau hierzu abzielen, dann hätte sie dem Item überhaupt nicht zugestimmt:

**Interview Nr. 25 (weiblich): „stimme eher zu“**

**Id. 25:** (spontane Äußerung vor Nachfrage) Das ist 'ne dumme Frage.

**Int.:** Warum?

**Id. 25:** Weil die Sozialisation ... Natürlich, wenn die Frauen zu Hause sind, dann ist es so, dass sie „einen besseren Blick haben“, aber das heißt nicht, dass die Männer die Fähigkeit nicht haben, es zu können. Also so, wie es jetzt formuliert ist, die heute sozialisierten Frauen ... Ja, das glaube ich schon, weil die damit einfach routinierter sind. (...) Wenn es auf die Frauen als Geschlecht „Frau“ bezogen ist, würde ich sagen „stimme ich überhaupt nicht zu“. Obwohl, in dieser Welt ist es, glaube ich, so, dass Frauen das durch die Sozialisation immer zugeteilt wurde und deswegen machen die das jetzt auch mehr. (...) Grundsätzlich gehe ich davon aus, dass es [einen besseren Blick] ein Mann genauso haben kann, wenn er es machen muss. Das würde ich ihm nie absprechen. Aber in der Gesellschaft ist es im Moment so, dass sie es einfach haben durch die Beschäftigung damit. Das muss man wirklich ... die Frage ist blöd.

Da das Item somit nicht von allen Befragten in der beabsichtigten Weise verstanden wurde, war eine Überarbeitung notwendig. In Anlehnung an das Item *Hausarbeit und Kinderbetreuung* wurde das Item *Blick Hausarbeiten* um den Zusatz „von Natur aus“ ergänzt. Diese Alternativformulierung war bereits während der zweiten Runde der kognitiven Interviews mitgetestet worden und erwies sich als eindeutig.

Endgültige Itemversion *Blick Hausarbeiten*

„Frauen haben von Natur aus einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.“

### 3.4.5.3 Zusammenfassung

Wie eingangs erläutert, wurden im vorangegangenen Abschnitt nur beispielhaft die Ergebnisse einiger Items vorgestellt (darunter u.a. diejenigen Items, die letztendlich auch in die FWSEM-Skala aufgenommen wurden). Insgesamt wurden allerdings 41 Items mittels kognitiver Techniken untersucht, für die im Rahmen der Interviews ähnliche Formulierungsdefizite identifiziert werden konnten. Es lassen sich folgende zentrale Ergebnisse festhalten: Mit Hilfe der kognitiven Interviews wurden mehrere Frage-Probleme identifiziert, die dazu führten, dass die Items nicht in der intendierten Weise von den Befragten verstanden wurden. In erster Linie wurden Formulierungen oder Schlüsselbegriffe ausfindig gemacht, die sich als zu unspezifisch/vage erwiesen (z.B. „kleine Kinder“ oder „wenige Wochen“). In mehreren Fällen zeigte sich, dass den InterviewteilnehmerInnen genauere Angaben fehlten, um sich für eine Antwortkategorie entscheiden zu können. Dazu zählten u.a. das genaue Alter des Kindes, die Art der außerhäuslichen Betreuung der Kinder, der Umfang der Erwerbstätigkeit der Mutter sowie der Beweggrund für die Erwerbstätigkeit. Des Weiteren waren einige Begrifflichkeiten nicht bekannt oder sie wurden anders ausgelegt, als dies beabsichtigt war (z.B. „Kinderkrippe“, „erfolgreiche

berufstätige Frau“ oder „Talent“). Auch die Verwendung eines doppelten Stimulus stellte sich mitunter als problematisch heraus. Auf Basis dieser Erkenntnisse wurden die Fragen umformuliert und präzisiert.

Nicht alle Items, die in Runde 3 der kognitiven Interviews noch geprüft und überarbeitet worden waren, wurden jedoch auch in Pretest-Phase 3 quantitativ untersucht. Unter den getesteten Aussagen befanden sich mehrere Items, die inhaltlich sehr ähnlich waren und deren unterschiedliche Formulierungsalternativen mit Hilfe der kognitiven Methoden überprüft werden sollten. Von diesen Items wurden diejenigen ausgewählt, die sich im Rahmen der kognitiven Interviews am besten bewährten. Anzumerken ist auch, dass es bei einigen wenigen Items auch nach einer (mehrfachen) Überarbeitung nicht gelang, eine zufriedenstellende Frageformulierung zu finden (z.B. Item *Verständnis*). Diese Fragen wurden (größtenteils)<sup>47</sup> aus dem Itempool entfernt, so dass sich letztendlich die Zahl der Items nach Pretest-Phase 2 auf 28 verkleinert hatte (siehe Tabelle 5 und Anhang 1 für einen Überblick über die vorgenommenen Veränderungen und Reduzierungen des Itempools im Anschluss an die einzelnen Pretest-Phasen).

### 3.5 Pretest-Phase 3: Auswahl der Items und Einsatz der FWSEM-Skala

#### 3.5.1 Überblick

In den beiden quantitativen Studien von Pretest-Phase 3 sollten nun anhand ihrer psychometrischen Güte die besten Items für die endgültige FWSEM-Skala ausgewählt sowie ihre Konstrukt- und Übereinstimmungsvalidität analysiert werden. Hierfür wurden eine Online-Befragung (Studie 6) und eine schriftliche Befragung (Studie 7) durchgeführt. In Studie 8 wurde die neu entwickelte FWSEM-Skala im Rahmen einer telefonischen Bevölkerungsumfrage eingesetzt, um zu prüfen, ob sich die Skala auch im Feld bewährt. Hierfür wurden abermals psychometrische Analysen durchgeführt.

#### 3.5.2 Stichprobe 6 und 7

In Studie 6 beteiligten sich 207 Personen (145 Frauen und 62 Männer) im Alter zwischen 18 und 63 Jahren ( $M = 26.81$ ,  $SD = 8.51$ ) an der Online-Befragung. Der dazugehörige Link wurde über das Adressbuch der Autorin sowie verschiedene Verteiler-Listen der Philipps-Universität Marburg weitergeleitet. Darüber hinaus wurden alle Personen angeschrieben, die sich auf die Anzeige zur Rekrutierung von Befragten für die kognitiven

---

<sup>47</sup> Das Item *Verständnis* wurde beispielsweise noch nicht unmittelbar nach Pretest-Phase 2 aus dem Itempool herausgenommen, da es inhaltlich einen zusätzlichen Aspekt ansprach, der in den anderen Items nicht explizit enthalten war: die finanzielle Absicherung der Familie auch ohne ein zusätzliches Einkommen der Mutter. In Pretest-Phase 3 wiesen die Itemkennwerte das Item auch nicht als problematisch aus, aufgrund der Bedenken hinsichtlich der überarbeiteten Itemformulierung, wurde es aber letztlich doch nicht in die FWSEM-Skala aufgenommen.

Interviews hin gemeldet hatten, aber nicht daran teilgenommen hatten. Als Anreiz für die Teilnahme wurden drei Gutscheine im Wert von je 40 Euro unter allen Befragten verlost.

Insgesamt nahmen 90 Personen (53 Frauen und 35 Männer; zwei Personen machten bezüglich des Geschlechts keine Angaben) auf freiwilliger Basis an der schriftlichen Befragung in Studie 7 teil. Die Altersspanne der Befragten reichte von 20 bis 49 Jahren ( $M = 25.31$ ,  $SD = 3.92$ ). Die Gelegenheitsstichprobe setzte sich aus (ehemaligen = 8.9%) Studierenden der Justus-Liebig-Universität Gießen und der Philipps-Universität Marburg zusammen. Die TeilnehmerInnen wurden im Sommersemester 2009 auf freiwilliger Basis teilweise aus Lehrveranstaltungen, überwiegend jedoch auf dem Campusgelände der jeweiligen Universitäten rekrutiert. Dementsprechend breit gestreut waren auch die einzelnen Studienfächer der Befragten: Politikwissenschaft (10%), Wirtschaftswissenschaften (6.7%), Pädagogik (5.6%), Soziologie (5.6%), Jura (3.3%), Physik (3.3%), Psychologie (3.3%), Ökotoxikologie (3.3%), Sozialwissenschaften (3.3%) und Sonstige (46.7%) (für die exakte Zusammensetzung beider Stichproben siehe Tabelle 6). Als Incentive wurde den TeilnehmerInnen ein Schokoladenriegel angeboten.

### 3.5.3 Stichprobe 8

Die Befragung in Studie 8 wurde zwischen Mai und August 2010 als computerunterstütztes telefonisches Interview (Computer Assisted Telephone Interview, CATI-Methode) von dem professionellen Umfrageinstitut Infratest (mit 23 InterviewerInnen) durchgeführt. Die Auswahl der Haushalte erfolgte nach dem „Gabler-Häder-Verfahren“ (siehe hierzu Häder, 2000). Um eine Zufallsauswahl der Befragten innerhalb der Haushalte zu erhalten, wurde der „Schwedenschlüssel“ verwendet. Ein Teil der Befragten (4.8%) wurde über das Mobilfunknetz kontaktiert, um auch jüngere Menschen zu erreichen, die häufig über keinen Festnetzanschluss mehr verfügen.

Insgesamt nahmen 397 Personen (darunter 242 Frauen und 155 Männer) im Alter zwischen 16 und 90 Jahren ( $M = 54.66$ ,  $SD = 15.89$ ) an der Umfrage teil. Knapp die Hälfte (54.7%) der Befragten war zum Zeitpunkt der Befragung erwerbstätig, etwa ein Drittel (32.5%) befand sich bereits im Altersruhestand, und die übrigen Personen waren arbeitslos, Hausfrauen bzw. Hausmänner, Studierende, SchülerInnen oder gaben an, einer anderen Tätigkeit nachzugehen. Hinsichtlich des höchsten Bildungsabschlusses war die Stichprobe heterogen zusammengesetzt: 30% der Befragten hatte ein Studium absolviert, 21.4% hatten das Abitur abgelegt, 30.5% verfügten über die Mittlere Reife, 14.1% besaßen einen Hauptschulabschluss, 0.8% der Befragten konnten (noch) keinen Abschluss vorweisen und die verbliebenen 2.8% gaben an, einen anderen Abschluss, als die bereits genannten, zu besitzen (für die exakte Zusammensetzung der Stichprobe siehe Tabelle 6).

### 3.5.4 Erhebungsinstrumente in Stichprobe 6

Das Hauptanliegen von Studie 6 war es, die Items für die endgültige Version der FWSEM-Skala zu ermitteln sowie deren Konstruktvalidität zu analysieren. Neben den 28 Items des überarbeiteten Itempools nach Pretest-Phase 2 wurden daher in Studie 6 sieben weitere Skalen zur Validierung eingesetzt, darunter vier Sexismus-Skalen, eine Autoritarismus-Skala, eine Skala zur Messung der Sozialen Dominanzorientierung und eine Skala zur Messung sozialer Erwünschtheit (Angaben zu den Skalennamen, Itembeispiele und die Werte der internen Konsistenz sind Tabelle 9 zu entnehmen. Für eine genaue Itemauflistung aller verwendeter Skalen siehe Anhang 6. Der dort aufgeführte Fragebogen diente als Vorlage für die Online-Erhebung und stimmt mit dieser inhaltlich überein).<sup>48</sup>

Die Items wurden allen Befragten in gleicher Reihenfolge dargeboten (zunächst wurden die FWSEM-Items und anschließend die übrigen sieben Skalen in der Abfolge erhoben, in der sie auch in Tabelle 9 aufgelistet sind). Zur Bewertung der einzelnen Aussagen stand das sechsstufige Antwortformat aus Phase 1 und 2 zur Verfügung (von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“, bzw. bei der Skala zur Messung der sozialen Erwünschtheit von 1 „trifft überhaupt nicht zu“ bis 6 „trifft voll und ganz zu“).

Sowohl für die FWSEM-Gesamtskala als auch die FS-Subskala und die WS-Subskala wurden als Hinweis für die konvergente Validität positive Korrelationen mit den vier Sexismus-Skalen erwartet (ASS, MSS, GRB, AMMSA), da ihnen eine Abwertung des weiblichen Geschlechts und die Festlegung auf die traditionell weibliche Rolle zugrunde liegen. Derartige Zusammenhänge wurden schon früher von verschiedenen AutorInnen aufgezeigt (vgl. u.a. Eckes & Six-Materna 1998, 1999; Gerger, Kley, Bohner & Siebler, 2007; Glick & Fiske, 1996). Auch für die beiden gruppenbezogenen Ideologien (RWA und SDO) wurden positive Korrelationen angenommen, da RWA u.a. durch ein Festhalten an gesellschaftlichen Konventionen gekennzeichnet ist und SDO einer Befürwortung gruppenbasierter Ungleichheiten und hierarchisch strukturierter Beziehungen entspricht (vgl. Swim et al., 2010, S. 161). Für andere Sexismus-Skalen konnte in der Vergangenheit bereits mehrfach eine Verbindung zu autoritären Einstellungen und zur Sozialen Dominanzorientierung belegt werden (vgl. u.a. Eckes & Six-Materna 1998, 1999; Schmidt-Mummendey, Schmerl, & Schmidt, 1973). Im Zusammenhang mit der Sozialen Erwünschtheits-Skala sollten hingegen keine signifikanten Korrelationen mit der FWSEM-Skala gefunden werden (Hinweis auf die diskriminante Validität).

---

<sup>48</sup> Da Cronbachs Alpha für alle Skalen (mit Ausnahme der FT-Skala) in einem akzeptablen Bereich lag, wurde für jede Skala ein Gesamtscore über alle Items hinweg berechnet. Hierfür wurden alle Itemwerte einer Skala aufaddiert und durch die Anzahl der Items geteilt. Aufgrund der unzufriedenstellenden Reliabilität der FT-Skala von  $\alpha = .26$  wurde sie in den folgenden Analysen nicht miteinbezogen.



Tabelle 9: Itembeispiele und Reliabilitäten der Skalen, die zur Konstruktvalidierung der FWSEM-Skala in Stichprobe 6 eingesetzt wurden

Skala (Anzahl der Items)	Itembeispiele	Cronbachs Alpha
MSS (10)	„Wenn Frauen tatsächlich einmal schlechter bezahlt werden als Männer, dann nur deshalb, weil sie einfachere Arbeit zu leisten haben.“	.87
ASS (22)		.93
Hostiler Sexismus (11)	„Die meisten Frauen interpretieren harmlose Äußerungen oder Handlungen als frauenfeindlich.“	.93
Benevolenter Sexismus (11)	„Egal wie erfolgreich ein Mann auch sein mag, ohne eine Frau, die ihn liebt, fehlt ihm etwas ganz Wichtiges.“	.89
GRB (10)	„Einen Beruf zu haben ist ja ganz schön, aber das, was die meisten Frauen wirklich wollen, sind ein Heim und Kinder.“	.80
RWA (9)	„Das wichtigste, was Kinder lernen sollten, ist Gehorsam und Respekt gegenüber Eltern und Vorgesetzten.“	.77
SDO (3)	„Es gibt Gruppen in der Bevölkerung, die weniger wert sind als andere.“	.68
AMMSA (9)	„Frauen zieren sich gerne. Das bedeutet nicht, dass sie keinen Sex wollen.“	.94
BIDR		
FT (3)	„Ich habe schon mal zu viel Wechselgeld zurückbekommen und nichts gesagt.“	.26
ST (3)	„Mein erster Eindruck von Menschen stellt sich gewöhnlich als richtig heraus.“	.58

Anmerkungen. MSS = Skala zur Erfassung von Vorurteilen gegenüber Frauen auf der Basis des Konzepts des modernen Sexismus (Eckes & Six-Materna, 1998). ASS = Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999) mit den beiden Subskalen zum hostilem (HS) und zum benevolenten Sexismus (BS). GRB = Geschlechterrollen-Batterie des International Social Survey Programme-Modul 2002 Familie in Deutschland. RWA = Autoritarismus-Kurzskala (Petzel, Wagner, Nicolai & van Dick, 1997). SDO = Kurzversion der Social Dominance Orientation Scale aus dem Projekt Gruppenbezogene Menschenfeindlichkeit (Heitmeyer, 2002, 2010). AMMSA = Kurzversion der Skala zur Erfassung der Akzeptanz moderner Mythen über sexuelle Aggression (Gerger, Kley, Bohner & Siebler, 2007). BIDR = Balanced Inventory of Desirable Responding, Kurzskala zur Erfassung der Tendenz der Befragten, sozial erwünschte Antworten zu geben (Winkler, Kroh & Spiess, 2006). FT = Fremdtäuschung (impression management) und ST = Selbsttäuschung (self-deceptive enhancement) sind Subkomponenten der BIDR-Kurzskala.

### 3.5.5 Erhebungsinstrumente in Stichprobe 7

Mit Hilfe von Studie 7 sollte insbesondere die Übereinstimmungsvalidität der neuen FWSEM-Skala beleuchtet werden. Zu diesem Zweck wurden zu Beginn des Fragebogens die ambivalent-sexistischen Einstellungen der Befragten mit Hilfe der ASS (Eckes & Six-Materna, 1999) sowie mit den 28 Items des überarbeiteten Itempools aus Pretest-Phase 2, aus dem schließlich auch die endgültige 12-Item-Version der FWSEM-Skala hervorging, erhoben (sechsstufige Antwortskala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“).<sup>49</sup> Anschließend sollten die TeilnehmerInnen einen nicht-traditionellen Frauentyp (Karrierefrauen) und einen traditionellen Frauentyp (Hausfrauen) anhand einer geschlechterstereotypen Merkmalsskala beurteilen.<sup>50</sup> Hierzu wurde der „German Extended Personal Attributes Questionnaire“ (GEPAQ) von Runge, Frey, Gollwitzer, Helmreich und Spence (1981) eingesetzt. Der GEPAQ umfasst insgesamt 32 geschlechterstereotype Merkmale, von denen aus zeitökonomischen Gründen jedoch in Anlehnung an Eckes (2001) nur die Hälfte abgefragt wurde. Die ausgewählten 16 Merkmale verteilen sich gleichmäßig auf die folgenden vier Kategorien: (a) feminin-positiv (sanft, fähig auf andere einzugehen, verständnisvoll und hilfsbereit), (b) feminin-negativ (leichtgläubig, weinerlich, nervös und unterwürfig), (c) maskulin-positiv (unabhängig, kann Druck gut standhalten, wettbewerbsorientiert und selbstsicher) und (d) maskulin-negativ (überheblich, prahlerisch, nur auf sich selbst bedacht und feindselig). Die Befragten wurden gebeten, für die beiden Subgruppen jeweils nacheinander einzuschätzen, auf wie viel Prozent der Karriere- bzw. Hausfrauen ihrer Meinung nach die Merkmalsbeschreibungen zutreffen.<sup>51</sup>

Im Zusammenhang mit der FS-Subskala wurde erwartet, dass sie mit negativen Stereotypen über Karrierefrauen zusammenhängen. Das heißt, Personen, die zum feindseligen Sexismus neigen, sollten Karrierefrauen maskulin-negative Eigenschaften zuschreiben und gleichzeitig feminin-positive Eigenschaften absprechen. Die WS-Subskala sollte

<sup>49</sup> Da Cronbachs Alpha für beide Skalen akzeptable Werte aufwies (HS = .87, BS = .85; FS = .90, WS = .89), wurde jeweils ein Gesamtscore über alle Items hinweg berechnet. Hierfür wurden alle Itemwerte einer Skala aufaddiert und durch die Anzahl der Items geteilt.

<sup>50</sup> Diese beiden Gruppen wurden ausgewählt, da sie sich in der bisherigen Forschung als Prototypen für nicht-traditionelle bzw. traditionelle Subgruppen von Frauen erwiesen haben und in mehreren Studien (siehe hierzu z.B. Eckes, 1994; Glick, Diebold, Bailey-Werner & Zhu, 1997) bereits vergleichend untersucht wurden (vgl. Eckes, 2001).

<sup>51</sup> Die Reihenfolge der beiden Gruppen wurde systematisch variiert: der einen Hälfte der Befragten wurde zuerst die Gruppe der Hausfrauen und danach die Gruppe der Karrierefrauen zur Einschätzung vorgelegt, für die andere Hälfte war die Abfolge genau umgekehrt. Die Reihenfolge hatte keinen Effekt. Einzige Ausnahme: Wurde den Befragten zuerst die Gruppe der Karrierefrauen und anschließend die Gruppe der Hausfrauen vorgelegt, so wurden den Hausfrauen in signifikant höherem Maße feminin-negative Eigenschaften zugeschrieben als in der Gruppe mit umgekehrter Abfolge.  $M$  (Karrierefrauen-Hausfrauen) = 40.86;  $M$  (Hausfrauen-Karrierefrauen) = 33.35,  $t(84) = -2.01$ ,  $p < .05$ . Da jedoch keine Hypothese über einen Zusammenhang der beiden FWSEM-Subskalen und der Zuschreibung von feminin-negativen Eigenschaften bestand und sich diesbezüglich auch keine signifikanten Korrelationen ergaben, spielt der beobachtete Reihenfolgeeffekt im Rahmen dieser Untersuchungen keine Rolle und kann vernachlässigt werden.

hingegen mit positiven Stereotypen über Hausfrauen korrelieren. Folglich sollten Personen, die zum wohlwollenden Sexismus neigen, Hausfrauen mit feminin-positiven Eigenschaften in Verbindung bringen. Diese Hypothesen stützen sich auf mehrere Untersuchungen (vgl. Eckes, 2002; Fiske, Cuddy, Glick, & Xu, 1999, 2002) im Zusammenhang mit dem Stereotype Content Model (siehe Unterkapitel 2.2.3.4) in denen herausgefunden wurde, dass nicht-traditionelle Frauen typischerweise als emotional kalt angesehen und ihnen somit feminin-positive Eigenschaften wie Warmherzigkeit abgesprochen werden. Gleichzeitig wird Karrierefrauen in der Regel auch eine hohe Kompetenz zugeschrieben, was jedoch auf hostile Sexisten nicht zutreffen dürfte, da sie ein komplett negatives Bild von nicht-traditionellen Frauen haben und ihnen sogar mangelnde Kompetenz unterstellen. Im Gegensatz dazu werden traditionelle Frauen als warmherzig erachtet, so dass ihnen feminin-positive Eigenschaften zugesprochen werden sollten.

### 3.5.6 Erhebungsinstrumente in Stichprobe 8

In Studie 8 wurde die neu konstruierte FWSEM-Skala im Rahmen einer Bevölkerungsumfrage eingesetzt, die sich mit verschiedenen gesellschaftlichen Themen befasst.<sup>52</sup> Den Befragten wurde mitgeteilt, dass es bei der Umfrage darum gehe, die Meinung der Bevölkerung zum sozialen Zusammenleben in Deutschland zu erfassen.

Von Relevanz für die Validierung der FWSEM-Skala waren v.a. die Items zur Erfassung der Sozialen Dominanzorientierung (SDO), des Autoritarismus (RWA), der Akzeptanz moderner Mythen über sexuelle Aggression (AMMSA) sowie zur Gruppenbezogenen Menschenfeindlichkeit (GMF). Letztere drückt sich in der „Abwertung, Ablehnung und Diskriminierung von gesellschaftlichen Minderheiten“ (Heitmeyer, 2002, 2010) aus und setzt sich aus den folgenden Elementen zusammen: (Traditioneller und Moderner) Sexismus, Fremdenfeindlichkeit, Rassismus, Etabliertenvorrechte, Antisemitismus, Islamophobie, Homophobie, Abwertung von Obdachlosen, Abwertung von Behinderten sowie Abwertung von Langzeitarbeitslosen. Einzelne Itembeispiele sowie die Werte der internen Konsistenz sind für die eben genannten Skalen in Tabelle 10 aufgeführt.<sup>53</sup>

---

<sup>52</sup> Die Befragung wurde im Rahmen des „GMF (Gruppenbezogene Menschenfeindlichkeit) Graduierten-Surveys“ 2010 (Prof. Dr. Andreas Zick) durchgeführt und durch die Universität Bielefeld gefördert. Die verschiedenen Themenschwerpunkte (wie z.B. die Wirtschafts- und Finanzkrise, die Wahrnehmung und Bewertung politischer Gestaltungsentwürfe) spiegeln folglich die Forschungsinteressen derjenigen WissenschaftlerInnen wider, die an dem Graduierten-Survey teilgenommen haben.

<sup>53</sup> Da Cronbachs Alpha für alle Skalen akzeptable Werte aufwies, wurde jeweils ein Gesamtscore über alle Items hinweg berechnet. Hierfür wurden alle Itemwerte einer Skala aufaddiert und durch die Anzahl der Items geteilt. Im Zusammenhang mit der FWSEM-Gesamtskala und den beiden Subskalen wurde gleichermaßen vorgegangen.

Tabelle 10: Itembeispiele und Reliabilitäten der Skalen, die zur Konstruktvalidierung der FWSEM-Skala in Stichprobe 8 eingesetzt wurden

Skala (Anzahl der Items)	Itembeispiele	Cron- bachs Alpha
RWA (4)	„Verbrechen sollten härter bestraft werden.“	.75
SDO (3)	„Es gibt Gruppen in der Bevölkerung, die weniger wert sind als andere.“	.59
AMMSA (9)	„Frauen zieren sich gerne. Das bedeutet nicht, dass sie keinen Sex wollen.“	.78
GMF		
Traditioneller Sexismus (2)	„Frauen sollen sich wieder mehr auf die Rolle der Ehefrau und Mutter besinnen.“	.82
Moderner Sexismus (2)	„Die Diskriminierung von Frauen ist in Deutschland immer noch ein Problem.“	.67
Fremdenfeindlichkeit (2)	„Es leben zu viele Ausländer in Deutschland.“	.69
Rassismus (2)	„Die Weißen sind zu Recht führend in der Welt.“	.53
Islamophobie (2)	„Muslimen sollte die Zuwanderung nach Deutschland untersagt werden.“	.70
Antisemitismus (2)	„Juden haben in Deutschland zuviel Einfluss.“	.67
Etabliertenvorrechte (2)	„Wer irgendwo neu ist, sollte sich erst mal mit weniger zufrieden geben.“	.54
Homophobie (3)	„Es ist ekelhaft, wenn Homosexuelle sich in der Öffentlichkeit küssen.“	.82
Abwertung Obdachloser (3)	„Die Obdachlosen in den Städten sind unangenehm.“	.66
Abwertung Behinderter (3)	„Behinderte erhalten zu viele Vergünstigungen.“	.81
Abwertung Langzeitarbeitsloser (2)	„Die meisten Langzeitarbeitslosen sind nicht wirklich daran interessiert, einen Job zu finden.“	.68

Anmerkungen. GMF = Itembatterie des Syndroms Gruppenbezogene Menschenfeindlichkeit (Heitmeyer, 2002, 2010). RWA = Autoritarismus-Kurzskala (Petzel, Wagner, Nicolai, & van Dick, 1997). SDO = Kurzversion der Social Dominance Orientation Scale aus dem Projekt Gruppenbezogene Menschenfeindlichkeit (Heitmeyer, 2002, 2010). AMMSA = Kurzversion der Skala zur Erfassung der Akzeptanz moderner Mythen über sexuelle Aggression (Gerger, Kley, Bohner & Siebler, 2007).

Die Items wurden allen Befragten von den InterviewerInnen in gleicher Reihenfolge vorgelesen (siehe Anhang 7). Zur Bewertung der Aussagen stand den TeilnehmerInnen eine vierfach abgestufte Ratingskala von 1 „stimme voll und ganz zu“ bis 4 „stimme überhaupt nicht zu“ zur Verfügung. Gleich zu Beginn und auch zum Ende des Interviews wurden umfangreiche demografische Angaben der Befragten erfasst (z.B. das Geschlecht, die Nationalität, die Religionszugehörigkeit, das Geburtsjahr, der Familienstand, der Bildungsgrad, die Erwerbstätigkeit sowie die Höhe des Einkommens).

Wie bereits für Studie 6 wurde auch für Studie 8 eine positive Korrelation der drei FWSEM-Skalen mit der Autoritarismus-Skala (RWA) und der Skala zur Erfassung der Sozialen Dominanzorientierung (SDO) sowie mit den drei Sexismus-Skalen (TS, MS und AMMSA) angenommen. Gleiches galt neben dem traditionellen und modernen Sexismus auch für die neun übrigen Elemente des GMF-Syndroms, deren gemeinsamer Kern die Ideologie der Ungleichheit/Ungleichwertigkeit ist. So konnte bereits in früheren Arbeiten Sexismus als Bestandteil eines größeren Einstellungssyndroms identifiziert werden, welches verschiedene Arten von Vorurteilen erfasst. Schmidt-Mummendey, Schmerl und Schmidt (1973) fanden beispielsweise einen Zusammenhang von misogynen (darunter fielen neben offen sexistischen auch scheinbar ritterliche) Einstellungen mit anderen konservativ autoritären Vorurteilen wie Autoritarismus, Ethnozentrismus, Dogmatismus und Konformismus. Sie sprechen daher von einem autoritär-konservativ-misogynen Einstellungssyndrom. Aosved, Long, und Voller (2009) konnten eine Verbindung von Sexismus (traditionell und modern) mit Rassismus, sexuellen Vorurteilen, Klassismus (classism) und Intoleranz aufgrund einer Glaubenzugehörigkeit (religious intolerance) feststellen. Folglich wurde vermutet, dass auch für den ambivalenten Sexismus ein positiver Zusammenhang mit den einzelnen Elementen des Syndroms Gruppenbezogener Menschlichkeit besteht.

### 3.5.7 Ergebnisse aus Studie 6, 7 und 8

#### 3.5.7.1 Endgültige Auswahl der Items für die FWSEM-Skala

Um die endgültige Auswahl der Items für die FWSEM-Skala zu treffen, wurden in Studie 6 alle 28 Items des überarbeiteten Itempools erneut auf ihre Itemschwierigkeit, ihre Trennschärfen sowie ihre Faktorladungen hin untersucht. Sie sollten dabei denselben Kriterien genügen wie in Studie 1 und 2:

- (a) *mittlere Itemschwierigkeiten* (definiert als Itemmittelwerte größer 2.0 und kleiner 5.0),
- (b) *hohe Trennschärfen* (definiert als eine korrigierte Itemwert-Skalenwert-Korrelation größer .40),
- (c) *möglichst reine Faktorladungen* (definiert als Hauptladungen größer .55 und Nebenladungen kleiner .30.) auf dem theoretisch zu verortenden Faktor.

Darüber hinaus wurden die Reliabilitäten verschiedener Skalenversionen getestet. Sowohl die Gesamtskala als auch beide Subskalen sollten ein zufriedenstellendes Niveau erzielen. Dementsprechend wurden die drei oben genannten Kriterien durch ein viertes ergänzt:

(d) hohe *interne Konsistenz* (definiert als Cronbachs- $\alpha > .80$ )

In einem ersten Schritt wurden die Itemschwierigkeiten sowie die Trennschärfen analysiert. Es zeigte sich, dass fast alle Items, die sich nach Abschluss der Pretest-Phase 2 noch im überarbeiteten Itempool befanden, zufriedenstellende Itemschwierigkeiten (FS-Subskala:  $2.00 \leq M \leq 3.49$ ; WS-Subskala:  $2.61 \leq M \leq 4.00$ ) und Trennschärfen (Gesamtskala:  $.42 \leq r_{it} \leq .76$ ; FS-Subskala:  $.42 \leq r_{it} \leq .80$ ; WS-Subskala:  $.48 \leq r_{it} \leq .73$ ) besaßen. Lediglich zwei Items wiesen einen Mittelwert kleiner 2.0 auf und wurden aus dem Itempool entfernt.

In einem zweiten Schritt wurden die Faktorladungen sowie die interne Konsistenz geprüft. Hinsichtlich der Faktorenanalyse wurden insgesamt drei verschiedene Tests durchgeführt: Eine Faktorenanalyse nur mit den 26 verbliebenen Items aus dem überarbeiteten Itempool der FWSEM-Skala, eine simultane Analyse gemeinsam mit der ASS (Eckes & Six-Materna, 1999) sowie eine abschließende Analyse mit allen Items der vier verschiedenen Sexismus-Skalen, die in Studie 6 eingesetzt wurden.

Es wurden folgende Items aus dem Itempool ausgeschlossen:

- (1) Items, die keine reinen Ladungen auf einem Faktor besaßen (siehe Punkt (c)).
- (2) Items, die empirisch auf einem anderen Faktor luden, als sie theoretisch verortet wurden (z.B. Items, die inhaltlich dem hostilem Sexismus zugeordnet waren, jedoch auf einem Faktor des benevolenten Sexismus luden oder umgekehrt (siehe Punkt (c))).
- (3) Items, die auf einem Faktor der bereits bestehenden ASS oder einer anderen Sexismus-Skala (MSS, GRB und AMMSA) luden, da dies die neuen Items hätte redundant werden lassen.

So weit dies die Daten aus Studie 7 zuließen, wurden die gleichen Analysen durchgeführt und mit den Ergebnissen aus Studie 6 verglichen. Unter Berücksichtigung der genannten Kriterien wurde eine Skalenversion mit 12 Items (je 6 pro Subskala) für die endgültige FWSEM-Skala ausgewählt, die in beiden Gesamtstichproben (Stichprobe 6 und Stichprobe 7) gute Ergebnisse erzielte.

Diese Resultate werden nun der Einfachheit halber im Folgenden gemeinsam mit den Befunden aus Studie 8 präsentiert, um so einen Gesamteindruck zur Güte der neu entwickelten FWSEM-Skala zu vermitteln.

### 3.5.7.2 Itemkennwerte der FWSEM-Skala

In Tabelle 11 sind die einzelnen Itemformulierungen der FWSEM-Skala (getrennt nach feindseligem und wohlwollendem Sexismus) gemeinsam mit ihren Mittelwerten und Standardabweichungen in drei verschiedenen Stichproben aufgeführt.

Tabelle 11: Mittelwerte und Standardabweichungen der FWSEM-Items in drei Stichproben

Itemkürzel	Mittelwerte			Standardabweichungen		
	S6	S7	S8	S6	S7	S8
FS-Subskala						
<i>Rabenmütter</i>	2.15	2.00	1.77	1.21	1.17	0.97
<i>Schlechtes Gewissen</i>	2.20	2.14	1.84	1.26	1.24	1.04
<i>Zu wenig Aufmerksamkeit</i>	2.71	2.59	2.28	1.21	1.23	1.10
<i>Bedürfnisse vernachlässigen</i>	2.50	2.52	2.15	1.10	1.23	1.00
<i>Selbstverwirklichung</i>	2.56	2.15	1.96	1.33	1.46	1.09
<i>Berufliche Anerkennung</i>	2.30	2.17	1.77	1.16	1.30	0.90
WS-Subskala						
<i>Spüren, was Kinder brauchen</i>	2.92	3.12	2.47	1.25	1.53	1.11
<i>Größere Fürsorglichkeit</i>	2.82	2.99	2.59	1.24	1.49	1.09
<i>Auf Bedürfnisse eingehen</i>	2.91	2.92	2.69	1.17	1.48	1.12
<i>Sanft und Warmherzig</i>	2.61	2.66	2.26	1.04	1.38	1.09
<i>Hausarbeit und Kinderbetreuung</i>	2.96	2.89	2.83	1.41	1.71	1.15
<i>Blick Hausarbeiten</i>	2.76	2.90	2.88	1.52	1.60	1.15

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. S6 = Studie 6. S7 = Studie 7. S8 = Studie 8. Höhere Werte zeigen eine sexistischere Einstellung an (S6 und S7: Skala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“, S8: recodierte Skala von 1 „stimme überhaupt nicht zu“ bis 4 „stimme voll und ganz zu“).

*Mittelwerte.* In Studie 6 und 7 bewegen sich sowohl für die FS-Subskala als auch für die WS-Subskala alle Mittelwerte unterhalb der Skalenmitte von 3.5 (sechsstufige Skala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“). Die Befragten tendierten im Durchschnitt somit eher dazu, den Items nicht zuzustimmen. Den niedrigsten Mittelwert besitzt das Item *Rabenmütter* aus der FS-Subskala ( $M = 2.00$ ,  $SD = 1.17$ ), den höchsten das Item *Spüren, was Kinder brauchen* aus der WS-Subskala ( $M = 3.12$ ,  $SD = 1.53$ ) in Stichprobe 7.

In Studie 8 (Bevölkerungsstichprobe) liegen alle Werte des feindseligen Sexismus unter der Skalenmitte von 2.5 (vierstufige Skala) und somit eher im verneinenden Bereich ( $1.77 \leq M \leq 2.28$ ;  $0.90 \leq SD \leq 1.10$ ). Für den wohlwollenden Sexismus rangieren die Mittelwerte in einem Bereich von  $2.26 \leq M \leq 2.88$  ( $1.09 \leq SD \leq 1.15$ ). Insgesamt 4 Items

weisen dabei einen Mittelwert größer 2.5 auf (*Größere Fürsorglichkeit, Auf Bedürfnisse eingehen, Hausarbeit und Kinderbetreuung* und *Blick Hausarbeiten*) (siehe Tabelle 11). Diese unterschiedlichen Ergebnisse für die beiden Subskalen bestätigen bisherige Beobachtungen, dass offene Vorurteile gegenüber Frauen, wie sie der feindselige Sexismus erfasst, als sozial unerwünscht gelten und daher auch seltener geäußert werden (siehe Unterkapitel 1.1 und 2.4.2.3) (vgl. Campbell, Schellenberg, & Senn, 1997, S. 89; Eckes & Six-Materna, 1998, S. 226 ff.; Swim et al., 2010, S. 138). Der wohlwollende Sexismus hingegen stellt eine verdeckte, subtile Form des Sexismus dar, wird dementsprechend häufig auch nicht als solcher erkannt und eher offen mitgeteilt (vgl. Glick & Fiske, 2001a). (Für einen Vergleich der Mittelwerte der weiblichen und männlichen Befragten siehe Unterkapitel 3.5.7.9).

*Schiefe und Kurtosis.* Die Verteilungen der Skalenwerte der FWSEM-Skalen sind in Abbildung 4 dargestellt, die exakten Angaben für die Schiefen und die Kurtoses können Tabelle 12 entnommen werden. Es fällt auf, dass die FS-Subskala tendenziell eine links-gipflige Verteilung besitzt und somit leicht rechtsschief ist. Dies spiegelt sich auch im Zusammenhang mit den extremen Antworttendenzen einiger Befragter wider. So besaßen relativ viele Personen ( $60 \triangleq 16.9\%$ ) in Stichprobe 8 einen Gesamtscore von 1.00 für die FS-Subskala, d.h. diese Befragten stimmten durchgängig allen Items „überhaupt nicht zu“ und sind daher berufstätigen Müttern gegenüber in keiner Weise feindselig-sexistisch eingestellt. Nur sehr wenige InterviewteilnehmerInnen ( $6 \triangleq 1.69\%$ ) äußerten sich hingegen extrem feindselig-sexistisch und stimmten durchgängig allen Aussagen „voll und ganz zu“ (Gesamtscore 4.00). In Stichprobe 6 und 7, in denen eine sechsstufige Antwortskala eingesetzt wurde, ist eine ähnliche Tendenz zu erkennen. (Gesamtscore 1.00:  $20 \triangleq 9.7\%$  (Studie 6) und  $9 \triangleq 10.5\%$  (Studie 7); Gesamtscore 6.00:  $0\%$  (Studie 6) und  $1 \triangleq 1.2\%$  (Studie 7)).

Die WS-Subskala besitzt hingegen negative Werte bei der Kurtosis, die darauf hinweisen, dass die Verteilung flacher ist als eine Normalverteilung. Im Hinblick auf die Verteilung der Skalenwerte der WS-Subskala in Stichprobe 8 fällt folglich auch auf, dass alle Antwortkategorien nahezu gleich stark besetzt sind. So stimmten 18 Befragte ( $\triangleq 4.8\%$ ) dem Item durchgängig „überhaupt nicht zu“ (Gesamtscore von 1.00) und 29 Personen ( $\triangleq 7.8\%$ ) stimmten hingegen bei allem sechs Items der Skala „voll und ganz zu“ (Gesamtscore von 4.00). In Bezug auf Stichprobe 6 und 7 ergibt sich allerdings ein anderes Bild, wenn man sich die Randverteilungen im positiven Skalenbereich anschaut. So fällt auf, dass sich in beiden Stichproben (im Gegensatz zur Bevölkerungsstichprobe in Studie 8) nur eine einzige Person durchgängig wohlwollend-sexistisch äußerte und folglich allen Items „voll und ganz zustimmte“ (Gesamtscore 6.00:  $0\%$  (Studie 6) und  $1 \triangleq 1.2\%$  (Studie 7)).

Im Hinblick auf die Abweichung von der Normalverteilung lassen sich für beide Subskalen folgende Aussagen treffen: Der Kolmogorov-Smirnov-Test ergibt, dass die Abweichungen überwiegend zu vernachlässigen sind. Signifikante Abweichungen liegen nur in drei Fällen in der Bevölkerungsstichprobe (Studie 8) vor: für die FS-Subskala ( $Z = 2.27$ ,  $p$



< .001) und die WS-Subskala ( $Z = 1.53, p = .02$ ) in der Gesamtstichprobe sowie für die FS-Subskala in der Teilstichprobe der Frauen ( $Z = 1.95, p < .001$ ). Für Studie 6 und 7 kann durchgängig (für alle drei FWSEM-Skalen und alle Teilstichproben) von einer Normalverteilung der Skalenwerte ausgegangen werden. Gleiches gilt für die männliche Teilstichprobe in Studie 8.

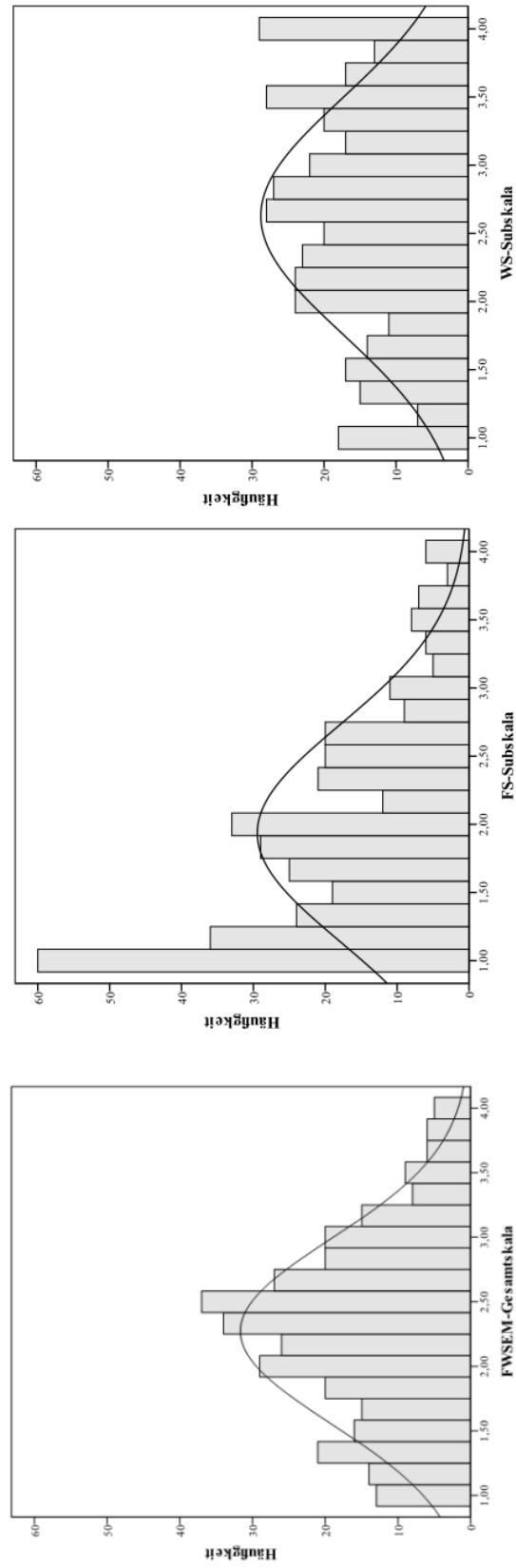


Abbildung 4. Häufigkeitsverteilungen der FWSEM-Skalen in Stichprobe 8.

Tabelle 12: Mittelwerte, Standardabweichungen, Schiefen und Kurtoses der FWSEM-Skalen in drei Stichproben

FWSEM-Skalen	Mittelwerte			Standardabweichungen			Schiefen		Kurtoses			
	S6	S7	S8	S6	S7	S8	S6	S7	S6	S7	S8	
Gesamtskala	2.62	2.60	2.27	0.88	0.96	0.72	-0.02	0.56	0.16	-0.64	0.84	-0.51
FS-Subskala	2.41	2.26	1.94	1.02	1.04	0.80	0.56	1.09	0.70	-0.16	1.61	-0.31
WS-Subskala	2.83	2.91	2.63	1.01	1.23	0.86	-0.05	0.30	-0.15	-0.60	-0.67	-0.96
	Gesamtstichprobe											
	Frauen											
Gesamtskala	2.60	2.55	2.24	0.87	1.11	0.75	-0.08	0.76	0.22	-0.62	0.51	-0.68
FS-Subskala	2.36	2.20	1.90	1.02	1.17	0.80	0.64	1.31	0.73	-0.04	1.86	-0.34
WS-Subskala	2.84	2.90	2.60	1.01	1.34	0.89	-0.10	0.37	-0.13	-0.58	-0.79	-1.03
	Männer											
Gesamtskala	2.67	2.71	2.33	0.92	0.64	0.65	0.07	-0.49	0.09	-0.68	-0.15	-0.06
FS-Subskala	2.52	2.39	1.99	1.02	0.83	0.80	0.40	0.32	0.67	-0.28	-0.49	-0.23
WS-Subskala	2.80	2.98	2.68	1.03	1.06	0.82	0.08	0.07	-0.16	-0.59	-0.63	-0.84

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. S6 = Studie 6. S7 = Studie 7. S8 = Studie 8. Höhere Mittelwerte zeigen eine sexistischere Einstellung an (S6 und S7: Skala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“, S8: recodierte Skala von 1 „stimme überhaupt nicht zu“ bis 4 „stimme voll und ganz zu“).

### 3.5.7.3 Reliabilitäten und Itemwert-Skalenwert-Korrelationen der FWSEM-Skalen

*Cronbachs Alpha.* Die Werte der internen Konsistenz sind jeweils für die Gesamtskala und die beiden Subskalen in Tabelle 13 zusammengestellt. Die Alpha Koeffizienten erreichen in allen drei Gesamtstichproben (Studie 6, 7 und 8) sowohl für die FWSEM-Gesamtskala als auch beide Subskalen ein sehr zufriedenstellendes Niveau. Den niedrigsten Wert erzielt Cronbachs Alpha für die WS-Subskala in Studie 8 mit einem Wert von  $\alpha = .87$ , den höchsten für die FS-Subskala in Studie 6 mit einem Wert von  $\alpha = .92$ . Auch in den weiblichen und männlichen Teilstichproben sind die Werte der internen Konsistenz durchweg als gut zu bezeichnen.

Tabelle 13: Reliabilitäten (Cronbachs Alpha) der FWSEM-Gesamtskala und der beiden Subskalen in drei Stichproben

FWSEM-Skalen	Studie		
	S6	S7	S8
	Gesamtstichprobe		
Gesamtskala	.91 (206)	.89 (86)	.89 (341)
FS-Subskala	.92 (206)	.90 (86)	.88 (354)
WS-Subskala	.88 (207)	.89 (88)	.87 (374)
	Frauen		
Gesamtskala	.90 (145)	.93 (52)	.91 (213)
FS-Subskala	.92 (145)	.92 (52)	.89 (217)
WS-Subskala	.88 (145)	.91 (52)	.88 (232)
	Männer		
Gesamtskala	.92 (61)	.73 (32)	.84 (128)
FS-Subskala	.91 (61)	.83 (32)	.86 (137)
WS-Subskala	.89 (62)	.84 (34)	.84 (142)

*Anmerkungen.* FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. S6 = Studie 6. S7 = Studie 7. S8 = Studie 8. Neben den Reliabilitäten sind in Klammern die jeweiligen Fallzahlen angegeben.

*Korrigierte Itemwert-Skalenwert-Korrelationen.* Insgesamt liegen die Trennschärfen für beide Subskalen in allen drei Stichproben (S6, S7 und S8) in einem sehr guten Bereich (siehe Tabelle 14). In Studie 6 ergeben sich Werte zwischen  $.72 \leq r_{it} \leq .83$  für den feindseligen Sexismus und Werte zwischen  $.64 \leq r_{it} \leq .74$  für den wohlwollenden Sexismus. In Studie 7 liegen die korrigierten Itemwert-Skalenwert-Korrelationen für die FS-Subskala in einem Bereich zwischen  $.67 \leq r_{it} \leq .80$  und für die WS-Subskala zwischen  $.54 \leq r_{it} \leq .79$ . Auch in der Bevölkerungsstichprobe (S8) fallen die Ergebnisse der Trennschärfe-Koeffizienten beider Subskalen zufriedenstellend aus (FS-Subskala:  $.64 \leq r_{it} \leq .77$ ; WS-Subskala:  $.57 \leq r_{it} \leq .74$ ). In der Teilstichprobe der Frauen weisen die Itemwert-Skalenwert-Korre-

lationen für beide Subskalen in allen drei Studien sehr gute Werte auf. In der Teilstichprobe der Männer besitzen lediglich in Studie 7 zwei Items eine Trennschärfe  $< .40$  (*schlechtes Gewissen* und *Blick Hausarbeiten*  $r_{it} = .38$ ), alle übrigen Trennschärfen sind als gut zu bezeichnen.

Tabelle 14: Korrigierte Itemwert-Skalenwert-Korrelationen der FWSEM-Items in drei Stichproben

Itemkürzel	Gesamtstichprobe						Frauen			Männer		
	S6	S7	S8	S6	S7	S8	S6	S7	S8	S6	S7	S8
FS-Subskala												
<i>Rabenmütter</i>	.78	.80	.71	.78	.84	.73	.79	.68	.67	.79	.68	.67
<i>Schlechtes Gewissen</i>	.77	.72	.77	.79	.85	.78	.74	.38	.75	.74	.38	.75
<i>Zu wenig Aufmerksamkeit</i>	.74	.69	.67	.74	.77	.72	.76	.50	.59	.76	.50	.59
<i>Bedürfnisse vernachlässigen</i>	.72	.67	.67	.77	.68	.67	.59	.62	.66	.59	.62	.66
<i>Selbstverwirklichung</i>	.74	.73	.66	.72	.73	.68	.80	.76	.64	.80	.76	.64
<i>Berufliche Anerkennung</i>	.83	.80	.64	.84	.83	.65	.81	.72	.61	.81	.72	.61
WS-Subskala												
<i>Spüren, was Kinder brauchen</i>	.72	.78	.74	.72	.85	.76	.72	.68	.71	.72	.68	.71
<i>Größere Fürsorglichkeit</i>	.70	.70	.71	.69	.82	.77	.71	.47	.61	.71	.47	.61
<i>Auf Bedürfnisse eingehen</i>	.66	.74	.68	.65	.83	.71	.69	.62	.63	.69	.62	.63
<i>Sanft und Warmherzig</i>	.72	.79	.70	.70	.79	.71	.78	.81	.68	.78	.81	.68
<i>Hausarbeit und Kinderbetreuung</i>	.74	.69	.59	.75	.67	.58	.74	.74	.61	.74	.74	.61
<i>Blick Hausarbeiten</i>	.64	.54	.57	.63	.61	.64	.66	.38	.49	.66	.38	.49

Anmerkungen. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. S6 = Studie 6, S7 = Studie 7, S8 = Studie 8. Die Trennschärfen wurden jeweils für beide Subskalen getrennt berechnet.

### 3.5.7.4 Explorative Faktorenanalyse der FWSEM-Skala

Um die Dimensionalität der FWSEM-Skala zu überprüfen, wurde mit SPSS eine Hauptachsenanalyse mit anschließender Promax-Rotation durchgeführt.<sup>54</sup> Für Stichprobe 6 ergaben sich nach dem Kaiser-Kriterium zwei Faktoren mit einem Eigenwert größer Eins (diese Anzahl wurde auch durch den Scree-Test bestätigt), von denen der erste 47.62% und der zweite 14.08% der Varianz erklären. Äquivalent zur inhaltlichen Einteilung der Items fällt auch die empirische Analyse aus. So laden alle Items, die der Dimension „Feindseliger Sexismus“ zuzuordnen sind, eindeutig ( $> .70$ ) auf dem ersten und die Items des „Wohlwollenden Sexismus“ eindeutig ( $> .56$ ) auf dem zweiten Faktor (siehe Tabelle 15).

Tabelle 15: Faktorladungen der FWSEM-Items in drei Stichproben

Itemkürzel	S6		S7		S8	
	FS	WS	FS	WS	FS	WS
<i>Schlechtes Gewissen</i>	<b>.85</b>		<b>.75</b>		<b>.86</b>	
<i>Rabenmütter</i>	<b>.89</b>	-.11	<b>.88</b>		<b>.83</b>	
<i>Berufliche Anerkennung</i>	<b>.88</b>		<b>.88</b>		<b>.75</b>	
<i>Selbstverwirklichung</i>	<b>.77</b>		<b>.82</b>	-.11	<b>.71</b>	
<i>Bedürfnisse vernachlässigen</i>	<b>.70</b>		<b>.67</b>		<b>.66</b>	.12
<i>Zu wenig Aufmerksamkeit</i>	<b>.76</b>		<b>.69</b>	.10	<b>.65</b>	.14
<i>Spüren, was Kinder brauchen</i>		<b>.83</b>		<b>.83</b>		<b>.84</b>
<i>Größere Fürsorglichkeit</i>		<b>.81</b>	-.14	<b>.82</b>		<b>.79</b>
<i>Sanft und Warmherzig</i>		<b>.74</b>		<b>.87</b>		<b>.75</b>
<i>Auf Bedürfnisse eingehen</i>	.31	<b>.56</b>		<b>.79</b>		<b>.71</b>
<i>Hausarbeit und Kinderbetreuung</i>		<b>.79</b>	.19	<b>.65</b>		<b>.67</b>
<i>Blick Hausarbeiten</i>		<b>.68</b>		<b>.56</b>		<b>.62</b>
Eigenwert	5.71	1.69	5.26	2.02	5.40	2.10
Erklärte Varianz	47.62	14.08	43.80	16.86	41.40	13.86

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. Fett markierte Werte geben die Hauptladungen der einzelnen Items an. Faktorladungen  $< .10$  sind in der Tabelle nicht angeführt, um diese übersichtlicher zu gestalten.

<sup>54</sup> Es wurde eine Hauptachsenanalyse mit Promax-Rotation ausgewählt, da sie oblique ist und somit eine Korrelation der Faktoren erlaubt. In Anlehnung an die Ergebnisse von Glick und Fiske (1996) sowie von Eckes und Six-Materna (1999) wurde davon ausgegangen, dass es sich beim „feindseligen Sexismus“ und beim „wohlwollenden Sexismus“ um zwei distinkte, aber dennoch miteinander korrelierte Facetten von Sexismus handelt. Die Wahl der Rotationsmethode sollte folglich eine Verzerrung der Ergebnisse verhindern.

Auch in Stichprobe 7 und 8 zeigt sich eine zweifaktorielle Struktur. Erneut laden alle Items des feindseligen Sexismus gemeinsam auf dem ersten und die des wohlwollenden Sexismus auf dem zweiten Faktor (siehe Tabelle 15). Dabei fallen alle Ladungen durchweg sehr rein aus, d.h. sie besitzen hohe Hauptladungen auf ihrem eigenen Faktor ( $> .56$ ) und nur geringe Nebenladungen ( $< .19$ ) auf dem anderen Faktor. Das einzige Item, welches eine etwas höhere Nebenladung mit einem Wert von  $.31$  besitzt, ist das Item *Auf Bedürfnisse eingehen*. Dies trifft jedoch nur auf Stichprobe 6 zu. In Stichprobe 7 und 8 besitzt auch dieses Item eine eindeutige, reine Faktorladung.

Eine Analyse getrennt für die weibliche und die männliche Teilstichprobe führte in allen drei Studien (Studie 6 bis Studie 8) zu den gleichen Ergebnissen: Auch hier zeigte sich eine Zwei-Dimensionalität der Skala und eine Eindeutigkeit der Itemladungen (siehe beispielhaft die Ergebnisse aus Stichprobe 8 in Tabelle 16).

Tabelle 16: Faktorladungen der FWSEM-Items für die Gesamtstichprobe und die beiden Teilstichproben in Studie 8

Itemkürzel	Frauen		Männer	
	FS	WS	FS	WS
<i>Schlechtes Gewissen</i>	<b>.87</b>		<b>.83</b>	
<i>Rabenmütter</i>	<b>.81</b>		<b>.82</b>	-.15
<i>Berufliche Anerkennung</i>	<b>.77</b>	-.11	<b>.73</b>	
<i>Selbstverwirklichung</i>	<b>.76</b>		<b>.67</b>	
<i>Bedürfnisse vernachlässigen</i>	<b>.61</b>	.16	<b>.70</b>	.10
<i>Zu wenig Aufmerksamkeit</i>	<b>.67</b>	.16	<b>.61</b>	
<i>Spüren, was Kinder brauchen</i>		<b>.86</b>		<b>.78</b>
<i>Größere Fürsorglichkeit</i>		<b>.86</b>		<b>.66</b>
<i>Sanft und Warmherzig</i>	.16	<b>.70</b>		<b>.78</b>
<i>Auf Bedürfnisse eingehen</i>	.16	<b>.66</b>		<b>.72</b>
<i>Hausarbeit und Kinderbetreuung</i>		<b>.68</b>		<b>.67</b>
<i>Blick Hausarbeiten</i>		<b>.71</b>		<b>.54</b>
Eigenwert	6.01	1.79	4.43	2.69
Erklärte Varianz	46.77	11.41	32.97	18.57

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. Fett markierte Werte geben die Hauptladungen der einzelnen Items an. Faktorladungen  $< .10$  sind in der Tabelle nicht angeführt, um diese übersichtlicher zu gestalten.

### 3.5.7.5 Konfirmatorische Faktorenanalyse der FWSEM-Skala

Mit den Daten aus Stichprobe 6, 7 und 8 wurde neben einer explorativen auch eine konfirmatorische Faktorenanalyse mit AMOS 20.0 durchgeführt, um zu überprüfen, welches

Modell die beste Anpassung an die Daten liefert. Die gewählten Modelle stützten sich dabei auf die theoretischen Annahmen sowie auf die Resultate der explorativen Faktorenanalyse in SPSS. So wurde neben einem „Ein-Faktormodell“ ein „Zwei-Faktoren-Modell“ getestet. Beim Ein-Faktormodell wurde von einem einzigen übergeordneten Sexismusfaktor ausgegangen, auf dem alle Items laden. Beim Zwei-Faktoren-Modell wurde hingegen zwischen einem feindseligen und einem wohlwollenden Sexismusfaktor unterschieden, denen jeweils die entsprechenden Items der Subskalen zugeordnet wurden.

Um die Güte der Modelle zu beurteilen, wurden die folgenden Fit-Indizes herangezogen: Verhältnis von  $\chi^2$ -Wert zu den Freiheitsgraden ( $\chi^2/df$ ), Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA),  $p$ -value of Close fit index ( $p$ -Close), Bayesian Information Criterion (BIC) und Consistent Akaike Information Criterion (CAIC). Dabei spricht man von einer akzeptablen Anpassung an die Daten, wenn  $\chi^2/df < 3$ , CFI  $> .95$ , RMSEA  $< .08$  (vgl. Moosbrugger & Schermelleh-Engel, 2008, S. 319) und  $p$ -close  $> .50$  ist. BIC und CAIC erlauben einen Modellvergleich. Kleinere Werte deuten darauf hin, dass das Modell die Daten besser vorhersagen kann und somit zu bevorzugen ist

Schaut man sich die Fitmaße beider Modelle an, so zeigt sich, dass das Zwei-Faktoren-Modell in allen drei Studien eine bessere Anpassung an die Daten liefert als das Ein-Faktormodell (siehe Tabelle 17). Abbildung 5 sind die standardisierten Ladungen in Stichprobe 8 zu entnehmen. Die Korrelation beider latenten Dimensionen liegen bei  $r = .46$ . Folglich stellen feindseliger und wohlwollender Sexismus in der Tat zwei distinkte, aber dennoch miteinander korrelierte Facetten von Sexismus dar (für eine genauere Analyse des Zusammenhangs von feindseligem und wohlwollendem Sexismus siehe Unterkapitel 3.5.7.8).

Da jedoch auch das Zwei-Faktoren-Modell noch nicht durchgängig zufriedenstellende Fitmaße besaß, wurden drei Residuenkorrelationen zugelassen, die sich durch thematisch ähnliche Schwerpunktsetzungen der Items erklären lassen (siehe modifiziertes Zwei-Faktoren-Modell in Abbildung 6). Im Falle der FS-Subskala handelt es sich um drei Items, die Karrierefrauen als schlechte Mütter hinstellen (*Schlechtes Gewissen*, *Zu wenig Aufmerksamkeit* und *Bedürfnisse vernachlässigen*), bezüglich der WS-Subskala dreht es sich um die beiden Items, die sich auf die klassischen Hausarbeiten beziehen (*Blick Hausarbeiten* und *Hausarbeit und Kinderbetreuung*). Abbildung 6 zeigt die standardisierten Faktorladungen, die erklärte Varianz sowie die jeweiligen Gütemaße für das modifizierte Zwei-Faktoren-Modell der FWSEM-Skala in Studie 8. Es weist nun gute Anpassungsmaße auf.



Tabelle 17: Fitmaße des Ein- und des Zwei-Faktoren-Modells in drei Stichproben

Fitmaße	Studie		
	S6	S7	S8
	Ein-Faktor-Modell		
$\chi^2$	510.75	295.90	815.59
df	54	54	54
$\chi^2/df$	9.46	5.48	15.10
CFI	.70	.61	.65
RMSEA	.20	.23	.19
<i>p</i> -Close	.00	.00	.00
BIC	638.62 (415.57)	403.63 (350.11)	958.78 (465.36)
CAIC	662.62 (493.57)	427.63 (428.11)	982.78 (543.36)
	Zwei-Faktoren-Modell		
$\chi^2$	167.06	84.08	160.36
df	53	53	53
$\chi^2/df$	3.15	1.59	3.03
CFI	.93	.95	.95
RMSEA	.10	.08	.07
<i>p</i> -Close	.00	.07	.00
BIC	300.26 (415.57)	196.30 (350.11)	309.51 (465.36)
CAIC	325.26 (493.57)	221.30 (428.11)	334.51 (543.36)

Anmerkungen.  $\chi^2/df$  = Verhältnis vom Chi-Quadrat-Wert zu den Freiheitsgraden. CFI = Comparative Fit Index. RMSEA = Root Mean Square Error of Approximation. *p*-Close = *p*-value of Close fit index. BIC = Bayesian Information Criterion. CAIC = Consistent Akaike Information Criterion. Bei BIC und CAIC sind jeweils in Klammern die Werte für das saturierte Modell angegeben.

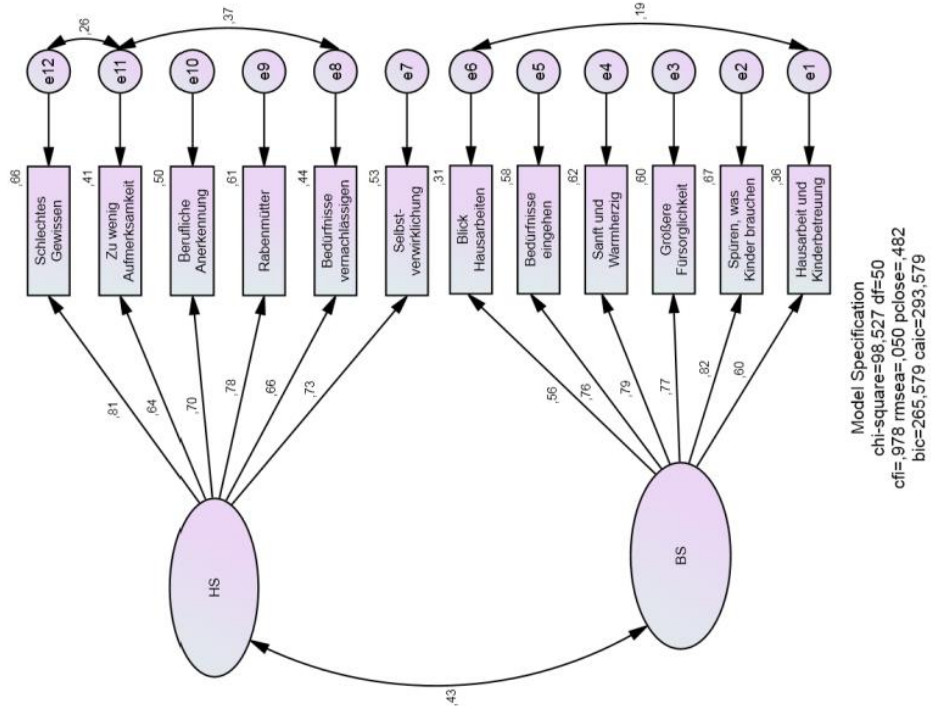


Abbildung 6. Modifiziertes Zwei-Faktoren-Modell der FWSEM-Skala in Stichprobe 8 mit den entsprechenden Gütemaßen.

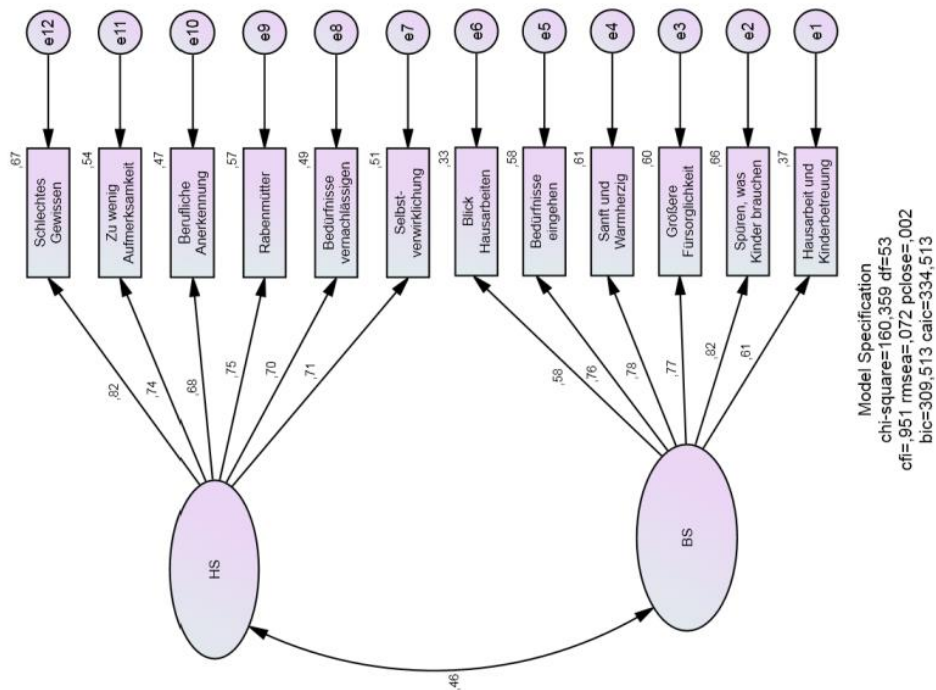


Abbildung 5. Ursprüngliches Zwei-Faktoren-Modell der FWSEM-Skala in Stichprobe 8 mit den entsprechenden Gütemaßen.

### 3.5.7.6 Explorative Faktorenanalyse über alle Sexismus-Items in Stichprobe 6

Abgesehen von einer Faktorenanalyse mit den 12 Items der FWSEM-Skala wurde eine weitere explorative Faktorenanalyse (Hauptachsenanalyse mit anschließender Promax-Rotation) mit allen 63 Sexismus-Items, die in Studie 6 erhoben wurden, berechnet. Auf diese Weise sollte sichergestellt werden, dass die FWSEM-Skala auch tatsächlich etwas Neues abbildet und ihre einzelnen Items nicht nur auf den Faktoren bereits etablierten Skalen mitladen, da dies die neuen Fragen hätte redundant werden lassen. Folglich sollte sowohl für die feindselige Sexismus-Subskala als auch die wohlwollende Sexismus-Subskala ein eigener neuer Faktor bestehen. Bezugnehmend auf die Theorien und bisherige empirische Ergebnisse wurden neben diesen beiden noch sieben weitere Faktoren erwartet: jeweils ein Faktor für die Skala zur Erfassung von Vorurteilen gegenüber Frauen auf der Basis des Konzepts des modernen Sexismus (MSS), für die Skala zur Erfassung der Akzeptanz moderner Mythen über sexuelle Aggression (AMMSA) sowie für die Geschlechterrollen-Batterie des International Social Survey Programme-Modul 2002 Familie in Deutschland (GRB) sowie vier Faktoren für die Skala zur Erfassung des ambivalenten Sexismus (einen hostilen Sexismus-Faktor (HS) und drei Faktoren für den benevolenten Sexismus: protektiver Paternalismus (PP), komplementäre Geschlechterdifferenzierung (KGD) sowie die intime Heterosexualität (IHS).

Die Faktorenanalyse ergab 11 Faktoren mit einem Eigenwert größer Eins, die gemeinsam 67.68% der Varianz erklären (siehe Anhang 8). Während die Faktorenstruktur für die MSS, die GRB und den protektiven Paternalismus nicht den Erwartungen entspricht<sup>55</sup>, haben sich für die FS-Subskala und die WS-Subskala hypothesenkonform zwei eigenständige Faktoren herauskristallisiert. Dabei laden alle Items des feindseligen Sexismus eindeutig ( $> .62$ ) auf dem ersten und die Items des wohlwollenden Sexismus eindeutig ( $> .52$ ) auf dem sechsten Faktor. Lediglich das Item *Auf Bedürfnisse eingehen* („Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.“) besitzt eine etwas höhere Nebenladung von  $.36$  auf dem ersten Faktor (Feindseliger Sexismus), alle anderen Items der FWSEM-Skala weisen nur geringe Nebenladungen ( $< .24$ ) auf.

Anzumerken ist in diesem Zusammenhang, dass die drei Items der Geschlechterrollen-Batterie des ISSP, die sich mit den Konsequenzen der Erwerbstätigkeit der Mutter für die Familie befassen, ebenfalls auf dem Feindseligen Sexismus-Faktor laden. Hier scheint es eine inhaltliche Überschneidung zwischen der neu entwickelten FS-Subskala und einem Teil der GRB-Items zu geben. Wie bereits in Unterkapitel 2.4.2.4 und 3.2.1 erwähnt, ist die GRB die einzige der etablierten Skalen, die sich explizit auf berufstätige Mütter und

---

<sup>55</sup> So zerfällt der Moderne Sexismus (MSS) in zwei Faktoren, der Protektive Paternalismus (PP) teilt sich ebenfalls und lädt teilweise mit Items der Intimen Heterosexualität (IHS) sowie mit Items der Geschlechterrollen-Batterie (GRB). Die GRB wiederum spaltet sich in vier Faktoren auf: zwei eigene Faktoren sowie einen Faktor gemeinsam mit Items der IHS und einen Faktor mit der Feindseligen Sexismus-Skala (FS).

die negativen Konsequenzen der Berufstätigkeit für die Familie bezieht. Vollkommen deckungsgleich sind beide Skalen dennoch nicht, in diesem Fall hätten alle Items gemeinsam auf einem Faktor laden müssen

Der Wohlwollende Sexismus, der von einem durchweg positiven Frauen- bzw. Mutterbild bestimmt ist, scheint hingegen in den bisherigen Sexismus-Skalen überhaupt noch nicht abgebildet worden zu sein. So befassen sich zwar einige Items der GRB auch mit der Aufgabenteilung von Familien- und Berufsarbeit, allerdings wird hier nicht die besondere Eignung der Mutter für die traditionell weiblichen Tätigkeitsbereiche hervorgehoben, sondern eine normative Zuschreibung vorgenommen („Frauen sollten dies und Männer jenes tun“) ohne eine positive Begründung abzugeben. Folglich besteht der wohlwollende Sexismus-Faktor auch nur aus den Items der WS-Subskala.

### 3.5.7.7 Explorative Faktorenanalyse über alle Sexismus-Items in Stichprobe 8

Wie bereits in Studie 6 so wurde auch in Studie 8 eine explorative Faktorenanalyse mit allen 25 Sexismus-Items (aus insgesamt 4 Skalen) durchgeführt. Zieht man die verschiedenen Sexismus-Theorien, auf deren Basis die einzelnen Skalen entwickelt wurden, heran, so sind fünf Faktoren zu erwarten: jeweils einen Faktor für den feindseligen Sexismus (FS), den wohlwollenden Sexismus (WS), den traditionellen Sexismus (TS), den modernen Sexismus (MS) sowie für die Akzeptanz moderner Mythen über sexuelle Aggression (AMMSA).

Die Hauptachsenanalyse mit anschließender Promax-Rotation ergab 6 Faktoren mit einem Eigenwert größer Eins (erklärte Gesamtvarianz = 62.36%) (siehe Tabelle 18). Für den feindseligen, den wohlwollenden und den modernen Sexismus ergibt sich der Hypothese entsprechend jeweils ein Faktor. Entgegen den Erwartungen zerfällt die AMMSA-Skala allerdings in zwei eigenständige Faktoren sowie einen weiteren gemeinsamen Faktor mit dem traditionellen Sexismus.

Zu den Itemladungen ist anzumerken, dass sowohl die Items des feindseligen als auch die des wohlwollenden Sexismus jeweils mit Hauptladungen  $> .53$  und Nebenladungen  $< .23$  als eindeutig einzustufen sind. Lediglich das Item *Hausarbeit und Kinderbetreuung* besitzt mit einem Wert von  $.35$  eine etwas höhere Nebenladung auf einem der AMMSA-Faktoren. Somit konnten die Ergebnisse aus Studie 6 repliziert werden: Auch in Studie 8 zeigt sich, dass die beiden Subskalen der FWSEM-Skala im Rahmen einer simultanen Analyse aller Sexismus-Skalen jeweils einen eigenständigen Faktor bilden und somit inhaltlich etwas Neues repräsentieren.

Tabelle 18: Faktorladungen aller Sexismus-Items in Stichprobe 8

Items	Faktor					
	FS	WS	AMMSA	AMMSA	AMMSA + TS	MS
Frauen denken meiner Meinung nach nur an sich selbst, wenn sie ein Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um berufliche Anerkennung zu bekommen.	.84					
Frauen sollten ein schlechtes Gewissen haben, wenn sie aus Karrieregründen ihr Kind mit einem halben Jahr tagsüber zur Betreuung zu einer Tagesmutter geben.	.83	-.11				
Frauen, die aus Karrieregründen ihr Kind mit einem halben Jahr in eine Kindertagesstätte geben, sind für mich Rabenmütter.	.76				.11	
Ich finde Frauen egoistisch, die bereits ein halbes Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um sich selbst zu verwirklichen.	.69			-.11	.14	
Viele Karrierefrauen denken meiner Meinung nach nur an ihr berufliches Vorankommen und vernachlässigen die Bedürfnisse ihrer Kinder.	.64	.19	-.13	.12		
Ich denke, dass Karrierefrauen ihren Kindern häufig zu wenig Aufmerksamkeit schenken.	.64	.16	.22	-.17	.11	
Eine Mutter spürt besser als ein Vater, was die eigenen Kinder brauchen.		.88				
Eine Mutter ist vom Wesen her sanfter und warmerziger als ein Vater und daher auch besser für die Betreuung eines Kindes geeignet.		.81	-.12		.13	
Mütter besitzen eine größere Fürsorglichkeit als Väter.		.80				
Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.		.66			.11	
Frauen besitzen eine größere natürliche Begabung als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.		.54	.35	-.21	-.12	
Frauen haben von Natur aus einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.		.53	.23			

Items	FS	WS	AMMSA	AMMSA	AMMSA + TS	MS
Frauen zieren sich gerne. Das bedeutet nicht, dass sie keinen Sex wollen.			.66			
Wenn eine Frau einen Mann nach dem Discobesuch auf eine Tasse Kaffee in ihre Wohnung einlädt, dann ist sie auf ein sexuelles Abenteuer aus.	-.12		.54		.11	.13
Bei Männern ist es biologisch bedingt, dass sie von Zeit zu Zeit sexuellen Druck ablassen müssen.	.14		.47	.12		
Viele Frauen neigen dazu, eine nett gemeinte Geste zum „sexuellen Übergriff“ hochzuspielen.			-.13	.73		
Harmlose Gesten als „sexuelle Belästigung“ auszulegen ist eine beliebte Waffe im Kampf der Geschlechter.	.13		.14	.57	-.18	
Frauen bezichtigen ihre Männer häufig einer Vergewaltigung in der Ehe, um sich für eine gescheiterte Beziehung zu rächen.			.16	.49	.14	
Viele Frauen beschweren sich schon aus nichtigem Anlass über sexuelle Belästigung, nur um als emanzipiert zu gelten.			.18	.29	.19	-.11
Für eine Frau sollte es wichtiger sein, ihrem Mann bei seiner Karriere zu helfen, als selbst Karriere zu machen. (TS)	.16		-.11		.75	
Frauen sollen sich wieder mehr auf die Rolle der Ehefrau und Mutter besinnen. (TS)	.37				.55	
Wenn eine Frau mit einem Mann eine Beziehung eingeht, muss sie sich darüber im Klaren sein, dass der Mann sein Recht auf Sex einfordern wird. (AMMSA)	-.15		.24	.15	.44	
Wer als Frau so unvorsichtig ist, nachts durch „dunkle Gassen“ zu gehen, trägt eine gewisse Mitschuld an der eigenen Vergewaltigung. (AMMSA)			.23	.11	.39	
Die jetzige Beschäftigungspolitik benachteiligt die Frauen.						.91
Die Diskriminierung von Frauen ist in Deutschland immer noch ein Problem.						.52
Eigenwert	7.27	2.61	2.02	1.47	1.18	1.05
Erklärte Varianz	29.06	10.44	8.07	5.87	4.71	4.21

Anmerkungen. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. AMMSA = Akzeptanz moderner Mythen über sexuelle Aggression. TS = Traditioneller Sexismus. MS = Moderner Sexismus. Fett markierte Werte geben die Hauptladungen der einzelnen Items an. Faktorladungen < .10 sind in der Tabelle nicht angeführt, um sie übersichtlicher zu gestalten.

### 3.5.7.8 Beziehung zwischen der FS- und der WS-Subskala

Ausgehend von der Theorie des ambivalenten Sexismus wurden positive Korrelationen der beiden Subskalen erwartet. So stellen hostiler und benevolenter Sexismus zwei Aspekte einer Sexismus-Ideologie dar, die miteinander verbunden sind (Glick & Fiske, 1996; Nadler & Morrow, 1959). Diese Hypothese kann in allen drei Studien (S6, S7 und S8) bestätigt werden: Sowohl in der Gesamtstichprobe als auch in den beiden Teilstichproben der Frauen und der Männer zeigen sich signifikante, positive Korrelationen (siehe Tabelle 19). Eine Ausnahme bildet die männlich Teilstichprobe in Studie 7: Hier ist die Korrelation der beiden Subskalen nicht signifikant.

Vergleicht man die Höhe der Korrelationen für die weiblichen und die männlichen Teilstichproben, so zeigt sich folgendes Bild: Während in Studie 6 der Zusammenhang zwischen der FS- und der WS-Subskala bei den Männern tendenziell höher ausfällt (ohne dass der Geschlechtsunterschied jedoch statistisch bedeutsam ist), sind die Korrelationen beider Skalen in den beiden übrigen Stichproben signifikant niedriger als bei den Frauen (Studie 7:  $z = -3.00$ ,  $p < .01$ ; Studie 8:  $z = -3.34$ ,  $p < .01$ ). Diese Ergebnisse decken sich mit den Befunden zur deutsch- und englischsprachigen Variante der ASS: Auch hier liegen die Korrelationen des hostile und des benevolent Sexismus für die Teilstichprobe der Männer konsistent unter denen der Frauen. Signifikant sind diese Geschlechtsunterschiede in den Korrelationen bei Eckes und Six-Materna (1999) jedoch nur in einer von fünf Stichprobe, bei Glick und Fiske (1996) in drei von sechs Stichproben.

Insgesamt betrachtet, bewegen sich die Korrelationen für die FS- und die WS-Subskala ( $-.07 \leq r \leq .56$ ) auf einem ähnlich hohen Niveau wie diejenigen der HS- und der BS-Subskala von Glick und Fiske (1996) ( $-.12 \leq r \leq .61$ ). Der Durchschnittswert der Korrelationen betrug im Falle der ASS  $r = .51$  (vgl. Eckes und Six-Materna, 1999), für die FWSEM-Skala  $r = .43$ ).

Tabelle 19: Korrelationen zwischen der FS- und der WS-Subskala in drei Stichproben

	Studie		
	S6	S7	S8
Gesamtstichprobe	.50** (206)	.41** (86)	.45** (341)
Frauen	.46** (145)	.56** (52)	.56** (213)
Männer	.59** (61)	-.07 (32)	.25** (128)

Anmerkungen. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. S6 = Studie 6. S7 = Studie 7. S8 = Studie 8. Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*\*  $p < .01$ .

### 3.5.7.9 Geschlechtsunterschiede in den Mittelwerten

Um einen weiteren Hinweis auf die Konstruktvalidität der neu entwickelten FWSEM-Skala zu erhalten, wurden die Mittelwerte der Skalen auf Geschlechtsunterschiede hin untersucht. Folgt man den gängigen Kriterien im Bereich der Sexismusforschung, so sollten Männer stärker zum Sexismus neigen als Frauen und daher auch im Durchschnitt höhere Mittelwerte erzielen als diese (vgl. z.B. Swim, Aikin, Hall & Hunter, 1995; McHugh & Frieze, 1997). Diese Annahme bezieht sich auf Skalen, die sexistische Einstellungen gegenüber der Gesamtgruppe der Frauen erfassen. Im vorliegenden Fall handelt es sich jedoch um eine Skala, die sexistische Einstellungen in Bezug auf die Subgruppe der Mütter erfasst. Folglich wurde erwartet, dass die Geschlechtsunterschiede für die FWSEM-Skala zwar auftreten, aber nicht so deutlich ausfallen wie bei anderen Sexismus-Skalen. So ist anzunehmen, dass Frauen im Durchschnitt ihrer eigenen Geschlechtergruppe wohlgesonnener sind, als dies Männer Frauen gegenüber sind. Gleichzeitig ist es durchaus denkbar, dass Hausfrauen beispielsweise berufstätigen Müttern in ihren Lebensweisen gegenüber eher negativ eingestellt sind oder umgekehrt, so dass es hier auch zu feindseligen Einstellungen von Frauen gegenüber anderen Frauen kommen kann (siehe Unterkapitel 3.2.1).

Die Daten in Tabelle 20 lassen auf den ersten Blick erkennen, dass die Mittelwertsdifferenzen zwischen Frauen und Männer sowohl für die Gesamtskala als auch für beide Subskalen gering ausfallen. Auch der *t*-Test für unabhängige Stichproben bestätigt, dass für die FWSEM-Skalen in keiner der drei Stichproben signifikante Geschlechtsunterschiede vorliegen.<sup>56</sup> Da in Studie 8 die FS-Subskala in der Teilstichprobe der Frauen nicht normalverteilt ist, wurde zudem der nicht-parametrische Mann-Whitney-U-Test durchgeführt. Auch er ergab, dass keine Geschlechtsunterschiede im Antwortverhalten im Hinblick auf den feindseligen Sexismus bestehen ( $Z = -1.18, p = .24$ ).

---

<sup>56</sup> Bortz (2005) ist zu entnehmen, dass der *t*-Test auf eine Verletzung der Normalverteilungsannahme relativ robust reagiert und die Ergebnisse des Tests als präzise angenommen werden können, wenn die Stichprobenumfänge gleich groß sind und/oder die Varianzen des Merkmals in beiden unabhängigen Stichprobe gleich ausfallen (vgl. Bortz, 2005, S. 141). Der Levene-Test ergab, dass mit hoher Wahrscheinlichkeit davon auszugehen ist, dass im Hinblick auf den feindseligen Sexismus die Varianzen in der weiblichen und männlichen Teilstichprobe in Studie 8 als gleich angenommen werden können ( $F = 0.2, p = 0.88$ ). Folglich können auch in diesem Fall die Ergebnisse des *t*-Test herangezogen werden, um die Geschlechtsunterschiede zu untersuchen. Nichtsdestotrotz wurde zur Absicherung der Ergebnisse zusätzlich der nicht-parametrische Mann-Whitney-U-Test durchgeführt.



Tabelle 20: Mittelwerte der FWSEM-Skalen für Frauen und Männer in drei Stichproben

Studie	Geschlecht		<i>t</i>	<i>df</i>
	w	m		
Gesamtskala				
S6	2.60 (.87)	2.67 (.92)	0.50	204
S7	2.55 (1.11)	2.71 (.64)	0.84	81.73
S8	2.24 (.75)	2.33 (.65)	1.16	296.26
FS-Subskala				
S6	2.36 (1.02)	2.52 (1.02)	1.01	204
S7	2.20 (1.17)	2.39 (.83)	0.78	82
S8	1.90 (.80)	1.99 (.80)	1.09	352
WS-Subskala				
S6	2.84 (1.01)	2.80 (1.03)	0.24	205
S7	2.90 (1.34)	2.98 (1.1)	0.30	84
S8	2.59 (.89)	2.68 (.82)	0.97	372

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. w = weiblich, m = männlich. *t* = *t*-Wert. *df* = Freiheitsgrad. Höhere Mittelwerte zeigen eine sexistischere Einstellung an (S6 und S7: Skala von 1 „stimme überhaupt nicht zu“ bis 6 „stimme voll und ganz zu“, S8: recodierte Skala von 1 „stimme überhaupt nicht zu“ bis 4 „stimme voll und ganz zu“). Neben den Mittelwerten sind in Klammern die entsprechenden Standardabweichungen angegeben.

### 3.5.7.10 Ost-/West-Unterschiede in den Mittelwerten

Neben den Geschlechtsunterschieden wurden in Stichprobe 8 die Mittelwertsunterschiede zwischen west- und ostdeutschen Befragten untersucht, um die Skala auf ihre Validität hin zu prüfen. Aufgrund der beiden unterschiedlichen Mütter(leit)bilder, die in beiden Teilen Deutschlands vermittelt wurden (siehe Unterkapitel 2.1.1.3), wurde von folgenden Hypothesen ausgegangen:

- (1) Feindseliger Sexismus: *Westdeutsche* sollten hinsichtlich des feindseligen Sexismus *höhere Mittelwerte* aufweisen als Personen aus Ostdeutschland und somit berufstätigen Mütter und der außerhäuslichen Kinderbetreuung gegenüber negativer eingestellt sein.
- (2) Benevolenter Sexismus: *Westdeutsche* sollten im Vergleich zu Ostdeutschen hinsichtlich des wohlwollenden Sexismus ebenfalls *höhere Mittelwerte* besitzen und somit die besondere Eignung der Frau für die Kinderbetreuung unterstützen.

Wie Tabelle 21 zu entnehmen ist, konnte die Hypothese in Bezug auf den feindseligen Sexismus bestätigt werden. Die Befragten aus Westdeutschland ( $M = 2.01$ ;  $SD = .85$ ) besaßen signifikant höhere Mittelwerte als die Befragten aus Ostdeutschland ( $M = 1.79$ ;

$SD = .67$ ),  $t(286.30) = 2.62$ ,  $p = .01$ , d.h. sie waren berufstätigen Müttern gegenüber feindseliger eingestellt. Hinsichtlich des benevolenten Sexismus ergaben sich ebenfalls signifikante Mittelwertsunterschiede, allerdings fielen diese nicht in der erwarteten Richtung aus. So besaßen die Ostdeutschen ( $M = 2.76$ ;  $SD = .83$ ) höhere Mittelwerte als die Westdeutschen ( $M = 2.55$ ;  $SD = .88$ ),  $t(372) = -2.19$ ,  $p = .03$ , und attestierten Frauen somit in einem stärkeren Maße eine besondere Befähigung für die traditionell weiblichen Familienaufgaben (Kinderbetreuung).

Erneut wurden die Ergebnisse des  $t$ -Tests mit Hilfe des Mann-Whitney-U-Tests abgesichert, da sowohl die FS- als auch die WS-Subskala in der Gesamtstichprobe in Studie 8 nicht normalverteilt waren.<sup>57</sup> Auch der nicht-parametrische Test ergab, dass signifikante Unterschiede im Antwortverhalten zwischen west- und ostdeutschen Befragten bestehen. Während Westdeutsche im Zusammenhang mit der FS-Subskala signifikant höhere Werte aufweisen ( $Z = -1.94$ ,  $p = .05$ ), gilt das Gleiche für Ostdeutsche bezüglich der WS-Subskala ( $Z = -2.23$ ,  $p = .03$ ).

Tabelle 21: Mittelwerte der FWSEM-Skalen für West- und Ostdeutsche in Stichprobe 8

Skala	Bundesgebiet		$t$	$df$
	West	Ost		
Gesamtskala	2.27 (.75)	2.27 (.64)	0.04	266.94
FS-Subskala	2.01 (.85)	1.79 (.67)	2.62*	286.30
WS-Subskala	2.55 (.88)	2.76 (.83)	- 2.19*	372

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus.  $t = t$ -Wert.  $df$  = Freiheitsgrad. Höhere Mittelwerte zeigen eine sexistischere Einstellung an (recodierte Skala von 1 „stimme überhaupt nicht zu“ bis 4 „stimme voll und ganz zu“). Neben den Mittelwerten sind in Klammern die entsprechenden Standardabweichungen angegeben. \*  $p < .05$ .

### 3.5.7.11 Korrelationen der FWSEM-Skala mit der ASS

Um die Konstruktvalidität (konvergente Validität) der neu entwickelten FWSEM-Skala zu überprüfen, wurden in Studie 6 und 7 die Korrelationen der Gesamtskala und der beiden Subskalen mit der Skala zur Erfassung des ambivalenten Sexismus (ASS) von Eckes und Six-Materna (1999) berechnet. Die einzelnen Korrelationen sind in Tabelle 22 aufgelistet. Für beide Subskalen sind darüber hinaus die Partialkorrelationen angegeben,

<sup>57</sup> Laut Bortz (2005) ist bei ungleichen Stichprobenumfängen und ungleichen Varianzen mit einer höheren Zahl an Fehlentscheidungen im Zusammenhang mit dem  $t$ -Test zu rechnen, wenn die Normalverteilungsannahme verletzt ist. Die Berechnung des Mann-Whitney-U-Tests sollte daher im Falle der FS-Subskala definitiv erfolgen, da für sie nicht angenommen werden kann, dass die Varianzen der west- und ostdeutschen Befragten gleich sind ( $F = 12.79$ ,  $p < .01$ ). Im Hinblick auf die WS-Subskala kann die Annahme der Varianzhomogenität beibehalten werden, da der Levene-Test nicht signifikant ist ( $F = 1.16$ ,  $p = .28$ ).

bei denen der Effekt der jeweils anderen Subskala kontrolliert wurde, um den reinen Einfluss der einzelnen Komponenten erfassen zu können.

Wie bereits angeführt, wurde angenommen, dass die FWSEM-Skala mit der ASS konsistent mittlere bis hohe positive Korrelationen<sup>58</sup> aufweist, da beiden Messinstrumenten das gleiche theoretische Konstrukt zugrunde liegt. Diese Erwartungen wurden bestätigt. Sowohl für die Gesamtskalen als auch für die jeweiligen Subskalen zeigen sich in Studie 6 und 7 mittlere bis sehr hohe positive Korrelationen (siehe Tabelle 22). Einzige Ausnahme bildet die Partialkorrelation der FS-Subskala mit der BS-Subskala in Studie 6: für sie besteht zwar ein positiver Zusammenhang, dieser ist jedoch nicht signifikant. In den weiblichen und männlichen Teilstichproben bestehen ebenfalls erwartungsgemäß signifikante positive Korrelationen zwischen beiden Skalen.

Tabelle 22: Korrelationen zwischen den FWSEM-Skalen und den ASS-Skalen in Stichprobe 6 und 7

FWSEM-Skalen	ASS		
	Gesamtskala	HS-Subskala	BS-Subskala
	S6		
Gesamtskala	.66** (199)	.60** (200)	.56** (203)
FS-Subskala	.48** (199)	.47** (200)	.38** (203)
FS, kontrolliert für WS	.24**	.26**	.13
WS-Subskala	.66** (200)	.57** (201)	.59** (204)
WS, kontrolliert für FS	.55**	.44**	.50**
	S7		
Gesamtskala	.74** (76)	.63** (80)	.63** (80)
FS-Subskala	.59** (76)	.56** (80)	.45** (80)
FS, kontrolliert für WS	.49**	.48**	.30**
WS-Subskala	.63** (76)	.50** (80)	.59** (80)
WS, kontrolliert für FS	.55**	.36**	.50**

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. ASS = Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999). HS = Hostiler Sexismus. BS = Benevolenter Sexismus. S6 = Studie 6. S7 = Studie 7. Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*\* $p < .01$

<sup>58</sup> Kühnel und Krebs (2001) zufolge gelten Korrelationen mit einem Wertebereich von  $0.05 < r < 0.20$  als gering, von  $0.20 < r < 0.50$  als mittel,  $r > 0.50$  als hoch und  $r > 0.70$  als sehr hoch (vgl. Kühnel & Krebs, 2001, S. 405).

### 3.5.7.12 Korrelationen der FWSEM-Skala mit anderen Sexismus-Skalen

Neben dem Zusammenhang mit der ASS wurden in Studie 6 und in Studie 8 Pearson-Produkt-Moment-Korrelationen (zweiseitig) mit drei weiteren Sexismus-Skalen (Moderner Sexismus, Geschlechterrollen-Orientierung bzw. traditioneller Sexismus sowie die Akzeptanz moderner Mythen über sexuelle Aggression) analysiert, um die Konstruktvalidität der FWSEM-Skala zu überprüfen. In Tabelle 23 und 24 sind alle Korrelationen der FWSEM-Skalen für die Gesamtstichprobe sowie für die Teilstichproben der Frauen und Männer aufgeführt.

Tabelle 23: Korrelationen zwischen den FWSEM-Skalen und drei anderen Sexismus-Skalen in Stichprobe 6

FWSEM-Skalen	Sexismus-Skalen		
	MSS	GRB	AMMSA
	Gesamtstichprobe		
Gesamtskala	.48** (202)	.61** (204)	.60** (199)
FS-Subskala	.41** (202)	.64** (204)	.52** (199)
FS, kontrolliert für WS	.23*	.55**	.34**
WS-Subskala	.42** (203)	.42** (205)	.53** (200)
WS, kontrolliert für FS	.29**	.16*	.37**
	Frauen		
Gesamtskala	.42** (141)	.61** (144)	.59** (139)
FS-Subskala	.35** (141)	.65** (144)	.48** (139)
FS, kontrolliert für WS	.20**	.59**	.31**
WS-Subskala	.37** (141)	.40** (144)	.53** (139)
WS, kontrolliert für FS	.28**	.15	.40**
	Männer		
Gesamtskala	.59** (61)	.62** (60)	.64** (60)
FS-Subskala	.53** (61)	.60** (60)	.61** (60)
FS, kontrolliert für WS	.28*	.42**	.42**
WS-Subskala	.51** (62)	.48** (61)	.54** (61)
WS, kontrolliert für FS	.33*	.23	.27*

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. MSS = Skala zur Erfassung von Vorurteilen gegenüber Frauen auf der Basis des Konzepts des modernen Sexismus (Eckes & Six-Materna, 1998). GRB = Geschlechterrollen-Batterie des International Social Survey Programme-Modul 2002 Familie in Deutschland. AMMSA = Kurzversion der Skala zur Erfassung der Akzeptanz moderner Mythen über sexuelle Aggression (Gerger et al., 2007). Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*  $p < .05$ , \*\*  $p < .01$ .

In Stichprobe 6 zeigen sich, der Hypothese entsprechend, durchgängig signifikante, positive Korrelationen – dies gilt sowohl für die FWSEM-Gesamtskala als auch für beide

Subskalen. Der stärkste Zusammenhang ( $r = .64$ ) besteht zwischen der feindseligen Sexismus-Skala (FS) und der Geschlechterrollen-Batterie (GRB), welche beide offen sexistische Einstellungen erfassen. Dieser Befund wird weiter gestützt, wenn man sich die jeweiligen Partialkorrelationen beider Subskalen anschaut – während der Wert für die FS-Subskala kaum sinkt, verringert sich die Korrelation der WS-Subskala bei Konstanthaltung des Einflusses des feindseligen Sexismus deutlich. Im Falle der weiblichen und der männlichen Teilstichprobe ist der Zusammenhang zwischen wohlwollendem Sexismus und der Befürwortung traditioneller Geschlechterrollen sogar nicht mehr signifikant.

*Tabelle 24: Korrelationen zwischen den FWSEM-Skalen und drei anderen Sexismus-Skalen in Stichprobe 8*

FWSEM-Skalen	Sexismus-Skalen		
	MS	TS	AMMSA
	Gesamtstichprobe		
Gesamtskala	.15** (341)	.53** (340)	.49** (340)
FS-Subskala	.20** (354)	.52** (352)	.37** (352)
FS, kontrolliert für WS	.22**	.49**	.23**
WS-Subskala	.06 (374)	.36** (373)	.46** (373)
WS, kontrolliert für FS	-.01	.22**	.33**
	Frauen		
Gesamtskala	.15* (213)	.50** (212)	.50** (213)
FS-Subskala	.19** (217)	.45** (216)	.36** (216)
FS, kontrolliert für WS	.22**	.36**	.11
WS-Subskala	.09 (232)	.39** (231)	.52** (232)
WS, kontrolliert für FS	-.00	.27**	.38**
	Männer		
Gesamtskala	.13 (128)	.59** (128)	.47** (127)
FS-Subskala	.20* (137)	.61** (136)	.38** (136)
FS, kontrolliert für WS	.19	.59**	.32**
WS-Subskala	.00 (142)	.31** (142)	.36** (141)
WS, kontrolliert für FS	-.09	.22*	.27**

*Anmerkungen.* FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. MS = Moderner Sexismus (Kurzskala des GMF-Syndroms mit 2 Items, siehe Heitmeyer, 2002,2010). TS = Traditioneller Sexismus (Kurzskala des GMF-Syndroms mit 2 Items, siehe Heitmeyer, 2002,2010). AMMSA = Kurzversion der Skala zur Erfassung der Akzeptanz moderner Mythen über sexuelle Aggression (Gerger et al., 2007). Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*  $p < .05$ , \*\*  $p < .01$ .

In Studie 8 sehen die Ergebnisse ganz ähnlich aus: Auch hier weisen die drei FWSEM-Skalen signifikante, positive Korrelationen mit dem traditionellen Sexismus auf, die durchgängig höher ausfallen als die Zusammenhänge mit den anderen Skalen.

Mit der Akzeptanz moderner Mythen über sexuelle Aggression bestehen ebenfalls signifikante, positive Korrelationen. Einzige Ausnahme bildet die Partialkorrelation der FS-Subskala mit der AMMSA-Skala in der weiblichen Teilstichprobe: Sie ist bei Kontrolle des Einflusses der WS-Subskala nicht mehr signifikant.

Im Hinblick auf den Modernen Sexismus ist folgendes zu erwähnen: Entgegen den Erwartungen besteht zwischen dem wohlwollenden Sexismus und dem modernen Sexismus weder in der Gesamt-Stichprobe noch in der weiblichen oder der männlichen Teilstichprobe ein signifikanter Zusammenhang. Auch für die Gesamtskala und die FS-Subskala liegen die Korrelationen mit der MS eher im niedrigen Bereich. Wird der Einfluss des benevolenten Sexismus konstant gehalten, so besteht für die männlichen Befragten sogar überhaupt kein signifikanter Zusammenhang mehr zwischen dem feindseligen und dem modernen Sexismus. Insgesamt gesehen konnte somit die Verbindung zwischen dem modernen Sexismus und dem feindseligen Sexismus in Studie 8 nicht so eindeutig belegt werden bzw. bezüglich des wohlwollenden Sexismus überhaupt nicht belegt werden.

#### ***3.5.7.13 Korrelationen der FWSEM-Skalen mit Autoritarismus und Sozialer Dominanzorientierung***

Um die konvergente Validität der FWSEM-Skala noch eingehender zu untersuchen, wurden in Studie 6 und 8 neben den Sexismus-Skalen auch eine Skala zur Messung autoritärer Einstellungen (RWA) und eine zur Erfassung der Sozialen Dominanzorientierung (SDO) eingesetzt. In Tabelle 25 können die Korrelationen beider Studien eingesehen und miteinander verglichen werden.

Wie erwartet, korrelieren in Stichprobe 6 alle FWSEM-Skalen mit der RWA-Skala und der SDO-Skala signifikant positiv. Insbesondere der Zusammenhang zwischen dem wohlwollenden Sexismus und RWA erweist sich sowohl in der in Gesamt-Stichprobe als auch in beiden Teilstichproben als vergleichsweise hoch (dies gilt auch für die Partialkorrelationen). Während dies bei den weiblichen Befragten auch für Korrelation zwischen dem wohlwollenden Sexismus und der Sozialen Dominanzorientierung gilt, zeigt sich bei den männlichen Befragten ein deutlicher Unterschied: Hier weist die WS-Subskala mit der SDO-Skala überhaupt keine signifikante Korrelation auf. Für Männer scheinen wohlwollend-sexistische Einstellungen gegenüber Müttern folglich nicht mit einer Sozialen Dominanzorientierung einherzugehen.

Stattdessen zeigt sich eine Verbindung zu feindselig-sexistischen Einstellungen. So ist in der männlichen Teilstichprobe die FS-Subskala positiv mit der SDO-Skala korreliert, was auch noch gilt, wenn der Einfluss der WS-Subskala kontrolliert wird. In der Gesamtstichprobe und der Teilstichprobe der Frauen ist der positive Zusammenhang zwischen

dem feindseligen Sexismus und dem Autoritarismus bzw. der Sozialen Dominanzorientierung hingegen zu einem Großteil auf den Einfluss des benevolenten Sexismus zurückzuführen. Bei dessen Konstanzhaltung sind die Korrelationen zwischen der FS-Subskala mit RWA und SDO nicht mehr signifikant.

Auch in Studie 8 bestehen erwartungsgemäß durchweg signifikante, positive Korrelationen zwischen den FWSEM-Skalen und den beiden gruppenbezogenen Ideologien (RWA und SDO), die zur Validierung herangezogen wurden. Erneut erweist sich die Verknüpfung von wohlwollendem Sexismus und RWA als besonders beständig (wie ein Blick auf die Partialkorrelationen zeigt). Für den Zusammenhang von wohlwollendem Sexismus und sozialer Dominanzorientierung, wird abermals das Korrelationsmuster aus Studie 6 repliziert: Während in der Gesamtstichprobe und in der Teilstichprobe der Frauen signifikante, positive Korrelationen für beide Skalen bestehen, trifft dies für die Männer nicht ausnahmslos nicht zu. Schaut man sich die Partialkorrelationen an, so fällt auf, dass der Zusammenhang zwischen der WS-Subskala und der SDO-Skala nicht mehr signifikant von Null verschieden ist, sobald der feindselige Sexismus auspartialisiert wird.

Im Hinblick auf den feindseligen Sexismus ergibt sich hingegen folgendes Bild: Betrachtet man in der Teilstichprobe der Frauen die Partialkorrelationen der FS-Subskala, bei der der Einfluss der WS-Subskala kontrolliert worden ist, zeigt sich, dass kein signifikanter Zusammenhang mehr mit der RWA-Skala besteht. (Dieses Ergebnis ist mit den Befunden aus Studie 6 deckungsgleich. Allerdings bestand hier auch kein Zusammenhang mehr mit SDO, der in Studie 8 allerdings gegeben ist). In der männlichen Teilstichprobe besteht hingegen durchgängig ein signifikanter, positiver Zusammenhang zwischen feindseligem Sexismus und autoritären Einstellungen bzw. Sozialer Dominanzorientierung. Entgegen den Ergebnissen der weiblichen Teilstichprobe und auch der männlichen Teilstichprobe in Studie 6 ist auch bei Konstanzhaltung des Einflusses der WS-Subskala der Zusammenhang zwischen der FS-Subskala und RWA signifikant. Nicht so stark ausgeprägt wie in Studie 6, aber dennoch signifikant, ist die Partialkorrelation zwischen der FS-Subskala und der SDO-Skala.

Tabelle 25: Korrelationen zwischen den FWSEM-Skalen und der RWA- bzw. SDO-Skala in Stichprobe 6 und 8

FWSEM-Skalen	Gruppenbezogene Ideologien			
	RWA1	SDO	RWA2	SDO
	Gesamtstichprobe			
	Studie 6		Studie 8	
Gesamtskala	.51** (202)	.38** (204)	.52** (341)	.35** (341)
FS-Subskala	.35** (202)	.30** (204)	.36** (354)	.28** (354)
FS, kontrolliert für WS	.11	.11	.22**	.21**
WS-Subskala	.51** (203)	.33** (205)	.50** (374)	.30** (374)
WS, kontrolliert für FS	.43**	.29**	.41**	.22**
	Frauen			
	Studie 6		Studie 8	
Gesamtskala	.47** (141)	.36** (143)	.52** (213)	.41** (213)
FS-Subskala	.31** (141)	.21* (143)	.36** (217)	.33** (217)
FS, kontrolliert für WS	.08	.01	.10	.20*
WS-Subskala	.49** (141)	.40** (143)	.54** (232)	.37** (232)
WS, kontrolliert für FS	.42**	.37**	.46**	.24**
	Männer			
	Studie 6		Studie 8	
Gesamtskala	.59** (61)	.44** (61)	.52** (128)	.26** (128)
FS-Subskala	.46** (61)	.50** (61)	.36** (137)	.20* (137)
FS, kontrolliert für WS	.18	.38**	.31**	.20*
WS-Subskala	.54** (62)	.19 (62)	.42** (142)	.19* (142)
WS, kontrolliert für FS	.42**	.01	.34**	.18

*Anmerkungen.* FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. RWA1 und 2 = Autoritarismus-Kurzskala (Petzel et al., 1997) (RWA1 in S6: Kurzskala mit 9 Items; RWA2 in S8: Kurzskala mit 4 Items). SDO = Kurzversion der Social Dominance Orientation Scale aus dem Projekt Gruppenbezogene Menschenfeindlichkeit (Heitmeyer, 2002, 2010) (S6 und S8: identische Kurzskala mit 3 Items). S6 = Studie 6. S8 = Studie 8. Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*\* $p < .01$

### 3.5.7.14 Korrelationen der FWSEM-Skalen mit dem GMF-Syndrom

Neben den Korrelationen der FWSEM-Skalen mit anderen Sexismus-Skalen, sowie mit der RWA- und der SDO-Skala wurden die Zusammenhänge mit einer Reihe anderer Vorurteile untersucht, um die Validität der neu entwickelten Skala zu überprüfen. Dabei wurde angenommen, dass ambivalenter Sexismus positiv mit allen Elementen des Einstellungssyndroms Gruppenbezogener Menschenfeindlichkeit korreliert.



In der Gesamtstichprobe ergeben sich durchweg signifikante, positive Korrelationen, deren Werte zwischen  $r = .12$  (FS-Subskala (Partialkorrelation) und Abwertung Obdachloser) und  $r = .50$  (FWSEM-Gesamtskala und Etabliertenvorrechte) liegen (siehe Tabelle 26).

Für die Teilstichprobe der Frauen kann die Hypothese ebenfalls bestätigt werden. Auch hier zeigen sich sowohl für die Gesamtskala als auch für beide Subskalen signifikant positive Zusammenhänge mit allen Elementen des Syndroms. Schaut man sich jedoch die Partialkorrelationen an, wird deutlich, dass der Zusammenhang des feindseligen Sexismus mit den übrigen erfassten Vorurteilen überwiegend auf den Einfluss des wohlwollenden Sexismus zurückzuführen ist. Wird dieser nämlich konstant gehalten, ergeben sich lediglich für Homophobie, Etabliertenvorrechte und Islamophobie signifikante, positive Korrelationen (siehe Tabelle 26).

Für die Teilstichprobe der Männer liegt der umgekehrte Fall vor. Während für den feindseligen Sexismus auch bei Konstanthaltung des Einflusses des wohlwollenden Sexismus die Korrelationen signifikant bleiben, führt die Auspartialisierung des feindseligen Sexismus im Falle des wohlwollenden Sexismus dazu, dass vier der Korrelationen nicht mehr signifikant sind: die Abwertung Obdachloser, die Abwertung Langzeitarbeitsloser, Fremdenfeindlichkeit und Homophobie (siehe Tabelle 26).

Für die weiblichen Befragten scheint somit eher der wohlwollende Sexismus mit den einzelnen Elementen des Syndroms Gruppenbezogener Menschenfeindlichkeit zusammenzuhängen, für die männlichen Befragten trifft dies hingegen eher auf den feindseligen Sexismus zu.

Tabelle 26: Korrelationen zwischen den FWSEM-Skalen und dem GMF-Syndrom in Stichprobe 8

FWSEM-Skalen	GMF-Syndrom								
	Rassismus	Fremdenfeindlichkeit	Antisemitismus	Homophobie	Etablierten- vorrechte	Islamophobie	Abwertung Behinderter	Abwertung Obdachloser	Abwertung LA
	Gesamtstichprobe								
Gesamtskala	.48** (341)	.49** (341)	.32** (325)	.48** (340)	.50** (341)	.42** (338)	.26** (339)	.27** (341)	.39** (341)
FS-Subskala	.32** (354)	.39** (354)	.29** (338)	.46** (353)	.44** (354)	.35** (351)	.23** (352)	.20** (354)	.31** (354)
FS, kontrolliert für WS	.14*	.27**	.18**	.37**	.33**	.24**	.14*	.12*	.21**
WS-Subskala	.44** (374)	.42** (373)	.24** (354)	.34** (373)	.37** (374)	.36** (369)	.20** (372)	.27** (374)	.34** (374)
WS, kontrolliert für FS	.38**	.30**	.17**	.16**	.23**	.24**	.15**	.17**	.23**
	Frauen								
Gesamtskala	.46** (213)	.52** (213)	.23** (199)	.47** (212)	.46** (213)	.42** (212)	.21** (211)	.27** (213)	.40** (213)
FS-Subskala	.32** (217)	.37** (217)	.17* (203)	.42** (216)	.40** (217)	.34** (216)	.19** (215)	.17* (217)	.27** (217)
FS, kontrolliert für WS	.09	.11	.03	.27**	.23**	.15*	.05	-.01	.07
WS-Subskala	.46** (232)	.53** (231)	.21** (215)	.39** (231)	.39** (232)	.41** (230)	.18** (230)	.30** (232)	.41** (232)
WS, kontrolliert für FS	.36**	.42**	.18*	.20**	.22**	.27**	.14*	.25**	.34**
	Männer								
Gesamtskala	.52** (128)	.46** (128)	.46** (126)	.50** (128)	.56** (128)	.42** (126)	.35** (128)	.29** (128)	.39** (128)
FS-Subskala	.33** (137)	.44** (137)	.41** (135)	.52** (137)	.50** (137)	.36** (135)	.28** (137)	.26** (137)	.38** (137)
FS, kontrolliert für WS	.24**	.43**	.37**	.48**	.46**	.33**	.27**	.27**	.36**
WS-Subskala	.41** (142)	.26** (142)	.26** (139)	.26** (142)	.34** (142)	.29** (139)	.24** (142)	.21* (142)	.22** (142)
WS, kontrolliert für FS	.44**	.17	.23**	.16	.30**	.20*	.20*	.10	.13

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*  $p < .05$ , \*\*  $p < .01$ .

### 3.5.7.15 Korrelationen der FWSEM-Skalen mit Sozialer Erwünschtheit

Um die diskriminante Validität der FWSEM-Skala zu untersuchen, kam in Studie 6 eine Kurzskaala zur Erfassung der Tendenz der Befragten, sozial erwünschte Antworten zu geben von Winkler, Kroh und Spiess (2006) zum Einsatz. Sie untergliedert sich in zwei Subkomponenten: Zum einen die Fremdtäuschung (impression management) und zum anderen die Selbsttäuschung (self-deceptive enhancement). Die Werte der FWSEM-Skalen sollten idealerweise unabhängig von der Tendenz sein, ein sozial erwünschtes Antwortverhalten zu zeigen. Dementsprechend sollten Pearson-Produkt-Moment-Korrelationen mit Werten nahe Null gefunden werden, die nicht signifikant sind.

Aufgrund der unzufriedenstellenden Reliabilität von  $\alpha = .26$  konnte die Skala zur Erfassung der Fremdtäuschung (wie bereits erwähnt) leider nicht in die folgenden Analysen einbezogen werden, so dass nur Aussagen im Zusammenhang mit der Selbsttäuschung (ST)-Skala getroffen werden können. Wie erwartet, korrelieren die FWSEM-Gesamtskala sowie die beiden Subskalen nicht mit der ST-Subskala. Dies gilt sowohl für die Gesamtstichprobe als auch für die weibliche und männliche Teilstichprobe (siehe Tabelle 27). Das Antwortverhalten der Befragten im Hinblick auf den FWSEM-Gesamtskala sowie die FS- und die WS-Subskala scheint somit nicht systematisch durch eine Tendenz verzerrt zu werden, ein sozial erwünschtes Antwortverhalten (in Sinne einer Selbsttäuschung) an den Tag zu legen.

Tabelle 27: Korrelationen zwischen den FWSEM-Skalen und einer Skala zur sozialen Erwünschtheit in Stichprobe 6

FWSEM-Skalen	ST		
	Gesamtstichprobe	Frauen	Männer
Gesamtskala	-.04 (206)	-.09 (145)	.09 (61)
FS-Subskala	-.08 (206)	-.14 (145)	.08 (61)
FS, kontrolliert für WS	-.10	-.15	.03
WS-Subskala	.03 (207)	-.01 (145)	.10 (62)
WS, kontrolliert für FS	.04	.02	.03

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. ASS = Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999). ST = Selbsttäuschung (self-deceptive enhancement) ist eine Subkomponenten der Kurzskaala zur Erfassung der Tendenz der Befragten, sozial erwünschte Antworten zu geben (Winkler et al., 2006). Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. \*  $p < .05$ , \*\*  $p < .01$ .

### 3.5.7.16 Übereinstimmungsvalidität der FWSEM-Skalen im Vergleich zur ASS

Ziel der FWSEM-Skala ist es, ambivalent-sexistische Einstellungen gegenüber Müttern zu erfassen. Die Subkomponente des feindseligen Sexismus soll mit dabei negative Einstel-

lungen gegenüber nicht-traditionellen Frauen zusammenhängen, wohingegen wohlwollender Sexismus mit positiven Einstellungen gegenüber traditionellen Frauen einhergehen soll. Um die Übereinstimmungsvalidität der FWSEM-Skala zu überprüfen, wurden die Befragten in Studie 7 gebeten, die beiden Subgruppen Karrierefrauen und Hausfrauen anhand einer geschlechtsstereotypen Merkmalskala zu beurteilen.

*Stereotype über Karrierefrauen.* Die Korrelationen der FWSEM-Skalen und den Stereotypen über Karrierefrauen sind wie vorhergesagt (siehe Tabelle 28). Feindseliger Sexismus weist der Hypothese entsprechend eine positive Korrelation im Zusammenhang mit maskulin-negativen Eigenschaften und eine negative Korrelation im Zusammenhang mit feminin-positiven Eigenschaften auf. Gleiches trifft auch für die Gesamtskala zu (wobei dies auf den Einfluss der FS-Subskala zurückzuführen ist, da zwischen der WS-Subskala und den feminin-positiven Eigenschaften kein signifikanter Zusammenhang besteht). Dies bedeutet Folgendes: Personen, die zum feindseligen Sexismus neigen, sprechen Karrierefrauen somit feminin-positive Merkmalszuschreibungen wie etwa ein Empathievermögen ab und gleichzeitig maskulin-negative Merkmalszuschreibungen wie etwa eine Überheblichkeit zu.

*Stereotype über Hausfrauen.* Die Korrelationen der FWSEM-Skalen und den Merkmalszuschreibungen von Hausfrauen sind in Tabelle 28 eingetragen. Für den wohlwollenden Sexismus zeigt sich, wie erwartet, ein positiver Zusammenhang mit der Zuschreibung von feminin-positiven Eigenschaften zur Subgruppe der Hausfrauen.<sup>59</sup> Die Partialkorrelation, bei der feindseliger Sexismus kontrolliert wird, ist allerdings nicht signifikant. Die FS-Subskala ist hingegen unerwarteterweise signifikant positiv mit den feminin-positiven Merkmalszuschreibungen korreliert (dies trifft auch auf die Partialkorrelationen zu). Dies lässt darauf schließen, dass die positive Korrelation der FWSEM-Gesamtskala, welche ebenfalls besteht, auf den Einfluss der FS-Subskala zurückzuführen ist. Insgesamt bleibt festzuhalten: Personen, die zum wohlwollenden Sexismus, aber insbesondere auch zum feindseligen Sexismus tendieren, bewerten Hausfrauen durchweg positiv im Hinblick auf „typisch weibliche“ Eigenschaften.

Um beide ambivalente Sexismus-Skalen (FWSEM und ASS) bezüglich der Übereinstimmungsvalidität miteinander zu vergleichen und so Aussagen über die diskriminante Validität treffen zu können, wurden die oben beschriebenen Analysen auch mit der Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999) durchgeführt. Da beiden Skalen dieselbe Theorie zugrunde liegt, wurde bezüglich der Zusammenhänge der ASS-Skalen mit den Stereotypen über Karrierefrauen und Hausfrauen auch dasselbe Korrelationsmuster erwartet.

*Stereotype über Karrierefrauen.* Es zeigte sich, dass auch die HS-Subskala der Hypothese entsprechend positive Korrelationen mit maskulin-negativen Eigenschaften und

---

<sup>59</sup> Dieses Ergebnis deckt sich mit dem eingangs beschriebenen „women are wonderful“ effect (Eagly & Mladinic, 1994), bei dem Frauen, die die traditionelle Rolle erfüllen, feminin-positiv Merkmale zugesprochen werden.

eine negative Korrelation im Zusammenhang mit feminin-positiven Eigenschaften aufweist (siehe Tabelle 29). Unerwarteterweise ergibt sich im Falle der Partialkorrelationen darüber hinaus ein signifikanter positiver Zusammenhang zwischen hostilem Sexismus und der Zuschreibung feminin-*negativer* Eigenschaften. Karrierefrauen werden von hostilem Sexisten folglich nicht nur als überheblich und egoistisch (maskulin-negativ), sondern auch als weinerlich und leichtgäufig (feminin-negativ) charakterisiert, während ihnen u.a. Sanftheit und Hilfsbereitschaft (feminin-positiv Eigenschaften) abgesprochen werden.

Auch hinsichtlich des benevolenten Sexismus bestehen signifikante Partialkorrelationen, die nicht vorhergesagt wurden: So ist die BS-Subskala signifikant negativ mit feminin-negativen Eigenschaftszuschreibungen und positiv mit maskulin-positiven Eigenschaftszuschreibungen korreliert. Personen, die benevolent sexistisch eingestellt sind, besitzen demzufolge ein positives Bild von Karrierefrauen, das sie u.a. als wenig unterwürfig und wenig leichtgläubig, sondern stattdessen vielmehr als unabhängig und selbstsicher beschreibt.

*Stereotype über Hausfrauen.* Erwartungsgemäß ist benevolenter Sexismus signifikant positiv mit feminin-positiven Merkmalszuschreibungen korreliert (siehe Tabelle 29). Darüber hinaus ergeben sich signifikant negative Partialkorrelationen mit maskulin-negativen Stereotypen. Benevolente Sexisten sehen Hausfrauen somit u.a. als verständnisvoll und empathisch und gleichzeitig als wenig überheblich und egoistisch an.

Da die Korrelationsmuster somit nicht vollständig übereinstimmen, scheinen die FWSEM-Skala und die ASS eine gewisse diskriminante Validität aufzuweisen.

Tabelle 28: Korrelationen zwischen den FWSEM-Skalen und den Stereotypen über Karrierefrauen und Hausfrauen in Stichprobe 7

FWSEM-Skalen	Feminine Eigenschaften		Maskuline Eigenschaften	
	Positiv	Negativ	Positiv	Negativ
Karrierefrauen				
Gesamtskala	<b>-.30**</b> (83)	.10 (80)	.05 (83)	<b>.55**</b> (82)
FS-Subskala	<b>-.32**</b> (83)	.20 (80)	.09 (83)	<b>.56**</b> (82)
FS, kontrolliert für WS	<b>-.27**</b>	.22	.05	<b>.48**</b>
WS-Subskala	-.19 (84)	-.01 (81)	-.00 (84)	<b>.38**</b> (83)
WS, kontrolliert für FS	-.09	-.09	-.03	.21
Hausfrauen				
Gesamtskala	<b>.32**</b> (83)	.18 (82)	-.12 (82)	-.09 (81)
FS-Subskala	<b>.31**</b> (83)	.14 (82)	.00 (82)	-.11 (81)
FS, kontrolliert für WS	<b>.25**</b>	.09	.09	-.10
WS-Subskala	<b>.22**</b> (84)	.13 (84)	-.15 (84)	-.04 (82)
WS, kontrolliert für FS	<b>.11</b>	.12	-.20	-.01

Anmerkungen. FWSEM = Feindselig- und wohlwollend-sexistische Einstellungen in Bezug auf Mütter. FS = Feindseliger Sexismus. WS = Wohlwollender Sexismus. Neben den Produkt-Moment-Korrelationen sind in Klammern die jeweiligen Fallzahlen angegeben. Die Korrelationen, von denen erwartet wurde, dass sie signifikant sind, sind fett dargestellt. \*  $p < .05$ , \*\*  $p < .01$ .

Tabelle 29: Korrelationen zwischen den ASS-Skalen und den Stereotypen über Karrierefrauen und Hausfrauen in Stichprobe 7

FWSEM-Skalen	Feminine Eigenschaften		Maskuline Eigenschaften	
	Positiv	Negativ	Positiv	Negativ
Karrierefrauen				
Gesamtskala	<b>-.18</b> (74)	.01 (72)	.12 (74)	<b>.48**</b> (73)
HS-Subskala	<b>-.30**</b> (78)	.18 (75)	-.05 (78)	<b>.49**</b> (77)
HS, kontrolliert für BS	<b>-.29*</b>	.27*	-.12	<b>.47**</b>
BS-Subskala	-.06 (78)	-.13 (76)	.20 (78)	<b>.29**</b> (77)
BS, kontrolliert für HS	.11	-.25*	.24*	.08
Hausfrauen				
Gesamtskala	<b>.35**</b> (74)	.16 (73)	-.05 (73)	-.10 (73)
HS-Subskala	.20 (78)	.17 (77)	.00 (77)	.06 (77)
HS, kontrolliert für BS	.10	.13	.07	.15
BS-Subskala	<b>.37**</b> (78)	.15 (77)	-.15 (77)	-.19 (76)
BS, kontrolliert für HS	<b>.28*</b>	.05	-.11	-.25*

Anmerkungen. ASS = Skala zur Erfassung des ambivalenten Sexismus (Eckes & Six-Materna, 1999). HS = Hostiler Sexismus. BS = Benevolenter Sexismus. Weitere Anmerkungen siehe Tabelle 28.

## 4 Fazit und Diskussion

Ziel dieser Arbeit war es, eine deutschsprachige Skala zu generieren und zu validieren, die ambivalent-sexistische Einstellungen in Bezug auf Mütter erfasst. Dabei wurden zwei zentrale Anliegen verfolgt: ein inhaltliches und ein methodisches. Thematisch lag bei der Skalenentwicklung der Fokus auf einer Subgruppe von Frauen, die in bisherigen Messinstrumenten vernachlässigt wurde: die Gruppe der Mütter. Um die Ansichten der Befragten differenziert erfassen zu können, wurden der Theorie des ambivalenten Sexismus (Glick & Fiske, 1996, 1997) folgend neben offen feindseligen auch (scheinbar) wohlwollende Einstellungen berücksichtigt. Die *feindselige* Subkomponente der Skala zeichnet ein *negatives* Frauenbild und wertet Mütter ab, die aus der traditionellen Rolle ausbrechen, ihren Beruf kurz nach der Geburt weiter verfolgen möchten und ihre Kinder außerhäuslich betreuen lassen. Sie verstoßen somit gegen den Muttermythos und werden folglich als „Rabemütter“ charakterisiert. Die *wohlwollende* Subkomponente bezieht sich hingegen auf subjektiv *positiv* empfundene Zuschreibungen, die Frauen für die Übernahme der Kinderbetreuung als unersetzbar erscheinen lassen. Dabei werden Frauen insbesondere feminine Eigenschaften attestiert und ihre Anwesenheit für das kindliche Wohlergehen als essentiell herausgestellt. Auf Grund seines subtilen Charakters wird der wohlwollende Sexismus nicht sofort als solcher erkannt und ist auch eher gesellschaftlich akzeptiert. Für Untersuchungen in diesem Themenfeld hat dies den Vorteil, dass die Befragten sich weniger scheuen, derartige Ansichten in Umfragen offen zu äußern – im Hinblick auf den feindseligen Sexismus ist dies nicht immer der Fall. Letztendlich dienen die scheinbar wohlwollenden Zuschreibungen jedoch auch dazu, Diskriminierungen gegenüber Frauen zu rechtfertigen, die traditionelle Arbeitsteilung der Geschlechter zu festigen und somit die männliche Vorrangstellung in der Gesellschaft fortzuschreiben.

In methodischer Hinsicht kam bei der Skalentestung neben zwei weiteren Methoden (Analyse von Pretest-Daten und ExpertInnen-Review) ein Pretest-Verfahren zum Einsatz, welches in den vergangenen Jahren international zunehmend an Bedeutung gewonnen hat, aber in der deutschen Forschungslandschaft bislang noch zu selten verwendet wird: die Technik der kognitiven Interviews. Wie in Unterkapitel 2.5.2 erläutert, ist es Befragten im Rahmen einer standardisierten Befragung nicht möglich, ihr Frageverständnis durch ein Grounding abzusichern. Dies kann dazu führen, dass die Items nicht im intendierten Sinne aufgefasst und Antworten gegeben werden, die nicht dem tatsächlich erfragten Verhalten oder der tatsächlichen Einstellung entsprechen. Folglich ist die Formulierung von „guten“ Fragen, deren pragmatischer Sinn sich den Befragten von alleine erschließt, für die Qualität der Daten entscheidend. Die Methode der kognitiven Interviews ist im Vergleich zu anderen Pretest-Verfahren am besten dazu geeignet, bei einer größeren Anzahl von Items mögliche Verständnisprobleme aktiv aufzuspüren und deren Ursachen detailliert auf den Grund zu gehen (siehe Unterkapitel 2.6).

Fowler (2001) nennt neben dem pragmatischen Verständnis (kognitiver Standard) sieben weitere Standards, die eine gute Frage erfüllen sollte (siehe Unterkapitel 2.6). Im Zuge der Entwicklung und Testung der FWSEM-Skala wurden neben den kognitiven Standards (zu denen das pragmatische Frageverständnis zählt) vorrangig die inhaltlichen sowie die psychometrischen Standards in den Blick genommen, aber auch die Multi-Mode-Capability wurde berücksichtigt.<sup>60</sup> Da keine Pretest-Methode in der Lage ist, alle dieser vier anvisierten Standards gleichermaßen zu überprüfen, wurde eine Kombination verschiedener Verfahren gewählt.

Ziel von *Pretest-Phase 1* war es, die Itemanzahl des ursprünglichen Itempools, der als Ausgangspunkt für die Entwicklung der FWSEM-Skala diente, zu reduzieren. Hierfür wurde eine psychometrische Analyse der Pretest-Daten (quantitative Field-Technik) durchgeführt. In *Pretest-Phase 2* wurde im Rahmen von drei Runden kognitiver Interviews (qualitative Pre-Field-Technik) das intendierte Frageverständnis der verbliebenen Items untersucht. Neben den inhaltlichen und kognitiven Standards sollte in dieser Pretest-Phase die Multi-Mode-Capability der Items getestet werden. In *Pretest-Phase 3* erfolgte erneut eine Analyse von Pretest-Daten. Vorrangiges Anliegen dieser Phase war es, die besten Items für die endgültige FWSEM-Skala auszuwählen und die psychometrischen Standards zu analysieren. Da neben einer Online-Umfrage auch eine schriftliche Befragung in dieser Pretest-Phase durchgeführt wurde, konnte zudem die Multi-Mode-Capability der Skala getestet werden.

#### 4.1 Fazit zur methodischen Vorgehensweise

Insgesamt erwies sich die oben beschriebene Kombination von Pretest-Verfahren (psychometrische Analyse der Pretest-Daten und kognitive Interviews mit vorgeschaltetem ExpertInnen-Review) als effektiv. Bezüglich der einzelnen Phasen können für zukünftige Untersuchungen folgende Empfehlungen gegeben werden:

In der ersten Pretest-Phase sollte die Analyse der Pretest-Daten genutzt werden, um die ursprüngliche Itemanzahl des Itempools noch deutlicher zu verringern. Idealerweise sollte am Ende von Pretest-Phase 1 bereits eine mögliche Vorversion der Skala bestehen. Dies hätte den Vorteil, dass man deutlich weniger Items im Rahmen der kognitiven Interviews auf etwaige Frage-Probleme hin untersuchen müsste und alle Items gleichermaßen intensiv hinterfragt werden könnten. Zudem verkürzt sich der zeitliche Aufwand, der für die Durchführung und v.a. die Analyse der kognitiven Interviews nötig ist, auf diese Weise erheblich. Auch die Wahrscheinlichkeit, dass mögliche Frage-Probleme übersehen werden (siehe Unterkapitel 2.6.1.6), ist bei einer kleineren Anzahl von Items geringer.

---

<sup>60</sup>Die vier übrigen Standards (Kosteneffektive Nutzung der Umfragezeit, Interpersonelle Standards, Benutzerfreundlichkeit und Multi-Language-Capability), die eher die Konstruktion und Testung eines finalen Fragebogens für den Einsatz im Rahmen einer (internationalen) Studie mit unterschiedlichen Skalen betreffen, wurden nicht überprüft.



Zur Pretest-Phase 2 ist Folgendes festzuhalten: Mit Hilfe der kognitiven Interviews konnte erfolgreich eine Reihe von Problemen identifiziert werden, die mit dem Frage-Antwort-Prozess zusammenhingen – dies traf insbesondere auf Verständnisprobleme zu. Aufgrund der aktiven Vorgehensweise der Technik befanden sich darunter auch solche Probleme, die den Befragten selbst nicht bewusst waren. Des Weiteren erwiesen sich einige Itemformulierungen als problematisch, von denen dies im Rahmen des ExpertInnen-Reviews nicht erwartet wurde. Auch in den Itemkennwerten aus Pretest-Phase 1 zeigten sich diese Defizite nicht. So wurden manche Items von den Befragten komplett anders verstanden, als es intendiert war, obwohl sie zufriedenstellende psychometrische Eigenschaften aufwiesen. Dieser Befund ist zugleich ein Beleg für die Schwäche der Methode der (quantitativen) Analyse von Pretest-Daten. Zufriedenstellende Itemkennwerte sind folglich nicht gleichbedeutend mit einem intendierten Itemverständnis auf Seiten der Befragten.

Im Hinblick auf die einzelnen Nachfragetechniken hat sich eine flexible Vorgehensweise mit einer Kombination aus *spontaneous* und *scripted probes* (siehe Unterkapitel 2.6.1.5) besonders bewährt. Selbstverständlich war es im Rahmen dieser Untersuchung von Vorteil, dass ein Großteil der Interviews von der Verfasserin persönlich durchgeführt wurde, da die exakte Intention der Items somit bekannt war und Missverständnisse leichter erfasst werden konnten. Folglich war es viel besser möglich, spontan auf die Antworten der Befragten zu reagieren und möglichen Frage-Problemen tiefer auf den Grund zu gehen. Zudem konnten diejenigen Items, die offensichtlich Formierungsdefizite aufwiesen, bereits während des Interviews markiert werden. Dies erleichterte wiederum die Auswertung der kognitiven Interviews, da bereits erste Hinweise auf problematische Items vorlagen. Wenn es die zeitlichen und finanziellen Mittel einer Studie erlauben, scheint daher die intensive Schulung der InterviewerInnen, die über die Ziele und Intentionen der zu testenden Fragen aufklärt, im Vorfeld durchaus ratsam und gewinnbringend.

Darüber hinaus wäre es sinnvoll, mehrere Personen in die Analyse der kognitiven Interviews mit einzubeziehen, da verschiedene AuswerterInnen zu unterschiedlichen Ergebnissen gelangen und es folglich ratsam ist, verschiedene Perspektiven zu berücksichtigen. Conrad und Blair (2009) geben hierzu folgende Einschätzung ab: “Although it cannot correct the fundamental problem of ambiguity in verbal reports, the use of multiple reviewers who work together to reach consensus on problem identification may to some extent mitigate the weakness in cognitive interview data” (Conrad & Blair, 2009, S. 53). So könnte innerhalb der Gruppe über die einzelnen Ergebnisse diskutiert und eine gemeinsame Entscheidung getroffen werden, ob bei einem identifizierten Problem ein tatsächliches Frage-Defizit vorliegt oder ob es nur ein „falscher Alarm“ ist (siehe Unterkapitel 2.6.1.6). Stehen einem diese Möglichkeiten nicht zur Verfügung, ist ein zeitlicher Abstand hilfreich, um das Material mit einer gewissen Distanz betrachten zu können.

In einigen wenigen Fällen war es nicht möglich, eine geeignete Umformulierung der Items zu finden (z.B. Item *Verständnis*). Während also ganz offensichtlich ein Frageproblem aufgespürt werden konnte („*finding a problem*“), gelang es nicht, es zufriedenstellend zu beheben („*fixing a problem*“, siehe Unterkapitel 2.6.1.6). Folglich wurden diese Items letztendlich aus dem Itempool entfernt. Auch vor diesem Hintergrund wäre es vorteilhaft, mit anderen ExpertInnen der Itemformulierung zusammenzuarbeiten, um Ideen über mögliche Formulierungsalternativen durchzusprechen zu können.

In Runde 3 der kognitiven Interviews wurde eine telefonische Befragung simuliert, um zu testen, ob das Frageverständnis auch bei einer ausschließlich akustischen Darbietung der Items gegeben ist. Hierfür wurden den Befragten die Items vorgelesen. Lediglich die sechsstufige Antwortskala, die in den vorhergehenden Interviews ebenfalls zum Einsatz kam und der Einheitlichkeit halber beibehalten werden sollte, wurde ihnen schriftlich vorgelegt. Diese Testung der Multi-Mode-Capability hat offensichtlich gut funktioniert. So erwies sich die FWSEM-Skala beim Einsatz im Feld im Zuge einer Telefonbefragung (Studie 8) als reliabel und valide. Dabei besaß die Skala ebenso gute Kennwerte wie auch zuvor bei den schriftlichen Befragungen (siehe Unterkapitel 3.5.7). Um eine noch stärkere Ähnlichkeit zu einer telefonischen Befragung herzustellen, könnte bei zukünftigen kognitiven Interviews auf ein vierstufiges Antwortformat zurückgegriffen werden.<sup>61</sup> Im Rahmen der kognitiven Interviews könnten die Skalenpunkte den TeilnehmerInnen dann nur mündlich erläutert werden, ohne sie schriftlich darzubieten. Diese Vorgehensweise ist insbesondere dann empfehlenswert, wenn mit mehreren unterschiedlichen Antwortoptionen gearbeitet wird (z.B. (a) „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“; (b) „sehr unwahrscheinlich“ bis „sehr wahrscheinlich“ oder (c) „sehr selten“ bis „sehr häufig“) und man testen möchte, ob der Wechsel der Skalen den Befragten Schwierigkeiten bereitet.

Im Zusammenhang mit Pretest-Phase 3 lief die Skalengenerierung und -testung in zufriedenstellender Weise ab. Mit Hilfe der Analyse der Pretest-Daten konnte erfolgreich eine 12-Item-Skala ausgewählt und ihre psychometrische Güte sowie ihre Multi-Mode-Capability analysiert werden.

## 4.2 Fazit zu den inhaltlichen Ergebnissen

Insgesamt sprechen die Ergebnisse dafür, dass es gelungen ist, mit der FWSEM-Skala ein zweidimensionales, hoch reliables und valides Instrument zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in Bezug auf Mütter zu entwickeln.

Dass die Skala dabei auch tatsächlich etwas Neues abbildet, konnte mit Hilfe zweier simultaner explorativer Faktorenanalysen aufgezeigt werden, die neben den FWSEM-Items auch verschiedene andere Sexismus-Items umfassten: Sowohl in Stichprobe 6 als

---

<sup>61</sup> In einer telefonischen Befragung werden üblicherweise lediglich vier Antwortoptionen verwendet, da eine größere Anzahl an Kategorien oftmals die Behaltensleistung der Befragten übersteigt.

auch in Stichprobe 8 bildeten die beiden Subskalen der FWSEM-Skala jeweils einen eigenständigen Faktor. In Bezug auf den feindseligen Sexismus ist dabei anzumerken, dass die drei Items der Geschlechterrollen-Batterie des ISSP, die sich mit den Konsequenzen der Erwerbstätigkeit der Mutter für die Familie befassen, gemeinsam mit den Items der FS-Subskala auf einem Faktor laden. Hier scheint es eine inhaltliche Überschneidung zu geben, vollkommen deckungsgleich sind die beiden Skalen aber nicht, da sonst alle Items gemeinsam auf einem Faktor hätten laden müssen. Der Wohlwollende Sexismus, der die besondere Eignung der Mutter für die traditionell weiblichen Tätigkeitsbereiche hervorhebt und eine subjektiv positiv empfundene Begründung für die traditionelle Arbeitsteilung liefert, scheint hingegen etwas komplett Neues zu repräsentieren. Keines der anderen Items besitzt auf diesem Faktor substantielle Ladungen.

Da der feindselige Sexismus offene Vorurteile erfasst, die als sozial unerwünscht gelten und daher auch seltener geäußert werden (vgl. Campbell, Schellenberg, & Senn, 1997, S. 89; Eckes & Six-Materna, 1998, S. 226 ff.; Swim et al., 2010, S. 138), ist es nicht verwunderlich, dass die Skalenmittelwerte in allen drei Stichproben für die FS-Subskala unterhalb der Skalenmitte, also im verneinenden, nicht-sexistischen Bereich liegen. Dies wird auch in den Besetzungen der Randverteilung erkennbar: Während nur sehr wenige (Studie 7 und 8) bzw. überhaupt keine Befragten (Studie 6) durchgängig den Items des FS-Subskala „voll und ganz zustimmten“, wählten relativ viele Personen bei sämtlichen Items die Antwortkategorie „stimme überhaupt nicht zu“. Die Verteilung ist folglich eher links-gipflig und somit leicht rechtsschief. Signifikant war die Abweichung von der Normalverteilung allerdings nur in Studie 8 im Falle der Gesamtstichprobe sowie in der weiblichen Teilstichprobe. In Studie 6 und 7 sowie in der männlichen Teilstichprobe in Studie 8 kann von einer Normalverteilung der Skalenwerte ausgegangen werden.

Der wohlwollende Sexismus stellt eine verdeckte, subtile Form des Sexismus dar. Dementsprechend wird er auch von den Befragten in der Regel nicht als etwas Negatives wahrgenommen und eher offen mitgeteilt (vgl. Glick & Fiske, 2001a). Dies spiegelt sich auch in den Skalenmittelwerten der WS-Subskala wider, die höher ausfallen als bei der FS-Subskala. In der Bevölkerungsstichprobe (Studie 8), neigten die Befragten eher dazu den Items zuzustimmen, so dass der Skalenmittelwert oberhalb der Skalenmitte lag. In Stichprobe 6 und 7 war dies hingegen nicht der Fall. Eine Analyse der Kurtosis-Werte ergab, dass die WS-Subskala tendenziell flacher ausfällt als eine Normalverteilung (signifikant war die Abweichung jedoch nur in der Gesamtstichprobe von Studie 8). Dies kommt auch in der Verteilung der Skalenwerte der WS-Subskala zum Ausdruck. In Stichprobe 8 fällt auf, dass alle Antwortkategorien relativ gleich stark besetzt sind. Betrachtet man allerdings nur die Randverteilungen so zeigt sich in Stichprobe 8 (Bevölkerungsstichprobe) ein anderes Bild als in Stichprobe 6 und 7. So fällt auf, dass sich in den beiden letztgenannten Stichproben nur eine einzige Person durchgängig wohlwollend-sexistisch äußerte und folglich allen Items „voll und ganz zustimmte“. Im Gegensatz dazu traf dies in Studie 8 auf 7.8% der Befragten zu (das waren 3% mehr als die Personen, die durchweg allen Aussagen „überhaupt nicht zustimmten“ und sich somit nicht-sexistisch äußerten).

Während sich also folglich in der Bevölkerungsstichprobe die Befragten tendenziell eher wohlwollend-sexistisch als nicht-sexistisch äußerten (siehe hierzu auch Ergebnisse in Bezug auf die Skalenmittelwerte), gilt dies für Stichprobe 6 (Online-Befragung) und 7 (schriftliche Befragung) nicht. Dieses Ergebnis ist vermutlich auf die unterschiedliche Zusammensetzung der Stichproben zurückzuführen. So waren die Befragten in Studie 6 und 7 (1) *hoch gebildet* (über 90% der Befragten besaßen in Stichprobe 6 das Abitur oder ein abgeschlossenes Hochschulstudium; Stichprobe 7 war eine Studierenden-Stichprobe), (2) *deutlich jünger* (Studie 6:  $M_{Alter} = 26.81$ ; Studie 7:  $M_{Alter} = 25.31$ ; Studie 8:  $M_{Alter} = 54.66$ ), (3) *seltener Erwerbstätigkeit* (Studie 6: 26.1%; Studie 8: 54.7%) und (4) *häufiger kinderlos* (Studie 6: 86%; Studie 7: 94.4%; Studie 8: 19.4%) als die TeilnehmerInnen in Studie 8. Insbesondere im Hinblick auf den Bildungsgrad und das Alter der Befragten in Studie 6 und 7 ist anzunehmen, dass dies einen Einfluss auf die Antworttendenzen in der Form hat, dass ältere Personen und solche mit einem niedrigerem Abschluss eher sexistische Einstellungen äußern.

Diese Vermutung wurde anhand der Daten von Stichprobe 8 überprüft. Es zeigte sich, dass Bildung ( $r = -.33, p < .01$ ), Alter ( $r = .18, p < .01$ ) und Erwerbstätigkeit ( $r = -.16, p < .01$ ) signifikant in den erwarteten Richtungen mit wohlwollendem Sexismus korreliert. Auf die Variable Elternschaft trifft dies nicht zu ( $r = .07, p = .16$ ). Eine Regressionsanalyse mit wohlwollendem Sexismus als abhängiger Variable und Bildung, Alter und Elternschaft als unabhängige Variablen, wies lediglich Bildung als signifikanten Prädiktor aus ( $\beta = -.31, t(354) = -6.17, p < .01$ ). Dieser Befund lässt sich folgendermaßen interpretieren: Je niedriger der Bildungsabschluss einer Person ist, desto stärker neigt sie zum wohlwollenden Sexismus. Bildung erklärte dabei einen signifikanten Anteil der Varianz von wohlwollendem Sexismus ( $R^2 = .12, F(3, 354) = 16.69, p < .01$ ). Da Bildung dasjenige Merkmal ist, in dem sich Stichprobe 6 und 7 von Stichprobe 8 am deutlichsten unterscheiden, stützt dieses Ergebnis die oben genannte Vermutung.

Die Reliabilitäten der FWSEM-Gesamtskala und der beiden Subskalen erzielten sowohl in den drei Gesamtstichproben als auch in den weiblichen und männlichen Teilstichproben ein sehr zufriedenstellendes Niveau. Gleiches trifft auch auf die Trennschärfen der einzelnen Items zu. Lediglich in der Teilstichprobe der Männer weisen in Studie 7 zwei Items eine Trennschärfe  $< .40$  auf. Alle übrigen korrigierten Item-Skalenwert-Korrelationen liegen über diesem Wert und sind demnach als gut einzustufen.

In explorativen Faktorenanalysen erwies sich die Skala als zweidimensional, mit einem feindselig- und einem wohlwollend-sexistischen Faktor. Dabei sind alle Ladungen als rein zu bezeichnen, d.h. sie besitzen hohe Hauptladungen auf dem eigenen und nur geringe Nebenladungen auf dem anderen Faktor. Auch im Rahmen der konfirmatorischen Faktorenanalysen zeigte das Zwei-Faktoren-Modell eine bessere Anpassung an die Daten als das Ein-Faktor-Modell. Feindseliger und wohlwollender Sexismus weisen dabei in allen Stichproben eine signifikante, positive Korrelation auf. Einzige Ausnahme bildet die männlich Teilstichprobe in Studie 7: Hier ist die Korrelation der beiden Subskalen

nicht signifikant von Null verschieden.<sup>62</sup> Insgesamt gesehen deuten die Ergebnisse darauf hin, dass feindseliger und wohlwollender Sexismus zwei distinkte, aber dennoch verbundene Elemente eines komplementären Einstellungssystems darstellen.

Im Hinblick auf die Geschlechtsunterschiede in den Skalenmittelwerten wurde erwartet, dass Männer signifikant höhere Werte aufweisen und somit sexistischer eingestellt sind als Frauen. Allerdings wurde vermutet, dass die Differenzen nicht so deutlich ausfallen wie dies bei anderen Sexismus-Skalen der Fall ist, da sich die FWSEM-Skala nicht auf die Gesamtgruppe der Frauen bezieht, sondern auf zwei weibliche Subgruppen (berufstätige vs. nicht-berufstätige Mütter). Folglich ist es durchaus vorstellbar, dass auch Frauen anderen Frauen gegenüber sexistisch eingestellt sind, wenn sie nicht der gleichen Subgruppe angehören. So könnten Vollzeit-Mütter beispielsweise berufstätige Mütter abwerten oder umgekehrt.

Entgegen den Erwartungen erwies sich keine der Mittelwertsdifferenzen zwischen weiblichen und männlichen Befragten als signifikant. Im Zusammenhang mit diesen Ergebnissen ist Folgendes anzumerken: Eine Berechnung der Geschlechtsunterschiede für die vier bzw. drei Sexismus-Skalen, die in Stichprobe 6 und 8 zur Konstruktvalidierung eingesetzt wurden, ergaben sich jeweils nur für die Moderne Sexismus-Skala signifikante Unterschiede in der erwarteten Richtung. D.h. hier leugneten männliche Befragte in stärkerem Maße als weibliche Befragte, dass die Diskriminierung von Frauen immer noch besteht (S6:  $t(201) = -2.12, p < .05$ ; S8:  $t(395) = -3.44, p < .01$ ). In Stichprobe 7 wurde nur die ASS mit ihren beiden Subskalen neben der FWSEM-Skala eingesetzt. Dort besaßen die Männer lediglich für die ASS-Gesamtskala und die HS-Subskala signifikant höhere Mittelwerte als die Frauen, nicht jedoch in Bezug auf die BS-Subskala. Demnach ist es durchaus möglich, dass die Ergebnisse für die FWSEM-Skala auf Merkmale der Stichproben zurückzuführen sind. So waren die Befragten in Stichprobe 6 und 7 hoch gebildet und demzufolge auch egalitärer eingestellt (siehe Ausführungen oben). Dies könnte dazu geführt haben, dass ein Unterschied im Antwortverhalten zwischen den weiblichen und männlichen StudienteilnehmerInnen nicht mehr erkennbar war.

Ebenso denkbar ist es, dass Geschlechtsunterschiede im Falle der FWSEM-Skala kein geeigneter Indikator für die Konstruktvalidität sind, da sich die Skala eben nicht auf Frauen im Allgemeinen bezieht, sondern auf unterschiedliche Subgruppen von Müttern. Wie bereits in Unterkapitel 3.2.1 und 3.2.2 erläutert, ist anzunehmen, dass die sexistischen Einstellungen folglich nicht strikt entlang der Geschlechtergrenzen verlaufen und sich nicht nur auf Männer beschränken, sondern sich durchaus auch Frauen sexistisch äußern können. Vermutlich spielt also weniger das Geschlecht die ausschlaggebende Rolle, son-

---

<sup>62</sup> Dieses Ergebnis geht mit bisherigen Analysen von Glick und Fiske (1996) sowie von Eckes und Six-Materna (1999) einher, die in mehreren männlichen Teilstichproben ebenfalls keinen signifikanten Zusammenhang von feindseligem und wohlwollendem Sexismus fanden.

dern eher die Tatsache, ob eine Person moderne, egalitäre Ansichten oder eben konservative Vorstellungen hinsichtlich der geschlechtsspezifischen Arbeitsteilung und insbesondere in Bezug auf die Kinderbetreuung vertritt.

Der Hypothese entsprechend besaßen die Befragten aus Westdeutschland im Hinblick auf den feindseligen Sexismus signifikant höhere Mittelwerte als die Befragten aus Ostdeutschland. Personen aus den alten Bundesländern waren somit berufstätigen Müttern gegenüber negativer eingestellt als Personen aus den neuen Bundesländern. Hinsichtlich des benevolenten Sexismus kehrte sich das Bild unerwarteterweise um: So wiesen die Ostdeutschen signifikant höhere Mittelwerte auf als die Westdeutschen und sprachen Frauen somit in einem stärkeren Maße eine besondere Prädisposition für die traditionell weiblichen Familienaufgaben (insbesondere für die Kinderbetreuung) zu. Dieses Ergebnis lässt sich möglicherweise dadurch erklären, dass eine durchgängige Erwerbstätigkeit von Müttern dem ostdeutschen Leitbild entsprach, ihre Zuständigkeit für die Kindererziehung jedoch nicht in Frage gestellt wurde. Schäfgen (2000) zufolge hielt man auch in den neuen Bundesländern an der Hauptverantwortung der Frauen für den familiären Bereich fest.

Für die konvergente Validität der FWSEM-Skalen sprechen signifikante, positive Korrelationen mit der ASS und ihren beiden Subskalen sowie mit drei weiteren Sexismus-Skalen (MSS, GRB und AMMSA). Am höchsten fallen die Korrelationen zwischen der FS-Subskala und der Geschlechterrollen-Batterie (GRB) bzw. der Traditionellen Sexismus-Skala aus, welche beide ebenfalls offen sexistische Einstellungen erfassen. Nicht eindeutig bestätigt werden konnte der Zusammenhang zwischen wohlwollendem und modernem Sexismus: Während in Studie 6 sowohl in der Gesamtstichprobe als auch in beiden Teilstichproben signifikante Zusammenhänge vorlagen, war dies in Studie 8 nicht der Fall. Denkbar wäre, dass dieses Ergebnis mit der verwendeten Kurzversion der MS-Skala in Studie 8 zusammenhängt, die lediglich aus zwei Items besteht (Item 1 „Die Diskriminierung von Frauen ist in Deutschland immer noch ein Problem“ und Item 2 „Die jetzige Beschäftigungspolitik benachteiligt die Frauen“). Möglicherweise kann sie Vorurteile gegenüber Frauen nicht in angemessener Weise erfassen. In Unterkapitel 2.4.2.2 wurde am Konzept des Modernen Sexismus kritisiert, dass es mitunter zu Fehleinschätzungen kommen könnte, wenn die Befragten eine fortgesetzte Diskriminierung von Frauen leugnen, da sie einfach nicht wissen, dass beispielsweise weibliche Angestellte immer noch weniger Lohn erhalten als männliche Angestellte. Folglich würde man den Befragten auf Grund ihrer Unwissenheit eine modern sexistische Einstellung unterstellen, ohne dass sie auch ihre Überzeugung treffend widerspiegelt. Derartige Fehleinstufungen treten im Falle von zwei Items vermutlich häufiger auf, da die Bandbreite der angesprochenen Themen einfach deutlich geringer ist. Für die Vermutung einer geringen Validität der Skala spricht die Tatsache, dass die verwendete MS-Skala in Studie 8 mit der FS-Subskala nur eine vergleichsweise schwache Korrelation besitzt ( $r = .20, p < .01$ ) und mit der AMMSA-Skala überhaupt kein signifikanter Zusammenhang besteht ( $r = .05, p = .40$ ). Gleiches trifft auch auf acht der zehn Elemente des Syndroms

Gruppenbezogener Menschenfeindlichkeit zu: Eine signifikante Korrelation konnte lediglich mit klassischem Sexismus ( $r = .25, p < .01$ ) und der Abwertung Homosexueller ( $r = .15, p < .01$ ) gefunden werden.

Gleichzeitig deckt sich der Befund aus Studie 8 sowohl mit den Ergebnissen von Glick und Fiske (1996) als auch mit denen von Eckes und Six-Materna (1999). In ihren Untersuchungen fanden die AutorInnen überhaupt keine signifikanten (Partial-) Korrelationen zwischen der BS-Subskala und der MSS. Für zukünftige Untersuchungen wäre es daher interessant, sich die Verbindung dieser beiden Sexismus-Facetten (evtl. auch im Rahmen von qualitativen Interviews) noch genauer anzusehen.

Neben den Sexismus-Skalen wurden in Studie 6 und 8 eine Autoritarismus-Skala (RWA) und eine Skala zur Sozialen Dominanzorientierung (SDO) eingesetzt, um die konvergente Validität der FWSEM-Skala zu testen. Der Hypothese entsprechend ergaben sich sowohl für die Gesamtskala als auch beide Subskalen signifikante positive Zusammenhänge mit der RWA- und der SDO-Skala. In der Teilstichprobe der Frauen zeigten sich insbesondere in Bezug auf den wohlwollenden Sexismus signifikante Korrelationen mit RWA und SDO. Der Zusammenhang zwischen feindseligem Sexismus und Right-Wing-Autoritarismus bzw. Sozialer Dominanzorientierung scheint dahingegen fast vollständig auf den Einfluss des wohlwollenden Sexismus zurückzuführen zu sein (wie die Partialkorrelationen verdeutlichen). In der männlichen Teilstichprobe besteht ebenfalls in beiden Studien ein signifikanter positiver Zusammenhang zwischen der WS-Subskala und der RWA-Skala, nicht jedoch mit der SDO-Skala. Wohlwollender Sexismus scheint folglich bei Männern eher mit dem Festhalten an konservativen Werten verbunden zu sein als mit einer Befürwortung gruppenbasierter Ungleichheiten. Im Gegensatz zu den Frauen hängt bei den Männern auch der feindselige Sexismus eindeutig mit beiden Einstellungsformen zusammen. So erweisen sich auch bei Konstanthaltung des Einflusses des wohlwollenden Sexismus die Mehrzahl der Korrelationen zwischen feindseligem Sexismus und Right-Wing-Autoritarismus bzw. Sozialer Dominanzorientierung als signifikant.

Des Weiteren wurden in Studie 8 die Korrelationen der FWSEM-Gesamtskala und der beiden Subskalen mit einer Reihe anderer Vorurteils-Skalen untersucht, die gemeinsam das „Syndrom Gruppenbezogener Menschenfeindlichkeit“ bilden (Heitmeyer, 2002, 2010). Es wurde davon ausgegangen, dass auch feindseliger und wohlwollender Sexismus mit diesen Ideologien der Ungleichwertigkeit zusammenhängen und dementsprechend signifikante positive Korrelationen mit allen Elementen des Syndroms aufweisen. Für die Gesamtstichprobe konnte diese Hypothese für alle drei FWSEM-Skalen bestätigt werden. Abermals lieferte eine getrennte Analyse nach Geschlecht aufschlussreiche Ergebnisse im Hinblick auf die Partialkorrelationen. Während in der weiblichen Teilstichprobe eher der wohlwollende Sexismus mit den einzelnen Elementen des Syndroms Gruppenbezogener Menschenfeindlichkeit zusammenhing, gilt dies in der männlichen Teilstichprobe eher für den feindseligen Sexismus. Folglich neigen Männer, die berufstätigen Müttern gegenüber

eine Antipathie hegen, auch eher dazu, andere Personengruppen (z.B. Zuwanderer, Juden, Muslime, Homosexuelle) abzuwerten. Bei Frauen trifft dies hingegen auf diejenigen zu, die die Einstellung vertreten, Mütter – und eben nicht die Väter – sollten aufgrund ihrer natürlichen Prädisposition die traditionell weibliche Geschlechterrolle der Kinderbetreuung übernehmen.

Die zuvor geschilderten Ergebnisse im Zusammenhang mit gruppenbezogenen Ideologien (RWA, SDO und GMF) deuten darauf hin, dass für Frauen vor allem der wohlwollende Sexismus für Männer hingegen insbesondere der feindselige Sexismus mit anderen Vorurteilen und einer Ideologie der Ungleichwertigkeit zusammenhängt. Dieser Befund sollte in zukünftigen Studien tiefergehend untersucht werden. Eine mögliche Erklärung wäre, dass im Falle der Männer der wohlwollende Sexismus mitunter auf ein mangelndes Selbstvertrauen in Bezug auf die Kinderbetreuung zurückzuführen ist, und weniger dem Wunsch entspricht, die bestehende Vorrangstellung des eigenen Geschlechts aufrechtzuerhalten. Auch wenn diese Denkweise letztendlich Frauen die Verantwortung für den familiären Bereich zuschreibt und somit die bestehende Geschlechterhierarchie festigt, würde dies erklären, weshalb der benevolente Sexismus nicht für alle männlichen Befragten mit der Abwertung und einem Ungleichheitsdenken in Bezug auf andere Personengruppen einhergeht. Diese Interpretation wird durch die nicht signifikante Korrelation von WS und SDO in der männlichen Teilstichprobe gestützt.

Bei Frauen hingegen ist es durchaus vorstellbar, dass der wohlwollende Sexismus mit einem Ungleichheitsdenken zu tun hat, der in ein konservatives Einstellungssystem eingebettet ist. So wird durch den wohlwollenden Sexismus die Überlegenheit des weiblichen gegenüber dem männlichen Geschlecht im häuslich-familiären Bereich herausgestellt. Die durchweg signifikant positiven Korrelationen des wohlwollenden Sexismus mit Autoritarismus und sozialer Dominanzorientierung in der Teilstichprobe der Frauen bestärken diese Vermutung. Es scheint folglich nicht verwunderlich, dass der benevolente Sexismus deutliche Zusammenhänge mit den einzelnen Elementen des Syndroms Gruppenbezogener Menschenfeindlichkeit aufweist.

Im Hinblick auf die Tendenz von Befragten, ein sozial erwünschtes Antwortverhalten zu zeigen, konnte keine systematische Verzerrung im Zusammenhang mit den FWSEM-Skalen festgestellt werden. So ergaben sich, der Hypothese entsprechend, weder für die Gesamtskala noch für beide Subskalen signifikante Korrelationen mit der ST(Selbsttäuschungs)-Skala. Im Hinblick auf eine etwaige Fremdtäuschung (FT) kann leider keine Aussage getroffen werden, da die FT-Skala aufgrund ihrer sehr geringen Reliabilität nicht in die Analysen miteinbezogen wurde. In zukünftigen Untersuchungen, sollte daher das bisherige Ergebnis mit Hilfe anderer Skalen zur Messung eines sozial erwünschten Antwortverhaltens noch weiter abgesichert werden.

Studie 7 diente der Überprüfung der Übereinstimmungsvalidität der FWSEM-Skala. Wie erwartet, korrelierte der feindselige Sexismus mit einem komplett negativen Bild von Karrierefrauen. Unerwarteterweise trifft dies darüber hinaus auch für die positiven Einstel-



lungen in Bezug auf Hausfrauen zu. So ist feindseliger Sexismus (auch bei Konstanthaltung des Einflusses des wohlwollenden Sexismus) signifikant positiv mit den feminin-positiven Merkmalszuschreibungen korreliert. Folglich scheint die FS-Subskala sowohl in Bezug auf Karrierefrauen als auch in Bezug auf Hausfrauen eine Übereinstimmungsvalidität zu besitzen. Je stärker eine Person zum feindseligen Sexismus neigt, desto stärker spricht sie (1) *Karrierefrauen* feminin-positive Eigenschaften ab und gleichzeitig maskulin-negative Eigenschaften zu und desto stärker attestiert sie (2) *Hausfrauen* feminin-positive Eigenschaften. Dieses Ergebnis ist insofern nachvollziehbar, als dass beide Ansichten von Frauen gut miteinander vereinbar sind: nicht-traditionelle Frauentypen werden „bestraft“ und traditionelle Frauentypen „belohnt“ (vgl. hierzu Glick & Fiske, 2001a).

Für den wohlwollenden Sexismus zeigt sich ebenfalls, der Hypothese entsprechend, ein positiver Zusammenhang mit der Zuschreibung von feminin-positiven Eigenschaften zur Subgruppe der Hausfrauen. Allerdings ist anzumerken, dass die Partialkorrelation für die WS-Subskala, bei der der Einfluss der FS-Subskala auspartialisiert wurde, nicht mehr signifikant war. Möglicherweise ist dieses Ergebnis darauf zurückzuführen, dass in Studie 7 nach der Bewertung von Hausfrauen gefragt wurde, sich die WS-Subskala jedoch auf Vollzeit-Mütter bezieht. Denkbar wäre in diesem Zusammenhang auch, dass insbesondere für Studierende der Begriff „Hausfrau“ vielleicht sogar negative Assoziationen geweckt hat und sie eher an ein „Heimchen am Herd“ gedacht haben, als eine Frau, die vielleicht nach der Geburt ihres Kindes eine Zeit lang aus dem Berufsleben ausgestiegen ist, um sich der Betreuung des Kindes voll und ganz widmen zu können. Folglich sollte die Übereinstimmungsvalidität der WS-Subskala in zusätzlichen Studien noch näher untersucht werden, in denen auch explizit eine Bewertung von (nicht-)berufstätigen *Müttern* vorgenommen wird. Aufschlussreich wäre es in diesem Kontext auch, Vergleiche mit der ASS von Eckes & Six-Materna (1999) anzustellen, um Hinweise auf die diskriminante Validität zu erhalten.

Für die Bewertung der Karriere- und Hausfrauen konnten in Studie 7 bereits beide ambivalenten Sexismus-Skalen miteinander verglichen werden. Es zeigte sich, dass sowohl die FWSEM-Skala als auch die ASS die Korrelationen aufweisen, die den Hypothesen entsprechen. Die Korrelationsmuster unterscheiden sich allerdings in folgenden Punkten:

- (1) Im Falle der ASS geht hostiler Sexismus (HS-Subskala) mit einem komplett negativen Bild in Bezug auf Karrierefrauen einher, das unerwarteterweise auch die Zuschreibung von feminin-negativen Eigenschaften beinhaltet. Hostile Sexisten sehen Karrierefrauen somit nicht nur u.a. als egoistisch, sondern auch zugleich als unterwürfig und leichtgläubig an.
- (2) Im Hinblick auf die Hausfrauen zeigt sich für die HS-Subskala, im Gegensatz zur FWSEM-Skala, keine Übereinstimmungsvalidität mit der Zuschreibung von feminin-positiven Merkmalen.

- (3) Der benevolente Sexismus (BS-Subskala) korreliert bezüglich der Karrierefrauen signifikant negativ mit feminin-negativen Eigenschaftszuschreibungen und positiv mit maskulin-positiven Eigenschaftszuschreibungen. Folglich besitzen Personen, die benevolent sexistisch eingestellt sind, ein positives Bild von Karrierefrauen, das sie u.a. als unabhängig und selbstsicher auszeichnet. Dieses Ergebnis widerspricht der Theorie des benevolenten Sexismus insofern, als dass benevolenter Sexismus Frauen belohnt, die die traditionell weibliche Geschlechterrolle erfüllen – Karrierefrauen gehören allerdings eindeutig einer Subgruppe von Frauen an, die aus der traditionellen Rolle ausbrechen.
- (4) Bezüglich der Hausfrauen hängt benevolenter Sexismus nicht nur mit einer Zuschreibung feminin-positiver Merkmale zusammen, wie dies vorhergesagt wurde, sondern es ergeben sich darüber hinaus signifikante negative Partialkorrelationen mit maskulin-negativen Eigenschaftszuschreibungen (prahlerisch, überheblich, egoistisch und feindselig).

Beide Skalen scheinen somit eine gewisse diskriminante Validität zu besitzen, da die Korrelationsmuster nicht vollständig einander gleichen. Für die Zukunft wären ergänzende Studien bezüglich der Vorhersagevalidität beider Skalen im Vergleich wünschenswert, um den Neuigkeitsgehalt der FWSEM-Skala weiter zu unterstreichen.

Abschließend wird auf eine Studie von Süßenbach und Bohner (2011) verwiesen, bei der die neu entwickelte FWSEM-Skala zum Einsatz kam.<sup>63</sup> Anhand einer Bevölkerungstichprobe untersuchten die Autoren die gegenwärtige Akzeptanz von Vergewaltigungsmythen. Es zeigte sich, dass ambivalenter Sexismus (FWSEM-Gesamtskala) verglichen mit einer Reihe weiterer Ideologien der Ungleichwertigkeit der stärkste Prädiktor für die Akzeptanz von Vergewaltigungsmythen war ( $\beta = .17$ ,  $t(385) = 3.40$ ,  $p = .001$ ). Eine getrennte Analyse für beide Subskalen zeigte auf, dass insbesondere der wohlwollende Sexismus (WS-Subskala) ein starker Prädiktor für die Akzeptanz von Vergewaltigungsmythen ist ( $\beta = .19$ ,  $t(384) = 4.25$ ,  $p = .001$ ), wohingegen der feindselige Sexismus (FS-Subskala) allein kein signifikanter Prädiktor war ( $\beta = .00$ ,  $t(384) = -0.00$ ,  $p = .99$ ). Dieses Ergebnis steht im Einklang mit den Befunden von Abrams, Viki, Masser und Bohner (2003), die in ihrer Studie die ASS einsetzten. Dort sagte benevolenter Sexismus, nicht jedoch hostiler Sexismus, Verantwortungszuschreibungen an das weibliche Vergewaltigungsoffer vorher, wenn dieses sich nicht gemäß der traditionellen Rolle verhalten hatte.

Alles in allem sprechen die Ergebnisse der acht vorgestellten Studien dafür, dass es gelungen ist, mit der FWSEM-Skala ein zweidimensionales, hoch reliables und valides Messinstrument zur Erfassung feindselig- und wohlwollend-sexistischer Einstellungen in

---

<sup>63</sup> Dabei handelt es sich um dieselbe Befragung wie in Studie 8, die im Rahmen des „GMF (Gruppenbezogene Menschenfeindlichkeit) Graduierten-Surveys“ 2010 (Prof. Dr. Andreas Zick) durchgeführt und durch die Universität Bielefeld gefördert wurde.

Bezug auf Mütter zu entwickeln. Die Skala hat sich bei verschiedenen Erhebungsmodalitäten (schriftlich, telefonisch, online) bewährt und ist nicht nur in Studierenden-, sondern auch in Bevölkerungsstichproben einsetzbar. Dabei zeigte die Skala durchweg ein hohes Maß an Stabilität bezüglich der ermittelten Item- und Skalenkennwerte.

### 4.3 Forschungsperspektiven und Ausblick

Die FWSEM-Skala lässt sich in einer Vielzahl von Forschungskontexten anwenden. Dies trifft auf Einzeluntersuchungen zu, in denen beispielsweise negative wie auch subjektiv positive Einstellungen in Bezug auf (nicht-)erwerbstätige Mütter erfasst werden. Gerade dieser Forschungsbereich dürfte angesichts der aktuellen Krippendebatte und des gesellschaftlichen Diskurses im Zusammenhang mit der Einführung eines Betreuungsgeldes in Deutschland von besonderem Interesse sein. Folgende Fragestellungen könnten in diesem Kontext untersucht werden: Wie werden erwerbstätige Mütter gesehen, die ihre unter Dreijährigen in einer Kindertagesstätte betreuen lassen? Bestehen ihnen gegenüber Vorurteile, wenn sie dies nicht aus einer finanziellen Notwendigkeit heraus tun? Gelten Sie als Mütter zweiter Klasse? Wandeln sich mit dem Ausbau an öffentlichen Kinderbetreuungseinrichtungen die Vorbehalte im Westen Deutschlands gegenüber der außerhäuslichen Betreuung oder bleibt die Überzeugung bestehen, dass die Mutter nach wie vor im Idealfall selbst die Betreuung des Kindes zu Hause übernehmen sollte?

Letztere Frage knüpft an eine weitere Einsatzmöglichkeit der FWSEM-Skala an. So kann die Skala auch genutzt werden, um gesellschaftliche Veränderungen in den Einstellungen gegenüber Müttern zu beobachten und zu analysieren (gesellschaftliche Dauerbeobachtung bzw. Social Monitoring). Bislang werden in den großen (inter-) nationalen Umfragen wie der Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften (ALLBUS), dem Sozio-oekonomischen Panel (SOEP) oder dem International Social Survey Programme (ISSP) nur Geschlechterrollen-Skalen eingesetzt, in denen scheinbar wohlwollende Erklärungen für die Präferenz der traditionellen Arbeitsteilung fehlen. Subtile Einstellungen und Vorurteile können somit nicht erfasst werden – gerade dies ist jedoch wünschenswert, da offen sexistische Einstellungen aufgrund des gesellschaftlichen Diskurses bezüglich der Gleichberechtigung heutzutage immer seltener offen geäußert werden (siehe Unterkapitel 1.2).

Im Kontext repräsentativer Bevölkerungsumfragen wären zudem weiterführende Analysen zu Ost-/West-Unterschieden interessant – insbesondere im Zusammenhang mit dem wohlwollenden Sexismus, da diese Ergebnisse in Studie 8 nicht den Erwartungen entsprachen (siehe Unterkapitel 3.5.7.10). Denkbar wäre hier beispielsweise auch die Durchführung qualitativer Interviews, die die zugrundeliegenden Ansichten und Einstellungen der Befragten noch tiefergehend aufdecken und somit möglicherweise erklären könnten, weshalb Befragte aus den neuen Bundesländern wohlwollend sexistischen Aussagen stärker zustimmen als Befragte aus den alten Bundesländern. Auch die Analyse

möglicher Nord-/Südgefälle sowie die Untersuchung von Stadt- und Landunterschieden hinsichtlich der Verbreitung beider Formen von Sexismus erscheinen im Kontext mütterlicher Erwerbstätigkeit wünschenswert.

Wie bereits in Unterkapitel 4.2. erwähnt, wäre es ebenfalls aufschlussreich, die sexistischen Einstellungen von Frauen tiefergehend zu untersuchen. Es ist davon auszugehen, dass Frauen sich über andere Frauen sexistisch äußern, wenn sie nicht der Eigengruppe angehören. So könnten Vollzeit-Mütter (Eigengruppe) berufstätige Mütter (Fremdgruppe) abwerten oder umgekehrt. Es wäre interessant, diese Annahme mit Hilfe einer geeigneten Stichprobe eingehender zu prüfen. Dabei sollten die Daten idealerweise Aufschluss darüber geben, (a) welche innerfamiliären Arbeitsteilungsmuster nach der Geburt eines Kindes gewählt wurden, (b) ob diese Aufteilung auch gewünscht war und (c) wie lange die Erwerbstätigkeit für die Kinderbetreuung unterbrochen wurde.

Des Weiteren wäre eine Längsschnittuntersuchung interessant, in der etwaige Veränderungen in den Einstellungen der Befragten nach dem Übergang zur Elternschaft analysiert werden. Stimmen Frauen wohlwollend- und evtl. auch feindselig-sexistischen Aussagen in stärkerem Maße zu, nachdem sie zu einer Mutter geworden sind? Wie sieht dies bei den Vätern aus?

Auch im Rahmen von Studien, die die Unterrepräsentanz von Frauen bzw. Müttern in Führungspositionen untersuchen, kann der Einsatz der FWSEM-Skala wichtige Hinweise liefern. Der Theorie der Maternal Wall zufolge (siehe Unterkapitel 2.3.2) bewirken die Einstellungen von Personalverantwortlichen bezüglich Mutterschaft und ihren normativen Vorstellungen bezüglich einer (Nicht-)Erwerbstätigkeit verschiedene Diskriminierungsmuster. So könnte die Ansicht, dass berufstätige Mütter nicht gleichzeitig den Anforderungen an eine gute Arbeitnehmerin und an eine gute Mutter gerecht werden können in der Verhaltenserwartung münden, dass die Mutter ihre Karriere aufgeben und sich lieber auf die Betreuung des Kindes konzentrieren sollte. Es wäre aufschlussreich diese und andere Zusammenhänge näher zu beleuchten.

Als nächster Schritt ist die Übersetzung der FWSEM-Skala ins Englische und eine anschließende Testung und Validierung geplant. Dies würde den Einsatz der Skala im Rahmen internationaler Vergleiche ermöglichen. Spannend wäre es beispielsweise, die Einstellungen der Befragten bezüglich (nicht-)berufstätiger Mütter in Ländern wie etwa Frankreich, Großbritannien, Schweden und/oder Italien zu untersuchen, in denen unterschiedliche kulturelle (Mütter-)Leitbilder und strukturelle Rahmenbedingungen (beispielsweise bezüglich der Kinderbetreuungsmöglichkeiten) vorherrschen (siehe Theorie des Geschlechter-Arrangements von Pfau-Effinger (2000) in Unterkapitel 2.1.1).

Um die eingangs aufgeworfene Frage zum Stand der Geschlechtergerechtigkeit beantworten zu können, sind nach Waltraud Cornelißen (Leiterin der Abteilung Geschlechterforschung am Deutschen Jugendinstitut) aktuelle Statistiken und Umfragewerte erforderlich:

*„Eine Gesellschaft, die sich ihrem Verfassungsanspruch auf Gleichberechtigung ehrlich stellen will, braucht (...) ein fundiertes öffentlich zugängliches Wissen über den Stand der Geschlechtergerechtigkeit. Deshalb ist es wichtig, die Berichterstattung zur Geschlechtergerechtigkeit aufrecht zu erhalten und differenzierend auszubauen. So kann einer kritischen Öffentlichkeit aufgezeigt werden, auf welchem Stand sich die Gleichberechtigung in welchen Teilen der bundesdeutschen Gesellschaft tatsächlich befindet. Denjenigen, die gleichstellungspolitisch aktiv werden wollen, sind aktuelle Zahlen eine wichtige Argumentationshilfe“ (Deutsches Jugendinstitut, 2006, o.S.).*

Die FWSEM-Skala Skala kann hierzu einen relevanten Beitrag leisten. So kann sie offene, feindselige sowie subtile, (scheinbar) wohlwollende Ansichten in der Bevölkerung hinsichtlich (nicht-)berufstätiger Mütter erfassen, die wiederum individuelle Rollenbilder und Geschlechterstereotype (deskriptive wie auch präskriptive) widerspiegeln und die die Arbeitsteilung in Paarbeziehung beeinflussen. Dabei kann sie zum einen Entwicklungen über die Zeit abbilden und zum anderen den aktuellen Stand der Geschlechtergerechtigkeit auf Einstellungsebene verdeutlichen und somit aufzeigen, wo noch Handlungsbedarf besteht. Sie hat sich empirisch als hoch reliables und valides Messinstrument erwiesen, dass sich bei verschiedenen Erhebungsmodalitäten und in verschiedenen Stichproben bewährt hat. Die zuvor beschriebenen vielfältigen Einsatzmöglichkeiten der FWSEM-Skala lassen somit hoffen, dass sie die Forschungslandschaft zukünftig bereichern wird.

## 5 Literatur

- Abele, A. E., & Wojciszke, B. (2007). *Agency and communion from the perspective of self versus others*. *Journal of Personality and Social Psychology*, 93 (5), S. 751-763.
- Abrams, D., Viki, G. T., Bohner, G., & Masser, B. (2003). Perceptions of Stranger and Acquaintance Rape: The Role of Benevolent and Hostile Sexism in Victim Blame and Rape Proclivity. *Journal of Personality and Social Psychology*, 84(1), 111-125.
- Allmendinger, J., Leuze, K., & Black, J. M. (2008). 50 Jahre Geschlechtergerechtigkeit und Arbeitsmarkt. *Aus Politik und Zeitgeschichte*, 24/25, 18-25.
- Allport, G. (1954). *The Nature of Prejudice*. Reading, MA: Addison-Wesley.
- Aosved, A., Long, P., & Voller, E. (2009). Measuring Sexism, Racism, Sexual Prejudice, Ageism, Classism, and Religious Intolerance: The Intolerant Schema Measure. *Journal of Applied Social Psychology*, 39(10), 2321-2354.
- Asendorpf, J. B., & Neyer, F. J. (2012). *Psychologie der Persönlichkeit* (5., vollst. überarb. Aufl.). Berlin: Springer.
- Athenstaedt, U. (2000). Normative Geschlechtsrollenorientierung: Entwicklung und Validierung eines Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 21(1), 91-104.
- Auth, D. (2007). Pronatalistischer Aktionismus: von der bevölkerungspolitischen Instrumentalisierung und Ökonomisierung der Familienpolitik in Deutschland. In D. Auth, & B. Holland-Cunz (Hrsg.), *Grenzen der Bevölkerungspolitik: Strategien und Diskurse demographischer Steuerung* (S. 79-102). Opladen: Barbara Budrich.
- Beatty, P. (2004): Paradigms of cognitive interviewing practice, and their implications for developing standards of best practice. *ZUMA-Nachrichten Spezial: QUEST 2003 Proceedings of the 4th Conference on Questionnaire Evaluation Standards 21-23 October 2003*, 9, 8-25.
- Beatty, P., & Schechter, S. (1994). *An Examination of Mode Effects in Cognitive Laboratory Research*. Abgerufen am 30. Dezember 2012, von [http://www.amstat.org/sections/srms/proceedings/papers/1994\\_224.pdf](http://www.amstat.org/sections/srms/proceedings/papers/1994_224.pdf).
- Beatty, P. C., & Willis, G. B. (2007). Research Synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, 71(2), 287-311.
- Beere, C. A., King, D. W., Beere, D. B., & King, L. A. (1984). The Sex-Role Egalitarianism Scale: A Measure of Attitudes Toward Equality Between the Sexes. *Sex Roles*, 10(7/8), 563-576.
- Bielefeldt, H. (2010). Das Diskriminierungsverbot als Menschenrechtsprinzip. In U. Hormel, & A. Scherr (Hrsg.), *Diskriminierung: Grundlagen und Forschungsergebnisse* (S. 21-34). Wiesbaden: VS Verlag für Sozialwissenschaften.

- Blair, J., Conrad, F., Castellano Ackermann, A., & Claxton, G. (May, 2006). *The Effect of Sample Size on Cognitive Interview Findings. Proceedings of the American Statistical Association, Section on Survey Research Methods*. Alexandria, VA: American Statistical Association. Abgerufen am 10. Juni 2009, von Institute for Social Research (University of Michigan) [http://www.abtassociates.com/presentations/AAPOR06\\_Sample\\_Size\\_Cognitive\\_Interviews.pdf](http://www.abtassociates.com/presentations/AAPOR06_Sample_Size_Cognitive_Interviews.pdf).
- Blanke, K., Gauckler, B., & Sattelberger, S. (2008). Fragebogen auf dem Prüfstand: Testmethoden und deren Einsatz in der amtlichen Statistik, *Wirtschaft und Statistik*, 8, S. 641-649,
- Blohm, M, Harkness, J., Klein, S., & Scholz, E. (2003). *Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2002*. ZUMA-Methodenbericht, 12, Mannheim: Zentrum für Umfragen, Methoden und Analysen. Abgerufen am 20. Juni 2011, von <http://www.za.uni-koeln.de/data/en/issp/questionnaires/q2002/DE2002.pdf>.
- BMFSFJ (2006). *Tagesbetreuungsausbaugesetz: Optimale Betreuung und frühe Förderung*. Abgerufen am 27. November 2012, von <http://www.bmfsfj.de/BMFSFJ/Service/Archiv/16-legislatur,did=72380.html>.
- BMFSFJ (2010a). *Kinderförderungsgesetz (KiföG)*. Abgerufen am 27. November 2012, von <http://www.bmfsfj.de/BMFSFJ/kinder-und-jugend,did=133282.html>.
- BMFSFJ (2010b). *Kristina Schröder: „Der Bund hält seine Zusagen beim Kita-Ausbau ein“*. Abgerufen am 27. November 2012, von <http://www.bmfsfj.de/BMFSFJ/kinder-und-jugend,did=161776.html>.
- BMFSFJ (2013). *Bundeselterngeld- und Elternzeitgesetz*. Abgerufen am 20. Januar 2013, von <http://www.bmfsfj.de/BMFSFJ/gesetze,did=93110.html>.
- BMFSFJ (Hrsg.) (2003). *Die Familie im Spiegel der amtlichen Statistik. Lebensformen, Familienstrukturen, wirtschaftliche Situation der Familien und familiendemografische Entwicklungen in Deutschland* (Erweiterte Neuauflage 2003). Erstellt im Auftrag des Bundesministeriums für Familie, Senioren, Frauen und Jugend ( BMFSFJ ) in Zusammenarbeit mit dem Statistischen Bundesamt. Berlin: DruckVogt GmbH. Abgerufen am 10. Januar 2011, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Broschuerenstelle/Pdf-Anlagen/PRM-24184-Gesamtbericht-Familie-im-Spieg,property=pdf.pdf>.
- BMFSFJ (Hrsg.) (2004). *Frauen in Deutschland. Von der Frauen- zur Gleichstellungspolitik*. Abgerufen am 13. April 2011, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Broschuerenstelle/Pdf-Anlagen/Frauen-in-Deutschland,property=pdf,bereich=bmfsfj,sprache=de,rwb=true.pdf>.
- BMFSFJ (Hrsg.) (2008a). *Beruflicher Wiedereinstieg nach der Familiengründung. Bedürfnisse, Erfahrungen, Barrieren*. Berlin. Abgerufen am 19. Januar 2009, von

- <http://www.bmfsfj.de/RedaktionBMFSFJ/Pressestelle/Pdf-Anlagen/beruflicher-wiedereinstieg-nach-familiengruendung-sinus-studie,property=pdf,rwb=true.pdf>.
- BMFSFJ (Hrsg.) (2008b). *Perspektive Wiedereinstieg. Ziele, Motive und Erfahrungen von Frauen vor, während und nach dem beruflichen Wiedereinstieg. Quantitative Repräsentativuntersuchung von Sinus Sociovision im Auftrag des BMFSFJ*. Berlin. Abgerufen am 14. Januar 2013, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Broschuerenstelle/Pdf-Anlagen/Perspektive-Wiedereinstieg-Ziele-Motive-Erfahrungen,property=pdf,bereich=bmfsfj,sprache=de,rwb=true.pdf>.
- BMFSFJ (Hrsg.) (2012). *Familienreport 2011. Leistungen, Wirkungen, Trends*. Berlin. Abgerufen am 27. November 2012, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Broschuerenstelle/Pdf-Anlagen/Familienreport-2011,property=pdf,bereich=bmfsfj,sprache=de,rwb=true.pdf>.
- BMFSFJ, BMBF, BMVBS, BMAS, & BMWI (Hrsg.) (2006). *2. Bilanz Chancengleichheit. Frauen in Führungspositionen. Zweite Bilanz der Vereinbarung zwischen der Bundesregierung und den Spitzenverbänden der Deutschen Wirtschaft zur Förderung der Chancengleichheit von Frauen und Männern in der Privatwirtschaft*. Abgerufen am 11. Mai 2011, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Abteilung4/Pdf-Anlagen/bilanz-chancengleichheit,property=pdf,bereich=,rwb=true.pdf>.
- BMFSFJ, Statistisches Bundesamt (Hrsg.) (2003). *Wo bleibt die Zeit? Die Zeitverwendung der Bevölkerung in Deutschland 2001/02*. Abgerufen am 10. Januar 2011, von [https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/WobleibtdieZeit5639101029004.pdf;jsessionid=9BFFF-F352C28A29A2E4677706DCFEEE5.cae1?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/WobleibtdieZeit5639101029004.pdf;jsessionid=9BFFF-F352C28A29A2E4677706DCFEEE5.cae1?__blob=publicationFile).
- Bohner, G. (2003). Einstellungen. In W. Stroebe, K. Jonas, & M. Hewstone (Hrsg.), *Sozialpsychologie. Eine Einführung* (4. Aufl., S. 265-315). Berlin [u.a.]: Springer.
- Bohner, G., & Wänke, M. (2002). *Attitudes and attitude change*. Hove, UK: Psychology Press.
- Bortz, J. (2005). *Statistik für Human- und Sozialwissenschaftler* (6. Aufl.). Heidelberg: Springer.
- Bradburn, N., Sudman, S., & Wansink, B. (2004). *Asking Questions. The Definitive Guide to Questionnaire Design – For Market Research, Political Polls, and Social and Health Questionnaires*. San Francisco: Jossey-Bass.
- Braun, M. (2006). *Funktionale Äquivalenz in interkulturell vergleichenden Umfragen: Mythos und Realität*. Mannheim: Zentrum für Umfragen, Methoden und Analysen (ZUMA).



- Braun, M. (2008). Using Egalitarian Items to Measure Men's and Women's Family Roles. *Sex Roles*, 59(8/9), 644-656.
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, 86(2), 307-324.
- Brewer, M. B. (2001). Ingroup identification and intergroup conflict. When does ingroup love become outgroup hate? In R. D. Ashmore, L. Jussim, & D. Wilder (Hrsg.), *Social identity, intergroup conflict, and conflict reduction* (S. 17-41). New York: Oxford University Press.
- Bridges, J. S., & Etaugh, C. (1995). College Students' Perceptions of Mothers: Effects of Maternal Employment-Childrearing Pattern and Motive for Employment. *Sex Roles*, 32 (11/12), 735-751.
- Brogan, D., & Kutner, N. G. (1976). Measuring sex-role orientations: A normative approach. *Journal of marriage and the family*, 38, 31-40.
- Cameron, C. (1977). Sex-role attitudes. In S. Oskamp (Ed.), *Attitudes and opinions* (pp. 339-359). Englewood Cliffs, NJ: Prentice Hall.
- Campbell, B. E., Schellenberg, G., & Senn C. Y. (1997). Evaluating measures of contemporary sexism. *Psychology of Women Quarterly*, 21, 89-102.
- Cannell, C., Oksenberg, L., Kalton, G., Bischooping, K., und Fowler, F. J. (1989). *New Techniques for Pretesting Survey Questions. Final Report*. Ann Arbor: The University of Michigan; Survey Research Center. Boston: University of Massachusetts; Center for Survey Research.
- Caspari, L. (2011, 30. März). *Der Glaubenskrieg um die Frauenquote*. Zeit Online. Abgerufen am 10. November 2011, von <http://www.zeit.de/politik/deutschland/2011-03/frauenquote-spitzentreffen-schroeder>.
- Center for Work Life Law (Work Life Law UC Hastings College of the Law) (2006). *Issue Brief. Current Law Prohibits Discrimination Based on Family Responsibilities and Gender Stereotyping*. Abgerufen am 11. Februar 2011, von <http://www.worklifelaw.org/pubs/IssueBriefFRD.pdf>.
- Clark, H. H., & Schober, M. F. (1992). Asking questions and influencing answers. In J. M. Tanur (Hrsg.), *Questions about questions: Inquiries into the cognitive bases of surveys* (S. 15-48). New York: Russell Sage.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*. 12, 229-238.
- Conrad, F., & Blair, J. (1996). From impressions to data: Increasing the objectivity of cognitive interviews. *Proceedings of the Section on Survey Research Methods. American Statistical Association*, 1-9.

- Conrad, F. G., & Blair, J. (2009). Sources of Error in Cognitive Interviews. *Public Opinion Quarterly*, 73(1), 32-55.
- Cordes, M. (2008). Gleichstellungspolitiken: Von der Frauenförderung zum Gender Mainstreaming. In R. Becker, & B. Kortendiek (Eds.), *Geschlecht & Gesellschaft: Bd. 35. Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie* (2. Aufl., S. 916-924). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Cornelißen, W. (Hrsg.) (2005). *Gender-Datenreport. 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland*. München.
- Cornelißen, W., Dressel, C., & Lohel, V. (2005). Erwerbseinkommen von Frauen und Männern. In W. Cornelißen (Hrsg.), *Gender-Datenreport. 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland* (S. 159-223). München.
- Crawford, M., & Unger, R. K. (2004). *Women and gender: A feminist psychology* (4. Aufl.). Boston: McGraw-Hill.
- Crosby, F. J., Williams, J. C., & Biernat, M. (2004). The Maternal Wall. *Journal of Social Issues*, 60(4), 675-682.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2004). When Professionals Become Mothers, Warmth Doesn't Cut the Ice. *Journal of Social Issues*, 60(4), 701-718.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS Map: Behaviors From Intergroup Affect and Stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631-641.
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and Competence as Universal Dimension of Social Perception: The Stereotyp Content Model and the BIAS Map. In M. P. Zanna (Hrsg.), *Advances in Experimental Social Psychology*, (S. 61-137). Academic Press.
- DeMaio, T. J., & Landreth, A. (2004). Do Different Cognitive Interview Techniques Produce Different Results? In S. Presser (Hrsg.), *Wiley series in survey methodology. Methods for testing and evaluating survey questionnaires* (S. 89-108). Hoboken, NJ: Wiley.
- DeMaio, T., Mathiowetz, N., Rothgeb, J., Beach, M. E., & Durant, S. (1993). *Protocol for Pretesting Demographic Surveys at the Census Bureau*, Census Bureau Monograph, Washington, DC: U.S. Bureau of the Census.
- Deutsche Telekom (2010). *Deutsche Telekom führt als erstes Dax-30-Unternehmen Frauenquote für die Führung ein*. Abgerufen am 19. Mai 2011, von <http://www.telekom.com/dtag/cms/content/dt/de/829454>.
- Deutscher Bundestag (2011). *Abgeordnete in Zahlen: Frauen und Männer*. Abgerufen am 14. Januar 2013, von [http://www.bundestag.de/bundestag/abgeordnete17/mdb\\_zahlen/frauen\\_maenner.html](http://www.bundestag.de/bundestag/abgeordnete17/mdb_zahlen/frauen_maenner.html).

- Deutsches Jugendinstitut (2006). „Von der Frauenbewegung zur Genderforschung“. Interview mit PD Dr. Waltraud Cornelißen, DJI Leiterin der Abteilung Geschlechterforschung. DJI Online Thema 2006/05 Gleichstellung auf dem Prüfstand. Abgerufen am 01. März 2011, von <http://www.dji.de/cgi-bin/projekte/output.php?projekt=570&Jump1=RECHTS&Jump2=20>.
- Die Bundesregierung (2006). 2. Bilanz Chancengleichheit. Frauen in Führungspositionen. Abgerufen am 19. Mai 2011, von <http://www.bmfsfj.de/RedaktionBMFSFJ/Abteilung4/Pdf-Anlagen/bilanz-chancengleichheit,property=pdf,bereich=rwb=true.pdf>.
- Dressel, C. (2005). Erwerbstätigkeit – Arbeitsmarktintegration von Frauen und Männern. In BMFSFJ (Hrsg.), *Gender-Datenreport. 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland* (S. 99-158). München.
- Dressel, C., Cornelißen, W., & Wolf, K. (2005). Vereinbarkeit von Familie und Beruf. In BMFSFJ (Hrsg.), *Gender-Datenreport. 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland* (S. 266-341). München.
- Dressel, K., & Wanger, S. (2008). Erwerbsarbeit: Zur Situation von Frauen auf dem Arbeitsmarkt. In R. Becker, & B. Kortendiek (Hrsg.), *Geschlecht & Gesellschaft: Bd. 35. Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie* (2. Aufl., S. 481–490). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Eagly, A. H. (1987). *Sex Differences in Social Behavior: A Social-role Interpretation*. Hillsdale, NJ: Erlbaum
- Eagly, A. H., & Carli, L. L. (2007). *Through the labyrinth: The truth about how women become leaders*. Boston, Mass: Harvard Business School Press.
- Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H., & Mladinic, A. (1989). Gender Stereotypes and Attitudes Toward Women and Men. *Personality and Social Psychology Bulletin*, 15(4), 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., & Steffen, V. J. (1986). Gender stereotypes, occupational roles, and beliefs about part-time employees. *Psychology of Women Quarterly*, 10(3), 252-262.
- 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). London: Lawrence Erlbaum Associates.

- Eckes, T. (1994). Features of men, features of women: Assessing stereotypic beliefs about gender subtypes. *British Journal of Social Psychology*, 33, 107–123.
- Eckes, T. (2001). Ambivalenter Sexismus und die Polarisierung von Geschlechterstereotypen. *Zeitschrift für Sozialpsychologie*, 32(4), 235–247.
- Eckes, T. (2002). Paternalistic and Envious Gender Stereotypes: Testing Predictions From the Stereotype Content Model. *Sex Roles*, 47(3/4), 99–114.
- Eckes, T. (2008). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker, & B. Kortendiek (Hrsg.), *Geschlecht & Gesellschaft: Bd. 35. Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie* (2. Aufl., S. 171–182). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Eckes, T., & Six-Materna, I. (1998). Leugnung von Diskriminierung: Eine Skala zur Erfassung des modernen Sexismus. *Zeitschrift für Sozialpsychologie*, 29(3), 224–238.
- Eckes, T., & Six-Materna, I. (1999). Hostilität und Benevolenz: Eine Skala zur Erfassung des ambivalenten Sexismus. *Zeitschrift für Sozialpsychologie*, 30(4), 211–228.
- Faulbaum, F., Prüfer, P., & Rexroth, M. (2009). *Was ist eine gute Frage? Die systematische Evaluation der Fragenqualität* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A Model of (Often Mixed) Stereotype Content: Competence and Warmth Respectively Follow From Perceived Status and Competition. *Journal of personality and social psychology*, 82(6), 878–902.
- Fiske, S. T., Xu, J., Cuddy, A. J. C., & Glick, P. (1999). (Dis)respecting versus (Dis)liking: Status and Interdependence Predict Ambivalent Stereotypes of Competence and Warmth. *Journal of Social Issues*, 55(3), 473–489.
- Fitzgerald, R., Collins, D., Widdop, S., & Gray, M. (2009). *Testing for equivalence using cross-national cognitive interviewing. Centre for Comparative Social Surveys Working Paper series, Paper no. 01*. City University London; Centre for Comparative Social Surveys. Abgerufen am 23. August 2010, von [https://www.city.ac.uk/\\_data/assets/pdf\\_file/0014/125132/CCSS-Working-Paper-No-01.pdf](https://www.city.ac.uk/_data/assets/pdf_file/0014/125132/CCSS-Working-Paper-No-01.pdf).
- Forsyth, B., Rothgeb, J., & Willis, G. B. (2004). Does Pretesting Make a Difference? An Experimental Test. In S. Presser (Hrsg.), *Wiley series in survey methodology. Methods for testing and evaluating survey questionnaires* (S. 525–547). Hoboken, NJ: Wiley.
- Fowler, F. J., Jr. (2001). Why it is easy to write bad questions. *ZUMA-Nachrichten*, 48, 49–66.
- Freitag, J. (2011, 3. Juli). *Komm, ich zeig' dir meinen Kängurubauch. Judith Holofernes im Gespräch*. Frankfurter Allgemeine Sonntagszeitung, S. 26.

- Fthenakis, W. E., Kalicki, B., & Peitz, G. (2002). *Paare werden Eltern. Die Ergebnisse der LBS-Familien-Studie*. Opladen: Leske und Budrich.
- Fuchs, G., & Bothfeld, S. (2011). Gleichstellung in Deutschland im europäischen Vergleich. *Aus Politik und Zeitgeschichte*, 37/38, 7-18.
- Fuegen, K., Biernat, M., Haines, E., & Deaux, K. (2004). Mothers and Fathers in the Workplace: How Gender and Parental Status Influence Judgments of Job-Related Competence. *Journal of Social Issues*, 60(4), 737-754.
- Gerger, H., Kley, H., Bohner, G., & Siebler, F. (2007). The Acceptance of Modern Myths About Sexual Aggression Scale: Development and Validation in German and English. *Aggressive Behavior*, 33, 422-440.
- Gerlach, I. (2010). *Familienpolitik*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gille, M., & Marbach, J. (2004). Arbeitsteilung von Paaren und ihre Belastung mit Zeitstress. In Statistisches Bundesamt (Hrsg.), *Forum der Bundesstatistik Band 43. Alltag in Deutschland. Analysen zur Zeitverwendung. Beiträge zur Ergebniskonferenz der Zeitbudgeterhebung 2001/02 am 16./17. Februar 2004 in Wiesbaden* (S. 86-113). Abgerufen am 23. November 2011, von [https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/Alltag1030443049004.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/Thematisch/EinkommenKonsumLebensbedingungen/Zeitbudgeterhebung/Alltag1030443049004.pdf?__blob=publicationFile).
- Glick, P., Diebold, J., Bailey-Werner, B., & Zhu, L. (1997). The Two Faces of Adam: Ambivalent Sexism and Polarized Attitudes Toward Women. *Personality and Social Psychology Bulletin*, 23(12), 1323-1334.
- Glick, P., & Fiske, S. T. (1996). The Ambivalent Sexism Inventory: Differentiating hostile and benevolent sexism. *Journal of Personality and Social Psychology*, 70(3), 491-512.
- Glick, P., & Fiske, S. T. (1997). Hostile and Benevolent Sexism: Measuring Ambivalent Sexist Attitudes Toward Women. *Psychology of Women Quarterly*, 21, 119-135.
- Glick, P., & Fiske, S. T. (1999). The Ambivalence Toward Men Inventory: Differentiating Hostile and Benevolent Beliefs About Men. *Psychology of Women Quarterly*, 23, 519-536.
- Glick, P., & Fiske, S. T. (2001a). An ambivalent alliance. Hostile and benevolent sexism as complementary justifications for gender inequality. *The American psychologist*, 56(2), 109-118.
- Glick, P., & Fiske, S. T. (2001b). Ambivalent Sexism. In M. P. Zanna (Hrsg.), *Advances in Experimental Social Psychology* (Vol. 33, S. 115-188). San Diego: Academic Press.
- Glick, P., Fiske, S. T., Mladinic, A., Saiz, J., Abrams, D., Masser, B., Adetoun, B., Osagie, J.E., Akande, A., Alao, A., Brunner, A., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Materna, I., Expósito, F., Moya, M.,

- Foddy, M., Kim, H. J., Lameiras, M., Sotelo, M. J., Mucchi-Faina, A., Romani, M., Sakalli, N., Udegbe, B., Yamamoto, M., Ui, M., Ferreira, M. C., López López, W. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, 79, 763-775.
- Glick, P., & Hilt, L. (2000). Combative Children to Ambivalent Adults: The Development of Gender Prejudice. In T. Eckes, & H. M. Trautner (Hrsg.), *The Developmental Social Psychology of Gender* (S. 243–272). London: Lawrence Erlbaum Associates.
- Göring-Eckardt, K. (2006, 24. August). „Selbst bei den Grünen gilt Frauenpolitik als Loser-Thema“. Abgerufen am 10. Mai 2011, von <http://www.zeit.de/2006/35/Feminismus-Goering-Eckhardt>.
- Gomolla, (o.J.) *Institutionelle Diskriminierung im Bildungs- und Erziehungssystem*. Abgerufen am 19. Dezember 2011, von <http://www.egora.uni-muenster.de/ew/personen/medien/gomolla.pdf>.
- Goodwin, S. A., & Fiske, S. T. (2004). Power and gender: The double-edged sword of ambivalence. In R. K. Unger (Hrsg.), *Handbook of the psychology of women and gender* (S. 358.366). New York: Wiley.
- Gray, M., Blake, M., & Campanelli, P. (forthcoming 2014). The Use of Cognitive Interviewing Methods to Evaluate Mode Effects in Survey Questions, *Field Methods*, 26 (2).
- Grice, H. P. (1975). Logic and conversation. In P. Cole, & J. L. Morgan (Hrsg.), *Syntacs and semantics, Vol.3, Speech acts* (S. 41-58). New York: Academic Press.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey Methodology* (1. Aufl.). Hoboken, N.J.: Wiley.
- Groves, R. M., Fultz, N. H., & Martin, E. (1992). Direct questioning about comprehension in a survey setting. In J. M. Tanur (Hrsg.), *Questions about Questions: Inquiries into the Cognitive Bases of Surveys* (S. 49-61). New York: Russel.
- Häder, S. (2000). Telefonstichproben. *ZUMA-How-to-Reihe*, 6, 1-13.
- Heitmeyer, W. (Hrsg.) (2002). *Deutsche Zustände. Folge 1*. Frankfurt am Main: Suhrkamp.
- Heitmeyer, W. (Hrsg.) (2010). *Deutsche Zustände. Folge 9*. Berlin: Suhrkamp.
- Herman, E. (2007). *Das Eva-Prinzip: Für eine neue Weiblichkeit*. München: Goldmann.
- Hoecker, B. (1999). *Frauen, Männer und die Politik*. Bonn: Dietz.
- Hoecker, B. (2008). 50 Jahre Frauen in der Politik - Späte Erfolge, aber nicht am Ziel. *Aus Politik und Zeitgeschichte*, 24-25, 10-18.

- Hormel, U., & Scherr, A. (2004). *Bildung für die Einwanderungsgesellschaft. Perspektiven der Auseinandersetzung mit struktureller, institutioneller und interaktioneller Diskriminierung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hormel, U., & Scherr, A. (2010). *Diskriminierung: Grundlagen und Forschungsergebnisse*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hummelsheim, D. (2009). *Die Erwerbsbeteiligung von Müttern: Institutionelle Steuerung oder kulturelle Prägung? Eine empirische Untersuchung am Beispiel von Belgien, West- und Ostdeutschland*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kilianski, S. E., & Rudman, L. A. (1998). Wanting It Both Ways: Do Women Approve of Benevolent Sexism? *Sex Roles*, 39 (5/6), 333-352.
- Kinder, D. R., & Sears, D. O. (1981). Prejudice and politics: Symbolic racism versus racial threats to the good life. *Journal of Personality and Social Psychology*, 40, 414-431.
- Klauer, K. C. (2008). Soziale Kategorisierung und Stereotypisierung. In L.-E. Petersen, & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen* (S. 23-32). Weinheim: Beltz.
- Klement, C., & Rudolph, B. (2003). Auswirkungen staatlicher Rahmenbedingungen und kultureller Leitbilder auf das Geschlechterverhältnis. Deutschland und Finnland. *Aus Politik und Zeitgeschichte*, 44, 23-30.
- Koppetsch, C., & Burkart, G. (1999). *Die Illusion der Emanzipation: Zur Wirksamkeit latenter Geschlechtnormen im Milieuvvergleich*. Konstanz: Universitäts-Verlag Konstanz.
- Kortendiek, B. (2008). Familie: Mutterschaft und Vaterschaft zwischen Traditionalisierung und Modernisierung. In R. Becker, & B. Kortendiek (Hrsg.), *Geschlecht & Gesellschaft: Bd. 35. Handbuch Frauen- und Geschlechterforschung. Theorie, Methoden, Empirie* (2. Aufl., S. 434-445). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Krameyer, A., & Schmidt, H. D. (1973). Entwicklung von Skalen zur Messung misogynischer Einstellungen. In H. D. Schmidt, C. Schmerl, & A. Krameyer (Hrsg.), *Frauenfeindlichkeit. Sozialpsychologische Aspekte der Misogynie* (S. 40-55). München: Juventa-Verlag.
- Krameyer, A., & Steinbach, D. (1973). Ritterlichkeit - eine indirekte Form von Misogynie? In H. D. Schmidt, C. Schmerl, & A. Krameyer (Hrsg.), *Frauenfeindlichkeit. Sozialpsychologische Aspekte der Misogynie* (S. 56-79). München: Juventa-Verlag.
- Krampen, G. (1979). Eine Skala zur Messung der normativen Geschlechtsrollen-Orientierung (GRO-Skala). *Zeitschrift für Soziologie*, 8(3), 254-266.
- Krampen, G. (1983). Eine Kurzform der Skala zur Messung normativer Geschlechtsrollen-Orientierungen. *Zeitschrift für Soziologie*, 12(2), 152-156.

- Kühnel, S.-M., & Krebs, D. (2001). *Statistik für die Sozialwissenschaften: Grundlagen, Methoden, Anwendungen* (Orig.-Ausg.). Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Kurz, K., Prüfer, P., & Rexroth, M. (1999). Zur Validität von Fragen in standardisierten Erhebungen: Ergebnisse des Einsatzes eines kognitiven Pretestinterviews. *Zuma-Nachrichten*, 23(44), 83-107.
- Leitner, S. (2006). Von der indirekten zur direkten Förderung von Familienarbeit: Bekannte Enttäuschungen und neue (falsche) Hoffnungen. In U. Degener, & B. Rosenzweig (Hrsg.), *Die Neuverhandlung sozialer Gerechtigkeit. Feministische Analysen und Perspektiven* (S. 321-339). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Levin, K., Willis, G., Forsyth, B., Norberg, A., Stapleton Kudela, M., Stark, D., & Thompson, F. (2009). Using Cognitive Interviews to Evaluate the Spanish-Language Translation of Dietary Questionnaire. *Survey Research Methods*. Vol.3, No.1, 13-25.
- Lorber, J. (1994). *Paradoxes of Gender*. Binghampton: Vail-Ballou Press.
- Mason, K. O., Czajka, J. L., & Arber, S. (1976). Changes in U. S. Women's Sex-role Attitudes, 1964-1974. *American Sociological Review*, 41 (4), 573-596.
- McHugh, M. C., & Frieze, I. H. (1997). The Measurement of Gender-Role-Attitudes: A Review and Commentary. *Psychology of Women Quarterly*, 21, 1-16.
- Moosbrugger, H., & Schermelleh-Engel, K. (2008). Exploratorische (EFA) und Konfirmatorische Faktorenanalyse (CFA). In H. Moosbrugger, & A. Kelava (Hrsg.), *Testtheorie und Fragebogenkonstruktion* (S. 307-324). Heidelberg: Springer.
- Mühling, T., Rost, H., Rupp, M., & Schulz, F. (2006). *Kontinuität trotz Wandel. Die Bedeutung traditioneller Familienleitbilder für die Berufsverläufe von Müttern und Vätern*. Weinheim: Juventa Verlag.
- Mummendey, A., Kessler, T., & Otten, S. (2009). Sozialpsychologische Determinanten - Gruppenzugehörigkeit und soziale Kategorisierung. In A. Beelmann, & K. J. Jonas (2008). *Diskriminierung und Toleranz: Psychologische Grundlagen und Anwendungsperspektiven* (S. 43-60). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Nadler, E. B., & Morrow, W. R. (1959). Authoritarian Attitudes Toward Women, and their Correlates. *The Journal of social psychology*, 49 (1), 113-123.
- Nisbett, Richard, & Wilson, Timothy. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Obermeier, B. (2006). *Karriererisiko: Weiblich, 30 plus*. FAZ.NET. Abgerufen am 11. November 2006, von <http://www.faz.net/s/RubEC1ACFE1EE274C81BCD3621EF555C83C/Doc~EE908A266F8A24D04A3A675CE9B205216~ATpl~Ecommon~Sconten t.html>.



- Onnen-Isemann, C. (2007) Deutsche Perspektive: Kinderlosigkeit - französische Perspektive: Elternschaft? Familienpolitik und Fertilitätsunterschiede in Frankreich und Deutschland. In D. Auth, & B. Holland-Cunz (Hrsg.), *Grenzen der Bevölkerungspolitik: Strategien und Diskurse demographischer Steuerung* (S. 165-180). Opladen: Barbara Budrich.
- Petersen, L.-E., & Six, B. (Hrsg.) (2008). *Stereotype, Vorurteile und soziale Diskriminierung: Theorien, Befunde und Interventionen*. Weinheim: Beltz.
- Petzel, T., Wagner, U., Nicolai, K., & van Dick, R. (1997). Ein kurzes Instrument zur Messung der Auroitarimus-Neigung. *Gruppendynamik*, 28(3), 251-258.
- Pfau-Effinger, B. (2000). *Kultur und Frauenerwerbstätigkeit in Europa. Theorie und Empirie des internationalen Vergleichs*. Opladen: Leske + Budrich.
- Pfau-Effinger, B. (2001). Wandel wohlfahrtsstaatlicher Geschlechterpolitiken im soziokulturellen Kontext. In B. Heintz (Hrsg). *Geschlechtersoziologie. Sonderheft 41/2001 der Kölner Zeitschrift für Soziologie und Sozialpsychologie* (S. 487-511). Wiesbaden: Westdeutscher Verlag.
- Pfau-Effinger, B. (2005). *Wandel der Geschlechterkultur und Geschlechterpolitiken in konservativen Wohlfahrtsstaaten – Deutschland, Österreich und Schweiz*. gender...politik...online. Abgerufen am 04. Januar 2011, von [http://web.fu-berlin.de/gpo/pdf/tagungen/wandel\\_geschl\\_pfau\\_effinger.pdf](http://web.fu-berlin.de/gpo/pdf/tagungen/wandel_geschl_pfau_effinger.pdf).
- Pfau-Effinger, B., & Geissler, B. (2002). Cultural change and family policies in East and West Germany. In A. Carling, S. Duncan, & R. Edwards (Hrsg.), *Analysing Families: Morality and Rationality in Policy and Practice* (S. 77-83). London, New York: Routledge.
- Pinl, C. (2004). Wo bleibt die Zeit? Die Zeitbudgeterhebung 2001/02 des Statistischen Bundesamts. *Aus Politik und Zeitgeschichte*, 31-32, 19-25.
- Porst, R. (2000). Question Wording - Zur Formulierung von Fragebogen-Fragen. *ZUMA How-to-Reihe*, 2, 1-11.
- Porst, R. (2008). *Fragebogen: Ein Arbeitsbuch* (1. Aufl.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Porter, N. B. (2006). Re-defining superwoman: An Essay on Overcoming the "Maternal Wall" in the Legal Workplace. *Duke Journal of Gender Law & Policy*, 13(55), 55-84.
- Presser, S., & Blair, J. (1994). Survey Pretesting: Do different Methods produce different Results? *Sociological Methodology*, 24, 73-104.
- Presser, S., Couper, M. P., Lessler, J. T., Martin, E., Martin, J., Rothgeb, J. M., et al. (2004). Methods for testing and evaluating Survey Questions. *Public Opinion Quarterly*, 68 (1), 109-130.

- Prüfer, P., & Rexroth, M. (1996). Verfahren zur Evaluation von Survey-Fragen: Ein Überblick. *ZUMA-Nachrichten*, 39, 95-115.
- Prüfer, P., & Rexroth, M. (2000). Zwei-Phasen-Pretesting. *ZUMA-Arbeitsbericht*, 08, 1-21.
- Prüfer, P., & Rexroth, M. (2004). Paraphrasing can be dangerous: A little experiment. *ZUMA-Nachrichten Spezial: QUEST 2003 Proceedings of the 4th Conference on Questionnaire Evaluation Standards 21 – 23 October 2003*, 9, 52-59.
- Prüfer, P., & Rexroth, M. (2005). Kognitive Interviews. *ZUMA How-to-Reihe*, 15, 1-20.
- Rammstedt, B. (2004). Zur Bestimmung der Güte von Multi-Item-Skalen. Eine Einführung. *ZUMA How-to-Reihe*, 12, 1-25.
- Reich-Hilweg, I. (1979). *Männer und Frauen sind gleichberechtigt: Der Gleichberechtigungsgrundsatz (Art. 3, Abs. 2 GG) in der parlamentarischen Auseinandersetzung 1948 – 1957 und in der Rechtsprechung des Bundesverfassungsgerichts 1953 – 1975*. Frankfurt/Main: Europäische Verlagsanstalt.
- Rudman, L. A. (2005). Rejection of women? Beyond prejudice as antipathy. In J. F. Dovidio, P. Glick, & L. A. Rudman (Hrsg.), *On the nature of prejudice: Fifty years after Allport* (S. 106-120). Malden, MA: Blackwell.
- 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., Frey, D., Gollwitzer, P. M., Helmreich, R. L., & Spence, J. T. (1981). Masculine (Instrumental) and Feminine (Expressive) Traits: A Comparison between Students in the United States and West Germany. *Journal of Cross-Cultural Psychology*, 12 (2), 142-162.
- Sachverständigenkommission (Sachverständigenkommission zur Erstellung des Ersten Gleichstellungsberichtes der Bundesregierung, & Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e.V.) (Hrsg.) (2011), *Neue Wege – gleiche Chancen. Gleichstellung von Frauen und Männern im Lebensverlauf. Gutachten der Sachverständigenkommission an das Bundesministerium für Familie, Senioren, Frauen und Jugend für den ersten Gleichstellungsbericht der Bundesregierung*. Troisdorf: RMP Rautenberg Media & Print Verlag KG.
- Schaeffer, N. C., & Presser, S. (2003). The Science of Asking Questions. *Annual Review of Sociology*, 29, 65-88.
- Schäffgen, K. (2000). *Die Verdopplung der Ungleichheit: Sozialstruktur und Geschlechterverhältnisse in der Bundesrepublik und in der DDR*. Opladen: Leske + Budrich.

- Schmerl, C., & Steinbach, D. (1973). Ritterlichkeit - eine indirekte Form von Misogynie? In H. D. Schmidt, C. Schmerl, & A. Krameyer (Hrsg.), *Frauenfeindlichkeit. Sozialpsychologische Aspekte der Misogynie* (S. 56-79). München: Juventa-Verlag.
- Schmidt-Mummendey, A., Schmerl, C., & Schmidt, H. D. (1973). Ist das Vorurteil gegenüber der Frau solider Bestandteil eines autoritär-konservativen Syndroms? In H. D. Schmidt, C. Schmerl, & A. Krameyer (Hrsg.), *Frauenfeindlichkeit. Sozialpsychologische Aspekte der Misogynie* (S. 101-122). München: Juventa-Verlag.
- Schneider, N. F., & Dorbritz, J. (2011). Wo bleiben die Kinder? Der niedrigen Geburtenrate auf der Spur. *Aus Politik und Zeitgeschichte*, 10/11, 26-34.
- Schober, M. F. (1999). Making Sense of Questions: An Interactional Approach. In M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur, & R. Tourangeau (Hrsg.), *Cognition and Survey Research* (S. 77-93). New York: John Wiley & Sons.
- Scholz, E., Harkness, J., & Klein, S. (2003). *ISSP 2002 Germany. Family and Changing Gender Roles III. ZUMA Report on the German Study*. ZUMA Methodenbericht, 15, Mannheim: Zentrum für Umfragen, Methoden und Analysen. Abgerufen am 11. Juni 2009, von <http://www.za.uni-koeln.de/data/en/issp/questionnaires/q2002/DE2002.pdf>.
- Schüßler (2006). *Werden kognitive Pretests ihrem Anspruch gerecht? Eine empirische Untersuchung zu Sexismus und Geschlechterrollen*. Unveröffentlichte Diplomarbeit. Justus-Liebig-Universität Gießen.
- Schüßler, A., & Schmidt, P. (2008). *Does Cognitive Interviewing Matter? Testing the Effects of Cognitive Methods on Data Quality*. Publikation im Rahmen der 7<sup>th</sup> International Conference on Social Sciences Methodology (RC 33 of the ISA) in Neapel, Italien.
- Schwarz, N., Strack, F., & Hippler, H.-J. (1991). Kognitionspsychologie und Umfrageforschung: Themen und Befunde eines interdisziplinären Forschungsgebietes. *Psychologische Rundschau*, 42, 174-186.
- Szesny, S., & Stahlberg, D. (2002). Geschlechtsstereotype Wahrnehmung von Führungskräften. *Wirtschaftspsychologie*, 1, 36-41.
- Six, B. (2008). Soziale Dominanz und Diskriminierung. In L.-E. Petersen, & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen* (S. 172-183). Weinheim: Beltz.
- Six-Materna, I. (2008). Sexismus. In L.-E. Petersen, & B. Six (Hrsg.), *Stereotype, Vorurteile und soziale Diskriminierung. Theorien, Befunde und Interventionen* (S. 121-130). Weinheim: Beltz.
- Snijkers, G. (2004). Cognitive Laboratory Experiences and Beyond: Some ideas for future research. *ZUMA-Nachrichten Spezial: QUEST 2003 Proceedings of the 4th Conference on Questionnaire Evaluation Standards 21 – 23 October 2003*, 9, 190-203.

- Social Psychology Network (o.J.). *Understanding Prejudice. Frequently Asked Questions: Ambivalent Sexism*. Abgerufen am 30. Juni 2009, von <http://www.understandingprejudice.org/asi/faq>.
- Spence, J. T., & Hahn, E. D. (1997). The attitudes toward women scale and attitude change in college students. *Psychology of Women Quarterly*, 21, 17-34.
- Spence, J. T., Helmreich, R. L., & Stapp, J. (1973). A short version of the Attitudes toward Women Scale (AWS). *Bulletin of the Psychonomic Society*, 2(4), 219-220.
- Stachelhaus, R. (2010, 19. September). *Auf dem Weg nach oben. Her mit den Frauen!* FAZ.NET-Spezial. Die Frauenfalle. Abgerufen am 25. September 2010, von <http://www.faz.net/aktuell/wirtschaft/auf-dem-weg-nach-oben-her-mit-den-frauen-1574722.html>.
- Stahlberg, D., Dickenberger, D., & Szillis, U. (2009). Geschlechterdiskriminierung. In A. Beelmann, & K. J. Jonas (Hrsg.), *Diskriminierung und Toleranz. Psychologische Grundlagen und Anwendungsperspektiven* (S. 193-21). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Statistics Sweden (2004). *Design your question right. How to develop, test, evaluate and improve questionnaires*. Abgerufen am 16. November 2011, von [http://www.scb.se/statistik/\\_publikationer/OV9999\\_2004A01\\_BR\\_X97OP0402.pdf](http://www.scb.se/statistik/_publikationer/OV9999_2004A01_BR_X97OP0402.pdf).
- Statistisches Bundesamt (2009a). *2008: Verdienstunterschied von Männern und Frauen weiter bei 23%. Pressemitteilung Nr.428 vom 12.11.2009*. Wiesbaden. Abgerufen am 20. Mai 2011, von [http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2009/11/PD09\\_\\_428\\_\\_621,templateId=renderPrint.psml](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2009/11/PD09__428__621,templateId=renderPrint.psml).
- Statistisches Bundesamt (2009b). *Mütter arbeiten Häufiger in Teilzeit. Pressemitteilung Nr.391 vom 14.10.2009*. Wiesbaden. Abgerufen am 20. Mai 2011, von [http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2009/10/PD09\\_\\_391\\_\\_122.psml](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2009/10/PD09__391__122.psml).
- Statistisches Bundesamt (2010a). *Frauenanteil in Professorenschaft steigt weiter auf über 18%. Pressemitteilung Nr.235 vom 06.07.2010*. Abgerufen am 08. November 2011, von [http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2010/07/PD10\\_\\_235\\_\\_213,templateId=renderPrint.psml](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2010/07/PD10__235__213,templateId=renderPrint.psml).
- Statistisches Bundesamt (2010b). *Alles beim Alten: Mütter stellen Erwerbstätigkeit hinten an*. Wiesbaden. Abgerufen am 16. September 2010, von [https://www.destatis.de/DE/Publikationen/STATmagazin/Arbeitsmarkt/2010\\_03/2010\\_03PDF.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/STATmagazin/Arbeitsmarkt/2010_03/2010_03PDF.pdf?__blob=publicationFile).

- Statistisches Bundesamt (2011). *Elterngeld: Väterbeteiligung im Jahresverlauf 2009 weiter gestiegen. Pressemitteilung Nr. 195 vom 19.05.2011*. Wiesbaden. Abgerufen am 14. Januar 2013, von [https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2011/05/PD11\\_195\\_22922.html](https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2011/05/PD11_195_22922.html).
- Steinbach, A. (2004). Wie Paare sich die Arbeit teilen. In W. E. Fthenakis, & M. E. Textor (Hrsg.), *Das Online-Familienhandbuch*. Abgerufen am 18. Januar 2011, von <https://www.familienhandbuch.de/cms/Familienforschung-Arbeitsteilung.pdf>.
- Still, M. C. (2006) *Litigating the Maternal Wall: U.S. Lawsuits Charging Discrimination Against Workers with Family Responsibilities*. Center for Work Life Law. University of California Hastings College of the Law. Abgerufen am 11. Februar 2011, von [http://www.law.yale.edu/documents/pdf/FRD\\_report\\_FINAL1.pdf](http://www.law.yale.edu/documents/pdf/FRD_report_FINAL1.pdf).
- Stürzer, M. (2005). Bildung, Ausbildung und Weiterbildung. In Bundesministerium für Familie, Senioren, Frauen und Jugend (Hrsg.), *Gender-Datenreport. 1. Datenreport zur Gleichstellung von Frauen und Männern in der Bundesrepublik Deutschland* (S. 21-98). München.
- Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking about answers. The application of cognitive survey methodology*. San Francisco: Jossey-Bass.
- Süssenbach, P., & Bohner, G. (2011). Acceptance of sexual aggression myths in a representative sample of German residents. *Aggressive Behavior*, 37, 374-385.
- Swim, J. K., Aikin, K. J., Hall, W. S., & Hunter, B. A. (1995). Sexism and Racism: Old-Fashioned and Modern Prejudices. *Journal of personality and social psychology*, 68(2), 199-214.
- Swim, J. K., Becker, J., Lee, E., & Pruitt, E. R. (2010). Sexism reloaded: Worldwide evidence for its endorsement, expression, and emergence in multiple contexts. In H. Landrine, & N. Russo (Hrsg.), *Handbook of diversity in feminist psychology* (S. 137-172). Washington, DC: American Psychological Association.
- Swim, J. K., & Campbell, B. (2003). Sexism: Attitudes, Beliefs and Behaviors. In R. Brown, & S. L. Gaernter (Hrsg.), *Blackwell Handbook of Social Psychology: Intergroup Processes* (S. 218-237). Malden, MA: Blackwell.
- Swim, J. K., & Hyers, L. L. (2009). Sexism. In T. D. Nelson (Hrsg.), *Handbook of prejudice, stereotyping, and discrimination* (S. 407-430). New York: Psychology Press.
- Täubner, M. (2008) Unterbezahlt, übervorteilt, ausgenutzt. Frauen, tut was! *Hochschul-anzeiger*, 95, 16.
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behavior. *European Journal of Social Psychology*, 1, 149-178.

- Tajfel, H., & Turner, J. C. (1986). The social identity theory of intergroup behavior. In S. Worchel, & W. G. Austin (Hrsg.), *Psychology of intergroup relations* (S. 7-24). Chicago, IL: Nelson-Hall.
- Tougas, F., Brown, R., Beaton, A. M., & Joly, S. (1995). Neosexism: Plus Ça Change, Plus C'est Pareil. *Personality and Social Psychology Bulletin*, 21(8), 842-849.
- Tourangeau, R. (1984). Cognitive sciences and survey methods. In T. Jabine, M. Straf, J. Tanur, & R. Tourangeau (Hrsg.), *Cognitive Aspects of Survey Methodology: Building a bridge between disciplines* (S. 73-100). Washington, DC: National Academic Press.
- Träger, J. (2007). Neue Wege familialer Arbeitsteilung – Neuorientierung in der Familienpolitik? In D. Auth, & B. Holland-Cunz (Hrsg.), *Grenzen der Bevölkerungspolitik: Strategien und Diskurse demographischer Steuerung* (S. 145-163). Opladen: Barbara Budrich.
- Vinken, B. (2001). *Die deutsche Mutter: Der lange Schatten eines Mythos*. München: Piper.
- Wade, M. L., & Brewer, M. B. (2006). The Structure of Female Subgroups: An Exploration of Ambivalent Stereotypes. *Sex Roles*, 54 (11/12), 753-765.
- Walther, K., & Schaeffer-Hegel, B. (2007). Karriere mit Kindern?! *Aus Politik und Zeitgeschichte*, 7, 15-20.
- Weber, U., & Schaeffer-Hegel, B. (2000). *Geschlechterarrangements in der Bundesrepublik*. *Aus Politik und Zeitgeschichte*. 31/32, 5-10.
- Williams, J. (2001). *Unbending gender: Why family and work conflict and what to do about it*. Oxford: Oxford University Press.
- Williams, J. (2006a). *Long Time No See: Why Are There Still So Few Women in Academic Science and Engineering?* In Committee on Maximizing the Potential of Women in Academic Science and Engineering, National Academy of Sciences, National Academy of Engineering, Institute of Medicine (Hrsg.), *Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering: Report of a Workshop* (S. 149-166). Washington, DC: The National Academies Press.
- Williams, J. (2006b). Moving Beyond the “Chilly Climate” to a New Model for Spurring Organizational Change. In Committee on Maximizing the Potential of Women in Academic Science and Engineering, National Academy of Sciences, National Academy of Engineering, Institute of Medicine (Hrsg.), *Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering: Report of a Workshop* (S. 51-56). Washington, DC: The National Academies Press.
- Williams, J. C., & Segal, N. (2003). Beyond the maternal wall: Relief for family caregivers who are discriminated against on the job. *Harvard Women's Law Journal*, 26, 77-162.

- Williams, J. C., & Westfall, E. S. (2006). Deconstructing the maternal wall: Strategies for vindicating the civil rights of "carers" in the workplace. *Duke Journal of Gender Law & Policy*, 13(31), 31-53.
- Willis, G. B. (1999). *Cognitive Interviewing: A "How To" Guide*. Abgerufen am 12. Juni 2009, von <http://appliedresearch.cancer.gov/areas/cognitive/interview.pdf>.
- Willis, G. B. (2005). *Cognitive Interviewing: A tool for improving questionnaire design*. Thousand Oaks: Sage.
- Willis, G. B., & Lessler, J. T. (1999). *Question Appraisal System. QAS-99*. Rockville, MD: Research Triangle Institute. Abgerufen am 20. Januar 2013, von <http://appliedresearch.cancer.gov/areas/cognitive/qas99.pdf>.
- Winkler, N., Kroh, M., & Spiess, M. (2006). *Entwicklung einer deutschen Kurzsкала zur zweidimensionalen Messung von sozialer Erwünschtheit. Diskussionspapier 579*. Deutsches Institut für Wirtschaftsforschung: Berlin.
- Woolett, A., & Marshall, H. (2001). Motherhood and Mothering, In R. Unger (Hrsg.), *Handbook of the psychology of women and gender* (S. 170-182). New York: Wiley.
- Zerle, C., & Krok, I. (o.J.). *Null Bock auf Familie? Der schwierige Weg junger Männer in die Vaterschaft*. Kurzfassung. Abgerufen am 14. Januar 2013, von [http://www.dji.de/bibs/Vaeterstudie/news0811\\_1\\_vaeter\\_kurzfassung.pdf](http://www.dji.de/bibs/Vaeterstudie/news0811_1_vaeter_kurzfassung.pdf).

## Anhang

## Anhang 1: Überblick Veränderungen und Reduzierungen des Itempools

	Pretest- Phase 1	Pretest- Phase 2	Pretest- Phase 3
Tabelle A1 Überblick über die vorgenommenen Veränderungen und Reduzierungen des Itempools (Feindseliger Sexismus) im Anschluss an die einzelnen Pretest-Phasen			
Feindseliger Sexismus			
Ich gehe davon aus, dass viele berufstätige Frauen ihre weiblichen Reize gezielt einsetzen, um ihrer Karriere auf die Sprünge zu helfen.	✓		
<i>Ich kann mir vorstellen, dass viele berufstätige Frauen ihre weiblichen Reize gezielt einsetzen, um ihrer Karriere auf die Sprünge zu helfen. (Umformulierung des Items im Rahmen der kognitiven Interviews)</i>		✓	☞
Frauen können Druck schlechter standhalten als Männer und sind daher für leitende Positionen in der Berufswelt weniger gut geeignet.	☞		
Frauen haben häufig zu wenig Durchsetzungsvermögen um sich in der männlichen Arbeitswelt zu behaupten.	☞		
Frauen sind für Führungspositionen im Job weniger gut geeignet als Männer.	☞		
Viele erfolgreich berufstätige Frauen verlieren ihre Weiblichkeit.	✓	✓	☞
Viele Karrierefrauen sind kaltschnäuzig.			
Viele Karrierefrauen sind <i>kalthertzig</i> . (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓		
<i>Karrierefrauen</i> sind häufig kalthertzig. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
Wenn mehrere Frauen in einem Büro zusammenarbeiten, gibt es häufig „Zickenterror“.	✓	✓	☞
Viele erfolgreich berufstätige Frauen verdanken ihre berufliche Stellung allein der Frauenquote (festgelegter Frauenanteil bei der Besetzung von Gremien oder Stellen).	☞		



	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Feindseliger Sexismus			
Viele Frauen rufen im Berufsleben schon beim geringsten Anlass nach der Frauenbeauftragten.	☞		
Eine Frau ist für den Beruf einer Pilotin oder Ingenieurin genauso gut geeignet wie für den einer Kindergärtnerin oder Kosmetikerin.*	☞		
Frauen haben häufig zu wenig Biss, um in ihrem Beruf in eine leitende Position zu gelangen.	✓		
<i>Frauen fehlt es meiner Meinung nach häufig am nötigen Biss, um in ihrem Beruf in eine leitende Position aufzusteigen. (Umformulierung des Items im Rahmen der kognitiven Interviews)</i>		✓	☞
<i>Weiblichen Führungskräften fehlt es häufig an Durchsetzungskraft. (Neue Item-Formulierung im Rahmen der kognitiven Interviews)</i>		✓	☞
Viele berufstätige Frauen sind selbst daran schuld, dass sie keine Karriere machen.	✓	☞	
Viele Frauen beschwerten sich im Berufsleben schon aus nichtigem Anlass darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.	✓	☞	
<i>Schon bei Kleinigkeiten beschwerten sich viele Frauen im Berufsleben darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden. (Zusätzliche Alternativ-Formulierung im Rahmen des ExpertInnen-Reviews)</i>	✓	☞	
Viele Frauen wittern grundlos überall im Berufsleben Ungerechtigkeiten (z. B. bei der Jobvergabe oder beim Gehalt).	☞		
Viele Frauen sind im Berufsleben peinlich genau darauf bedacht, im Vergleich zu ihren männlichen Kollegen bloß nicht zu kurz zu kommen (z. B. in Bezug auf das Gehalt oder bei Beförderungen).	✓	☞	
Weibliche Führungskräfte verhalten sich häufig zickig.	✓	☞	
Frauen besitzen mindestens genauso gute Führungsqualitäten wie Männer.*	☞		

	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Feindseliger Sexismus			
Erfolgreich berufstätige Frauen sind häufig „Mannweiber“.	✓		
<i>Karrierefrauen</i> sind häufig „Mannweiber“. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
Viele erfolgreich berufstätige Frauen sind gefühllos.	✓	☞	
Viele Karrierefrauen scheinen vergessen zu haben, dass das Glück einer Frau keineswegs allein darin bestehen sollte, Geld und Ansehen zu verdienen.	☞		
Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine <i>finanziell</i> versorgen könnte.			
Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine <i>finanziell</i> versorgen könnte. (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓		
<i>Ich habe Verständnis</i> dafür, wenn die Mutter <i>eines 1-jährigen Kindes Vollzeit erwerbstätig sein möchte</i> , obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.* (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
Viele erfolgreich berufstätige Frauen sind derart auf ihre Karriere fixiert, dass ihre Kinder darunter leiden.	✓	☞	
Für viele berufstätige Frauen von heute zählt allein das Streben nach beruflichem Erfolg und Anerkennung.	✓	☞	
Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.	✓		
<i>Frauen</i> sollten ein schlechtes Gewissen haben, wenn sie <i>aus Karrieregründen ihr Kind mit einem halben Jahr tagsüber zur Betreuung zu einer Tagesmutter geben</i> . (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓

	Pretest- Phase 1	Pretest- Phase 2	Pretest- Phase 3
Feindseliger Sexismus			
Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen darüber die Bedürfnisse ihrer Familie.			
Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen <i>dabei</i> die Bedürfnisse ihrer Familie. (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓		
Viele <i>Karrierefrauen</i> denken <i>meiner Meinung nach</i> nur an ihr berufliches Vorankommen und <i>vernachlässigen</i> die Bedürfnisse ihrer <i>Kinder</i> . (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.	✓		
Frauen, die <i>aus Karrieregründen</i> ihr Kind mit <i>einem halbem Jahr</i> in eine <i>Kindertagesstätte</i> geben, sind für mich Rabenmütter. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Erfolgreich berufstätige Frauen vernachlässigen häufig wegen ihrem Beruf ihre Kinder.			
Erfolgreich berufstätige Frauen vernachlässigen häufig <i>ihre Kinder</i> wegen <i>ihrem Beruf</i> . (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓	☞	
Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.	✓		
<i>Frauen denken meiner Meinung nach nur an sich selbst, wenn sie ein Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um berufliche Anerkennung zu bekommen.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓

	Pretest- Phase 1	Pretest- Phase 2	Pretest- Phase 3
Feindseliger Sexismus			
In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass ihre Kinder zu kurz kommen.			
In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass sie ihren Kindern <i>zu wenig Aufmerksamkeit schenken</i> . (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓		
<i>Ich denke, dass Karrierefrauen ihren Kindern häufig zu wenig Aufmerksamkeit schenken.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Ich finde Frauen, die ihre Karriere für die Betreuung ihrer Kinder nicht hinten anstellen wollen, egoistisch.			
Ich finde Frauen <i>egoistisch</i> , die ihre Karriere für die Betreuung ihrer Kinder nicht <i>zurückstellen</i> wollen. (Umformulierung des Items im Rahmen des ExpertInnen-Reviews)	✓	✓	☞
<i>Ich finde Frauen egoistisch, die bereits wenige Wochen nach der Geburt ihres Kindes wieder in ihren Beruf zurückkehren, weil sie ihre Karriere weiter verfolgen möchten.</i> (Zusätzliche Alternativ-Formulierung im Rahmen des ExpertInnen-Reviews)	✓		
Ich finde Frauen egoistisch, die bereits <i>ein halbes Jahr nach der Geburt des Kindes wieder Vollzeit</i> in ihren Beruf zurückkehren, <i>um sich selbst zu verwirklichen</i> . (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Ich bin der Ansicht, dass es keineswegs alleine die Aufgabe der Frau ist, eine Lösung für die Vereinbarkeit der eigenen Berufstätigkeit mit der Betreuung der Kinder zu finden, sondern ebenso die des Mannes.*			☞

*Anmerkungen.* Die Itemreihenfolge stimmt mit derjenigen aus Pretest-Phase 1 überein. Eingerückte Items geben die endgültigen Um- bzw. Neuformulierungen im Rahmen des ExpertInnen-Reviews bzw. im Rahmen der kognitiven Interviews für diejenigen Items an, die in die nächste Pretest-Phase mitgenommen wurden (kursive Hervorhebung der Änderungen). \* Items mit Sternchen sind umgekehrt gepolt (d.h. eine hohe Zustimmung steht für eine *nicht*-sexistische Einstellung). ✓ Das Item blieb nach Abschluss der Pretest-Phase weiterhin im Itempool enthalten. ☞ Das Item wurde nach Abschluss der Pretest-Phase aus dem Itempool entfernt.

Tabelle A2  
 Überblick über die vorgenommenen Veränderungen und Reduzierungen des Itempools (Wohltuender Sexismus) im Verlauf der einzelnen Pretest-Phasen

Wohltuender Sexismus	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Frauen sind auf Grund ihrer Eigenschaften für typisch weibliche Berufe (wie z. B. Grundschullehrerin oder Krankenschwester) weitaus besser geeignet als Männer.	☞		
Frauen besitzen ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.	✓		
Frauen besitzen eine größere <i>natürliche Begabung</i> als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Frauen verfügen über mehr Geschmack als Männer, wenn es darum geht, die Wohnung ansprechend einzurichten.	☞		
Frauen haben einfach einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.	✓		
Frauen haben <i>von Natur aus</i> einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Viele Männer sind – im Gegensatz zu Frauen – damit überfordert, gleichzeitig den Haushalt zu führen und sich um die Kinder zu kümmern.	☞		
Frauen verfügen über ein größeres Talent als Männer, die Wohnung behaglich und gemütlich zu gestalten.	✓		
Frauen verfügen über <i>eine größere natürliche Begabung</i> als Männer, die Wohnung behaglich und gemütlich zu gestalten. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
<i>Frauen verstehen sich viel besser als Männer darauf, die Wohnung gemütlich und heimelig herzurichten.</i> (Zusätzliche Alternativ-Formulierung im Rahmen des ExpertInnen-Reviews)	✓	☞	

Wohltuender Sexismus	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Viele Frauen wünschen sich einen Mann, der der Behüter und Beschützer des gemeinsamen Heims ist.	✓	☞	
Insgesamt sehnen sich doch alle Frauen nach einem Mann, in dessen Armen sie Schutz und Geborgenheit finden.	✓		
Insgesamt sehnen sich <i>vielen Frauen</i> nach einem Mann, in dessen Armen sie Schutz und Geborgenheit finden. (Umformulierung des Items im Rahmen der kognitiven Interviews)	✓	✓	☞
Eine Frau braucht die starke Schulter eines Mannes, an die sie sich anlehnen kann.	✓		
Eine Frau <i>hat eher das Bedürfnis sich an die starke Schulter eines Mannes anzulehnen, als umgekehrt.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)	✓	✓	☞
Viele Frauen wünschen sich einen Mann, der ihr Herz im Sturm erobert.	✓	☞	
Frauen besitzen eine Fürsorglichkeit, über die Männer nicht im gleichen Ausmaß verfügen.			
Frauen besitzen <i>einfach eine viel größere Fürsorglichkeit als Männer.</i> (Umformulierung des Items im Rahmen der ExpertInnen-Reviews)	✓		
<i>Mütter</i> besitzen eine <i>größere</i> Fürsorglichkeit als <i>Väter.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Eine Mutter ist vom Wesen her sanfter und warmerherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.	✓		
Eine Mutter ist vom Wesen her sanfter und warmerherziger als ein Vater und daher auch besser für die Betreuung eines Kindes geeignet. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Als Mutter kann eine Frau ihre wertvollsten Eigenschaften entfalten.	☞		
Nur in ihrer Rolle als Mutter kann eine Frau eine umfassende Zufriedenheit und dauerhaftes Glück empfinden.	☞		

Wohltuender Sexismus	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Eine Mutter sollte sich in den ersten Jahren selbst um die Betreuung ihres Kindes kümmern, da sie eine ganz besondere Verbindung zum Kind besitzt.	✓		
<i>Meiner Meinung nach sollte die Mutter im ersten Jahr zu Hause bleiben und sich um die Betreuung des Kindes kümmern, da sie eine engere Verbindung zum Kind besitzt als der Vater.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
Ich denke, dass die Mutter insbesondere im ersten Lebensjahr für das Kind wichtiger ist als der Vater.	✓		
Ich denke, dass die Mutter <i>in den ersten beiden Lebensjahren</i> für das Kind wichtiger ist als der Vater. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	☞
Jede Frau besitzt einen natürlichen Mutterinstinkt, der jedoch einem Mann in vergleichbarer Weise fehlt.			☞
Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater oder eine andere Bezugsperson.			
Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater. (Umformulierung des Items im Rahmen des Experten-Reviews)	✓		
<i>Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.</i> (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als eine andere Bezugsperson (z.B. eine Tagesmutter). (Zusätzliche Alternativ-Formulierung im Rahmen des Experten-Reviews bzw. Items gesplittet)	✓		☞

Wohltuender Sexismus	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
Weder ein engagierter Vater noch eine ausgebildete Betreuerin können eine Mutter bei der Betreuung eines Kindes unter drei Jahren angemessen ersetzen.	✓	☞	
<i>Auch ein engagierter Vater wird eine Mutter bei der Betreuung eines Kindes unter drei Jahren kaum angemessen ersetzen können. (Zusätzliche Alternativ-Formulierung im Rahmen des Experten-Reviews)</i>	✓		
<i>Ich denke, dass ein Vater die Betreuung eines 1-jährigen Kindes kaum so gut übernehmen kann wie eine Mutter. (Umformulierung des Items im Rahmen der kognitiven Interviews)</i>		✓	☞
Die Person, die ein Kind unter einem Jahr am meisten braucht, ist meiner Meinung nach seine Mutter.	✓	✓	☞
Eine Mutter spürt viel besser als ein Vater, was die eigenen Kinder brauchen.	✓		
Eine Mutter spürt besser als ein Vater, was die eigenen Kinder brauchen. (Umformulierung des Items im Rahmen der kognitiven Interviews)		✓	✓
Eine Mutter mit kleinen Kindern sollte – wenn überhaupt – nur einen Teilzeit-Job haben, damit die Doppelbelastung aus Beruf und Familie für sie zu bewältigen ist.	☞		
Obwohl heutzutage viele Mütter arbeiten gehen, finde ich es eigentlich schöner, wenn sie sich ausschließlich der Betreuung ihrer Kinder widmen könnten.	☞		
Wenn es die finanzielle Lage der Familie erlaubt, sollte eine Mutter kleiner Kinder lieber keinen Job haben, da ansonsten die Doppelbelastung aus Beruf und Familie für sie vielleicht zu groß wird.	☞		
Junge Frauen sollten sich von dem gesellschaftlichen Anspruch befreien, unbedingt beruflichen Erfolg erzielen zu müssen, da ansonsten ihre Pläne in Bezug auf eine Familiengründung darunter leiden könnten.	☞		
Eine bereits berufstätige Mutter sollte es lieber vermeiden, sich auch noch mit einer leitenden beruflichen Position und den damit verbundenen Anforderungen zu belasten.	☞		
Für Mutter und Kind ist es am schönsten, wenn die Mutter den ganzen Tag zu Hause bei ihrem Kind bleiben kann.	☞		



Wohltuender Sexismus	Pretest-Phase 1	Pretest-Phase 2	Pretest-Phase 3
<p>Berufstätige Mütter sind heutzutage einer so enormen Doppelbelastung ausgesetzt, dass es eigentlich besser für sie wäre, wenn sie sich ausschließlich der Betreuung ihrer Kinder widmen würden.</p> <p><i>Anmerkungen.</i> Die Itemreihenfolge stimmt mit derjenigen aus Pretest-Phase 1 überein. Eingerückte Items geben die endgültigen Um- bzw. Neuformulierungen im Rahmen des ExpertInnen-Reviews bzw. im Rahmen der kognitiven Interviews für die nächste Pretest-Phase mitgenommen wurden (kursive Hervorhebung der Änderungen). ✓ Das Item blieb nach Abschluss der Pretest-Phase weiterhin im Itempool enthalten. ✗ Das Item wurde nach Abschluss der Pretest-Phase aus dem Itempool entfernt.</p>			

## Anhang 2: Schriftlicher Fragebogen aus Studie 1

### Herzlich Willkommen ...

... zu dieser Umfrage, die im Rahmen eines Dissertationsprojektes an der **Philipps-Universität Marburg** mit Unterstützung der **Deutschen Forschungsgemeinschaft** durchgeführt wird.

Wir freuen uns sehr, dass Sie an dieser Erhebung teilnehmen möchten. Der Fragebogen umfasst Aussagen zu verschiedenen gesellschaftlichen Themen. Lesen Sie sich bitte jede Aussage aufmerksam durch und überlegen Sie, ob Sie ihr zustimmen oder nicht zustimmen. Es gibt dabei **keine „richtigen“ oder „falschen“ Antworten**, sondern es interessiert uns **Ihre persönliche Meinung**. Es ist wichtig, dass Sie möglichst **alle** Fragen beantworten. Wenn Sie sich einmal nicht sicher sind, kreuzen Sie bitte die Antwortkategorie an, die am ehesten Ihrer Meinung entspricht. Zur Bewertung der Aussagen steht Ihnen eine sechsfach abgestufte Skala zur Verfügung.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Ihre Antworten und Daten bleiben selbstverständlich **anonym**.

**Beginnen möchten wir mit einigen Fragen zur Berufstätigkeit von Frauen. Inwieweit stimmen Sie den folgenden Aussagen zu oder nicht zu?**

		Stimme überhaupt nicht zu				Stimme voll und ganz zu
(1)	Ich gehe davon aus, dass viele berufstätige Frauen ihre weiblichen Reize gezielt einsetzen, um ihrer Karriere auf die Sprünge zu helfen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(2)	Frauen können Druck schlechter standhalten als Männer und sind daher für leitende Positionen in der Berufswelt weniger gut geeignet.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(3)	Frauen haben häufig zu wenig Durchsetzungsvermögen um sich in der männlichen Arbeitswelt zu behaupten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(4)	Frauen sind für Führungspositionen im Job weniger gut geeignet als Männer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(5)	Viele erfolgreich berufstätige Frauen verlieren ihre Weiblichkeit.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(6)	Viele Karrierefrauen sind kaltschnäuzig.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(7)	Wenn mehrere Frauen in einem Büro zusammenarbeiten, gibt es häufig „Zickenterror“.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(8)	Viele erfolgreich berufstätige Frauen verdanken ihre berufliche Stellung allein der Frauenquote (festgelegter Frauenanteil bei der Besetzung von Gremien oder Stellen).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(9)	Viele Frauen rufen im Berufsleben schon beim geringsten Anlass nach der Frauenbeauftragten.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(10)	Frauen sind auf Grund ihrer Eigenschaften für typisch weibliche Berufe (wie z. B. Grundschullehrerin oder Krankenschwester) weitaus besser geeignet als Männer.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(11)	Eine Frau ist für den Beruf einer Pilotin oder Ingenieurin genauso gut geeignet wie für den einer Kindergärtnerin oder Kosmetikerin.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(12)	Frauen haben häufig zu wenig Biss, um in ihrem Beruf in eine leitende Position zu gelangen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(13)	Viele berufstätige Frauen sind selbst daran schuld, dass sie keine Karriere machen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(14)	Viele Frauen beschwerten sich im Berufsleben schon aus nichtigem Anlass darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>







	Stimme überhaupt nicht zu				Stimme voll und ganz zu	
(73) Frauen versuchen, Macht zu erlangen, indem sie Männer immer mehr beherrschen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(74) Jeder Mann sollte eine Frau haben, die er wirklich liebt.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(75) Männer sind ohne Frauen unvollkommen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(76) Frauen übertreiben Probleme, die sie am Arbeitsplatz haben.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(77) Hat eine Frau erst einmal einen Mann „rumgekriegt“, dann versucht sie, ihn an die kurze Leine zu legen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(78) Wenn Frauen in einem fairen Wettbewerb gegenüber Männern den Kürzeren ziehen, behaupten sie gerne, sie seien diskriminiert worden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(79) Eine Frau sollte von ihrem Mann auf Händen getragen werden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(80) Viele Frauen haben Spaß daran, mit Männern zu „spielen“, indem sie sich zuerst verführerisch geben, dann aber die Annäherungsversuche der Männer zurückweisen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(81) Verglichen mit Männern haben Frauen ein besseres moralisches Empfinden.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(82) Ein Mann sollte bereit sein, sein eigenes Wohl zu opfern, um für seine Frau sorgen zu können.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(83) Feministinnen stellen an Männer vollkommen berechnete Forderungen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
(84) Verglichen mit Männern haben Frauen einen feineren Sinn für Kultur und einen besseren Geschmack.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Abschließend benötigen wir nur noch einige Angaben zu Ihrer Person.**

(85) **Ihr Geschlecht:**

Weiblich.....  Männlich.....

(86) **Ihr Alter:** \_\_\_\_\_ Jahre

(87) **Ihr Studiengang:** \_\_\_\_\_

(88) **Haben Sie eine/n feste/n Partner/in?**

Ja.....

Nein.....

(89) **Haben Sie Kinder?**

Ja.....

Nein.....

(90) **Ist Ihre Mutter zurzeit / oder war sie früher erwerbstätig?**

Ja, sie ist zurzeit / war früher erwerbstätig.....

Nein, sie war nie erwerbstätig.....

**Vielen Dank für Ihre Teilnahme!**

## Anhang 3: Itemliste für die Befragten in Runde 1 der kognitiven Interviews

Philipps-Universität Marburg  
Fachbereich 03 Gesellschaftswissenschaften  
und Philosophie  
DFG-Graduiertenkolleg  
Dipl.-Sozialwiss. Astrid Schübler



### Fragebogen kognitiver Pretest Runde 1

#### Antwortkategorien:

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

#### Aussagen:

1. Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen dabei die Bedürfnisse ihrer Familie.
2. Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.
3. Ich finde Frauen egoistisch, die ihre Karriere für die Betreuung ihrer Kinder nicht zurückstellen wollen.
4. Ich finde Frauen egoistisch, die bereits wenige Wochen nach der Geburt ihres Kindes wieder in ihren Beruf zurückkehren, weil sie ihre Karriere weiter verfolgen möchten.
5. Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.
6. Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.
7. Viele Frauen beschwerten sich im Berufsleben schon aus nichtigem Anlass darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.
8. Schon bei Kleinigkeiten beschwerten sich viele Frauen im Berufsleben darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.
9. Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater.

**Antwortkategorien:**Stimme über-  
haupt nicht zu

Stimme nicht zu

Stimme eher  
nicht zu

Stimme eher zu

Stimme zu

Stimme voll und  
ganz zu

10. Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als eine andere Bezugsperson (z.B. eine Tagesmutter).
11. Frauen verfügen über ein größeres Talent als Männer, die Wohnung behaglich und gemütlich zu gestalten.
12. Frauen verstehen sich viel besser als Männer darauf, die Wohnung gemütlich und heimelig herzurichten.
13. Frauen besitzen einfach eine viel größere Fürsorglichkeit als Männer.
14. Eine Mutter spürt viel besser als ein Vater, was die eigenen Kinder brauchen.
15. Die Person, die ein Kind unter einem Jahr am meisten braucht, ist meiner Meinung nach seine Mutter.
16. Auch ein engagierter Vater wird eine Mutter bei der Betreuung eines Kindes unter drei Jahren kaum angemessen ersetzen können.
17. Frauen besitzen ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.

Nun möchte ich Sie bitten, dass Sie sich die folgenden Fragen durchlesen und wieder überlegen, ob Sie ihnen zustimmen oder nicht zustimmen. **Kreuzen Sie dann jeweils die Antworten an, die Ihrer Meinung am besten entsprechen.** Sollten Sie mit einer Frage **Probleme** haben (z. B. mit bestimmten Begriffen oder Formulierungen oder sollten Sie nicht wissen, worauf sich die jeweilige Frage bezieht), dann markieren Sie sie bitte durch **einen Kreis am Rande des Fragebogens.**

**18. Viele berufstätige Frauen sind selbst daran schuld, dass sie keine Karriere machen.**Stimme über-  
haupt nicht zu

Stimme nicht zu

Stimme eher  
nicht zu

Stimme eher zu

Stimme zu

Stimme voll und  
ganz zu



19. Wenn mehrere Frauen in einem Büro zusammenarbeiten, gibt es häufig „Zickenterror“.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

20. Ich gehe davon aus, dass viele berufstätige Frauen ihre weiblichen Reize gezielt einsetzen, um ihrer Karriere auf die Sprünge zu helfen.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

21. Erfolgreich berufstätige Frauen sind häufig „Mannweiber“.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

22. Viele Karrierefrauen sind kaltherzig.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

23. Viele erfolgreich berufstätige Frauen sind gefühllos.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

24. Weibliche Führungskräfte verhalten sich häufig zickig.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

25. Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

26. In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass sie ihren Kindern zu wenig Aufmerksamkeit schenken.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

27. Erfolgreich berufstätige Frauen vernachlässigen häufig ihre Kinder wegen ihrem Beruf.

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**28. Viele erfolgreich berufstätige Frauen sind derart auf ihre Karriere fixiert, dass ihre Kinder darunter leiden.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**29. Frauen haben häufig zu wenig Biss, um in ihrem Beruf in eine leitende Position zu gelangen.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**30. Viele erfolgreich berufstätige Frauen verlieren ihre Weiblichkeit.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**31. Für viele berufstätige Frauen von heute zählt allein das Streben nach beruflichem Erfolg und Anerkennung.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**32. Ich denke, dass die Mutter insbesondere im ersten Lebensjahr für das Kind wichtiger ist als der Vater.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**33. Eine Mutter sollte sich in den ersten Jahren selbst um die Betreuung ihres Kindes kümmern, da sie eine ganz besondere Verbindung zum Kind besitzt.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**34. Frauen haben einfach einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**35. Ingeheim sehnen sich doch alle Frauen nach einem Mann, in dessen Armen sie Schutz und Geborgenheit finden.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**36. Viele Frauen wünschen sich einen Mann, der der Behüter und Beschützer des gemeinsamen Heims ist.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**37. Eine Frau braucht die starke Schulter eines Mannes, an die sie sich anlehnen kann.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**38. Viele Frauen wünschen sich einen Mann, der ihr Herz im Sturm erobert.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**39. Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**40. Viele Frauen sind im Berufsleben peinlich genau darauf bedacht, im Vergleich zu ihren männlichen Kollegen bloß nicht zu kurz zu kommen (z. B. in Bezug auf das Gehalt oder bei Beförderungen).**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Anhang 4: Evaluationsfragebogen in Runde 1 der kognitiven Interviews

Philipps-Universität Marburg  
Fachbereich 03 Gesellschaftswissenschaften  
und Philosophie  
DFG-Graduiertenkolleg  
Dipl.-Sozialwiss. Astrid Schüßler



### Fragebogen kognitiver Pretest Runde 1

INTERVIEW NR.:

DATUM:

ANFANGSZEIT:

#### INSTRUKTION

Zuerst einmal möchte ich mich für Ihr Kommen bedanken. Bevor wir mit der Befragung beginnen, möchte ich Sie gerne über die Ziele und den Ablauf dieser Befragung informieren.

Ziel dieser Befragung ist es, einen neu entwickelten Fragebogen zum Thema „Frauen und Männer und deren Verhältnis zueinander“ zu testen. Dabei soll herausgefunden werden, wie gut die einzelnen Fragen „funktionieren“, d. h. ob sie Verständnisprobleme bereiten oder ob sie klar und eindeutig formuliert sind.

Hierzu werde ich Ihnen verschiedene Fragen vorlegen, die Sie – genauso wie bei einer „normalen“ Befragung – beantworten. Hierfür wählen Sie bitte eine der 6 vorgegebenen Antwortkategorien aus, um anzugeben wie sehr Sie den einzelnen Aussagen zustimmen bzw. nicht zustimmen würden.

Im Anschluss daran wird die Verständlichkeit der Frageformulierung untersucht, indem ich Ihnen einige Nachfragen stelle. So werde ich Sie beispielsweise fragen, was Sie gedacht haben, als Sie einen bestimmten Begriff gelesen haben. Oder ich werde Sie fragen, was Sie denken, worum es in der Frage geht.

Dabei gibt es keine „richtigen“ oder „falschen“ Antworten. Es geht lediglich darum herauszufinden, ob die Frage so verstanden wird, wie es beabsichtigt war. Zögern Sie also nicht es anzusprechen, wenn Ihnen etwas unklar erscheint oder wenn eine Frage für Sie schwer zu beantworten bzw. zu verstehen ist. Diese Hinweise sind sehr wichtig für mich.

Wenn Sie damit einverstanden sind, würde ich das Interview gerne aufzuzeichnen und mir außerdem Notizen machen. Das hilft mir dabei im Nachhinein zu beurteilen, welche Fragen des Fragebogens umformuliert werden sollten. Alles was Sie sagen wird dabei natürlich vertraulich behandelt.

Haben Sie noch weitere Fragen, bevor wir beginnen?

[Wenn nein:] Dann lassen Sie uns nun mit den Fragen beginnen, die wir testen möchten.

**1. Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen dabei die Bedürfnisse ihrer Familie.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?
- b) Warum haben Sie sich für den Skalenwert ... entschieden?
- c) Als Sie eben die Formulierung „Bedürfnisse ihrer Familie“ gelesen haben, an was haben Sie dabei gedacht?

**2. Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?
- b) Von welchem Alter der Kinder sind Sie ausgegangen, als Sie den Begriff „kleine Kinder“ gelesen haben?
- c) Stellen Sie sich vor, Sie müssten jemandem den Begriff „Kinderkrippe“ erklären. Was würden Sie da sagen?

d) Gibt es für Sie einen Unterschied zwischen einer „Kinderkrippe“ oder einer „Kindertagesstätte“?

e) Als Sie das Wort „Rabenmutter“ gelesen haben, an was haben Sie dabei gedacht? Was verstehen Sie darunter?

**3. Ich finde Frauen egoistisch, die ihre Karriere für die Betreuung ihrer Kinder nicht zurückstellen wollen.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Warum haben Sie sich für den Skalenwert ... entschieden?

b) Wie würden Sie auf folgende Aussage antworten:

**4. Ich finde Frauen egoistisch, die bereits wenige Wochen nach der Geburt ihres Kindes wieder in ihren Beruf zurückkehren, weil sie ihre Karriere weiter verfolgen möchten.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**5. Erfolgreich berufstätige Frauen sollten ein schlechtes Gewissen haben, wenn sie ihre Kinder schon vor dem Kindergartenalter zur Betreuung in fremde Hände geben.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?

- b) Als Sie eben die Formulierung „zur Betreuung in fremde Hände geben“ gelesen haben, an was haben Sie dabei gedacht? Was verstehen Sie darunter?

**6. Ich habe kein Verständnis dafür, wenn eine Mutter kleiner Kinder einen Beruf ausübt, nur um sich selbst zu verwirklichen, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- a) Würden Sie bitte die Frage noch einmal mit Ihren eigenen Worten wiederholen. Was denken Sie warum es bei dieser Frage geht?
- b) Warum haben Sie sich für den Skalenwert ... entschieden?
- c) Von welchem Alter der Kinder sind Sie ausgegangen, als Sie den Begriff „kleine Kinder“ gelesen haben?

Zusatzfrage:

- a) Was verbinden Sie mit dem Begriff „Karrierefrau“? Was verstehen Sie darunter?
- b) Ist der Begriff „Karrierefrau“ für Sie irgendwie besonders positiv oder negativ besetzt? Wie sieht das im Vergleich zu einem „Karrieremann“ aus?
- c) Was bedeutet für Sie der Begriff „erfolgreiche berufstätige Frau“? Was verstehen Sie darunter?

d) Ist der Begriff „erfolgreiche berufstätige Frau“ für Sie irgendwie besonders positiv oder negativ besetzt?

e) Besteht für Sie ein Unterschied zwischen diesen beiden Begriffen oder nicht?  
[Wenn ja:] Worin besteht der Unterschied?

**7. Viele Frauen beschwerten sich im Berufsleben schon aus nichtigem Anlass darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?

b) Als Sie die Worte „aus nichtigem Anlass“ gelesen haben, an was haben Sie dabei gedacht? Können Sie mir ein paar Beispiele dafür nennen?

c) Wie würden Sie auf folgende Aussage antworten:

**8. Schon bei Kleinigkeiten beschwerten sich viele Frauen im Berufsleben darüber, sie seien im Vergleich zu einem männlichen Kollegen benachteiligt worden.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**9. Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als der Vater.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



**10. Insbesondere kleine Kinder brauchen ihre Mutter, da sie auf die Bedürfnisse des Kindes viel besser eingehen kann als eine andere Bezugsperson (z.B. eine Tagesmutter).**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?
- b) Von welchem Alter der Kinder sind Sie ausgegangen, als Sie den Begriff „kleine Kinder“ gelesen haben?
- c) Als Sie die Worte „Bedürfnisse des Kindes“ gelesen haben, an was haben Sie dabei gedacht? Können Sie mir ein paar Beispiele dafür nennen?
- d) Besteht für Sie ein Unterschied zwischen dem vorhergehendem und diesem Item oder nicht?  
[Wenn ja:] Worin besteht der Unterschied?

**11. Frauen verfügen über ein größeres Talent als Männer, die Wohnung behaglich und gemütlich zu gestalten.**

Stimme über- haupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

- a) Warum haben Sie sich für den Skalenwert ... entschieden?
- b) Als Sie das Wort „Talent“ gelesen haben, an was haben Sie dabei gedacht? Was bedeutet dieser Begriff hier in dieser Aussage für Sie?

c) Wie würden Sie auf folgende Aussage antworten:

**12. Frauen verstehen sich viel besser als Männer darauf, die Wohnung gemütlich und heimelig herzurichten.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**13. Frauen besitzen einfach eine viel größere Fürsorglichkeit als Männer.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Warum haben Sie sich für den Skalenwert ... entschieden?

b) Als Sie das Wort „Fürsorglichkeit“ gelesen haben, an was haben Sie dabei gedacht? Was bedeutet dieser Begriff hier in dieser Aussage für Sie?

**14. Eine Mutter spürt viel besser als ein Vater, was die eigenen Kinder brauchen.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?

b) Als Sie das Wort „spüren“ gelesen haben, an was haben Sie dabei gedacht? Können Sie ein paar Beispiele geben?

c) Von welchem Alter der Kinder sind Sie ausgegangen, als Sie die Aussage eben gelesen haben?

**15. Die Person, die ein Kind unter einem Jahr am meisten braucht, ist meiner Meinung nach seine Mutter.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Warum haben Sie sich für den Skalenwert ... entschieden?

**16. Auch ein engagierter Vater wird eine Mutter bei der Betreuung eines Kindes unter drei Jahren kaum angemessen ersetzen können.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Würden Sie bitte die Frage noch einmal mit Ihren eigenen Worten wiederholen. Was denken Sie worum es bei dieser Frage geht?

b) Warum haben Sie sich für den Skalenwert ... entschieden?

**17. Frauen besitzen ein größeres Talent als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.**

Stimme überhaupt nicht zu	Stimme nicht zu	Stimme eher nicht zu	Stimme eher zu	Stimme zu	Stimme voll und ganz zu
<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

a) Als Sie eben die Frage beantwortet haben, an was haben Sie da gedacht? Was ist Ihnen durch den Kopf gegangen bis Sie die Antwort gegeben haben?

b) Würden Sie bitte die Frage noch einmal mit Ihren eigenen Worten wiederholen. Was denken Sie worum es bei dieser Frage geht?

c) Wie haben Sie das Wort „Talent“ in dieser Frage aufgefasst?

Zusatzfrage:

Hatten Sie eben bei Beantwortung der Fragen ansonsten irgendwelche Probleme (mit einem Begriff, einer bestimmten Formulierung, usw.)?

[Wenn ja:] Welche waren das?

Nun möchte ich Sie bitten, dass Sie sich die folgenden Fragen durchlesen und wieder überlegen, ob Sie ihnen zustimmen oder nicht zustimmen. **Kreuzen Sie dann jeweils die Antworten an, die Ihrer Meinung am besten entsprechen.** Sollten Sie mit einer Frage **Probleme** haben (z. B. mit bestimmten Begriffen oder Formulierungen oder sollten Sie nicht wissen, worauf sich die jeweilige Frage bezieht), dann markieren Sie sie bitte durch **einen Kreis am Rande des Fragebogens.**

**18. Viele berufstätige Frauen sind selbst daran schuld, dass sie keine Karriere machen.**

**19. Wenn mehrere Frauen in einem Büro zusammenarbeiten, gibt es häufig „Zickenterror“.**

**20. Ich gehe davon aus, dass viele berufstätige Frauen ihre weiblichen Reize gezielt einsetzen, um ihrer Karriere auf die Sprünge zu helfen.**

- 21. Erfolgreich berufstätige Frauen sind häufig „Mannweiber“.**
- 22. Viele Karrierefrauen sind kaltherzig.**
- 23. Viele erfolgreich berufstätige Frauen sind gefühllos.**
- 24. Weibliche Führungskräfte verhalten sich häufig zickig.**
- 25. Frauen, die gerne weiter erwerbstätig sein möchten, auch wenn sie Kinder unter drei Jahren haben, denken meiner Meinung nach nur an sich.**
- 26. In der heutigen Zeit sind viele Frauen derart auf ihre Karriere fixiert, dass sie ihren Kindern zu wenig Aufmerksamkeit schenken.**
- 27. Erfolgreich berufstätige Frauen vernachlässigen häufig ihre Kinder wegen ihrem Beruf.**
- 28. Viele erfolgreich berufstätige Frauen sind derart auf ihre Karriere fixiert, dass ihre Kinder darunter leiden.**

29. Frauen haben häufig zu wenig Biss, um in ihrem Beruf in eine leitende Position zu gelangen.
30. Viele erfolgreich berufstätige Frauen verlieren ihre Weiblichkeit.
31. Für viele berufstätige Frauen von heute zählt allein das Streben nach beruflichem Erfolg und Anerkennung.
32. Ich denke, dass die Mutter insbesondere im ersten Lebensjahr für das Kind wichtiger ist als der Vater.
33. Eine Mutter sollte sich in den ersten Jahren selbst um die Betreuung ihres Kindes kümmern, da sie eine ganz besondere Verbindung zum Kind besitzt.
34. Frauen haben einfach einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.
35. Insgeheim sehnen sich doch alle Frauen nach einem Mann, in dessen Armen sie Schutz und Geborgenheit finden.
36. Viele Frauen wünschen sich einen Mann, der der Behüter und Beschützer des gemeinsamen Heims ist.

37. Eine Frau braucht die starke Schulter eines Mannes, an die sie sich anlehnen kann.

38. Viele Frauen wünschen sich einen Mann, der ihr Herz im Sturm erobert.

39. Eine Mutter ist vom Wesen her sanfter und warmherziger als ein Vater und daher auch besser für die Betreuung ihres Kindes geeignet.

40. Viele Frauen sind im Berufsleben peinlich genau darauf bedacht, im Vergleich zu ihren männlichen Kollegen bloß nicht zu kurz zu kommen (z. B. in Bezug auf das Gehalt oder bei Beförderungen).

Jetzt sind wir schon fast am Ende meiner Befragung angelangt. Ich benötige nur noch ein paar Angaben zu Ihrer Person.

41. [Von der Interviewerin selbst auszufüllen: ] **Das Geschlecht des/r Befragten?**

weiblich:

männlich:

42. **Wie alt sind Sie?** .....Jahre

43. **Welchen höchsten schulischen oder universitären Abschluss haben Sie bisher erreicht?**

Keinen Schulabschluss.

Volks-, Hauptschule.

Realschule/mittlere Reife.

Gymnasium/(Fach-)Abitur.

Abgeschlossenes Studium.

**44. Haben Sie eine/n feste/n Partner/in?**

ja  => Bitte weiter mit Frage 44

nein  => Bitte weiter mit Frage 46

**45. Wohnen Sie mit Ihrem/r Partner/in zusammen?**

ja

nein

**46. Ist Ihr/e Partner/in erwerbstätig?**

Ja, Vollzeit.

Ja, Teilzeit.

Nein.

**47. Sind Sie erwerbstätig?**

Ja, Vollzeit.

Ja, Teilzeit.

Nein.

**48. Haben Sie Kinder? (Ich meine damit eigene leibliche Kinder, auch solche die nicht mehr in Deinem Haushalt leben?)**

ja

nein

**49. Wie viele Kinder haben Sie?** .....

**50. Wie alt ist/sind Dein/e Kind/er?** .....

ENDZEIT:



## Anhang 5: „Problemtabelle“ aus Runde 1 der kognitiven Interviews

### Ergebnisse kognitiver Pretest Runde 1

#### Anmerkungen zum Ausfüllen:

(1) **AK** = Antwortkategorien, die der/die Befragte gewählt hat, sollen hier in der Tabelle eingetragen werden => Werte von 1-6:

- 1 = überhaupt nicht zustimmen
- 2 = nicht zustimmen
- 3 = eher nicht zustimmen
- 4 = eher zustimmen
- 5 = zustimmen
- 6 = voll und ganz zustimmen

(2) Folgende Fragen haben eine Spalte für „**Nein**“:

- a) Gibt es für Dich einen Unterschied zwischen...
- b) Hattest Du eben bei Beantwortung der Frage irgendwelche Probleme?

Trifft „nein“ zu, dann soll in der jeweiligen Spalte ein Kreuz eingetragen werden. Wird ein Unterschied gemacht oder gab es Probleme bei Beantwortung der Frage, dann wird in der Spalte „Ja“ die entsprechende Antwort der Befragten aufgenommen.

(3) **Kommentar:** Im Anschluss an jedes Interview das durchgeführt wurde, sollen die Interviewer hier ihre eigenen Kommentare, Anmerkungen, Beobachtungen, etc. notieren.

1. **Viele erfolgreiche berufstätige Frauen denken nur an ihr berufliches Vorankommen und vergessen dabei die Bedürfnisse ihrer Familie.**

Anmerkungen / Probleme / spontane Äußerungen der Befragten

ID	Anmerkungen/Probleme
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

- a) Als Du eben die Frage beantwortet hast, an was hast Du da gedacht? Was ist Dir durch den Kopf gegangen bis Du die Antwort gegeben hast?

ID	AK	Antworten
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

- b) Warum hast Du Dich für den Skalenwert ... entschieden?

ID	AK	Begründungen
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

- c) Als Du eben die Formulierung „Bedürfnisse ihrer Familie“ gelesen hast, an was hast Du dabei gedacht?

ID	Definitionen / Antworten
1	
2	
3	
4	
5	

6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

<b>Kommentare Interviewer</b>

**2. Frauen, die ihre kleinen Kinder aus Karrieregründen in eine Kinderkrippe geben, sind für mich Rabenmütter.**

Anmerkungen / Probleme / spontane Äußerungen der Befragten

ID	Anmerkungen/Probleme
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

a) Als Du eben die Frage beantwortet hast, an was hast Du da gedacht? Was ist Dir durch den Kopf gegangen bis Du die Antwort gegeben hast?

ID	AK	Antworten
1		
2		
3		
4		
5		

6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

b) Von welchem Alter der Kinder bist Du ausgegangen, als Du den Begriff „kleine Kinder“ gelesen hast?

ID	AK	Alter	Definitionen / Antworten
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			

c) Stell Dir vor, Du müsstest jemandem den Begriff „Kinderkrippe“ erklären. Was würdest Du da sagen?

ID	Definitionen / Antworten
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	

15
----

d) Gibt es für Dich einen Unterschied zwischen einer „Kinderkrippe“ oder einer „Kindertagesstätte“?

ID	Nein	Ja => Was ist der Unterschied? Definitionen / Antworten
1		
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		
15		

e) Als Du das Wort „Rabenmutter“ gelesen hast, an was hast Du dabei gedacht? Was verstehst Du darunter?

ID	Definitionen / Antworten
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	

<b>Kommentare Interviewer</b>

## Anhang

Merkmale der Befragten: Geschlecht, Alter, Bildung, Erwerbstätigkeit

ID	Geschlecht (w / m)	Alter (in Jahren)	Schulabschluss (kein, HS, MR, Abi, Studium)	Eigene Erwerbstätigkeit (VZ, TZ, Keine)
01				
02				
03				
04				
05				
06				
07				
08				
09				
10				
11				
12				
13				
14				
15				

Merkmale des/r Partners/in und der Kinder

ID	Feste/r Partner/in (Ja / Nein)	Zusammen wohnen (Ja / Nein)	Erwerbstätigkeit Partner/in (VZ, TZ, Keine)	Anzahl der Kinder	Alter der Kinder (in Jahren)
01					
02					
03					
04					
05					
06					
07					
08					
09					
10					
11					
12					
13					
14					
15					

Befragungsdauer

ID	Dauer (Minuten)
1	
2	

---

3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
	Mittelwert:
	Min.:
	Max.:







(34) Karrierefrauen sind häufig „Mannweiber“.

**Inwieweit stimmen Sie den folgenden Aussagen zu oder nicht zu?**

(35) Diskriminierung von Frauen ist in Deutschland immer noch ein Problem.

(36) Frauen und Männer haben in der heutigen Gesellschaft die gleichen Chancen, etwas zu erreichen.

(37) Die Forderungen von Frauen nach Gleichberechtigung sind leicht nachzuvollziehen.

(38) Wenn Frauen tatsächlich einmal schlechter bezahlt werden als Männer, dann nur deshalb, weil sie einfachere Arbeit zu leisten haben.

(39) Im Fernsehen gibt es häufig frauenfeindliche Darstellungen.

(40) Im Allgemeinen werden in unserer Gesellschaft Ehepartner gleich behandelt.

(41) In der Schule werden Mädchen immer noch benachteiligt.

(42) Heutzutage werden Frauen im Berufsleben fair behandelt.

(43) In den westlichen Ländern ist Gleichberechtigung von Frauen schon lange verwirklicht.

(44) Frauen finden häufig keine gut bezahlte Arbeit, weil sie diskriminiert werden.

**Im Folgenden werden Sie erneut verschiedene Aussagen über Frauen und Männer lesen. Wie ist Ihre Meinung dazu?**

(45) Egal wie erfolgreich ein Mann auch sein mag, ohne eine Frau, die ihn liebt, fehlt ihm etwas ganz Wichtiges.

(46) Viele Frauen versuchen unter dem Deckmantel der Gleichberechtigung, besondere Vergünstigungen zu erlangen, wie z. B. eine Bevorzugung bei der Besetzung von Arbeitsstellen.

(47) Bei einer Katastrophe sollten Frauen vor Männern gerettet werden.

(48) Die meisten Frauen interpretieren harmlose Äußerungen oder Handlungen als frauenfeindlich.

(49) Frauen sind zu schnell beleidigt.

(50) Man kann im Leben erst richtig glücklich sein, wenn man einen Partner hat, den man liebt.

(51) Was Feministinnen wirklich wollen, ist, dass Frauen mehr Macht bekommen als Männer.

(52) Viele Frauen haben eine Art von Ehrlichkeit, die nur wenige Männer besitzen.

(53) Frauen sollten von Männern umsorgt und geschützt werden.

(54) Die meisten Frauen sehen gar nicht, was Männer alles für sie tun.

(55) Frauen versuchen, Macht zu erlangen, indem sie Männer immer mehr beherrschen.

(56) Jeder Mann sollte eine Frau haben, die er wirklich liebt.





	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Abschließend benötigen wir nur noch einige Angaben zu Ihrer Person.**

**Ihr Geschlecht:**

Weiblich.....  Männlich.....

**Ihr Alter:** \_\_\_\_\_ Jahre

**Welches Fach bzw. welche Fächer studieren Sie?**

1. Fach: _____	als	Hauptfach ... <input type="checkbox"/>
		Hauptfach ... <input type="checkbox"/>
2. Fach: _____	als	Nebenfach... <input type="checkbox"/>
		Hauptfach ... <input type="checkbox"/>
3. Fach: _____	als	Nebenfach... <input type="checkbox"/>

**Welchen Abschluss streben Sie an? (nur 1 Nennung)**

Diplom.....   
 Magister.....   
 Bachelor.....   
 Master.....   
 Staatsexamen.....   
 Sonstiges, bitte eintragen: \_\_\_\_\_

**In welchem Semester sind Sie?**

Semester insgesamt: \_\_\_\_ Fachsemester: \_\_\_\_

**Haben Sie eine/n feste/n Partner/in?**

Ja.....   
 Nein.....

**Haben Sie Kinder?**

Ja.....   
 Nein.....

**Ist Ihre Mutter zurzeit / oder war sie früher erwerbstätig?**

Ja, sie ist zurzeit / war früher erwerbstätig.....   
 Nein, sie war nie erwerbstätig.....

**Vielen Dank für Ihre Teilnahme!**

## Anhang 7: Variablenliste für die telefonische Befragung in Studie 8

### GMF (Gruppenbezogene Menschenfeindlichkeit) Graduierten-Survey 2010

Variable			
	Recodiert	Skalen	Lable
GSCH			(nur nachfragen, wenn unklar: Darf ich fragen, welches Geschlecht Sie haben? Sind Sie ....)
AA01	AA01r		Bitte sagen Sie mir, inwiefern sich die Wirtschafts- und Finanzkrise bisher auf Ihr Leben ausgewirkt hat. Haben Sie durch die Krise Nachteile?
AA07	AA07g		Oder haben Sie durch die Krise Vorteile?
AA02	AA02r		Die Ursache der Wirtschafts- und Finanzkrise liegt...bei den Bankern und Spekulanten
AA03	AA03r		...bei den Amerikanern
AA04	AA04r		...bei den Ausländern
AA05	AA05r		...bei denen, die unseren Sozialstaat ausnutzen
AA06	AA06r		...im weltweiten Wirtschaftssystem
GELA			In welchem Land sind sie geboren?
STDE			Haben Sie die deutsche Staatsangehörigkeit?
STAA			Welche Staatsangehörigkeit haben Sie?
STAB			Welche Staatsangehörigkeit haben Sie außer der deutschen?
STEL			Sind ihre Eltern in Deutschland geboren?
STVA			In welchem Land ist ihr Vater geboren?
STMU			In welchem Land ist ihre Mutter geboren?
OSTWEST			In welchem Teil Deutschlands sind sie aufgewachsen? D.h. wo haben sie bis zu ihrem 15. Lebensjahr überwiegend gelebt?
RELI			Welcher Religionsgemeinschaft gehören Sie an?
DO01	DO01r		Die Gruppen, die in unserer Gesellschaft unten sind, sollen auch unten bleiben.
DO02	DO02r	Soziale Dominanzorientierung <b>mdo1</b> ( <b>do01r,do02r,do03r</b> )	Es gibt Gruppen in der Bevölkerung, die weniger wert sind als andere.
DO03	DO03r		Einige Bevölkerungsgruppen sind nützlicher als andere.

AU01	AU01r	Autoritarismus:	Verbrechen sollten härter bestraft werden.
AU02	AU02r	Gesamtskala <b>mau1 (au01r, au02r, au03r, au04r);</b>	Um Recht und Ordnung zu bewahren, sollte man härter gegen Außenseiter und Unruhestifter vorgehen.
AU03	AU03r	autoritäre Aggression <b>mau2 (au01r, au02r);</b>	Zu den wichtigsten Eigenschaften, die jemand haben sollte, gehören Gehorsam und Respekt vor dem Vorgesetzten.
AU04	AU04r	autoritäre Unterwürfigkeit <b>mau3 (au03, au04)</b>	Wir sollten dankbar sein für führende Köpfe, die uns sagen, was wir tun sollen.
SX03	SX03r		Frauen sollen sich wieder mehr auf die Rolle der Ehefrau und Mutter besinnen.
SX04	SX04r	Sexismus (klassisch) <b>msx2 (sx03r, sx04r)</b>	Für eine Frau sollte es wichtiger sein, ihrem Mann bei seiner Karriere zu helfen, als selbst Karriere zu machen.
SXA01	SXA01r	Sexismus (modern) <b>msx1 (sxa01r, sxa02r)</b>	Die Diskriminierung von Frauen ist in Deutschland immer noch ein Problem.
SXA02	SXA02r		Die jetzige Beschäftigungspolitik benachteiligt die Frauen.
SXA08_2	SXA08_2g		Väter besitzen eine größere Fürsorglichkeit als Mütter.
SXA14_2	SXA14_2g		Ein Vater spürt besser als eine Mutter, was die eigenen Kinder brauchen.
HE05M	HE05Mr	Islamophobie <b>mai2 (he05mr, he12mr)</b>	Durch die vielen Muslime hier fühle ich mich manchmal wie ein Fremder im eigenen Land.
HE12M	HE12Mr		Muslimen sollte die Zuwanderung nach Deutschland untersagt werden.
HE02H	HE02Hr	Heterophobie	Es ist ekelhaft, wenn Homosexuelle sich in der Öffentlichkeit küssen.
HE01H	HE01Hg	<b>mhe5 (he01hg, he02hr, he03hr)</b>	Ehen zwischen zwei Frauen bzw. zwischen zwei Männern sollten erlaubt sein.
HE03H	HE03Hr		Homosexualität ist unmoralisch.
HE02O	HE02Or	Abwertung Obdachloser <b>mhe6 (he01or, he02or, he03or)</b>	Die Obdachlosen in den Städten sind unangenehm.
HE01O	HE01Or		Bettelnde Obdachlose sollten aus den Fußgängerzonen entfernt werden.
HE03O	HE03Or		Die meisten Obdachlosen sind arbeitsscheu.
HE03B	HE03Br	Abwertung Behinderter <b>mhe7 (he03br, he04br, he05br)</b>	Für Behinderte wird in Deutschland zu viel Aufwand betrieben.
HE04B	HE04Br		Viele Forderungen von Behinderten finde ich überzogen.
HE05B	HE05Br		Behinderte erhalten zu viele Vergünstigungen.

<b>SVA01</b>	<b>SVA01g</b>		Wie oft haben Sie Sympathie für die hier lebenden Ausländer empfunden?
<b>SVA02</b>	<b>SVA02g</b>		Wie oft haben Sie Bewunderung für die hier lebenden Ausländer empfunden?
<b>FF04D</b>	<b>FF04Dr</b>		Es leben zu viele Ausländer in Deutschland.
<b>FF08D</b>	<b>FF08Dr</b>	Fremdenfeindlichkeit <b>mf1 (ff04dr, ff08dr)</b>	Wenn Arbeitsplätze knapp werden, sollte man die in Deutschland lebenden Ausländer wieder in ihre Heimat zurückschicken.
<b>RA01</b>	<b>RA01r</b>	Rassismus <b>mra2 (ra01r, ra03r)</b>	Aussiedler sollten besser gestellt werden als Ausländer, da sie deutscher Abstammung sind.
<b>RA03</b>	<b>RA03r</b>		Die Weißen sind zu Recht führend in der Welt.
<b>EV03</b>	<b>EV03r</b>	Etabliertenvorrechte <b>mev1 (ev03r, ev04r)</b>	Wer irgendwo neu ist, sollte sich erst mal mit weniger zufrieden geben.
<b>EV04</b>	<b>EV04r</b>		Wer schon immer hier lebt, sollte mehr Rechte haben, als die, die später zugezogen sind.
<b>AS01</b>	<b>AS01r</b>	Antisemitismus <b>mas1 (as01r, as02r)</b>	Juden haben in Deutschland zuviel Einfluss.
<b>AS02</b>	<b>AS02r</b>		Durch ihr Verhalten sind die Juden an ihren Verfolgungen mitschuldig.
<b>LZ01</b>	<b>LZ01r</b>	Abwertung Langzeitarbeitsloser <b>mlz2 (lz01r, lz03r)</b>	Die meisten Langzeitarbeitslosen sind nicht wirklich daran interessiert, einen Job zu finden.
<b>LZ03</b>	<b>LZ03r</b>		Ich finde es empörend, wenn sich die Langzeitarbeitslosen auf Kosten der Gesellschaft ein bequemes Leben machen.
<b>K38</b>	<b>K38r</b>		Für alte Menschen wird in Deutschland zu viel Aufwand betrieben.
<b>K39</b>	<b>K39r</b>		Alte Menschen erhalten in Deutschland zu viele Vergünstigungen.
<b>SXA07</b>	<b>SXA07r</b>		Frauen besitzen eine größere natürliche Begabung als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen
<b>SXA08</b>	<b>SXA08r</b>		Mütter besitzen eine größere Fürsorglichkeit als Väter.
<b>SXA09</b>	<b>SXA09r</b>		Ich finde Frauen egoistisch, die bereits ein halbes Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um sich selbst zu verwirklichen.
<b>SXA10</b>	<b>SXA10r</b>		Eine Mutter ist vom Wesen her sanfter und warmerherziger als ein Vater und daher auch besser für die Betreuung eines Kindes geeignet.
<b>SXA11</b>	<b>SXA11r</b>		Frauen, die aus Karrieregründen ihr Kind mit einem halbem Jahr in eine Kindertagesstätte geben, sind für mich Rabenmütter.



<b>SXA12</b>	<b>SXA12r</b>	Frauen haben von Natur aus einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.
<b>SXA13</b>	<b>SXA13r</b>	Viele Karrierefrauen denken meiner Meinung nach nur an ihr berufliches Vorankommen und vernachlässigen die Bedürfnisse ihrer Kinder.
<b>SXA14</b>	<b>SXA14r</b>	Eine Mutter spürt besser als ein Vater, was die eigenen Kinder brauchen.
<b>SXA15</b>	<b>SXA15r</b>	Frauen denken meiner Meinung nach nur an sich selbst, wenn sie ein Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um berufliche Anerkennung zu bekommen.
<b>SXA16</b>	<b>SXA16r</b>	Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.
<b>SXA17</b>	<b>SXA17r</b>	Frauen sollten ein schlechtes Gewissen haben, wenn sie aus Karrieregründen ihr Kind mit einem halben Jahr tagsüber zur Betreuung zu einer Tagesmutter geben.
<b>SXA18</b>	<b>SXA18r</b>	Ich denke, dass Karrierefrauen ihren Kindern häufig zu wenig Aufmerksamkeit schenken.
<b>SXA19</b>	<b>SXA19r</b>	Ich habe Verständnis dafür, wenn die Mutter eines 1-jährigen Kindes Vollzeit erwerbstätig sein möchte, obwohl ihr Partner die Familie auch alleine finanziell versorgen könnte.
<b>ISQ01</b>	<b>ISQ01r</b>	So wie wir jetzt in Deutschland zusammenleben, hat es schon seine Richtigkeit.
<b>ISQ02</b>	<b>ISQ02r</b>	Unsere Gesellschaft ist stark veränderungsbedürftig.
<b>DAP01</b>	<b>DAP01r</b>	Viele, die meinen, sie könnten die Welt verändern, sind in meinen Augen Träumer.
<b>DAP02</b>	<b>DAP02r</b>	Es ist nutzlos, sich die Welt anders vorzustellen, als sie wirklich ist.
<b>SZE01</b>	<b>SZE01r</b>	Die Probleme, die es früher gab, wird es auch in Zukunft geben.
<b>SZE02</b>	<b>SZE02r</b>	Vieles an der Art und Weise, wie wir leben, wird sich sowieso nicht ändern.
<b>K26</b>	<b>K26r</b>	Wie viele Ihrer Freunde und guten Bekannten sind Ausländer?
<b>K27</b>	<b>K27r</b>	Wie viele Ihrer Freunde und guten Bekannten sind Muslime?
<b>K28</b>	<b>K28r</b>	Wie viele Ihrer Freunde und guten Bekannten sind Langzeitarbeitslose?

<b>K29</b>	<b>K29r</b>	Wie viele Ihrer Freunde und guten Bekannten sind Frauen?
<b>K30</b>	<b>K30r</b>	Wie viele Ihrer Freunde und guten Bekannten sind behinderte Menschen?
<b>K31</b>	<b>K31r</b>	Wie viele Ihrer Freunde und guten Bekannten sind alte Menschen?
<b>K32</b>	<b>K32r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit Ausländern?
<b>K33</b>	<b>K33r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit Muslimen?
<b>K34</b>	<b>K34r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit Langzeitarbeitslosen?
<b>K35</b>	<b>K35r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit Frauen?
<b>K36</b>	<b>K36r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit behinderten Menschen?
<b>K37</b>	<b>K37r</b>	Wie viele Ihrer Freunde und guten Bekannten sind befreundet mit alten Menschen?
<b>VM42</b>	<b>VM42r</b>	Viele Frauen neigen dazu, eine nett gemeinte Geste zum „sexuellen Übergriff“ hochzuspielen.
<b>VM43</b>	<b>VM43r</b>	Bei Männern ist es biologisch bedingt, dass sie von Zeit zu Zeit sexuellen Druck ablassen müssen.
<b>VM44</b>	<b>VM44r</b>	Viele Frauen beschweren sich schon aus nichtigem Anlass über sexuelle Belästigung, nur um als emanzipiert zu gelten.
<b>VM45</b>	<b>VM45r</b>	Wer als Frau so unvorsichtig ist, nachts durch „dunkle Gassen“ zu gehen, trägt eine gewisse Mitschuld an der eigenen Vergewaltigung.
<b>VM46</b>	<b>VM46r</b>	Wenn eine Frau mit einem Mann eine Beziehung eingeht, muss sie sich darüber im Klaren sein, dass der Mann sein Recht auf Sex einfordern wird.
<b>VM47</b>	<b>VM47r</b>	Frauen bezichtigen ihre Männer häufig einer Vergewaltigung in der Ehe, um sich für eine gescheiterte Beziehung zu rächen.
<b>VM48</b>	<b>VM48r</b>	Harmlose Gesten als „sexuelle Belästigung“ auszulegen ist eine beliebte Waffe im Kampf der Geschlechter.
<b>VM49</b>	<b>VM49r</b>	Wenn eine Frau einen Mann nach dem Discobesuch auf eine Tasse Kaffee in ihre Wohnung einlädt, dann ist sie auf ein sexuelles Abenteuer aus.

VM50	VM50r	Frauen zieren sich gerne. Das bedeutet nicht, dass sie keinen Sex wollen.
ID15G (w)	ID15G(w)r	<b>Ein Mann / eine Frau</b> zu sein ist nebensächlich für mich.
ID16G (W)	ID16G(W)r	Für mein Selbstbild ist es wesentlich ein <b>Mann/ eine Frau</b> zu sein.
PO01		Wenn Sie an Ihre eigenen politischen Ansichten denken. würden Sie sich selbst als eher links. eher
PO02L		Wenn Links: 1. Etwas links/ 2. Ziemlich links /3. Sehr links.
PO02M		Wenn Mitte: 1. Links der Mitte/ 2. Genau in der Mitte 3. Rechts der Mitte
PO02R		Wenn Rechts: 1. Etwas rechts/ 2. Ziemlich rechts/ 3. Sehr rechts
GEBJ		Sagen Sie mir bitte, in welchem Jahr Sie geboren sind?
	alter_r	Alter in 5 Gruppen: 14-21/ 22-34/ 35-49/ 50-64/ ab 65 ( <b>es gibt einige unter 18 Jährige, die sind erstmal mit in der ersten Gruppe</b> )
FAMS		Welchen Familienstand haben Sie?
PART		Haben Sie eine feste Lebenspartnerschaft?
EI09		Wie viele Personen leben in Ihrem Haushalt, Kinder und Sie selbst mit eingeschlossen?
SXA21	SXA21g	Haben oder hatten Sie Kinder?
ZAHL_KI	ZAHL_Kig	Wie viele Kinder haben Sie?
ALTER_K1	ALTER_K1g	Wie alt ist ihr Kind/sind ihre Kinder? Erstes Kind:
ALTER_K2	ALTER_K2g	Wie alt ist ihr Kind/sind ihre Kinder? Zweites Kind:
ALTER_K3	ALTER_K3g	Wie alt ist ihr Kind/sind ihre Kinder? Drittes Kind:
ALTER_K4	ALTER_K4g	Wie alt ist ihr Kind/sind ihre Kinder? Viertes Kind:
ALTER_K5	ALTER_K5g	Wie alt ist ihr Kind/sind ihre Kinder? Fünftes Kind:
ALTER_K6	ALTER_K6g	Wie alt ist ihr Kind/sind ihre Kinder? Sechstes Kind:
ALTER_K7	ALTER_K7g	Wie alt ist ihr Kind/sind ihre Kinder? Siebtes Kind:
SXA22(w)	SXA22(w)g	<b>Bei Männern:</b> Haben Sie mindestens bis zum 3. Lebensjahr Ihres Kindes/Ihrer Kinder mit der Mutter gemeinsam in einem Haushalt gelebt? <b>Bei</b>

			<b>Frauen:</b> Haben Sie mindestens bis zum 3. Lebensjahr Ihres Kindes/Ihrer Kinder mit dem Vater gemeinsam in einem Haushalt gelebt?
SXA22A(w)	SXA22A(w/m)g		<b>Bei Männern:</b> Leben Sie aktuell mit der Mutter ihres Kindes/ihrer Kinder gemeinsam in einem Haushalt? <b>Bei Frauen:</b> Leben Sie aktuell mit dem Vater ihres Kindes/ihrer Kinder gemeinsam in einem Haushalt?
SXA23(w/m)	SXA23(w/m)g		<b>Bei Männern:</b> War Ihre (Ehe-)Partnerin bevor Sie Kinder hatten <b>Bei Frauen:</b> Waren Sie bevor Sie Kinder hatten
SXA24(w/m)	SXA24(w/m)g		<b>Bei Männern:</b> War Ihre (Ehe-)Partnerin als mindestens ein Kind noch nicht in den Kindergarten ging?/ <b>Bei Frauen:</b> Waren Sie als mindestens ein Kind noch nicht in den Kindergarten ging
SXA25(m)	SXA25(m)g		<b>Bei Männern:</b> War Ihre (Ehe-)Partnerin nachdem auch das jüngste Kind in den Kindergarten ging?/ <b>Bei Frauen:</b> Waren Sie nachdem auch das jüngste Kind in den Kindergarten ging?
SXA20	SXA20g		Und wie war es bei Ihnen selbst? War Ihre Mutter in der Zeit zwischen Ihrer Geburt und Ihrem 3. Lebensjahr mindestens ein Jahr lang erwerbstätig? à INTERVIEWER HINWEIS: Erwerbstätigkeit meint sowohl Teilzeit- als auch Vollzeit-Erwerbstätigkeit!
ZU01	zu01r	(zu01: 1=5; 2,3,4=4; 5,6=3; 7=2; 8=1; 9=9) = zu01r	Welchen höchsten schulischen oder universitären Abschluss haben Sie bisher erreicht? "niedriger Schulabschluss gruppiert": 1=Studium; 2=Abitur; 3=mittl. Reife; 4=Hauptschul.; 5=kein Abschluss; 9=anderer Abschluss
	zu01k	(zu01: 1-4=1; 5,6=2; 7,8=3; 9=9) = zu01k	"Schulabschluss, gruppiert in 3 Kategorien" 1=niedrig; 2=mittel; 3=hoch
ZU02			Haben Sie eine abgeschlossene Berufsausbildung?
ZU00			Sind sie zurzeit erwerbstätig?
ZU03E			Gibt es in Ihrem Haushalt eine Person, die ein höheres Einkommen hat als Sie?
ZU00N			Waren Sie in der Vergangenheit erwerbstätig?
ZU05N			Was machen Sie zurzeit? Sind Sie....
ZU03N			Gibt es in Ihrem Haushalt eine Person, die ein höheres Einkommen hat als Sie?
ZU00H			Ist diese Person erwerbstätig?
ZU05I			Was macht diese Person zurzeit? Ist sie

<b>ZU00P (w)</b>	Handelt es sich bei dieser Person um ihre/n Ehe- oder Lebenspartner/in?	
<b>EI08</b>	Wenn Sie mal alles zusammenrechnen: Wie hoch ist das monatliche Netto-Einkommen, das Sie alle zusammen in Ihrem Haushalt haben nach Abzug der Steuern und Sozialversicherung? Ich lese Ihnen jetzt einige Einkommensgruppen vor und Sie sagen mir bitte welche davon auf ihren Haushalt zutrifft.	
<b>WDH</b>	Vielen Dank für das Interview. Dürfen wir Sie in nächster Zeit noch einmal persönlich kontaktieren? Zu diesem Zweck müssen wir mit Ihrem Einverständnis Ihre Telefonnummer und Ihren	
<b>WDH2</b>	Für den Fall, dass wir Sie bei der nächsten Befragung nicht mehr unter Ihrer aktuellen Telefonnummer erreichen können, möchten wir Sie um folgende Informationen bitten: Besitzen Sie eine E-Mailadresse, die Sie regelmäßig oder zumindest gelegentlich nutzen?	
<b>EMAIL01</b>	Wie lautet Ihre E-Mailadresse?	
<b>HANDY01</b>	Haben Sie neben ihrem Festnetzanschluss ein Handy?	
<b>HANDY02</b>	Welche Rufnummer hat Ihr Handy?	
<b>FEST01</b>	Haben Sie neben ihrem Handy einen Festnetzanschluss?	
<b>FEST02</b>	Welche Rufnummer hat Ihr Festnetzanschluss?	
<b>BULA</b>	Bundesland	
<b>westost</b>	Bula 1-11= West; 12-16= OST	Befragte in West- und Ostdeutschland

*Anmerkung.* Endung „r“: Die Variable wurde recodiert, sodass hohe Werte auch eine hohe Ausprägung wiedergeben. Endung „g“: Es wurde nicht recodiert.

## Anhang 8: Faktorladungen aller Sexismus-Items in Stichprobe 6

## Anhang 8

## Tabelle A3

## Faktorladungen aller Sexismus-Items in Stichprobe 6

Items	Faktor											
	FS + GRB	AMMSA	HS	MSS	IHS + PP	WS	GRB (Rolle Männer)	KGD	PP + GRB	GRB	MSS	
Frauen denken meiner Meinung nach nur an sich selbst, wenn sie ein Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um berufliche Anerkennung zu bekommen.	.87											
Frauen sollten ein schlechtes Gewissen haben, wenn sie aus Karrieregründen ihr Kind mit einem halben Jahr tagsüber zur Betreuung zu einer Tagesmutter geben.	.87											
Frauen, die aus Karrieregründen ihr Kind mit einem halben Jahr in eine Kindertagesstätte geben, sind für mich Rabenmütter.	.87											
Ich finde Frauen egoistisch, die bereits ein halbes Jahr nach der Geburt des Kindes wieder Vollzeit in ihren Beruf zurückkehren, um sich selbst zu verwirklichen.	.79											
Ein Kind, das noch nicht zur Schule geht, wird wahrscheinlich darunter leiden, wenn seine Mutter berufstätig ist. (GRB)	.79											-.27
Ich denke, dass Karrierefrauen ihren Kindern häufig zu wenig Aufmerksamkeit schenken.	.73											
Alles in allem: Das Familienleben leidet darunter, wenn die Frau voll berufstätig ist. (GRB)	.70											
Viele Karrierefrauen denken meiner Meinung nach nur an ihr berufliches Vorankommen und vernachlässigen die Bedürfnisse ihrer Kinder.	.62											.21
Eine berufstätige Mutter kann ein genauso herzliches und vertrauensvolles Verhältnis zu ihren Kindern finden wie eine Mutter, die nicht berufstätig ist. (GRB)	.51						.22					-.22



Items	FS + GRB	AMMSA	HS	MSS	IHS + PP	WS	GRB (Rolle Männer)	KGD	PP + GRB	GRB	MSS
Die meisten Frauen sehen gar nicht, was Männer alles für sie tun.			<b>.64</b>								
Frauen übertreiben Probleme, die sie am Arbeitsplatz haben.	.20		<b>.61</b>								
Hat eine Frau erst einmal einen Mann "rumgekriegt", dann versucht sie, ihn an die kurze Leine zu legen.			<b>.60</b>								
Feministinnen stellen an Männer vollkommen berechnete Forderungen.			<b>.50</b>	.35							
Viele Frauen haben Spaß daran, mit Männern zu „spielen“, indem sie sich zuerst verführerisch geben, dann aber die Annäherungsversuche der Männer zurückweisen.			<b>.45</b>								
Diskriminierung von Frauen ist in Deutschland immer noch ein Problem.				<b>.74</b>							
Frauen und Männer haben in der heutigen Gesellschaft die gleichen Chancen, etwas zu erreichen.				<b>.72</b>							.24
Heutzutage werden Frauen im Berufsleben fair behandelt.				<b>.72</b>							
Im Fernsehen gibt es häufig frauenfeindliche Darstellungen.				<b>.70</b>							
Im Allgemeinen werden in unserer Gesellschaft Ehepartner gleich behandelt.				<b>.69</b>					.30		
In der Schule werden Mädchen immer noch benachteiligt.				<b>.62</b>							
Frauen finden häufig keine gut bezahlte Arbeit, weil sie diskriminiert werden.				<b>.57</b>							
Die Forderungen von Frauen nach Gleichberechtigung sind leicht nachzuvollziehen.				<b>.49</b>							
In den westlichen Ländern ist Gleichberechtigung von Frauen schon lange verwirklicht.				<b>.45</b>							
Man kann im Leben erst richtig glücklich sein, wenn man einen Partner hat, den man liebt.					<b>.86</b>						.22
Jeder Mann sollte eine Frau haben, die er wirklich liebt.					<b>.81</b>						.20
Egal wie erfolgreich ein Mann auch sein mag, ohne eine Frau, die ihn liebt, fehlt ihm etwas ganz Wichtiges.					<b>.78</b>						



Items	FS + GRB	AMMSA	HS	MSS	IHS + PP	WS	GRB (Rolle Männer)	KGD	PP + GRB	GRB	MSS
Männer sind ohne Frauen unvollkommen.					<b>.73</b>				.22		
Frauen sollten von Männern umsorgt und beschützt werden. (PP)		-.20			<b>.43</b>				.23	.21	
Eine Frau sollte von ihrem Mann auf Händen getragen werden. (PP)	-.21				<b>.39</b>	.24			.26		-.21
Eine Mutter spürt besser als ein Vater, was die eigenen Kinder brauchen.					<b>.91</b>						
Mütter besitzen eine größere Fürsorglichkeit als Väter.					<b>.87</b>					-.24	
Frauen besitzen eine größere natürliche Begabung als Männer, die Hausarbeit und die Betreuung der Kinder unter einen Hut zu bringen.					<b>.73</b>						
Eine Mutter ist vom Wesen her sanfter und warmerziger als ein Vater und daher auch besser für die Betreuung eines Kindes geeignet.					<b>.66</b>						
Ich denke, dass Mütter auf die Bedürfnisse von Kindern unter zwei Jahren besser eingehen können als Väter.					<b>.58</b>						
Frauen haben von Natur aus einen besseren Blick als Männer dafür, welche Hausarbeiten gerade getan werden müssen.					<b>.52</b>						
Männer sollten einen größeren Anteil an der Kinderbetreuung übernehmen, als sie es jetzt tun.							<b>.85</b>				
Männer sollten einen größeren Anteil an Hausarbeiten übernehmen, als sie es jetzt tun.							<b>.83</b>				
Viele Frauen haben eine Art von Ehrlichkeit, die nur wenige Männer besitzen.								<b>.68</b>			
Verglichen mit Männern haben Frauen ein besseres moralisches Empfinden.								<b>.63</b>			
Verglichen mit Männern haben Frauen einen feineren Sinn für Kultur und einen besseren Geschmack.							.29	<b>.47</b>			
Ein Mann sollte bereit sein, sein eigenes Wohl zu opfern, um für seine Frau sorgen zu können.				.33					<b>.65</b>		
Bei einer Katastrophe sollten Frauen vor Männern gerettet werden.								.21	<b>.57</b>		
Hausfrau zu sein ist genauso erfüllend wie gegen Bezahlung zu arbeiten. (GRB)		-.20					.27		<b>.47</b>		

Items	FS + GRB	AMMSA	HS	MSS	IHS + PP	WS	GRB (Rolle Männer)	KGD	PP + GRB	GRB	MSS
Einen Beruf zu haben ist ja ganz schön, aber das, was die meisten Frauen wirklich wollen, sind ein Heim und Kinder. (GRB)	.24								<b>.29</b>		
Einen Beruf zu haben ist das beste Mittel für eine Frau, um unabhängig zu sein.										<b>.70</b>	
Der Mann und die Frau sollten beide zum Haushaltseinkommen beitragen.											
Die Aufgabe des Mannes ist es, Geld zu verdienen, die der Frau, sich um Haushalt und Familie zu kümmern.											
Wenn Frauen tatsächlich einmal schlechter bezahlt werden als Männer, dann nur deshalb, weil sie einfachere Arbeit zu leisten haben.											
Eigenwert	20.72	4.05	3.76	2.99	2.52	1.83	1.57	1.46	1.35	1.25	1.13
Erklärte Varianz	32.89	6.43	5.97	4.74	4.00	2.91	2.50	2.32	2.15	1.98	1.79

*Anmerkungen.* FS = Feindseliger Sexismus. GRB = Geschlechterrollen. AMMSA = Akzeptanz moderner Mythen über sexuelle Aggression. HS = Hostiler Sexismus. MSS = Moderner Sexismus. IHS = Intime Heterosexualität. PP = Protektiver Paternalismus. WS = Wohlwollender Sexismus. KGD = Komplementäre Geschlechterdifferenzierung. Fett markierte Werte geben die Hauptladungen der einzelnen Items an. Faktorladungen < .10 sind in der Tabelle nicht angeführt, um sie übersichtlicher zu gestalten.

## Erklärung

„Ich erkläre: Ich habe die vorgelegte Dissertation selbständig, ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der ‚Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis‘ niedergelegt sind, eingehalten.“

Fernwald, Januar 2013

Astrid Schüßler