

Kommunikationswissenschaft

ZWISCHEN LIKE UND EMPÖRUNG

THEORETISCHE MODELLIERUNG UND EMPIRISCHE UNTERSUCHUNG  
VON REZEPTIONS- UND PERSUASIONSWIRKUNGEN VON SEXISTISCHEM  
HATE SPEECH

Inaugural-Dissertation

zur Erlangung des Doktorgrades Dr. phil.

im Fachbereich Erziehungswissenschaft und Sozialwissenschaften

der Westfälischen Wilhelms-Universität in Münster

vorgelegt von

Carla Schieb

2021

Dekanin	Prof. Dr. Ulrike Weyland
Vorsitzender des Promotionsausschusses	Prof. Dr. Johannes Bellmann
Erster Gutachter	Prof. Dr. Thorsten Quandt
Zweiter Gutachter	Prof. Dr. Volker Gehrau
Tag der mündlichen Prüfung (Disputation)	<u>22.04.2022</u>
Tag der Promotion	<u>22.04.2022</u>

# Inhalt

<b>ABBILDUNGSVERZEICHNIS .....</b>	<b>V</b>
<b>TABELLENVERZEICHNIS .....</b>	<b>VI</b>
<b>ABKÜRZUNGSVERZEICHNIS .....</b>	<b>VIII</b>
<b>1 EINLEITUNG .....</b>	<b>1</b>
1.1    RELEVANZBEGRÜNDUNG UND FRAGESTELLUNG .....	1
1.2    AUFBAU DER ARBEIT .....	9
<b>2 HATE SPEECH – THEORETISCHE FUNDIERUNG.....</b>	<b>11</b>
2.1    BEGRIFFLICHE UND DEFINITORISCHE EINORDNUNG .....	11
2.1.1 <i>Begriffsinflation</i> .....	11
2.1.2 <i>Internet als exklusiver Kommunikationsraum für Hate Speech?</i> .....	13
2.1.3 <i>Juristische Prägung früher Forschung</i> .....	15
2.1.4 <i>Kernmerkmale als definitorische Dimensionen</i> .....	16
2.1.4.1    Adressatengruppen .....	17
2.1.4.2    Kommunikative Inhalte .....	20
2.1.4.3    Kommunikative Wirkung .....	22
2.2    ZWISCHENFAZIT: HATE SPEECH ALS GRUPPENBEZOGENER ABWERTUNGSMECHANISMUS .....	30
2.3    BEGRIFFLICHE ABGRENZUNG .....	34
2.3.1 <i>Flaming</i> .....	34
2.3.2 <i>Trolling</i> .....	37
2.3.3 <i>Cybermobbing</i> .....	41
2.3.4 <i>Resümee</i> .....	44
2.4    EXKURS: BEFUNDE ZUR PSYCHOSOZIALEN PERSÖNLICHKEIT VON HATE SPEECH-PRODUZENTEN .....	47
<b>3 SEXISMUS ALS EINSTELLUNGSKONSTRUKT .....</b>	<b>57</b>
3.1    BEGRIFFSBESTIMMUNG UND ABGRENZUNG VON VERWANDTEN PHÄNOMENEN.....	57
3.2    FORSCHUNGSSTAND ZU SEXISMUS ALS EINSTELLUNGSKONSTRUKT .....	62
3.2.1 <i>Multikomponentenmodell</i> .....	63
3.2.1.1    Stereotype als kognitive Komponente .....	63
3.2.1.2    Emotionen als affektive Komponente.....	67
3.2.1.3    Diskriminierung als Verhaltenskomponente.....	71
3.2.2 <i>Eine integrative Einstellungsstruktur – Ambivalenter Sexismus als paradoxes Vorurteil</i> .....	74
3.3    SEXISTISCHE EINSTELLUNGEN ZU FRAUEN IN DER IT-BRANCHE .....	80
3.4    ZUSAMMENFASSUNG.....	84
<b>4 DAS ELABORATION LIKELIHOOD MODEL ALS ANSATZ ZUR ERKLÄRUNG VON EINSTELLUNGSÄNDERUNGEN .....</b>	<b>86</b>
4.1    EINFÜHRUNG IN DEN PERSUASIONSBEGRIFF .....	87
4.2    KLASSISCHE MODELLANNAHMEN .....	89
4.3    INFORMATIONSVERRARBEITUNG INNERHALB DES ELABORATIONSKONTINUUMS.....	91
4.3.1 <i>Einflussgrößen auf das Elaborationskontinuum</i> .....	91
4.3.1.1    Fähigkeit zur Verarbeitung.....	91
4.3.1.2    Motivation zur Verarbeitung .....	92
4.3.2 <i>Zwei Verarbeitungspfade</i> .....	94
4.3.2.1    Zentrale Route .....	94
4.3.2.2    Periphere Route.....	98
4.3.3 <i>Multiple Rollen von Variablen</i> .....	100
4.4    PERSUASION VON HATE SPEECH IN SOZIALEN NETZWERKSEITEN .....	102

4.5	ZWISCHENFAZIT: AUF DEM WEG ZU EINER MODELLIERUNG DER PERSUASIONSWIRKUNG VON SEXISTISCHEM HATE SPEECH ANHAND DES ELABORATION LIKELIHOOD MODELS.....	107
<b>5</b>	<b>ABLEITUNG DER HYPOTHESEN.....</b>	<b>113</b>
5.1	ELABORATION VON HATE SPEECH .....	113
5.2	ELABORATION UND WAHRGENOMMENE RELEVANZ .....	117
5.3	EINSTELLUNGEN ZUR BOTSCHAFT UND EINSTELLUNGEN ZU FRAUENKARRIEREN .....	118
5.4	VERHALTENSINTENTIONEN.....	120
<b>6</b>	<b>EMPIRISCHE UMSETZUNG.....</b>	<b>122</b>
6.1	STUDIE I – ELABORATION UND RELEVANZWahrnehmung VON HATE SPEECH .....	122
6.1.1	<i>Forschungsdesign und Methodik</i> .....	122
6.1.1.1	Studiendesign und Umsetzung .....	122
6.1.1.2	Aufbau des Fragebogens.....	125
6.1.1.3	Stimulusmaterial.....	126
6.1.1.4	Stichprobenbeschreibung .....	130
6.1.1.5	Operationalisierung der zentralen Konstrukte .....	133
6.1.2	<i>Ergebnisse</i> .....	143
6.1.2.1	Prüfung der Zufallsaufteilung .....	143
6.1.2.2	Auswirkung von Hate Speech auf die Elaboration .....	144
6.1.2.3	Auswirkung von Hate Speech auf die wahrgenommene persönliche Relevanz .....	163
6.1.2.4	Einflüsse auf die Verhaltensintention .....	170
6.1.3	<i>Einordnende Zusammenfassung der Befunde</i> .....	183
6.2	STUDIE II – ELABORATION, RELEVANZWahrnehmungen UND EINFLÜSSE AUF EINSTELLUNGEN .....	192
6.2.1	<i>Forschungsdesign und Methodik</i> .....	192
6.2.1.1	Anpassung des Studiendesigns und Stimulusmaterial .....	192
6.2.1.2	Aufbau des Fragebogens.....	193
6.2.1.3	Stichprobenbeschreibung .....	194
6.2.1.4	Operationalisierung zentraler Konstrukte .....	196
6.2.2	<i>Ergebnisse</i> .....	204
6.2.2.1	Prüfung der Zufallsaufteilung .....	204
6.2.2.2	Auswirkungen von Hate Speech auf Elaboration .....	205
6.2.2.3	Auswirkungen von Hate Speech auf die wahrgenommene Relevanz .....	214
6.2.2.4	Auswirkungen von Hate Speech auf den wahrgenommenen Meinungsfluss .....	218
6.2.2.5	Zusammenhang zwischen wahrgenommener Relevanz und Bewertung .....	219
6.2.2.6	Zusammenhang zwischen dem antizipierten Einfluss auf die Meinung und Bewertung... ..	227
6.2.2.7	Einflüsse auf Einstellungen zu Frauenkarrieren .....	229
6.2.2.8	Einflüsse auf Verhaltensintentionen.....	231
6.2.3	<i>Einordnende Zusammenfassung der Befunde</i> .....	235
<b>7</b>	<b>FAZIT.....</b>	<b>240</b>
7.1	ZUSAMMENFASSUNG DER ZENTRALEN ERGEBNISSE.....	240
7.2	METHODISCHE UND THEORETISCHE LIMITATIONEN .....	245
7.3	EINE FORSCHUNGSETHISCHE SCHLUSSBETRACHTUNG VON HATE SPEECH.....	249
	<b>LITERATURVERZEICHNIS.....</b>	<b>254</b>
	<b>ANHANG .....</b>	<b>299</b>

## Abbildungsverzeichnis

Abbildung 1: <i>Wissenschaftliche Publikationen zum Thema Hate Speech</i> .....	12
Abbildung 2: <i>Adressatengruppen</i> .....	18
Abbildung 3: <i>Kommunikativer Inhalt</i> .....	21
Abbildung 4: <i>Übersicht über den Zusammenhang zwischen den drei Kernmerkmalen von Hate Speech</i> .....	33
Abbildung 5: <i>Exemplarische weibliche Substereotype im Stereotypinhaltsmodell</i> .....	66
Abbildung 6: <i>Emotionen gegenüber exemplarischen weiblichen Substereotypen im SCM</i> ...	69
Abbildung 7: <i>Konstrukte der Theorie der individuellen Unterschiede in Geschlecht und IT</i> ..	82
Abbildung 8: <i>Hate Speech Kommentar (Hate Intensität) von Frauen (Geschlecht der Kommentierenden) mit vielen Likes (Reaktionsvolumen)</i> .....	128
Abbildung 9: <i>Persönlich wahrgenommene Relevanz in Abhängigkeit von der Treatment-Wahrnehmung</i> .....	134
Abbildung 10 <i>Zweidimensionale Anordnung der Verhaltensintention im Wahrnehmungsraum (Studie I)</i> .....	139
Abbildung 11: <i>Moderationseffekt von Hate Intensität in Abhängigkeit des Kognitionsbedürfnisses und Vorwissens</i> .....	150
Abbildung 12: <i>Konditionaler Effekt des Reaktionsvolumens auf den negativen Gedankenanteil in Abhängigkeit vom Kognitionsbedürfnis und Vorwissen</i> .....	153
Abbildung 13: <i>Einfluss der Hate Intensität, Geschlecht der Kommentierenden Reaktionsvolumen auf die wahrgenommene persönliche Relevanz</i> .....	164
Abbildung 14: <i>Einfluss von Hate Intensität auf die wahrgenommene persönliche Relevanz unter der Bedingung ambivalent sexistischer Einstellungen</i> .....	167
Abbildung 15: <i>Zweidimensionale Anordnung der Verhaltensintention im Wahrnehmungsraum (Studie II)</i> .....	201
Abbildung 16: <i>Zusammenhang zwischen Experimentalbedingungen und wahrgenommener persönlicher Relevanz</i> .....	217
Abbildung 17: <i>Zusammenhang zwischen den Experimentalbedingungen und dem wahrgenommenen Einfluss auf Meinungen</i> .....	219
Abbildung 18: <i>Kognitive Bewertung von Hate Speech</i> .....	220
Abbildung 19: <i>Einfaches Mediationsmodell zur Erklärung des Zusammenhangs von Geschlecht über die wahrgenommene Relevanz auf die kognitive Bewertung</i> .....	221
Abbildung 20: <i>Erweitertes Mediationsmodell zur Erklärung des Zusammenhangs von Geschlecht über die wahrgenommene Relevanz auf die kognitive Bewertung</i> .....	225

## Tabellenverzeichnis

Tabelle 1: Übersicht der kommunikativen Wirkung.....	25
Tabelle 2: Gegenüberstellung von Hate Speech und verwandten Online-Phänomenen auf Basis der drei Kernmerkmale. Der graue Druck weist auf vermutete, jedoch nicht ausreichend belegte Annahmen .....	46
Tabelle 3: Zusammenfassende Übersicht über psychosoziale Eigenschaften von Hate Speech-Produzierenden .....	56
Tabelle 4: Postulate des ELM .....	90
Tabelle 5: Schematische Notation der Versuchsanordnung .....	124
Tabelle 6: Stichprobengröße vor und nach Datenbereinigung .....	131
Tabelle 7: Deskriptive Kennwerte für die Items für wahrgenommene persönliche Relevanz .....	133
Tabelle 8: Deskriptive Kennwerte zur Messung der Verhaltensintentionen .....	138
Tabelle 9 Korrelationen der Items Verhaltensintention (VI) .....	138
Tabelle 10: Prüfung auf Randomisierung metrischer Variablen .....	143
Tabelle 11: Hierarchische Regressionsanalyse zur Erklärung des negativen Gedankenanteils .....	146
Tabelle 12: Konditionaler Effekt der Hate Intensität auf den Anteil an negativen Gedanken in Abhängigkeit des Kognitionsbedürfnisses und Vorwissens in einem additiven multiplen Moderationsmodell.....	149
Tabelle 13: Poissonregression mit mittelwertzentrierten Variablen zur Erklärung der Anzahl der Rezeptionsgedanken.....	155
Tabelle 14: Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung der kognitiven Auseinandersetzung mit dem Stimulusmaterial als Ausgangsbasis für Rezeptionsgedanken.....	157
Tabelle 15: Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung affektiver Reaktionen auf das Stimulusmaterial .....	159
Tabelle 16: Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung von Vorwissen und Erfahrungen in den Rezeptionsgedanken.....	160
Tabelle 17: Hurdle-Modell mit Poissonverteilung zur Erklärung sonstiger Gedanken.....	162
Tabelle 18: Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Zustimmung für Hate Intensität mittels Like-Button auszudrücken .....	172
Tabelle 19: Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Ablehnung für Hate Intensität mittels Reaktionsbutton auszudrücken .....	173
Tabelle 20: Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Zustimmung für Hate Intensität mittels sachlichen Kommentars auszudrücken.....	175
Tabelle 21: Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Widerspruch für Hate Intensität mittels sachlichen Kommentars auszudrücken.....	176

Tabelle 22: <i>Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention des Meldens in der Bedingung Hate Intensität</i> .....	180
Tabelle 23: <i>Deskriptive Kennwerte für die Relevanz-Items ("Solche Kommentare...")</i> .....	197
Tabelle 24: <i>Deskriptive Kennwerte für die kognitiven Bewertungs-Items ("Die Kommentare waren...")</i> .....	198
Tabelle 25: <i>Deskriptive Kennwerte für Einstellungen zu modernen Frauenkarrieren</i> .....	200
Tabelle 26: <i>Deskriptive Kennwerte für die Items Verhaltensintention (VI)</i> .....	201
Tabelle 27: <i>Prüfung auf Randomisierung zentraler metrischer Variablen</i> .....	204
Tabelle 28: <i>Einfluss auf den Anteil an negativen Gedanken mittels hierarchischer Regressionsanalyse</i> .....	206
Tabelle 29: <i>Poissonregression mit mittelwertzentrierten Kovariaten zur Erklärung der Anzahl von Rezeptionsgedanken (Studie II)</i> .....	208
Tabelle 30: <i>Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Kognitive Auseinandersetzung mit dem Stimulusmaterial)</i> .....	210
Tabelle 31: <i>Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Affektive Auseinandersetzung mit dem Stimulusmaterial)</i> .....	211
Tabelle 32: <i>Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Vorwissen)</i> .....	213
Tabelle 33: <i>Vereinfachtes Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Sonstige Gedanken)</i> .....	214
Tabelle 34: <i>Mediationsanalysen zur Erklärung des Einflusses von Relevanz auf die Bewertung von sexistischem Hate Speech</i> .....	226
Tabelle 35: <i>Mediationsanalysen zur Erklärung des wahrgenommenen Einflusses von Meinungen auf die Bewertung von sexistischem Hate Speech</i> .....	228

## Abkürzungsverzeichnis

ASI	Ambivalenter-Sexismus-Index
AV	Abhängige Variable
BEW	Kognitive Bewertung des Stimulusmaterials
CI	Konfidenzintervall
cvK	computervermittelte Kommunikation
DS	Dangerous Speech
EF	Einstellungen zu modernen Frauenkarrieren
ELM	Elaboration Likelihood Model
HC3	Heteroskedastizitätskonsistenter Schätzer der Standardfehler
HS	Hate Speech
HN	Habitualisierte Nutzung
IRR	Incidence Rate Ratio
IT	Informationstechnik
LL	Lower Level
KG	Kontrollgruppe
M	Mediator
MDS	Multidimensionale Skalierung
MINT	Mathematik, Informatik, Naturwissenschaft und Technik
NFC	Need for Cognition
OS	Offensive Speech
PI	Persönliches Involvement
RZ	Wahrgenommene persönliche Relevanz
RZG	Wahrgenommene gesellschaftliche Relevanz
RZN	Wahrgenommene Relevanz für Dritte
SCM	Stereotype Content Model
SE	Soziale Erwünschtheit

SK	Sachliche Kommentare
SNS	Soziale Netzwerkseite
ST	Stimulusmaterial
UL	Upper Level
UV	Unabhängige Variable
VI	Verhaltensintention
VIF	Variance Inflation Factor
VW	Vorwissen

# 1 Einleitung

## 1.1 Relevanzbegründung und Fragestellung

Im Bundestagswahlkampf 2021 wirkt die Kanzlerkandidatin von Bündnis 90/Die Grünen Annalena Baerbock nach den Fernsehdebatten zwar sympathischer und tatkräftiger, aber nicht zwangsläufig kompetenter als ihre beiden männlichen Mitbewerber (Infratest dimap, 2021). In den Vormonaten der letzten Wahlkampfphase vermischen sich in der Berichterstattung berechnete Kritikpunkte zu ihrem Fehlverhalten mit Verweisen auf ihre weibliche Geschlechterrolle, etwa durch Anspielungen auf die Vereinbarkeit von Familie und politischen Pflichten (Barthels, 2021). In sozialen Netzwerken ist die Kandidatin offener und massiver Feindseligkeit ausgesetzt: Es kursieren Falschnachrichten und Bildermontagen, die teilweise auch mit misogynen Hasskommentaren versehen sind (Görz y Moratalla et al., 2021; Klimpel, 2021). Baerbock ist kein Einzelphänomen, vielmehr zeigt eine Untersuchung von 2018 zu sexistischem *Hate Speech*, dass Politikerinnen in europäischen Parlamenten mit sexistischen und misogynen Anfeindungen sowohl offline, aber vor allem im Netz konfrontiert werden. Sie erhalten dabei auch Androhungen von sexueller Gewalt bis hin zu Mordankündigungen (Inter-Parliamentary Union, 2018).

Auch andere Frauen, die in der Öffentlichkeit stehen, wie z. B. Journalistinnen, Wissenschaftlerinnen oder Künstlerinnen sind von Hate Speech betroffen. Dabei wird ihnen entweder Kompetenz abgesprochen oder sie werden durch Falschbehauptungen diskreditiert, wobei der weibliche Körper gleichermaßen zum Gegenstand und Ziel der Attacken erklärt wird (Adams, 2018; Bovermann, 2021; Döring & Mohseni, 2020; OSZE, 2018; Springer & Troger, 2021). Die Betroffenen berichten von psychischen Beeinträchtigungen, die sich auf ihre Leistungsfähigkeit auswirken, fühlen sich z. T. durch aggressive Silencing-Strategien eingeschüchtert und ziehen sich manchmal auch gänzlich aus der Öffentlichkeit zurück (Geschke et al., 2019). Hate Speech wird damit zum dysfunktionalen Multiplikator für Diskriminierung, Polarisierung und Desensibilisierung, denn damit sind längst nicht nur individuelle, sondern auch gesellschaftliche Auswirkungen verbunden. Fragen der freien Meinungsäußerung und des Sagbaren bestimmen derzeit den öffentlichen Diskurs und schaukeln sich in der „digitalen Empörungsdemokratie“ (Piorkowski, 2018) immer höher und lassen kaum Raum für verständigungsorientierte Debatten. Herausfordernd gestaltet sich das Ausbalancieren von Ansprüchen auf Meinungsfreiheit auf der einen Seite und Schutz für Angefeindete auf der anderen, was in einer Demokratie Reibungsflächen erzeugt, weil die Erfüllung derartig komplexer und heterogener Ansprüche die ohnehin fragile soziale Ordnung stört.

Der Hate Speech-Begriff ist in der ursprünglichen Verwendung ein juristischer Terminus, der in der US-amerikanischen Rechtswissenschaft geprägt wurde (George, 2015) und in der frühen wissenschaftlichen Diskussion eng mit der Thematisierung von rassistischen Strukturen in der US-Gesellschaft verknüpft war (z. B. Sponholz, 2018). In den aktuellen (sozialwissenschaftlichen) Arbeiten gelten als primäre Adressaten von Hate Speech primär marginalisierte soziale Gruppen und ihre (wahrgenommenen) Mitglieder, wozu z. B. Ethnien, religiöse Zugehörigkeiten, aber auch Geschlechtsidentitäten gehören (Meibauer, 2013; Parekh, 2006). Im deutschsprachigen Raum lösten während der Flüchtlingskrise 2015 unzählige Hasskommentare in sozialen Netzwerken erstmals eine gesteigerte öffentliche Wahrnehmung von Hate Speech aus (Schieb & Preuss, 2016; Sponholz, 2020). Das Ausmaß an Hass, Feindseligkeit und menschenverachtenden Aussagen, das Nutzerinnen und Nutzer unter ihrem Klarnamen verbreiten, sind – zumindest die Rezipierenden – weder aus Leserbriefen noch aus interaktiven Call-in-Formaten im Rundfunk gewohnt. Das öffentliche Interesse an Hate Speech im Internet ist groß, was sich nicht nur in der Medienberichterstattung wegen dessen immanenten Skandalisierungspotenzials, sondern auch in einer verstärkten Forschungsaktivität zum Thema niederschlägt. Die Landesanstalt für Medien in Nordrhein-Westfalen veröffentlicht seit 2016 einen jährlichen Bericht zum Thema Hate Speech. Besonders davon betroffen sind die unter 25-Jährigen, denn 98 Prozent von ihnen gaben bei dieser aktuellen Befragung an, bereits Hate Speech im Internet begegnet zu sein, altersübergreifend sind es durchschnittlich immerhin noch 76 Prozent (Landesanstalt für Medien NRW, 2021). 2016 waren dies noch zwei Drittel der Befragten, die Hate Speech bereits im Netz wahrgenommen hatten (ebd.). Die Befunde lassen die Frage offen, ob Hate Speech insgesamt zugenommen hat oder lediglich die Sensibilisierung für diese Thematik gestiegen ist und in der Folge das Antwortverhalten beeinflusst.

Die Unklarheit bezüglich der Einordnung der Befunde sind auf die konzeptionelle Beliebigkeit des Begriffs im Alltagsverständnis und – schwerwiegender noch – auch in der wissenschaftlichen Forschung zurückzuführen (Schieb & Preuss, 2016; Sponholz, 2020). Aspekte von Hate Speech werden in Arbeiten behandelt, die sich etwa mit *Inzivilität* (Coe et al., 2014; B. T. Gervais, 2015; Rossini, 2020; Su et al., 2018) befassen oder synonyme Begriffe verwenden, wie z. B. *Extreme Speech* (Gagliardone, 2019; Pohjonen, 2019), *Fear Speech* (Buyse, 2014; Müller et al., 2014), *Offensive Speech* (Gagliardone et al., 2016) oder *Dangerous Speech* (Benesch, 2013; Howard, 2019). Die Umfrageforschung verweist zwar auf die Notwendigkeit übereinstimmender Verständnisse des Frageninhalts, um die Validität der Antworten bemessen zu können (Diekmann, 2007), aber auch in der LFM-Studie (Landesanstalt für Medien NRW, 2021) wird den Studienteilnehmenden folgende Frage gestellt: „Wie häufig haben Sie

persönlich schon Hate Speech bzw. Hasskommentare im Internet gesehen – z. B. auf Webseiten, in Blogs, in sozialen Netzwerken oder in Internetforen?“. Bei dieser Frageformulierung erschließt sich den Forschenden allerdings nicht, welches mentale Konzept von Hate Speech die Befragten haben, welches Vorwissen vorhanden ist und wie und ob sie die Rezeption aggressiver Inhalte im Netz als Hate Speech klassifizieren. Die vorliegende Arbeit verfolgt deshalb das Ziel, den Hate-Speech-Begriff kommunikationswissenschaftlich zu definieren. Prämisse ist es, dass Hate Speech sich auf keine bestimmten Kommunikationsräume beschränkt, sich aber deutlich sichtbar im Internet und in sozialen Netzwerken manifestiert. Mit dieser Annahme wird eine Reduktion von Hate Speech auf bestimmte empirische Sachverhalte verhindert (hier konkret Hasskommentare, Hasspostings im Internet). Dabei wird Hate Speech nicht mit Hasskommentaren und/oder Hasspostings im Internet gleichgesetzt, wobei sich Hate Speech durchaus in Kommentaren und Postings in sozialen Netzwerken finden lässt. Ja, Hate Speech ist in digitalen Räumen weit verbreitet, was sich auf die Nutzungsarchitektur von sozialen Netzwerken zurückführen lässt. Die Verbreitung von Hate Speech in sozialen Netzwerken ist vergleichsweise mühelos, denn es können potenziell hohe Reichweiten durch Nutzen der Share-Funktion erzielt werden. Zusätzlich belohnen Nutzeralgorithmen jede Interaktion mit ebensolchen Inhalten, d. h. jede Reaktion – ob als zustimmende oder ablehnende Interaktion (sei es durch Nutzung von Reaktionsbuttons oder der Kommentarfunktion) – verstärkt die Sichtbarkeit von Hate Speech innerhalb eines sozialen Netzwerks (Munn, 2020).

Synonym zu Hate Speech wird in der kommunikationswissenschaftlichen Forschungsliteratur der Begriff *Inzivilität* verwendet, manchmal auch in Angrenzung davon. Häufig bleibt dessen Konzeption jedoch unklar, weil terminologische Grenzziehungen fehlen. Es wird hier argumentiert, dass zwischen Inzivilität und Hate Speech distinkte Unterschiede bestehen, was sich empirisch belegen lässt. Während Dangerous Speech zusätzlich zu Hate Speech zu Gewalt aufruft und mobilisiert, beschränkt sich Inzivilität auf offensive, polemische oder grob unhöfliche *sprachliche* Ausdrucksweisen. In dieser Arbeit werden drei definitorische Dimensionen für Hate Speech auf der Basis der „Going-Against-Equation“ von Gagliardone und Kolleg:innen (2016) aus kommunikationswissenschaftlicher Perspektive reflektiert, um diese für das Verständnis von Hate Speech und die Abgrenzung zu verwandten Phänomenen fruchtbar zu machen. Der *Intergruppenbias* eignet sich als sozialpsychologischer Ansatz für das Verständnis von Hate Speech, um offen ausgetragene Aggressionen der Ingroup (Eigengruppe) gegenüber der Outgroup (Fremdgruppe) als Überlegenheit oder Angst und Wut bei tatsächlicher oder wahrgenommener Bedrohung zu erklären (Brewer, 2001). Die Outgroup wird im Hate-Speech-Kontext als *primäre Adressatin* verstanden, bei der es sich mehrheitlich um

marginalisierte soziale Gruppen oder ihre (wahrgenommenen) Gruppenmitglieder handelt. Davon abgegrenzt werden müssen *sekundäre Adressat:innen*, die Angehörige der Ingroup sind und – sofern sie nicht bereits Sympathisant:innen sind – von der Hetzkampagne überzeugt werden sollen.

Hate Speech lässt sich aufgrund seiner *kommunikativen Inhalte* nicht auf den sprachlichen Ausdruck von Hass reduzieren, obwohl vermutlich das Verbalisieren von Hate Speech im Internet am sichtbarsten ist (ob verschriftlicht in Form von Kommentaren, Postings, Blogbeiträgen oder verbalisiert in audiovisuellen Formaten). Allerdings erfolgt Abwertung und Diskriminierung auch mittels nichtsprachlicher, symbolischer Nutzung von Körpern und Objekten – online wie offline. Die Fotografie von Harald Ewert, der bei den Ausschreitungen in Rostock-Lichtenhagen 1992 den Arm zum Hitlergruß hob, steht exemplarisch für offen zur Schau gestellte Feindseligkeit (Hampel, 2002).

Die *kommunikative Wirkung*, d.h. die Rezeptionswirkung von Hate Speech leitet sich aus der Annahme von Kommunikation als sozialem Handeln ab, bei dem Sendende wie Empfangende gleichermaßen ein Verständnis über den Bedeutungsgehalt besitzen, was an sprechakttheoretische Herangehensweisen anschlussfähig ist (Sponholz, 2018; Topczewska, 2017). Aus sozialpsychologischen Ansätzen lassen sich Wirkungen von Hate Speech ableiten, die aus intrapsychischen, intra- und intergruppalen Interaktionen resultieren. In erster Linie dient Hate Speech der Erniedrigung, Verunglimpfung und Herabwürdigung von primären Adressat:innen<sup>1</sup> (also der wahrgenommenen Outgroups), um diese zu diskriminieren und ihr psychisches Wohlbefinden negativ zu beeinträchtigen (Marker, 2013). „Sprache wird oft intentional wie eine Waffe benutzt, um auf kommunikativem Wege Handlungen auszuführen, die anderen Menschen Schaden zufügen sollen“ (Schwarz-Friesel, 2013, S. 145).

Gleichzeitig werden damit Sympathisant:innen und Anhänger:innen in ihrem Selbstwert und ihrer sozialen Identität als sekundärer Adressatenkreis gestärkt (Blazak, 2001). Möglicherweise werden sie sogar dazu angestiftet, die attackierte Outgroup selbst verbal oder physisch anzugreifen (Gerstenfeld et al., 2003). Insbesondere das Gefahrenpotenzial physischer Gewalt aufgrund von Hate Speech in Verbindung mit Aufrufen zu Gewalttaten lässt sich am Begriff Dangerous Speech ablesen (Benesch, 2013; Maynard & Benesch, 2016). Zudem lenkt

---

<sup>1</sup> In der vorliegenden Arbeit wird auf die Verwendung des generischen Maskulinums zugunsten der Schreibweise des Gender-Doppelpunkts (:innen) bzw. einer genderneutralen Personenbezeichnung (z. B. Rezipierende) verzichtet. Werden männliche oder weibliche Personenbezeichnungen verwendet (z. B. Cuddy und Kolleginnen; die Autoren), dann handelt es sich um Angehörige des jeweiligen Geschlechts. Abstrakte Begriffe, die nicht einen Personenkreis i. e. S. meinen (z. B. in „Kommunikatorforschung“), werden nicht in gegenderter Schreibweise verwendet.

Hate Speech aufgrund seines tiefgreifenden Empörungspotenzials die öffentliche Aufmerksamkeit auf gesellschaftliche Spaltungsprozesse, verursacht bzw. verstärkt diese. Vielfach wird die Position vertreten, die Produktion und Reproduktion von Hate Speech stelle einen gesellschaftlichen Desintegrationsmechanismus dar, der soziale Stereotype, Vorurteile und Rechtfertigungen für Diskriminierung stetig verfestigt (Blee, 2004; Pérez-Escobar & Noguera-Vivo, 2021; Sorial, 2015; van Dijk, 1995).

Obwohl die „Going-Against-Equation“ die Perspektive der kommunizierenden Person weitgehend ausklammert, verhilft diese aus kommunikationswissenschaftlicher Sicht dazu, das Phänomen Hate Speech in seiner psychosozialen Gänze zu erfassen, ohne sich um eine definitorische Festlegung für Hate Speech zu bemühen. Um ein Profil von Hate-Speech-Produzierenden zu generieren, sind (sozial-)psychologische Ansätze und Konzepte geeignet, wie die *Right Wing Authoritarianism*-Skala, die *Social Dominance Orientation* oder die *Dark Triad/Tetrad* (Bilewicz et al., 2017; Buckels et al., 2019; Fox & Rooney, 2015). Diese Erklärungsansätze erweisen sich als prognostisch erklärungsstark und zeichnen gemeinsam mit der Erhebung politischer Einstellungen und Ideologien ein präzises Bild der Täter:innen. Darüber hinaus zeigt die *Social Identity Theory*, wie die Aktivierung der Salienz von Gruppenzugehörigkeiten zu kognitiven und emotionalen Ingroup-Outgroup-Vergleichen führt, wodurch das Selbstbild des Individuums verbessert bzw. ein positives Selbstbild (wieder-)hergestellt wird (K. M. Douglas et al., 2005). Die Forschungslage ist uneindeutig bezüglich der Frage, wann Gruppenzugehörigkeiten salient werden: Im Nachgang von Schlüsselereignissen wird Hate Speech bis hin zu Hassverbrechen vermehrt beobachtet, wobei nur Letztere in Kriminalitätsstatistiken dokumentiert sind (Chetty & Alathur, 2018; King & Sutton, 2013).

Die größte Sichtbarkeit für sekundäre Adressat:innen (zumindest im deutschsprachigen Raum) erlangt Hate Speech durch seine Verbreitung im Internet und hier insbesondere in den sozialen Netzwerkseiten (SNS). Dies liegt zum einen an der prinzipiell freien Zugänglichkeit mit relativ geringen Hürden für die Nutzung (z. B. Kosten, Einfachheit der Nutzung), der Anonymität der Nutzer:innen und den unzureichenden Regulierungspraktiken sowie der mangelnden Hinweisreize sowie Netzwerkeffekte. Als weitere Erklärerfaktoren für die Produktion und Verbreitung von Hate Speech dienen Nutzungspraktiken wie *Social Loafing* (Zurückhaltung bei der Formulierung von Gegenrede), *Online Disinhibition* (Joinson, 2007; Suler, 2004) oder die Orientierung an salienten Gruppennormen (*Social Identity Model of Deindividuation Effects*) (Postmes et al., 2001; Reicher et al., 1995).

Es stellt sich die Frage, wie die Geschlechtskategorie mit dem Intergruppenbias und folglich mit Hate Speech zusammenhängt. Im Gegensatz zu Männern werden Frauen deutlich

häufiger mit sexistischem Hate Speech attackiert (Fuchs & Schäfer, 2020; Henson et al., 2013; Jane, 2017; Megarry, 2014). Die Aufzählung der Geschlechtskategorie in neueren Definitionen von Hate Speech (Crowley, 2014; Delgado & Stefancic, 2009; Meibauer, 2013; Sorial, 2015) erscheint deshalb nur auf den ersten Blick als widersprüchlich. Unbestritten lässt sich Geschlecht als keine sinnvolle Kategorie konzipieren, die das Verhältnis zwischen sozial marginalisierten Outgroups und dominanten Ingroups beschreibt. In den vergangenen 100 Jahren wurden wichtige Errungenschaften in der Gleichberechtigung der Geschlechter erzielt und die Interaktionen zwischen Männern und Frauen im privaten wie beruflichen Kontext verlaufen insgesamt friedlicher, weil möglicherweise schneller öffentliche Debatten und rechtliche Sanktionen drohen. Andererseits zeigen aktuelle Statistiken, dass Frauen bis heute überproportional von häuslicher Gewalt bedroht sind (Bundeskriminalamt, 2020). Dazu besteht bis heute eine Statusdisparität zwischen Männern und Frauen (z. B. Einkommensunterschiede, unterschiedliches Prestige von Berufen, die als stereotyp „männlich“ bzw. „weiblich“ gelten), die auch mit der Geschlechterstereotypisierung zusammenhängt, nach der Männer als kompetent sowie führungsstark und Frauen als sympathisch und fürsorglich schematisiert werden (Barreto & Ellemers, 2005; Cuddy et al., 2009; Fiske et al., 2002; Fiske et al., 2006; Fiske & North, 2015).

Gleichzeitig entspricht dies auch Brewers (2001) Argumentation für die Ursachen gruppenbasierter Abwertung. Denn der soziale Wandel der vergangenen Jahrzehnte, der durch vielfältige familien-, bildungs-, arbeits- und wirtschaftspolitische Maßnahmen in Gang gesetzt wurde, bewirkt eine Erschütterung der sozialen Identität bei Menschen mit entsprechender Prädisposition (Hewstone et al., 2002), was dann zu kompensatorischen Emotionen führt. Dazu gehört nicht nur, dass diese ihre moralische Überlegenheit gegenüber dem „schwachen Geschlecht“ mittels Verachtung demonstrieren. Da der Zugang von Frauen zu Bildungs- und Arbeitsressourcen (vor allem in männlich dominierten Berufszweigen) die bisherige geschlechtstypische Arbeitsteilung auflöst, reagieren solche Männer mit Wut und kanalisieren diese in Anfeindungen, die ihre geschlechtsbezogene soziale Identität aus der disjunkten Unterscheidbarkeit zwischen Männern und Frauen speisen. Im Einklang mit Brewers Argumentation (2001) stehen Erkenntnisse aus der Sexismusforschung, nach denen sowohl Frauen als auch Männer Anfeindungen bei abweichendem Verhalten von der traditionellen Geschlechterrolle ausgesetzt sind (Connor et al., 2016; Rudman & Fairchild, 2004). Frauenkarrieren in der IT-Branche gelten als Paradebeispiel für die Divergenz vom Geschlechterstereotyp, denn diese beinhalten gleich zwei „Angriffsflächen“: Neben dem hartnäckigen Stereotyp, nach dem Frauen als inkompetent im mathematisch-naturwissenschaftlichen Bereich charakterisiert werden, erhöht sich in Verbindung mit dem Intergruppenbias auch die Vulnerabilität für

Hate Speech von Frauen, die in diesem Bereich beruflich tätig sind, da dies als Eindringen in eine männlich dominierte Berufssphäre gilt.

Die Etablierung des Sexismusbegriffs in der sozialpsychologischen Forschung und hier insbesondere in der Vorurteilsforschung zeigt sich daran, dass zahlreiche Arbeiten entstanden sind, die Sexismus als Einstellung konzipieren (Campbell et al., 1997; Connor et al., 2016; Dovidio et al., 2010; Glick & Fiske, 1996; Swim & Campbell, 2001). Messmethodisch bieten Einstellungen den Vorteil, dass damit kognitive, affektive und verhaltensbezogene Komponenten erfasst werden können, was den Nachweis von Effekten auf ebendiese Bestandteile von Einstellungen in Experimentaldesigns erlaubt. Gegenüber Konzepten wie Antifeminismus oder Frauenfeindlichkeit erweist sich der Sexismusbegriff als vorteilhaft, weil damit sowohl Frauen als auch Männer als Subjekt und Einstellungsobjekt sexistischer Aussagen gemeint sind (Glick & Fiske, 1996, 1999). Dadurch lässt sich ein wesentlich differenzierteres Bild von geschlechterstereotypen und damit sexistischen Einstellungen zeichnen. Das Konzept des *Ambivalenten Sexismus* spielt in dieser Arbeit eine zentrale Rolle, weil es benevolente und hostile Anteile von Sexismus integriert (Collani & Werner, 2002; Connor et al., 2016; Glick & Fiske, 1996; Viki & Abrams, 2003). Unter *Benevolenz* ist der präskriptive Erwartungsrahmen für die Konkordanz zwischen Verhalten und globalem Geschlechterstereotyp der bewerteten Person zu verstehen, während *Hostilität* die feindseligen Abwertungen aufgrund der wahrgenommenen Abweichung beinhaltet. Angenommen wird, dass Ambivalenter Sexismus sowohl im thematischen Rahmen der Arbeit („Frauenkarrieren in der IT-Branche“) als auch in Bezug auf Hate Speech als Abweichung vom Geschlechterstereotyp wahrgenommen wird.

Die Forschungsrelevanz ergibt sich sowohl aus den bisherigen sozialpsychologisch verankerten Erkenntnissen zur Prävalenz von sexistischen Einstellungen mit ihren Folgen sowie den Erklärungsansätzen für die Ursachen als auch dem Anstieg an Anfeindungen, Bedrohungen und Diskriminierung in Onlineumgebungen, der in der öffentlichen Debatte als zunehmend problematisch eingestuft wird. Wie sich diese Attacken auf die Rezeption bei primären Adressat:innen auswirken, ist verhältnismäßig gut in qualitativen Interviews (Crowley, 2014; Gelber & McNamara, 2015; Nielsen, 2002; Shepherd et al., 2015) oder quantitativen Befragungsstudien (Herek, 2009) erforscht. Vereinzelt finden sich dazu auch Untersuchungen mit Experimentaldesigns (Leets & Giles, 1997). Von Hate Speech geht nachgewiesenermaßen ein persuasives Potenzial auf Anhänger:innen und Sympathisant:innen aus (Oberschall, 2000; Straus, 2007). Wenige Studien existieren jedoch zu den kognitiven und affektiven Rezeptionswirkungen von Hate Speech bei nicht sympathisierenden sekundären Adressat:innen (Soral et al., 2018) bzw. diese sind aufgrund ihrer unklaren Konzeptionalisierung nicht eindeutig

dem Hate Speech-Effekt zuzuordnen (Alvarez-Benjumea & Winter, 2018; Kaakinen, Oksanen & Räsänen, 2018; Lambe, 2004; Ziegele et al., 2018). Auch der systematische Vergleich unterschiedlicher Intensitätsstufen von Hate Speech ist bislang kaum beforscht (Asbury & Haas, 2008). Deshalb liegt der zweite empirische Untersuchungsschwerpunkt der Arbeit auf den Persuasionseffekten von sexistischem Hate Speech bei sekundären Adressat:innen.

Aus theoretischer Sicht liegt die kommunikationswissenschaftliche Relevanz des Forschungsvorhabens in der Klärung des persuasiven Potenzials von Hate Speech in Abgrenzung zu Inzivilität und Dangerous Speech, indem die jeweiligen Wirkungsmechanismen auf kognitive, affektive und behaviorale Anteile von Einstellungen erarbeitet werden. Auf diese Weise kann der Sogeffekt von Hate Speech ermittelt werden, der aus demokratietheoretischer Perspektive als bedenklich einzustufen ist, wenn hierdurch gesellschaftliche Gruppen mobilisiert werden können, die ohne Hate Speech nicht empfänglich für gruppenbezogene Abwertungen wären. Die zentrale Fragestellung der Arbeit lautet folglich:

*Welches persuasive Potenzial geht von sexistischem Hate Speech in Abgrenzung zu Inzivilität und Dangerous Speech auf die kognitiven, affektiven und behavioralen Komponenten von Einstellungen zu Hate Speech und zu Frauenkarrieren in der IT-Branche als Einstellungsobjekte bei Rezipient:innen aus?*

Als theoretisches Rahmenwerk zur empirischen Umsetzung dient Petty und Cacioppo (1986a) *Elaboration Likelihood Model* (ELM). Dieses Zwei-Prozess-Modell schematisiert die Informationsverarbeitung in einem Kontinuum zwischen *zentraler* (aufmerksame Verarbeitung der verfügbaren Informationen) und *peripherer* Informationsverarbeitung (Ablenkung durch äußere Hinweisreize). Persuasionseffekte werden beständig auf der zentralen Route erwartet, während die periphere Route zwar unbeständige, dafür leicht erreichbare Wirkungen prognostiziert. Das Modell erlaubt einerseits differenzierte Aussagen zum Wirkmechanismus persuasiver Botschaften, z. B. bezüglich der Interaktionen aus Rezipient:innenmerkmalen und Eigenschaften der Botschaft, was für die Untersuchung grundlegender Wirkprozesse spricht. Andererseits ist das ELM in der Konsumentenforschung verwurzelt und geht von Annahmen aus, die sich mit Hate Speech nur unter bestimmten Bedingungen vereinbaren lassen. Die Anwendbarkeit des Modells wurde bereits im Zusammenhang mit der Persuasivität von extremistischen Websites (Dunbar et al., 2014; E. Lee & Leets, 2002) oder der Prognose zur Begegnung von Hate Speech mittels Counter Speech (Schieb & Preuss, 2018) getestet. Im Zusammenhang mit Annahmen, die aus dem ELM abgeleitet werden, bedarf es der Klärung, inwiefern sich das ELM auf sexistischen Hate Speech als Gegenstandsbereich übertragen lässt:

*In welchem Ausmaß ist das Elaboration Likelihood Model geeignet, Persuasionseffekte von sexistischer Hate Speech zu prognostizieren?*

Unter der Anwendung des Modells lassen sich sodann folgende nachgeordneten Fragestellungen formulieren, welche die kommunikationswissenschaftliche Perspektive auf Hate Speech mit der sozialpsychologischen Sexismus- und Persuasionsforschung zusammenführen:

*Welche Rezeptions- und Persuasionseffekte gehen von Botschaftsmerkmalen sexistischer Hate Speeches auf Rezipient:innen aus? Inwiefern wird die Persuasion durch Botschaftsmerkmale sexistischer Hate Speeches und Rezipient:innenmerkmalen beeinflusst?*

Im Rahmen dieser Arbeit wird zunächst der Hate-Speech-Begriff aus kommunikationswissenschaftlicher Perspektive reflektiert, um ein Verständnis für sexistische Hate Speech, seine Ursachen und Wirkungsweisen anhand von sozialpsychologischen Erkenntnissen aus der Sexismusforschung zu generieren, das der empirischen Analyse von Rezeptionseffekten unter Einsatz des ELM zugeführt wird.

## **1.2 Aufbau der Arbeit**

Die Bearbeitung der Fragestellungen erfolgt nach dem üblichen Schema sozialwissenschaftlicher empirischer Untersuchungen, das grob aus Vorstellung und Diskussion der theoretischen Ansätze, der methodischen Überlegungen, der Darstellung der Befunde und abschließender Diskussion besteht.

*Kapitel 2* nähert sich dem Hate Speech-Begriff zunächst mit einem Forschungsüberblick über die aktuellen v. a. sozialwissenschaftlichen und frühen juristischen Studien an. Aus kommunikationswissenschaftlicher Perspektive wird auf die definitorischen Dimensionen von Hate Speech eingegangen, um diese begrifflich von verwandten Phänomenen im Internet auf der Basis dieser Dimensionen abzugrenzen. Da mit dieser Arbeit eine möglichst umfassende Darstellung des Phänomens Hate Speech beabsichtigt wird, werden zusätzlich individual- und sozialpsychologische Erklärungsansätze für die Entstehungsbedingungen von Hate Speech vorgestellt.

In *Kapitel 3* wird Sexismus als Einstellungskonstrukt aus der Perspektive sozialpsychologischer Vorurteilsforschung vertiefend erläutert, um den Zusammenhang zwischen Hate Speech und der Geschlechtszugehörigkeit herauszuarbeiten. Da der Fokus der Arbeit auf „Frauen in der IT-Branche“ liegt, werden mögliche Ursachen und Ausprägungen von Sexismus gegenüber Frauen in der IT-Branche erläutert.

Als theoretisch-analytische Untersuchungsgrundlage von Rezeptions- und Persuasionseffekten dient das Elaboration Likelihood Model, dessen Forschungsstand sowie auch Eignung für die Bearbeitung der vorliegenden Fragestellung in *Kapitel 4* erörtert wird. In einem Zwischenfazit sind am Ende des Kapitels alle vorgestellten theoretischen Erkenntnisse zusammengefasst.

Zur Konkretisierung des Forschungsinteresses werden in *Kapitel 5* die theoretischen Erkenntnisse in zu prüfende Hypothesen überführt und detailliert begründet. *Kapitel 6* enthält neben der Beschreibung des empirischen Vorgehens auch die Ergebnisse aus zwei eigens durchgeführten experimentellen Primärstudien. Dazu ist das Kapitel in zwei große Abschnitte gegliedert und befasst sich in *Kapitel 6.1* mit Studie I und in *Kapitel 6.2* mit Studie II. Darin enthalten sind jeweils Darstellungen des Forschungsdesigns und der Methodik (*Kapitel 6.1.1* und *Kapitel 6.2.1*), Ergebnisse (*Kapitel 6.1.2* und *Kapitel 6.2.2*) und jeweils eine einordnende Zusammenfassung der Befunde (*Kapitel 6.1.3* und *Kapitel 6.2.3*).

Die Arbeit schließt mit einem Fazit in *Kapitel 7*, in dem nicht nur die zentralen Befunde gegenübergestellt und diskutiert, sondern auch die Limitationen der Studie mit einer forschungsethischen Schlussbetrachtung von Hate-Speech-Forschung im Allgemeinen sowie in den hier durchgeführten Studien im Speziellen genannt werden. Diese Reflexion dient dazu, Rezeptionsforschung im Hate Speech-Kontext kritisch zu würdigen.

## 2 Hate Speech – Theoretische Fundierung

Fraglich ist, welche Inhalte tatsächlich unter dem Hate Speech-Begriff subsumiert werden können. Denn vielfach bleibt der Hate Speech-Begriff undifferenziert und folgt einem Alltagsverständnis, das darunter eine Beleidigung oder Beschimpfung versteht, was unter den Begriff der Inzivilität fällt. Dass Hate Speech weit mehr als unhöfliche oder respektlose Bekundungen beinhaltet, wird in den nächsten Abschnitten behandelt. Im ersten Unterkapitel erfolgt eine begriffliche und definitorische Einordnung des Hate Speech-Phänomens. Dazu werden drei Kernmerkmale identifiziert, welche die eindeutige Identifikation von Hate Speech erlauben. In Kapitel 2.2 werden in einem Zwischenfazit die gewonnenen Erkenntnisse in eine Definition für Hate Speech überführt. Auf Basis der drei Kernmerkmale werden in Kapitel 2.3 drei besonders prominente Phänomene aggressiv wahrgenommener Sprache im Online-Kontext von Hate Speech konzeptuell abgegrenzt. Kapitel 2.4 behandelt in einem Exkurs psychosoziale Eigenschaften der Kommunikator:innen, die zwar nicht konzeptionell Teil der definitorischen Erfassung sind, aber die Relevanz von Hate Speech verdeutlichen.

### 2.1 Begriffliche und definitorische Einordnung

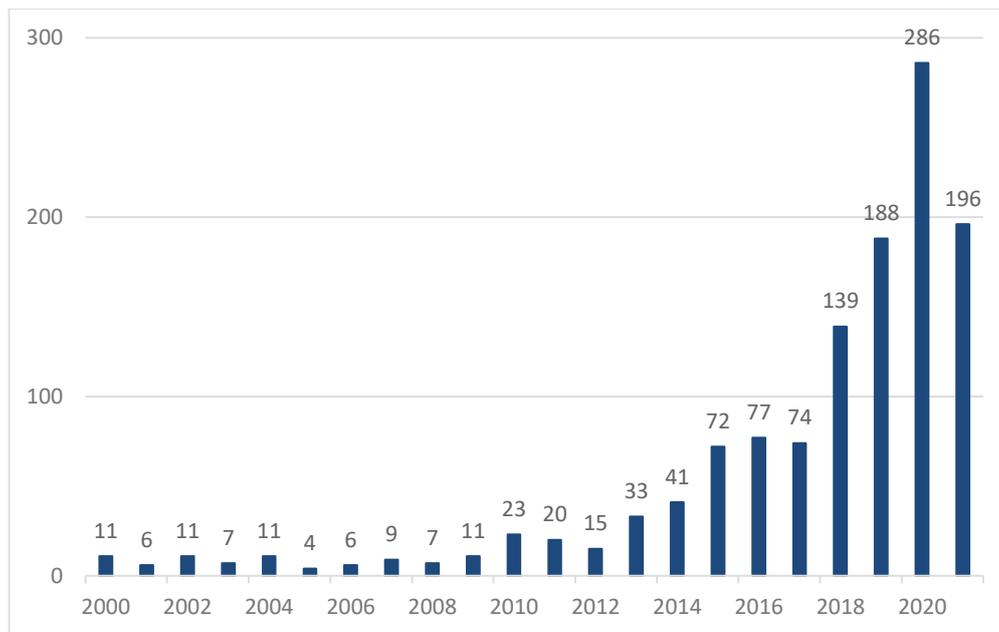
#### 2.1.1 Begriffsinflation

Wissenschaftliche Untersuchungen zu Hate Speech haben Konjunktur. So sind allein seit 2015 mehr als 1000 Publikationen zum Thema erschienen, ein Vielfaches des zwischen 2000 und bis 2014 veröffentlichten Forschungsoutputs (vgl. Abbildung 1). Untersuchungsobjekt sind z.B. *Hasskommentare* bei Facebook (Awan, 2016; Neumann & Arendt, 2016; Ziegele et al., 2018), in Blogforen (Anderson et al., 2014) oder im Kommentarbereich von Online-Nachrichtenseiten (Chandrasekharan et al., 2017; Coe et al., 2014; Santana, 2014) und *Hasspostings* in sozialen Netzwerken (Burnap & Williams, 2015; Groshek & Cutino, 2016; Kwok & Wang, 2013). Das Forschungsinteresse gilt z. B. Einstellungen zu Hate Speech (van Spanje & de Vreese, 2014), seiner Produktion (Costello & Hawdon, 2018; Erjavec & Kovacic, 2012), Rezeption (Costello et al., 2016) und Wirkung (u. a. Gelber & McNamara, 2015; Leets & Bowers, 1999).

Mit diesem Forschungsoutput ist jedoch auch eine Begriffsinflation verbunden (Schieb & Preuss, 2018), denn Aspekte von Hate Speech zeigen sich in Termini wie *(In-)civility/Impoliteness* (Alhabash, Baek et al., 2015; Groshek & Cutino, 2016; Lange, 2014; Megarry, 2014; Rösner et al., 2016), *Flaming* (D. Cho & Kwon, 2015), *Trolling* (Coles & West, 2016) oder *Cyberbullying* bzw. *Cybermobbing* (Räsänen et al., 2016). Darüber hinaus wird eine Vielzahl an Begrifflichkeiten synonym zu Hate Speech genutzt, so z. B. *Cyber Hate* (Burnap & Williams,

2015; Quandt & Festl, 2017), *E-Bile* (Jane, 2014), *Dangerous Speech* (Maynard & Benesch, 2016), *Fear Speech* (Buyse, 2014), *Extreme Speech* (Pohjonen & Udupa, 2017) oder *Hate Speech Online* bzw. *Hate Speech im Internet* als am häufigsten verwendeter Terminus (Ben-David & Matamoros-Fernández, 2016; Gagliardone et al., 2015; Gagliardone et al., 2016; Kaspar et al., 2017; Lev-On, 2018; Shepherd et al., 2015; Sponholz, 2018).

**Abbildung 1:** *Wissenschaftliche Publikationen zum Thema Hate Speech*



(Eigene Darstellung, Datenbasis: Web of Science, Stand: September 2021)

Die aktuelle Forschung zu Hate Speech konzentriert sich primär auf Hate Speech in seiner Ausprägung als „uncivil, hateful, and prejudiced comments“ (Hsueh et al., 2015, S. 558) – also auf Hasskommentare (Hughey & Daniels, 2013; Kalch & Naab, 2017; Lange, 2014; Loosen et al., 2017; Muddiman & Stroud, 2017; Neumann & Arendt, 2016; Rim & Song, 2016; Ziegele et al., 2017; Ziegele et al., 2018; Ziegele & Jost, 2016). In der Regel handelt es sich um Nutzerkommentare in sozialen Netzwerkseiten (SNS) oder unter journalistischen Beiträgen auf Online-Nachrichtenseiten (Ziegele, 2016). Dabei wird Hate Speech synonym zu Hasskommentaren verstanden: „Als ‚Hasskommentar‘ werden all jene Beiträge bezeichnet, die die Annahme einer Unterlegenheit anderer Personen aufgrund ihrer Rasse, ihres Geschlechts, ihrer Nationalität oder Religion beinhalten und/oder zu Gewalt, Diskriminierung und Hass aufrufen“ (Neumann & Arendt, 2016, S. 248, Herv. i. O.). Der nächste Abschnitt setzt sich mit dieser als unzureichend empfundenen Sicht auf Hate Speech auseinander und unterbreitet einen konzeptionellen Vorschlag, der im Folgenden ausführlich erläutert wird.

### 2.1.2 Internet als exklusiver Kommunikationsraum für Hate Speech?

Hate Speech in Online-Kommentaren in sozialen Netzwerkseiten stand nicht immer im Mittelpunkt eines breiten Forschungsinteresses. Vielmehr handelt es sich hierbei generell um das Aufgreifen von Trends und Moden in der Wissenschaft und speziell in der Erforschung von Hate Speech. Dafür spricht, dass Forschungsarbeiten, die nur 10 bis 20 Jahre älter sind, sich in erster Linie Inhalten auf *extremistischen Websites*, in *Blogs* oder *Chaträumen* widmen (Burris et al., 2000; Cammaerts, 2009; K. M. Douglas et al., 2005; Gerstenfeld et al., 2003; Glaser et al., 2002; Josey, 2010; E. Lee & Leets, 2002; Leets, 2001; B. Levin, 2002; Meddaugh & Kay, 2009; Pollock, 2009). Zeitgleich erscheinen generische Untersuchungen zu Weblogs (J. Schmidt, 2006), Chaträumen (Beißwenger, 2007; Whitty, 2002) oder Websites (van der Heijden, 2003; Vazire & Gosling, 2004), die sich mit diesen zu dem Zeitpunkt neuartigen Kommunikationsformen auseinandersetzen. Aktuelle Untersuchungen zu Hate Speech in diesen Publikations- und Kommunikationsräumen sind rar, obwohl auch weiterhin sogenannte Hate Sites wie Daily Stormer, Stormfront<sup>2</sup> oder die English Defence League existieren (Blakemore, 2016), also Websites, die als Veröffentlichungsplattform für Hate Speech dienen. Darüber hinaus ist Hate Speech im Internet so alt wie das Internet selbst. Bevor das World Wide Web 1989 am CERN in Genf entwickelt wurde, war das Usenet in den 1980er Jahren als Kommunikations- und Datenaustauschdienst unter den damaligen (wenigen) Internetnutzern weit verbreitet. Bereits dort wurden in Newsgroups Hassbotschaften veröffentlicht. Die vorhandenen Untersuchungen erläutern solche Vorfälle illustrierend aus historisch-einordnender Perspektive (Berlet, 2001; Citron & Norton, 2011; Delgado & Stefancic, 2014; Fleischhack, 2017).

Für die These spricht des Weiteren, dass die aktuelle Zuwendung zu Hate Speech in Online-Hasskommentaren sich nur bedingt mit veränderten Online-Nutzungsgewohnheiten erklären lässt, da der traditionell rezipierende Charakter dominiert (Hölig & Hasebrink, 2018). Die Sichtbarkeit von Hate Speech in Deutschland ist zwar hoch, denn drei von vier Befragten haben schon einmal Hasskommentare im Internet rezipiert (Landesanstalt für Medien NRW, 2021). Ursächlich hängt dies jedoch nicht mit der Anzahl der Hate Speech-Produzent:innen zusammen, da sich deren Zahlen im niedrigen einstelligen Prozentbereich befinden (Kaakinen, Räsänen et al., 2018; Kreißel et al., 2018). Vielmehr sorgen wenige Nutzende für einen hohen Output an entsprechendem Material (Erjavec & Kovacic, 2012), welches aufgrund

---

<sup>2</sup> In Deutschland aus rechtlichen Gründen wegen der dort verbreiteten Holocaustleugnung seit 2002 nicht mehr erreichbar (Quelle: <http://oj.is/95230>).

seiner Feindseligkeit und Aggressivität erhöhtes Sichtbarkeitspotenzial hat (Kaakinen, Räsänen et al., 2018).

Überdies unterscheidet sich Hate Speech im Internet nicht von Hate Speech in anderen Kommunikationsräumen<sup>3</sup>, denn Hate Speech stellt einen universellen Abwertungsmechanismus dar – unabhängig vom Kommunikationsraum<sup>4</sup> – wie in den folgenden Abschnitten ausführlich erörtert wird. Sicherlich spielen die besonderen Eigenschaften des Internets als Kommunikationsraum eine Rolle hinsichtlich der Ausprägung oder Gewichtung von Kommunikationsmodi (vertiefend Fraas et al., 2012)<sup>5</sup>. Für definitorische Zwecke sind sie jedoch vernachlässigbar, da eine generalisierbare Definition von Hate Speech angestrebt wird. Eine eingrenzende Sicht auf Hate Speech *im Internet* wäre der Generalisierbarkeit abträglich und vor allem anfällig für Technikdeterminismus. Immer dann, wenn in der vorliegenden Arbeit von Hate Speech die Rede ist, wird der Begriff im generischen Sinne eines Abwertungsmechanismus<sup>6</sup> verwendet. Wenn auf besondere Charakteristika von Online-Kommunikationsräumen hingewiesen werden soll, wird der Begriff *Hate Speech im Internet* verwendet. Private Räume, in denen sich Hate Speech mittels interpersonaler Kommunikation konstituiert (die sprichwörtlichen „eigenen vier Wände“) und halb-private („Stammtisch“, private E-Mails) sind von der Betrachtung im Rahmen dieser Arbeit ausgeschlossen, da diese einen geschlossenen Rezipientenkreis adressieren, ohne die gesellschaftliche Außenwirkung zu intendieren<sup>6</sup>.

Aus den genannten Gründen beziehen sich die folgenden Ausführungen ganz allgemein auf Hate Speech in der öffentlichen Sphäre, da nur hier das ‚öffentlich Machen‘ allgemeine Aufmerksamkeit erzeugt (Gerhards & Neidhardt, 1991) und nur die Öffentlichkeit als „Netzwerk

---

<sup>3</sup> Anders als die frühe Flaming-Forschung sind Untersuchungen zu Hate Speech im Internet weitgehend frei von technologiedeterministischen Annahmen. Online-Disinhibition und -Deindividuation werden zwar teilweise mit fehlenden sozialen Hinweisreizen oder technischer Anonymität in Verbindung gebracht (Reader, 2012; Santana, 2014), v. a. handelt es sich aber um soziale Kontexte, etwa soziale Anonymität oder die Ausrichtung an Gruppennormen (Kaspar, 2017). Die vorliegende Arbeit wird an mehreren Stellen Aspekte von Online-Disinhibition und -Deindividuation aufgreifen, z. B. in Kapitel 2.2.1, 4.4, um die sozialen Aspekte von enthemmtem Online-Verhalten zu beleuchten.

<sup>4</sup> Offline- und Online-Welten werden von jüngeren Internetnutzern kaum noch als getrennte Kommunikationsräume betrachtet (Räsänen et al., 2016; van Wilsem, 2011).

<sup>5</sup> Kapitel 4.4 greift bestimmte Aspekte computervermittelter Kommunikation heraus und diskutiert sie im Hinblick auf vermutete Persuasionseffekte von Hate Speech.

<sup>6</sup> Solche als privat wahrgenommene, geschlossene Gruppen wie die französische Facebook-Gruppe *Ligue Du LOL* müssen sich in der Öffentlichkeit rechtfertigen, wenn publik wird, dass sie Hate Speech produzieren oder für Cybermobbing verantwortlich sind. Die *Ligue du LOL* war als Netzwerk von hauptsächlich männlichen Journalisten und anderen Medienschaffenden v. a. zwischen 2009 und 2012 aktiv und koordinierte Hetzkampagnen gegen feministische Autor:innen, Journalist:innen und YouTuber:innen. Die Existenz der *Ligue du LOL* wurde am 08. Februar 2019 von der französischen Zeitung *Libération* aufgedeckt (Peteranderl, 2019).

für die Kommunikation von Inhalten und Stellungnahmen“ (Habermas, 1992, S. 436) einen aufmerksamkeitsstarken Problemdruck zu erzeugen vermag. Zur öffentlichen Sphäre zählt das Internet (Papacharissi, 2002, 2009) und damit potentiell wahrnehmbarer Hate Speech in Blogs, Foren, Social Media-Kommentaren etc. Des Weiteren ist auch die massenmedial vermittelte Öffentlichkeit nicht frei von Hate Speech-Diskursen (Sponholz, 2018). Und schließlich ist die Produktion von Hate Speech im öffentlichen Raum im Rahmen öffentlicher Veranstaltungen (z. B. Kundgebungen, Aufmärschen) zu nennen (Vorländer et al., 2016).

### 2.1.3 Juristische Prägung früher Forschung

US-amerikanische Rechtswissenschaftler setzen sich zuerst mit Hate Speech auseinander und betonen den Ursprung des Terminus indem sie ihn als „American expression“ (K. Boyle, 2001, S. 489) klassifizieren. Juristische Analysen dominieren bis Mitte der 2000er Jahre den wissenschaftlichen Diskurs, bevor Hate Speech immer mehr aus dem Blickwinkel sozialwissenschaftlicher Fragestellungen untersucht wird. Da die freie Meinungsäußerung durch den ersten Zusatzartikel in der Verfassung der *Vereinigten Staaten*, dem sogenannten First Amendment, besonderen Schutz genießt, ist ein Verbot von Hate Speech dort nicht gesetzlich normiert (K. Boyle, 2001; Citron & Norton, 2011; Reader, 2012). Das ist auch der Grund dafür, dass juristische Definitionen zwar Hate Speech als Tatbestand formulieren, aber die Rechtsfolgen in der Definition ausklammern. Im Wesentlichen werden Argumente erörtert, die Hate Speech zwar problematisieren, aber die Meinungsfreiheit als unantastbares Verfassungsgut bewerten (Hemmer, 2000; Rosenfeld, 2003). In vielen anderen Ländern (z. B. Frankreich, Deutschland) klammern *nationale Gesetzgebungen* Hate Speech und verwandte Sachverhalte wie die Holocaustleugnung aus dem Schutzbereich der Meinungsfreiheit aus (K. Boyle, 2001; Rosenfeld, 2012). Das deutsche Netzwerkdurchsetzungsgesetz (NetzDG) wird zwar umgangssprachlich als Hate Speech-Gesetz bezeichnet, es definiert aber nicht den Begriff als solchen, sondern verweist auf verwandte Straftatbestände, z. B. Verleumdung, Volksverhetzung, Beschimpfung von Bekenntnissen, Religionsgesellschaften und Weltanschauungsvereinigungen. Auch auf *supranationaler Ebene* ist Hate Speech vom fundamentalen Recht auf Meinungsfreiheit ausgenommen, etwa im Zivilpakt der Vereinten Nationen oder in der Europäischen Menschenrechtskonvention. Artikel 20, Absatz 2 des UN-Zivilpakts formuliert dieses Verbot explizit als: „Jedes Eintreten für nationalen, rassistischen oder religiösen Hass, durch das zu Diskriminierung, Feindseligkeit oder Gewalt aufgestachelt wird, wird durch Gesetz verboten.“

Erste juristische Erörterungen von Hate Speech beziehen sich auf rassistisch motivierte Äußerungen und Handlungen (sogenannte Hate Crimes), etwa in US-Colleges (Goldberg, 1995;

Lawrence, 1990). Wiederholte Hate Speech-Ausbrüche werden mit einem strukturell tief verankerten Rassismus in der US-amerikanischen Gesellschaft begründet (K. Boyle, 2001; Delgado & Stefancic, 2009; Delgado & Yun, 1995; Downs, 1985; Goldberg, 1995; Lawrence, 1990; Waldron, 2010; Watterson, 1991). Eng verwandt ist der Hate Speech-Begriff mit dem des *Ethnophaulismus*<sup>7</sup>, geprägt von Abraham Roback, und meint die Verunglimpfung von Ethnien (Roback, 1944). Diese enge Anbindung von Hate Speech an Rassismus zeigt sich z. B. in Georges Definition, der darunter „forms of expression aimed at persecuting people by vilifying their racial, ethnic, or other identities“ (George, 2015, S. 305) versteht. Ähnlich formuliert Rosenfeld und nimmt die religiöse Zugehörigkeit als Kategorie auf: „speech designed to promote hatred on the basis of race, religion, ethnicity or national origin“ (Rosenfeld, 2003, S. 1523). Die benannten Kategorien variieren zwar, aber im Kern handelt es sich um „race, religion, sexual orientation, immigration status, or gender“ (Delgado & Stefancic, 2009, S. 361). Gemeinsam ist allen Definitionen, dass wahrgenommene Outgroups Ziel von Hate Speech werden können, also jene gesellschaftlichen Gruppen, zu denen keine Identifikationsbasis besteht. Dazu gehören Mitglieder bestimmter ethnischer oder religiöser Minderheiten, Nationalitäten, Geschlechtszugehörigkeiten u. a. m., welche in Definitionen summarisch gelistet werden:

[...] speech that denigrates persons on the basis of their race or ethnic origin, religion, gender, age, physical condition, disability, sexual orientation (Leets & Giles, 1997, S. 261)

Hassrede kann sich richten gegen Personen oder Gruppen mit bestimmten Eigenschaften wie zum Beispiel Hautfarbe, Nationalität, Herkunft, Religionszugehörigkeit, Geschlecht, sexuelle Orientierung, sozialer Status, Gesundheit, Aussehen, oder Kombinationen davon (Meibauer, 2013, S. 2)

### 2.1.4 Kernmerkmale als definatorische Dimensionen

Wie die bisherigen Ausführungen gezeigt haben, wird Hate Speech als umstrittener Abwertungsmechanismus verstanden, der einer konzeptionellen und definatorischen Schärfung bedarf. Aus den vier Ebenen (Adressaten, Kommunikationsinhalt, Handlungsaufforderung und Machtungleichgewicht<sup>7</sup>) der „*Going-Against-Equation*“ (Gagliardone et al., 2016, S. 19) werden drei Kernmerkmale (Adressatengruppen, Kommunikationsinhalt und Wirkungsdimension) extrahiert, um dem Anspruch einer kommunikationswissenschaftlichen Definition von Hate Speech nachzukommen. Dies geschieht in Abweichung zu den Überlegungen von

---

<sup>7</sup> Das Machtungleichgewicht ist als Kontextgeber aufschlussreich, weil die Wirkmacht von Hate Speech von der Bekanntheit, Reichweite etc. der Sprechenden beeinflusst wird. Da zwischen Hate Speech-Produzent:innen und primären Adressatengruppen aber per Definition ein Machtungleichgewicht herrscht, wird dieser Aspekt aus der Definition ausgeklammert.

Gagliardone und Kolleg:innen, welche Hate Speech anhand der Analyse von Adressatengruppen und Kommunikationsinhalt ermitteln und die Handlungsaufforderung als Unterscheidungskriterium für *Dangerous Speech* mit und ohne Gewaltpotential verstehen. In den folgenden Abschnitten wird ausführlich erörtert, warum neben den beiden Kriterien von Gagliardone und Kolleg:innen (2016) auch die Wirkungsdimension maßgeblich für das Verständnis von Hate Speech ist.

### 2.1.4.1 Adressatengruppen

Hate Speech richtet sich gleichermaßen an völlig unterschiedliche Adressatengruppen, welche hier als primäre und sekundäre Adressat:innen bezeichnet werden. Diese Unterscheidung ist nicht nur notwendig, weil es sich um unterschiedliche Gruppen handelt, sondern auch weil sie von divergenten Konsequenzen betroffen sind, worauf später eingegangen wird. Die *primären Adressatengruppen* von Hate Speech sind Kollektive, also soziale Gruppen, aber auch Individuen als Repräsentierende einer sozialen Gruppe. Die angefeindeten sozialen Gruppen gehören in der Regel sozial marginalisierten, statusniedrigen Minderheiten an und werden als Outgroup, d. h. nicht als Teil der Mehrheitsgesellschaft wahrgenommen. Dazu zählen die in den vorigen Abschnitten genannten gesellschaftlichen Fremdgruppen. Auf Gruppenebene sind Ingroup und Outgroup geprägt von einem Machtungleichgewicht<sup>8</sup>, welches sämtliche Lebensbereiche umfasst (vgl. Kapitel 3.2.1.3) und in Form von Hate Speech seinen (extremen) sprachlichen Ausdruck findet. Neben den dominanten gesellschaftlichen Gruppen als mehrheitlich Produzierende und Verbreitende von Hate Speech gehören auch Minderheiten zum Täterkreis und richten Hate Speech gegen andere Minderheiten, aber auch gegen gesellschaftliche Majoritäten (Meibauer, 2013)<sup>9</sup>. Diese Forschungsperspektive bleibt jedoch bislang deutlich unterrepräsentiert und wird an dieser Stelle nicht weiter berücksichtigt.

Werden Individuen mit herabsetzenden Botschaften konfrontiert, ohne die Zugehörigkeit zu einer Fremdgruppe zu prononcieren, handelt es sich nicht um Hate Speech, sondern z. B. um Cybermobbing, Flaming oder Trolling, deren Abgrenzung in den nächsten Unterkapiteln vertieft wird. Die Aussage „Adem, du bist so eine Lachnummer“ ist je nach Kontext als Cybermobbing (Festl, 2015, 2016) oder als Flaming (O'Sullivan & Flanagan, 2003) zu bewerten. Richtet sich aber eine herabsetzende Botschaft gegen ein Individuum und wird es zusätzlich

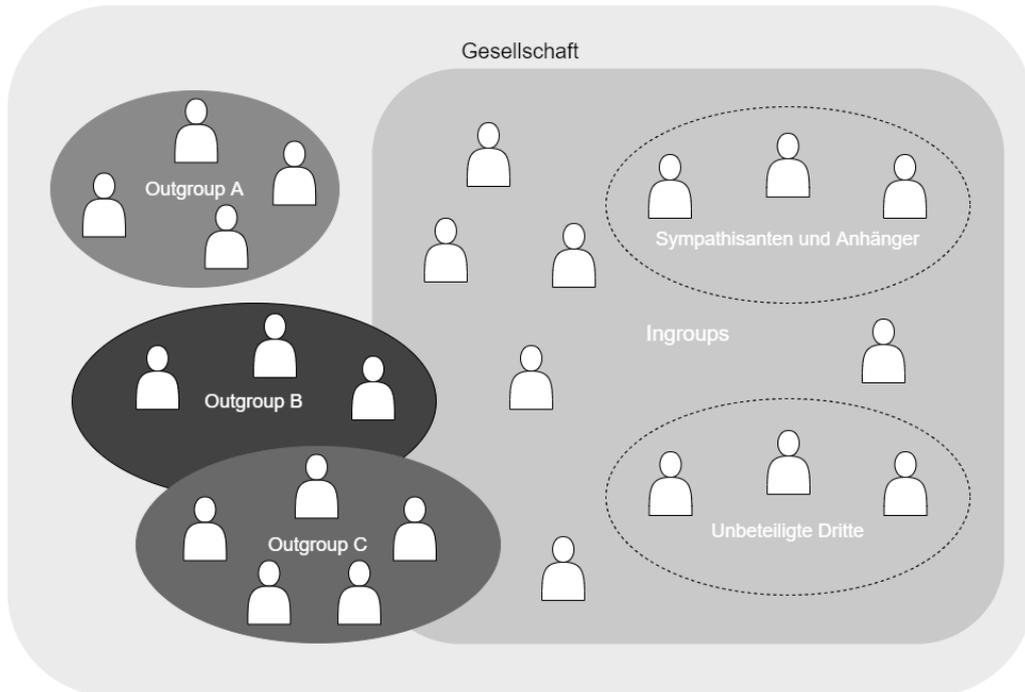
---

<sup>8</sup> Wie Gagliardone et al. (2016, S. 20) anmerken, tritt das Machtungleichgewicht noch deutlicher zu Tage, wenn Sprecher:innen nicht nur aufgrund der Zugehörigkeit zu einer dominanten sozialen Gruppe, sondern auch innerhalb dieser Gruppe einen hohen Status genießen.

<sup>9</sup> Kapitel 2.4 setzt sich vertiefend mit Hate Speech-Produzenten auseinander, weswegen an dieser Stelle deren Darstellung nur cursorisch erfolgt.

explizit als Vertreter:in einer bestimmten sozialen Gruppe angesprochen, wie in „Adem, ihr mordgeilen Moslems seid so eine Lachnummer“, lässt sich eine solche Aussage als Hate Speech kategorisieren. Hate Speech ist selbstredend nicht an faktische Zugehörigkeit zu einer sozialen Gruppe gebunden, sondern orientiert sich an attribuierten Hinweisreizen (Parekh, 2006; Sirsch, 2013), wie ein bestimmter Kleidungsstil oder Gesichtsbehaarung, um Angehörige einer bestimmten Religion, Ethnie oder sexuellen Orientierung zu identifizieren.

**Abbildung 2: Adressatengruppen**



(Quelle: Eigene Darstellung)

Im weiteren Sinne sind Mitglieder der Ingroup *sekundäre Adressat:innen* von Hate Speech, nämlich zum einen das nicht direkt tangierte Publikum, hier als *Unbeteiligte Dritte* bezeichnet. Deren Empörung wird von Hate Speech-Produzierenden billigend in Kauf genommen, sogar provoziert, mit dem Ziel, die Aufmerksamkeitslogik der Medien zu bedienen und damit die Reichweite ihrer Anliegen zu erhöhen (Erjavec & Kovacic, 2012). Zum anderen adressiert Hate Speech Sympathisant:innen und Anhänger:innen wegen seiner zugeschriebenen identitätsstiftenden Funktion (Gagliardone et al., 2015). Das dahinter liegende Ziel besteht in einer Verstärkung der Grenzziehung zwischen Ingroup und Outgroup, und zwar darüber „wer als legitimes Mitglied eines sozialen Zusammenhanges [...] gelten kann“ (Scherr, 2016, S. 11).

Generell sind intergrupale Positionskämpfe um Privilegien gesellschaftliche Realität (Scherr, 2016) und werden u. a. mittels Hate Speech ausgefochten. „Hassrede drückt aus, dass

Angehörige bestimmter Gruppen pauschal als weniger wertvoll und somit nicht als vollwertige und gleichberechtigte Mitglieder der Gesellschaft zu betrachten seien (Sirsch, 2013, S. 169). Bleiben die Attribute, die man einer sozialen Gruppe zuschreibt, unterkomplex (i. S. undifferenzierter Wissensbestände), ist das Maß der sozialen Ausgrenzung ungleich höher (Leader et al., 2009). Parekh (2006) expliziert die soziale Ausgrenzung:

[...] the target group is placed outside the pale of normal social relationships. It cannot be trusted to observe the rules governing society, to enter into meaningful relationships with other members of the community, to show loyalty to shared social institutions and practices, and is in general regarded as a hostile and unacceptable presence. It may therefore be legitimately exterminated, expelled, discriminated against, or at best tolerated as an unavoidable evil surviving on the margins of society. (Parekh, 2006, S. 214)

Diese Abwertung der Outgroup wird in der Sozialpsychologie als *Intergruppenbias* (Hewstone et al., 2002) bezeichnet<sup>10</sup>. Auf der Basis bestimmter sozialer Kategorien (z. B. Nationalität, Religionszugehörigkeit, Nachbarschaft, Geschlecht etc.) begreifen sich Menschen ihrer Ingroup zugehörig, also jener Gruppe, die ihre Kategoriemerkmale teilen (Tajfel, 1981). Alle anderen Menschen werden als Outgroup wahrgenommen, womit die Kategorisierung auf Basis der Gruppenzugehörigkeiten disjunkt ist. Brewers (1999) knappe Formel der „ingroup love“ vs. „outgroup hate“ fasst die Auswirkungen dieser Ingroup-/Outgroup-Kategorisierung zusammen: Im Allgemeinen wird die Ingroup gegenüber der Outgroup bevorzugt. Eine offene Abwertung der Outgroup tritt vergleichsweise seltener auf, denn meist leben Angehörige verschiedener Gruppen (z. B. Nationen, Ethnien) mehr oder minder friedlich nebeneinander (M. B. Brewer, 2001; weitergehend zu Intergruppenprozessen Brown & Gaertner, 2001). Ausbrechende Aggressionen, die sich in einer Abwertung der Outgroup (etwa in Form von Hate Speech) manifestieren, sind laut Marylann Brewer (2001) ursächlich an eine von zwei wahrgenommenen Bedrohungen gebunden: (1) Verlust der Zugehörigkeit zur Ingroup und (2) Verlust der Unterscheidbarkeit zwischen Ingroup und Outgroup. Das Verständnis von Hate Speech profitiert von dieser Argumentation, denn wer sich seiner Ingroup als nicht zugehörig fühlt, versucht sich enger an sie zu binden, fühlt sich der Outgroup moralisch überlegen und distanziert sich von ihr, indem er Verachtung und Ekel zeigt (M. B. Brewer, 2001). Scheint die eigene soziale Identität aufgrund von sozialen Wandlungsprozessen bedroht, etwa indem bislang marginalisierte Gruppen Zugang zu gesellschaftlicher Teilhabe erhalten o. ä., reagieren Mitglieder der Ingroup mit Angst und Wut gegenüber ihren Outgroups (M. B. Brewer, 2001; Cottrell & Neuberg, 2005; Neuberg & Cottrell, 2002). Kombiniert erscheinen die beiden Emotionen Verachtung aus einem Überlegenheitsgefühl gepaart mit Wut aus einer

---

<sup>10</sup> In Kapitel 3.2.1.3 folgt eine detaillierte Darstellung zentraler Theorien, die den Intergruppenbias erklären. In diesem Abschnitt geht es darum, den Zusammenhang zwischen Intergruppenbias und Hate Speech zu skizzieren.

subjektiven Bedrohungslage heraus als Erklärfaktoren für Hate Speech bis hin zu Hate Crimes und Genoziden: „[...] potent ingredients that are sufficient to kindle hatred, expulsion, and even ethnic cleansing“ (M. B. Brewer, 2001, S. 33).

#### 2.1.4.2 *Kommunikative Inhalte*

Gängige Definitionen von Hate Speech greifen zumeist den sprachlichen Aspekt auf:

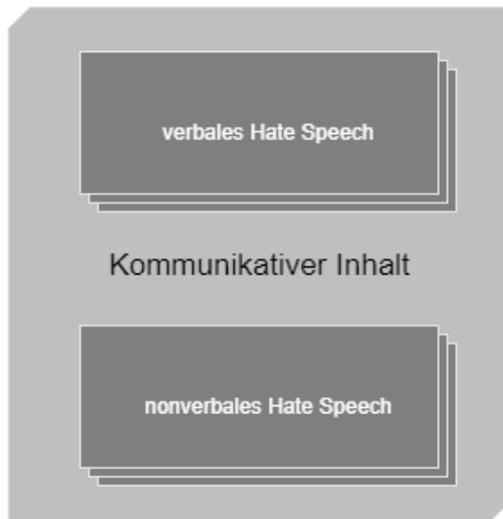
Hassrede ist eine Form von Rede, die inhaltlich demokratischen Grundwerten widerspricht und durch Einschüchterung – und der evtl. hieraus resultierenden Benachteiligung bestimmter Personengruppen – zu einer Unterminierung von Gleichheit und (positiver) Freiheit der Bürger führen kann. (Sirsch, 2013, S. 165)

Hate speech – that is speech designed to promote hatred [...] (Rosenfeld, 2003, S. 1523)

Kommunikationswissenschaftlich betrachtet ist Sprache sicherlich “immer Kommunikation, aber sie ist eine Kommunikationsform unter mehreren anderen“ (Maletzke, 1998, S. 44). Kommunikation selbst wird als dynamischer sozialer Prozess definiert (Beck, 2013b; Burkart, 2002; Kunczik & Zipfel, 2005), bei dem „seitens der Kommunikationspartner Sinn konstruiert, Information generiert und ausgetauscht“ (Pürer, 2014, S. 67) wird. Dies geschieht auf einer übergeordneten Ebene mittels sprachlicher, also verbaler, und/oder nonverbaler Kommunikation. *Verbale* Kommunikation erfolgt mündlich (z. B. face-to-face oder durch Medien vermittelt, etwa telefonisch, als Audio- oder Videomitschnitt etc.) oder schriftlich (z. B. in Blog-, Foren- und Chatbeiträgen, auf Handzetteln, in Büchern etc.) und transportiert den Inhalt und die Aussagekraft einer Botschaft (M. Maurer, 2016). Unter *nonverbaler Kommunikation* werden die drei Kommunikationsformen Kinesik, Figuretik und Akustik versammelt (Sager, 2018). *Kinesik* meint „alles aktive Körperverhalten mit semiotischem Charakter“ (Sager, 2018, S. 172) und gilt als wichtigste nonverbale Kommunikationsform (M. Maurer, 2016). Unter *Figuretik* werden Körpermanipulationen (Bartwuchs, Haartracht etc.) und die Körperausstattung (Tätowierungen, Kleidung) verstanden (Sager, 2018). Unter den *Akustik*begriff fällt die Prosodie (Tonhöhe, Akzent etc.) und das nicht-semantische vokale Display (z. B. Lachen, Räuspern) (ebd.).

Die Gesamtheit verbaler und nonverbaler Hate Speech-Äußerungen wird hier als *Kommunikativer Inhalt* bezeichnet und stellt nach den erörterten Adressatengruppen das zweite Kernmerkmal von Hate Speech dar. Das *verbale Kommunikationsrepertoire* bedient sich kollektiv verfügbarer Stereotype, wie im Folgenden ohne Anspruch auf Vollständigkeit beispielhaft anhand der Geschlechterkategorie illustriert wird.

Abbildung 3: Kommunikativer Inhalt



(Eigene Darstellung)

Phraseologismen, also idiomatische Redewendungen („Mädchen für alles“, „das Kind im Manne“), Syntagmen, d. h. zusammenhängende Ausdrücke („Mann gegen Mann“, „das schwache Geschlecht“, „eine frustrierte Emanze“), dehumanisierende Lexeme („Schlange“, „Esel“) oder pejorative Morpheme, wie -ling („Schönling“) sind übliche semantische Stilmittel (Meibauer, 2013; Schwarz-Friesel, 2013 für eine eingehende linguistische Auseinandersetzung am Beispiel Antisemitismus). Darüber hinaus werden Pejorativa, also verbale Abwertungen („Trulla“, „Lümmel“) und Schimpfwörter („Schlampe“, „Hurensohn“) verwendet. Negativ assoziierte Begriffe, sogenannte Disphemismen („Fräulein“, „Hungerhaken“, „Nerd“ und im Vergleich dazu das positiv assoziierte „Geek“), dienen ebenfalls der Abwertung, sind jedoch zeitlich oder situativ kontextabhängig (Bağ, 2017). Körperausscheidungen, Sexualität und Religion gelten als gesellschaftlich stark tabuisiert, da sie mit Ekel und Scham behaftet sind (Mancera, 2018), weswegen entsprechende Vulgarismen darauf abzielen, die Wirkung von Hate Speech zusätzlich zu verstärken. Mündliche Exempel für Hate Speech finden sich in herausgebrüllten Parolen auf PEGIDA-Demonstrationen (Vorländer et al., 2016), in Radioübertragungen während des ruandischen Genozids (Straus, 2007), oder in Audiomitschnitten auf extremistischen Seiten (Josey, 2010), um nur wenige Beispiele zu nennen. Wie bereits oben diskutiert, untersucht ein Großteil aktueller Hate Speech-Forschung den schriftlichen Ausdruck von Hate Speech im Onlinekontext. Primäre Forschungsobjekte sind – ohne Anspruch auf Vollständigkeit – Online-Hasskommentare (Coe et al., 2014; Ziegele et al., 2018), aber auch sogenannte Hate Sites, also extremistische Webseiten (Cammaerts, 2009;

Meddaugh & Kay, 2009) oder Hate Speech in Chat-Räumen (Glaser et al., 2002), in Briefen und E-Mails (Schwarz-Friesel, 2013).

*Nonverbale Kommunikationsformen* im Kontext von Hate Speech kommen ausschließlich dann in Betracht, wenn sie intentional verwendet werden. Außerdem wird bei der Analyse der situative Kontext berücksichtigt, damit nicht falsche Rückschlüsse auf isolierte, einzelne nonverbale Kommunikationshandlungen gezogen werden. Aus dem Bereich der Kinesik sind Gestiken zu nennen (also Hand- und Armbewegungen), die Aggressivität, Ablehnung oder Dominanz signalisieren, wie der ausgestreckte Mittelfinger, die symbolische Erschießung mit abgespreiztem Daumen und Zeigefinger, der Hitlergruß u. v. a. Diese sind häufig begleitet von Mimiken, die Verachtung oder Abscheu ausdrücken, wie das Augenkneifen, Naserümpfen oder Hochziehen eines Mundwinkels. Körperhaltungen und Körperbewegungen werden als Pantomimik bezeichnet und drücken im Hate Speech-Kontext Dominanz, Status aber auch Aggressivität aus. Sie zeigen sich beispielsweise in der demonstrativ herausgestreckten Brust oder in angedeuteten Trittbewegungen. Auch die Figuretik ist im Hate Speech-Kontext bedeutsam, wie z. B. kodifizierte subkulturelle Haartrachten, Kleidermoden von Rechtsradikalen (z. B. die Kapuzenkutten der Ku-Klux-Klan-Mitglieder) oder zur Schau getragene Fackeln bei rechtsradikalen Aufmärschen (Atkinson, 2018). Parasprachliche Kommunikation, die sogenannte Prosodie (M. Maurer, 2016), kann verbalen Hate Speech-Äußerungen Nachdruck verleihen, exemplifiziert durch eine Variation der Lautstärke und des Stimmregisters oder durch hämisches Auslachen.

Es besteht die Vermutung, dass eine Kumulation nonverbaler Formen von Hate Speech, vor allem in Verbindung mit entsprechenden sprachlichen Äußerungen dessen wahrgenommene Aggressivität verstärkt, was bislang eine empirische Frage bleibt. Die bisherige Hate Speech Literatur thematisiert nonverbale Hinweisreize allenfalls punktuell, meist als exemplifizierende Aufzählungen für Hate Speech (Delgado & Stefancic, 2009; Waldron, 2010). Die vielfältigen theoretischen Zugänge zu nonverbaler Kommunikation sind in der Hate Speech-Forschung bislang unbeachtet geblieben. Anschlussfähig an Hate Speech als Untersuchungsobjekt sind etwa solche Theorien, die von der Wirkung nonverbal geäußerter Dominanz, Macht und Statuszuschreibung (Burgoon et al., 2002) ausgehen oder eine Verhaltensanpassung an nonverbale Hinweisreize (Burgoon et al., 2010) propagieren.

### 2.1.4.3 *Kommunikative Wirkung*

Der niedrige Status diskriminierter oder marginalisierter sozialer Gruppen innerhalb der sozialen Hierarchie wird mittels Hate Speech weiter zementiert, was zu folgender These führt:

*Hate Speech ist kommunikatives Handeln.* Hate Speech als kommunikatives Handeln<sup>11</sup> aufzufassen, entspringt einer *handlungstheoretischen Perspektive*, deren Wurzeln sich in Max Webers Soziologie befinden (Burkart, 2002). Doch obwohl Weber eigentlich „*Handlung* erklären wollte – und Kommunikation dabei deutlich als Interaktionsmodus mitgedacht hat“ (Averbeck-Lietz, 2015, S. 36, Herv. i. O.), ist dessen Unterscheidung zwischen (Sich)Verhalten, Handeln und sozialem Handeln (Weber, 1980 [1922] für das Verständnis von Kommunikation im Allgemeinen (Burkart, 2002; Kunczik & Zipfel, 2005) und Hate Speech im Besonderen nützlich.

Während *Verhalten* als bloße physiologische Reaktion (z. B. Gähnen bei Müdigkeit) oder als unbewusstes oder routiniertes Sichverhalten (z. B. Auf-die-Uhr-schauen während eines langweiligen Vortrags) erfasst wird, ist *Handeln* „mit einem subjektiven Sinn“ (Weber, 1980 [1922] verbunden. Dass dieser Übergang zwischen unbewusstem Verhalten und sinnhaftem Handeln fließend ist, zeigt sich etwa, wenn man der:m Vortragenden mit dem Auf-die-Uhr-schauen signalisiert, dass die Redezeit abgelaufen ist. Damit ist Handeln bewusst und intentional (Beck, 2013a). Das bedeutet nach Burkart (2002), dass subjektiv sinnhafte (d. h. mit einer bestimmten Bedeutung für Handelnde) Handlungen bewusst ausgewählt werden, um ein gewünschtes Ziel zu erreichen. Ist eine sinnhafte Handlung auf andere Menschen ausgerichtet, spricht man von *sozialem Handeln* (ebd., Weber, 1980 [1922]). Insofern kann auch *Kommunikation als soziales Handeln* verstanden werden, „denn jedes gemeinsame Handeln von Individuen basiert auf geteilten, durch Kommunikation übermittelten Bedeutungen“ (Kunczik & Zipfel, 2005, S. 27). Diese ‚geteilten Bedeutungen‘ sind maßgeblich für die Verständigung zwischen den Kommunizierenden verantwortlich und begründen eine „allgemeine Intention“ (Burkart, 2002, S. 26). Burkart unterscheidet von der allgemeinen Intention die „spezielle Intention“, definiert als inhalts- oder situationsbezogene

---

<sup>11</sup> Kommunikatives Handeln folgt in diesem Zusammenhang (ähnlich argumentiert auch Burkart, 2002, S. 26) nur bedingt Habermas' Verständnis, formuliert in seiner *Theorie des kommunikativen Handelns* (Habermas, 1981). Darin arbeitet Habermas einen normativen Anspruch an kommunikativem Handeln heraus, welches Verständigung der Kommunikationspartner zum Ziel erklärt. In der idealtypischen Sprechsituation kennen und befolgen die Kommunikationspartner die vier Geltungsansprüche an gelungener Kommunikation: *Verständlichkeit* (geteilte Wort- und Sinnbedeutung), *Wahrheit* (Objekt- und Sinnbezug), *Wahrhaftigkeit* (subjektive Authentizität) und *Richtigkeit* (sozialer Bezug) (Averbeck-Lietz, 2015; Burkart, 2002, Habermas, 1981). Wer Hate Speech verbreitet, verletzt den Anspruch an Wahrheit (z. B. durch Aktivierung von Stereotypen und Vorurteilen über Outgroups), Wahrhaftigkeit (z. B. weil Dominanz und nicht bedrohte soziale Identität kommuniziert wird) sowie Richtigkeit (indem bewusst gesellschaftliche Normen und Werte verletzt werden und Grenzen des Sagbaren verschoben werden). In diesem Sinne sind Hate Speech-Produzierende nicht an Verständigung als primäres Ziel der Kommunikation interessiert, sondern verfolgen eigene Interessen, welche darin bestehen, im Konflikt die eigene soziale Identität durch Abwertung der Outgroup zu stärken. Damit ist Hate Speech in Habermas' Terminologie am ehesten als strategische Kommunikation zu bewerten.

Kommunikationsinteressen (Burkart, 2002, S. 27f.). In dem Vortragsbeispiel kennen Redner:in und Zuhörer:in die Bedeutung, die vom Auf-die-Uhr-schauen ausgeht (z. B. ablaufende Zeit) und mit dieser Handlung beabsichtigen Zuhörende ein Vortragsende zu initiieren (allgemeine Intention). Da Zuhörende als Mitglied des Organisationskomitees einer Tagung den planmäßigen Ablauf einhalten möchten (spezielle Intention), geben sie der vortragenden Person zu verstehen, dass der Vortrag beendet werden muss.

Die Frage nach der Intention von Hate Speech wird kontrovers diskutiert, da die individuelle Bewertung darüber variieren kann, ob eine Aussage vom Sprecher oder Adressaten als Hate Speech verstanden wird (vgl. auch den nächsten Abschnitt zur Abgrenzung von Hate Speech und Flaming). Konstruktivistische Perspektiven fordern eine holistische Sicht auf die unterstellte Intention, wozu die „sprachliche Handlung, die konventionalisierten und interaktiv immer wieder neu ausgehandelten Bedeutungen sprachlicher Zeichen, die Situation der sprachlichen Handlung mit den an dieser Kommunikation Beteiligten und die mit der sprachlichen Handlung appellierte Person und/oder aufgerufene Normalvorstellung“ (Hornscheidt, 2013, S. 39) gehören. Hornscheidt appelliert also dafür, sämtliche Dimensionen von Hate Speech integrativ zu berücksichtigen. Die vorliegende Arbeit versucht diesem Anforderungskatalog Rechnung zu tragen, denn sexistische Hate Speech-Aussagen (kommunikativer Inhalt) werden nicht nur ständig diskursiv (re-)produziert, sondern gerade im Kontext von weiblichen Karrierechancen (Frauen als Adressatinnen von Hate Speech) in der IT-Branche (situativer Kontext) persistieren diese.

Hate Speech wird als kommunikatives Handeln aufgefasst, denn Hate Speech ist sinnhaft auf andere Kommunikations“partner:innen“<sup>12</sup> ausgerichtet und entfaltet seine *kommunikative Wirkung* bei primären und sekundären Adressat:innen. Folgt man Burkarts (2002) Unterscheidung hinsichtlich der intentionalen Ausrichtung, so gilt auch für Hate Speech eine allgemeine Intention, die darin besteht, dass die genannten Kommunikations“partner:innen“ ein geteiltes Verständnis des Bedeutungsgehalts besitzen. Verbale (Schimpfworte, sprachliche Herabwürdigungen etc.) und nonverbale (Gesten und Mimiken oder Kleidung wie z. B. Nazi-uniformen etc.) Kommunikate werden in ihrer Bedeutung von allen verstanden. Aber auch indirekte, uneindeutige Formen von Hate Speech können sich mit der Zeit durchsetzen und von allen verstanden werden, wie das knappe „Go to work.“ (Benesch, 2013). Diese kodifizierte Aussage während des ruandischen Genozids 1994 wurde über das Radio an die Hutu

---

<sup>12</sup> Partner:in ist in Anführungszeichen hervorgehoben, um zu verdeutlichen, dass Hate Speech einen interaktionellen Kommunikationsmodus bewirkt, ohne den kooperativ aufeinander bezogenen Bedeutungsgehalt des Begriffs zu implizieren.

als Bevölkerungsmehrheit gerichtet und rief zur Ermordung von hunderttausenden Tutsi auf. Was die spezielle Intention von Hate Speech betrifft, so muss jeweils eine kontextabhängige Betrachtung erfolgen, um dessen inhaltlichen oder situativen Bezug im Sinne Burkarts (2002) zu berücksichtigen.

**Tabelle 1: Übersicht der kommunikativen Wirkung**

Adressatentyp	Wirkungsebene	Kommunikative Wirkung
Primäre Adressat:innen	Gesellschaftlich	Diskriminierung
	Intragruppal	Integration der Stereotype vs. Orientierung an Dominanzkultur
	Individuell	Beeinträchtigung des Wohlbefindens, verringertes Selbstwertgefühl, selbstschädigendes Verhalten
Sekundäre Adressat:innen (Sympathisant:innen und Anhänger:innen)	Intragruppal und individuell	Stärkung der sozialen Identität, Mobilisierung
Sekundäre Adressat:innen (Unbeteiligte Dritte)	Intergruppal und individuell	<ul style="list-style-type: none"> <li>In Opposition zu Hate Speech: Empörung, Enttäuschung, Ablehnung, Gegenrede</li> <li>In Konkordanz zu Hate Speech: Übernahme von Kognitionen und Emotionen, Desensibilisierung, Mobilisierung, sinkende Hilfsbereitschaft</li> </ul>

(Eigene Darstellung)

Aus linguistischer Perspektive ist die *Sprechakttheorie* geeignet, Hate Speech ebenfalls als kommunikative Handlung zu betrachten, die soziale Wirklichkeit wird (Butler, 2018; Sponholz, 2018; Topczewska, 2017), was in dem Buchtitel „Words that wound“ (Matsuda et al., 1993) und dem gleichnamigen Fachartikel von Richard Delgado (Delgado, 1982) prominent zur Sprache gebracht wird. Entwickelt von Austin (1975 [1955]) und weitergeführt von Searle (1969) besteht das Verdienst der Sprechakttheorie darin, sprachliche Äußerungen als Handlungen zu konzipieren: „talking is performing acts according to rules“ (Searle, 1969, S. 22). Sprachhandlungen lassen sich Searle (1969) zufolge analytisch in vier Teilakte untergliedern, und zwar in einen Äußerungs- (utterance act), propositionalen (propositional act)<sup>13</sup>, illokutionären (illocutionary act) und perlokutionären Sprechakt (perlocutionary act). Während der Äußerungsakt darin besteht, Morpheme, Wörter und Sätze hervorzubringen, nimmt der Sprecher in einem propositionalen Sprechakt Bezug auf Objekte und weist (prädiziert) diesen

<sup>13</sup> Bei Austin gehören der Äußerungs- und propositionale Akt zum lokutionären Akt, den er in drei weitere Teilakte (phonetisch, phatisch, rhetisch) auffächert (Austin, 1975). Searle (1969) akzentuiert hingegen die analytische Trennung des Äußerungs- vom Akt des Bezugnehmens und Bewertens.

bestimmte Eigenschaften zu. Ein illokutionärer Akt ist als vollständige Sprechhandlung definiert, indem der oder die Sprechende fragt, bittet, droht, empfiehlt, warnt, verspricht, zustimmt etc. „Illocutionary and propositional acts consist characteristically in uttering words in sentences in certain contexts, under certain conditions and with certain intentions“ (Searle, 1969, S. 24f.). Searle hebt den kontextuellen Bezug hervor, um zu betonen, dass eine Aussage mehrere Bedeutungen haben kann, wie sich an dem oben erwähnten Beispiel „Go to work“ verdeutlichen lässt. Die Aufforderung arbeiten zu gehen, kann nach einer Teambesprechung erfolgen, um das Ende der Sitzung zu deklarieren und den Beginn einer Arbeitsphase zu markieren. Während des ruandischen Genozids war diese Aussage als Aufruf zum Völkermord intendiert. In dem perlokutionären Sprechakt manifestiert sich die Wirkung des Sprechakts, indem der Sprecher sein Gegenüber überzeugt, verunsichert, kränkt, tröstet etc.

Sponholz (2018) betrachtet Hate Speech aus Austins (1975 [1955] Perspektive und arbeitet ein lokutionäres (Hate Speech als Aussage, Betrachtung des Wortschatzes), illokutionäres (Intention von Sprechenden) sowie ein perlokutionäres (Auswirkungen und Folgen von Hate Speech) Modell von Hate Speech heraus. Damit erweist sich der transdisziplinäre Zugang als teilweise fruchtbar für die Kommunikationswissenschaft, um einen theoretischen Bezugsrahmen für Hate Speech zu schaffen, und nicht auf Aussagenebene zu verharren. Die Ausklammerung nonverbaler Kommunikation innerhalb der Sprechakttheorie ist im Zusammenhang mit Hate Speech jedoch als Mangel zu bewerten, wird doch den gesprochenen Worten mit nonverbalen Signalen Nachdruck verliehen.

Dem sozialpsychologischen Ansatz dieser Arbeit folgend, wird der Forschungsstand zur Wirkung von Hate Speech im Folgenden entsprechend der vier sozialpsychologischen Analyseebenen dargestellt. Dabei handelt es sich zum einen um intrapsychische Prozesse und zum anderen um Effekte, die auf kollektiver Ebene angesiedelt sind, und zwar solche, die sich aus interindividuellen, intragruppalen und intergruppalen Interaktionen ergeben (Kessler & Fritsche, 2018).

Unter *intrapsychischen Prozessen* lassen sich innere Vorgänge, namentlich soziale Wahrnehmungen, kognitive Prozesse, Affekte und Motivationen zusammenfassen, die Individuen in sozialen Situationen erfahren. Diese müssen nicht notwendigerweise verborgen vor sich, also unbewusst, oder verborgen vor der sozialen Umwelt bleiben, sondern sind durchaus beobachtbar, sofern das Individuum seine Äußerung nicht willentlich unterdrückt. Bei primären Adressat:innen kann Hate Speech eine Beeinträchtigung des subjektiven Wohlergehens (Crowley, 2014; Keipi et al., 2017; Nielsen, 2002), psychisches Leid (Gee, 2002), selbstschädigendes Verhalten (Landrine & Klonoff, 1996) und eine starke Beeinträchtigung des

Selbstkonzepts (Gelber & McNamara, 2015; Tynes et al., 2010) auslösen. Emotionale Beeinträchtigungen bei Hate Speech Adressat:innen sind eher kurzfristiger Natur, langfristig sind Auswirkungen eher hinsichtlich der Einstellung zu Hate Speech und seinen Produzierenden feststellbar (Leets, 2002). Garcia und Kolleg:innen (1999) weisen bei Opfern von Hasskriminalität psychische Beeinträchtigungen nach. Dazu vergleichen sie die Antworten von 91 Opfern rassistisch motivierter Hasskriminalität mit denen von 45 Opfern schwerer Körperverletzung und zeigen, dass rassistisch motivierte Straftaten bei den Opfern signifikant häufiger unkontrollierte Gedanken rund um die Tat, um mögliche Racheakte bis hin zu Selbstmord auslösen. Weitere signifikante Auswirkungen zeigen sich in gesteigerter Nervosität, Konzentrationsschwierigkeiten und Depressionen. Betrachtet man mögliche Geschlechtsunterschiede, zeigt sich, dass weibliche Opfer von Hasskriminalität langfristig ein verringertes Selbstwertgefühl davontragen (ebd., ä. a. Barnes & Ephross, 1994). Vergleichbare Ergebnisse erzielen Herek, Cogan und Gillis (2002), die homosexuelle Opfer von Straftaten interviewen.

Unbeteiligte Dritte, d. h. Rezipierende, die nicht selbst von Hate Speech (bzw. inziviler Sprache in der empirischen Umsetzung) adressiert werden, reagieren darauf mit negativen Kognitionen und aggressiven Emotionen. Die Befunde sind nicht eindeutig, denn einerseits steht die Reaktion in Opposition zu feindseligen Äußerungen, indem sie Entrüstung und Ablehnung (B. T. Gervais, 2015) oder feindselige Kognitionen (Rösner et al., 2016) auslösen. Andererseits erhöhen inzivile Online-Kommentare die Risikowahrnehmung für Nanotechnologie (Anderson et al., 2014) oder senken die Spendenbereitschaft für Flüchtlinge (Ziegele et al., 2018), was einen Hinweis für mögliche persuasive Effekte auf Rezipierende liefert.

Effekte, die sich aus *interindividuellen Interaktionen* ergeben, werden von idiosynkratischen Merkmalen der beteiligten Individuen bestimmt (Kessler & Fritsche, 2018). Dazu gehören sämtliche Eigenschaften, welche das Selbst einer Person konstituieren und in der sozialen Interaktion beeinflusst werden (Morf & Koole, 2014). Diese Analyseebene ist nach Kenntnisstand der Autorin in der Hate Speech Forschung gänzlich ausgeklammert. Anschlussfähig erscheint die Untersuchung von Hate Speech-Wirkungen im Cyberbullying-Kontext: Studien, die das Mobbing sozialhierarchisch schwacher Gruppenangehöriger, z. B. von Menschen mit Behinderung (Didden et al., 2009) erklären, bewegen sich konzeptuell nahe am Hate Speech-Konzept.

*Intragruppale Interaktionen* dienen dazu, soziale Normen auszuhandeln, zu etablieren und ihre Einhaltung zu gewährleisten bzw. deren Verstoß zu sanktionieren (Kessler & Fritsche, 2018). Als soziale Normen werden Regeln und Standards ohne gesetzlich bindenden Charakter bezeichnet, welche das eigene Verhalten in einem realen, vorgestellten oder indirekten

Gruppenkontext lenken oder beschränken (Allport, 1985; Cialdini & Trost, 1998). Welchen Einfluss haben soziale Normen hinsichtlich der Bereitschaft Hate Speech Äußerungen zu tätigen? Die Befunde zum Einfluss von sozialen Normen auf Hate Speech im Onlinekontext sind widersprüchlich: So bilden sich soziale Normen in einem schwach regulierten Umfeld, wie in den Unterforen, sogenannten *Subreddits*, der User-Community *Reddit* aus. 2015 werden mehrere *Subreddits* mit ausgeprägtem Hate Speech gesperrt, was dazu führt, dass einige User:innen ihre Nutzerkonten bei *Reddit* löschen oder in thematisch verwandte *Subreddits* abwandern (Chandrasekharan et al., 2017). Obwohl in diesen *Subreddits* beleidigende, aggressive Sprache und Hate Speech ebenfalls verbreitet sind, kann durch die neu hinzugekommenen User:innen aus den gesperrten *Subreddits* kein signifikanter Anstieg von Hate Speech verzeichnet werden (ebd., ähnliche Befunde auch in Kaigo & Watanabe, 2007). Die Bereitschaft, User:innen dazu zu bringen, von Hate Speech gegen andere abzusehen oder primäre Adressat:innen zu verteidigen, ist jedoch gering (Costello et al., 2017).

Betreiber von Hate Sites versorgen die eigenen Anhänger:innen und Sympathisant:innen mit Sichtweisen, die vor allem Zweifel und Zwietracht zwischen sozialen Gruppen säen sollen (K. M. Douglas et al., 2005; Gerstenfeld et al., 2003). Ihre Strategie ist langfristig angelegt, denn indem Outgroups als Bedrohung stilisiert werden, erreichen sie eine festere Gruppenkohäsion der Ingroup (K. M. Douglas et al., 2005).

In einer experimentellen Untersuchung zeigen Alvarez-Benjumea und Winter (2018), dass die Bereitschaft Hate Speech in eigene Online-Kommentare einfließen zu lassen, in Kommentarspalten ohne Hate Speech am niedrigsten ist. Lesen Versuchspersonen vorurteilsbeladene Kommentare, neigen sie stärker dazu, selbst solche Kommentare zu verfassen (Hsueh et al., 2015). Die Autoren gehen davon aus, dass die soziale Umgebung als Hinweisgeber für die Akzeptanz solcher Aussagen wahrgenommen wird. Unhöfliche Kommentare (im Sinne eines groben Sprachstils) rufen bei Rezipierenden stärkere Reaktionen hervor, als inzivil (also demokratische Grundwerte in Frage stellende) Kommentare<sup>14</sup>, die in höflicher Weise vorgetragen werden (Kalch & Naab, 2017).

Primäre Adressat:innen von Hate Speech zeigen die Tendenz, sich von der eigenen Ingroup zu distanzieren und sich stärker an den sozialen Normen der Dominanzkultur zu orientieren, um weiterem Hate Speech oder einer Diskriminierung zu entgehen (Gelber & McNamara,

---

<sup>14</sup> Es gibt konzeptuelle Überschneidungen zu Hate Speech und Inzivilität (Ziegele, Köhler & Weber, 2018), wenn auch letzterer nicht eindeutig definiert ist (Coe et al. 2014). Papacharissi (2004) definiert Inzivilität als „negative collective face; that is, disrespect for the collective traditions of democracy. Civility [sic!] can then be operationalized as the set of behaviors that threaten democracy, deny people their personal freedoms, and stereotype social groups“ (Papacharissi, 2004, S. 267).

2015). Darüber hinaus internalisieren Minderheitengruppen die diskriminierenden Vorurteile und Stereotypen und zeigen auffällig viele Stresssymptome, die Meyer (2003) als Minderheitenstress bezeichnet.

Als *intergrupale Interaktionen* werden Interaktionen bezeichnet, welche zwischen verschiedenen Gruppen bzw. zwischen den Mitgliedern verschiedener Gruppen ablaufen. Blee (2004) zufolge entsteht und entfaltet sich (rassistisch motivierter) Hass in einem sozialen Gefüge. So können Konflikte zwischen Volksgruppen rasch in Gewalt münden, welche durch Hate Speech zusätzlich befeuert werden<sup>15</sup>. Exemplarisch zu nennen sind die Hate Radios in den gewaltsamen Auseinandersetzungen nach der Präsidentschaftswahl in Kenia 2007 (Somerville, 2011), während des Völkermordes 1994 in Ruanda (Kellow & Steeves, 1998; Yanagizawa-Drott, 2014), im Kongo (Vollhardt et al., 2006/2007), aber auch vom Fernsehen verbreiteter Hass und Propaganda während der Jugoslawienkriege in den 1990er Jahren (Bugarski, 2001; Oberschall, 2000). Mit der Verbreitung digitaler Technologien verlagern sich Intergruppenkonflikte zwischen Angehörigen unterschiedlicher Ethnien im kenianischen Parlamentswahlkampf 2013 auch in den digitalen Raum (Kimocho & Nyaga, 2016).

Eng verbunden mit der Erforschung von Hate Speech zwischen sozialen Gruppen ist das Konzept der *Dehumanisierung*, also der Abwertung und Aberkennung spezifischer menschlicher Eigenschaften, Emotionen oder moralischer Werte und Grundüberzeugungen gegenüber Outgroups (Bandura et al., 1975; N. Haslam, 2006; N. Haslam & Stratemeyer, 2016; Kteily & Bruneau, 2017; Leyens et al., 2000). Dieser Forschungszweig nimmt zwar nicht explizit Bezug auf Hate Speech, aber das Konzept der „*blatant dehumanization*“ im Vergleich zu ihrer subtilen, alltäglichen Variante, hat starke konzeptuelle Verbindungen zu Hate Speech: „*blatant dehumanization involves openly held beliefs about the inherent inferiority of other groups relative to the ingroup*“ (Kteily et al., 2015, S. 4). Im Gegenzug rekurriert die Hate Speech Forschung ausdrücklich auf Dehumanisierung als Mittel der Abwertung von Outgroups (s. o.). Dehumanisierung entsteht etwa durch Mediendarstellungen von Migrant:innen als Übertragende ansteckender Krankheiten, als Terrorverdächtige oder als Schwindler ohne echte Schutzbedürftigkeit (Esses et al., 2013).

Diskutiert werden die langfristigen Folgen von Hate Speech, welche gesellschaftliche Gruppen entzweien können, da es ein Klima der Feindseligkeit, Angst und des Misstrauens begünstigt (Leets & Bowers, 1999; Tsesis, 2001). Soral und Kollegen zeigen, dass die wiederholte Rezeption von Hate Speech zu einer Desensibilisierung gegenüber solchen Äußerungen und

---

<sup>15</sup> Die Nationalsozialisten greifen bei der Judenverfolgung ebenfalls zu Hate Speech, um die judeo-feindlichen Einstellungen in der Bevölkerung zu befeuern, wie Friesel zeigt (2013).

darüber hinaus dazu führt, dass Vorurteile gegenüber muslimischen oder homosexuellen Minderheiten bei polnischen Versuchspersonen verstärkt werden (Soral et al., 2018). Wer selbst von Hate Speech (etwa, weil er oder sie einer Minderheit angehört) betroffen ist, stuft Hate Speech eher als soziale Gefahr ein (Leets, 2001).

Zusammenfassend zeigt sich eine differenzierte Auseinandersetzung mit Wirkungen von Hate Speech. Ausgehend von einer handlungstheoretischen Perspektive auf Hate Speech-Wirkungen, zeugt die sozialpsychologische Perspektive von ergiebigen Anknüpfungspunkten für Forschung in diesem Bereich. Untersucht werden sowohl individuelle als auch kollektive Einflüsse mit einem Schwerpunkt hinsichtlich der Intergruppenebene und einer Vernachlässigung der interindividuellen Ebene, was sich auf die Anbindung an die Vorurteilsforschung und deren Fokus auf intergruppalen Zusammenhängen zurückführen lässt. Nachweisbar verfestigt Hate Speech Vorurteile und Stereotype zwischen sozialen Gruppen und extreme Formen von Intergruppenkonflikten können in offene Gewalt münden, welche durch Hate Speech noch befeuert werden. Innerhalb einer Gruppe sind wahrgenommene soziale Normen bedeutsam und entscheiden darüber, ob jemand zu Hate Speech greift, um Positionskämpfe auszufechten. Hate Speech stellt sich als gruppenbezogener Abwertungsmechanismus mit individuellen Effekten auf Psyche, Kognitionen und Einstellungen primärer und sekundärer Adressatengruppen dar.

### **2.2 Zwischenfazit: Hate Speech als gruppenbezogener Abwertungsmechanismus**

In den vorigen Ausführungen wurden drei Kernmerkmale identifiziert und kommunikationswissenschaftlich eingeordnet, namentlich (1) die Adressatengruppen, (2) die kommunikativen Inhalte und schließlich (3) die kommunikative Wirkung von Hate Speech. Vorgegangen war eine erste These, welche als grundlegende Prämisse zu begreifen ist, um Hate Speech nicht ausschließlich als Online-Phänomen misszuverstehen. Maßgeblich dafür sind nicht nur die terminologischen Wurzeln außerhalb des Online-Kontexts, sondern gerade der Variationsreichtum an Erscheinungsformen, welcher von einem definitiv eingegrenzten Fokus nicht abgedeckt würde. Eine breit gefasste konzeptionelle Prämisse erweist sich als epistemologisch vorteilhaft, weil auf diese Weise verallgemeinerbare Aussagen über Hate Speech getroffen werden können. Erst in der empirischen Untersuchung wird ein einzelner interessierender Aspekt (z. B. Hate Speech auf Facebook gegen Angehörige der Rohingya-Minderheit in Myanmar) identifiziert, ausgewählt und mit geeigneten Methoden geprüft.

Die Adressatengruppen, als erstes konstituierendes Kernmerkmal von Hate Speech, lassen sich in primäre und sekundäre Adressatengruppen unterscheiden. Bei den primären Adressatengruppen handelt es sich um Outgroups im Sinne gesellschaftlich marginalisierter Gruppen, deren Position im sozialen Hierarchiegefüge niedrig ist und die mittels Hate Speech angegriffen werden. Damit werden sie zur metaphorischen Zielscheibe für Hate Speech, was in Abbildung 4 verdeutlicht wird. Die sekundären Adressatengruppen sind Teil der Ingroup und allgemein gesprochen das Publikum, welches sich in Sympathisant:innen und Anhänger:innen der Hate Speech-Produzierenden sowie den Unbeteiligten Dritten einteilen lässt. Das zweite Kernmerkmal wird als kommunikativer Inhalt bezeichnet. Dabei folgt diese Arbeit einem weiten Verständnis, bei dem sämtliche Kommunikationsmodi (sprachliche wie nicht-sprachliche Hate Speech-Äußerungen) berücksichtigt werden, die gruppenbasierte Abwertung ausdrücken. Wie gezeigt wird, transportieren nicht nur schriftliche oder mündliche Aussagen Hate Speech, sondern auch eine ganze Bandbreite körpersprachlicher und figurativer Ausdrücke.

Hate Speech-Wirkungen werden als drittes Kernmerkmal identifiziert, und lassen sich aus dem Verständnis von Kommunikation als sozialem Handeln als intendiert auffassen, denn Hate Speech wird nicht „aus Versehen“ geäußert. Wirkungen bei primären und sekundären Adressatengruppen sind nicht als mechanistisch zu begreifen, und zwar dergestalt, dass Hate Speech eine bestimmte Wirkung (z. B. Diskriminierung) zur Folge hat. Vielmehr bedarf es einer integrativen Betrachtung, wie Hornscheidts (2013) Anforderungskatalog für sprachliche Diskriminierung fordert. Die Annahme einer zugrundeliegenden Intention von Hate Speech lässt sich mit Burkarts (2002) Feststellung einer allgemeinen Intention begründen, welche sich aus kollektiv geteilten Bedeutungen speist. Die vorhandene Marginalisierung und Diskriminierung primärer Adressatengruppen verstetigt sich also durch Hate Speech in der breiten gesellschaftlichen Wahrnehmung, aber insbesondere hinsichtlich der Selbstwahrnehmung der primär Adressierten. Die artikulierten Stereotype, Abwertungen und gezielten Feindseligkeiten werden in ihr Selbstbild integriert und führen zu einer Beeinträchtigung des Selbstwertgefühls sowie des subjektiven psychischen Wohlbefindens.

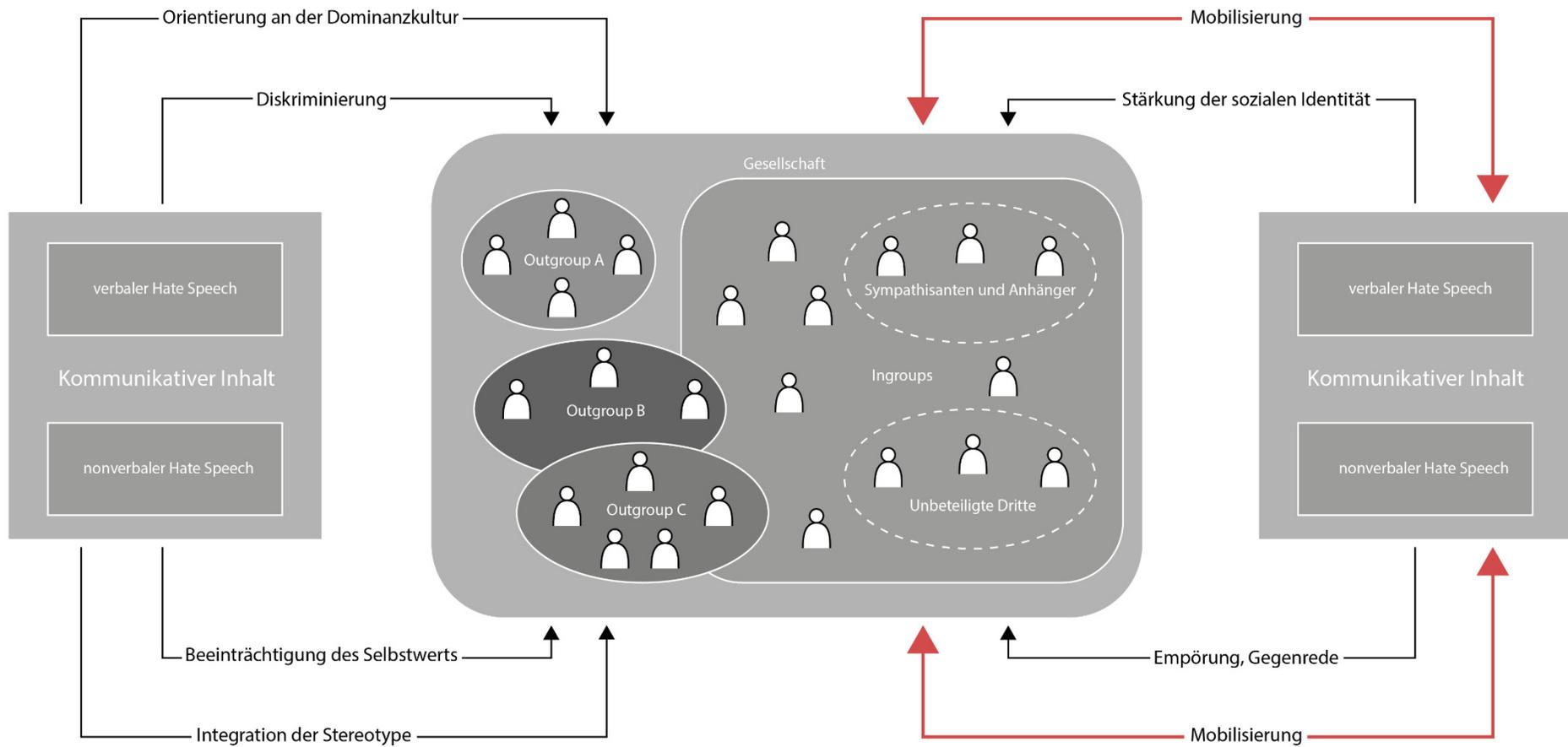
Bezogen auf die sekundären Adressatengruppen bewirkt Hate Speech bei Sympathisant:innen und Anhänger:innen eine Stärkung der sozialen Identität. Darüber hinaus dient es als Katalysator, um Stimmungen und die Artikulations- und Handlungsbereitschaft zu mobilisieren, d. h. die intendierte Wirkung für diesen Adressatenkreis besteht in der Mobilisierung, selbst Hate Speech zu produzieren, da sich Sprechende und gewogene Rezipienten als zusammengehörende soziale Gruppe begreifen. Die Empörungswirkung auf Seiten des

unbeteiligten Publikums wird von den Sprechenden billigend in Kauf genommen, wird gar bewusst provoziert, um die Reichweite der Sprechenden zu erhöhen und folglich Publizität zu erreichen. Eine mögliche Mobilisierungswirkung sekundärer Adressatengruppen ist zwar aus Sicht der Sprechenden eine erwünschte Wirkung, erfolgt aber allenfalls arbiträr. Abbildung 4 visualisiert die Zusammenhänge zwischen den drei Kernmerkmalen. Um die Komplexität des Modells möglichst gering zu halten, wurde auf zusätzliche mögliche Feedbackschleifen verzichtet. Ob die Annahmen, wie sie im Modell abgebildet sind, einer empirischen Überprüfung standhalten, soll im Rahmen dieser Arbeit geklärt werden.

Die drei Kernmerkmale liefern einen definitorischen kommunikationswissenschaftlichen Rahmen für Hate Speech, der sich wie folgt äußert. *Unter Hate Speech werden ganz allgemein sämtliche Kommunikationsformen verstanden, welche dazu dienen, gesellschaftshierarchisch niedrige kollektive Gruppen und deren wahrgenommene Vertreter:innen zu verunglimpfen, herabzuwürdigen und zu diskriminieren, öffentliche Empörung zu bewirken sowie Anhänger:innen und Sympathisant:innen zu gewinnen und zu mobilisieren.* Aus dieser allgemeinen Definition lässt sich noch eine spezifische formulieren, welche den Wortlaut nur marginal verändert, indem sie den Online-Kontext berücksichtigt. *Hate Speech im Internet bezeichnet sämtliche Kommunikationsformen, die in der öffentlichen Sphäre des Internets dargeboten werden, welche dazu dienen, gesellschaftshierarchisch niedrige kollektive Gruppen und deren wahrgenommene Repräsentant:innen zu verunglimpfen, herabzuwürdigen und zu diskriminieren, öffentliche Empörung zu bewirken sowie Anhänger:innen und Sympathisant:innen zu gewinnen und zu mobilisieren.* Da beide Definitionen jeweils die drei Kernmerkmale integrieren und sich nur der Kommunikationsraum ändert, wird im Lauf dieser Arbeit auf eine Unterscheidung von Hate Speech im allgemeinen Sinne und Hate Speech im Internet verzichtet. Hate Speech beinhaltet entsprechend sämtliche Kommunikationsformen, die der jeweiligen Kommunikationsumgebung zur Verfügung stehen. Wird eine bestimmte Kommunikationsumgebung explizit angesprochen und erscheint der Kontext bedeutsam, wird dies kenntlich gemacht, wie z. B. die Untersuchung von Hate Speech in E-Mails und Briefen an den Zentralrat der Juden in Deutschland und die Israelische Botschaft in Deutschland (Schwarz-Friesel, 2013).

Auf der Basis der diskutierten Kernmerkmale wird in den Unterabschnitten von Kapitel 2.3 Hate Speech von anderen verwandten Phänomenen im Online-Bereich abgegrenzt. Es zeigt sich nämlich, dass terminologische Konfusion herrscht, wenn es um dysfunktionale Formen der Online-Kommunikation geht. Dies ist insbesondere vor dem Hintergrund des Forschungs-

Abbildung 4: Übersicht über den Zusammenhang zwischen den drei Kernmerkmalen von Hate Speech



(Eigene Darstellung)

interesses dieser Arbeit relevant, die sich der Untersuchung von Rezeptions- und Persuasionseffekten von Hate Speech in SNS-Kommentaren widmet.

## 2.3 Begriffliche Abgrenzung

Nicht jede Obszönität oder Beleidigung fällt unter den Terminus Hate Speech, wie im Folgenden gezeigt wird. Betrachtet werden Flaming, Trolling und Cybermobbing, da diese konzeptionell nah am Hate Speech-Begriff verortet sind.<sup>16</sup> Während Flaming und Trolling in der wissenschaftlichen Auseinandersetzung häufig ausschließlich als Online-Phänomene diskutiert werden (Lea et al., 1992; Orengo Castellá et al., 2000; Siegel et al., 1986), gibt es sowohl für Hate Speech als auch Cybermobbing Offline Entsprechungen. In den nächsten Abschnitten wird nur das Internet als Kommunikationsraum berücksichtigt, so dass der Erörterungsfokus auf die eingeführten und diskutierten Kernmerkmale bzw. deren spezifische Ausprägung fällt mit dem Ziel einer eindeutigen begrifflichen Abgrenzung.

### 2.3.1 Flaming

Vielfach wird Flaming als Ausdruck verbaler Aggression im Online-Kontext definiert<sup>17</sup>. Charakteristisch sind demnach der Gebrauch obszöner Beschimpfungen und Beleidigungen, Einschüchterungen oder Drohungen, welche mittels exzessiver Interpunktion und weiteren typografischen Gestaltungsmitteln (durchgehende Großschreibung, farbliche Formatierungen, Dingbats, also grafische Symbole etc.) die wahrgenommene Aggressivität intensivieren (Lapidot-Lefler & Barak, 2012; Turnage, 2008). Im Wesentlichen besteht der *Kommunikationsinhalt* aus dem verfügbaren Kommunikationsrepertoire computervermittelter Kommunikation (cvK). So beschränken sich auch die meisten Definitionen darauf diesen Kommunikationsinhalt – als Flame bezeichnet – mehr oder minder erschöpfend aufzuzählen (Thompson, 1993).

flaming refers to a message sender's hostile emotional expressions characterized by using insulting, profane, or offensive languages (D. Cho & Kwon, 2015, S. 364)

hostility by insulting, swearing or using otherwise offensive language (Moor et al., 2010, S. 1537)

---

<sup>16</sup> Auf die Erörterung von Shitstorms, in der wissenschaftlichen Auseinandersetzung auch als negatives elektronisches Word-of-Mouth bezeichnet (Folger & Röttger, 2015), wird in dieser Arbeit verzichtet. Diese Entrüstungstürme sind reaktiv auf ein tatsächliches oder vermeintliches moralisches Fehlverhalten im Sinne einer *Moral Panic* gegenüber Unternehmen, Prominente und andere Personen gemünzt. Da Reaktivität nicht kennzeichnend für Hate Speech ist, kann diese Abgrenzung verkürzt erfolgen. Tiefergehend setzen sich beispielsweise Salzborn (2017) oder Johnen, Jungblut und Ziegele (2018) mit der Thematik auseinander.

<sup>17</sup> Neben Hate Speech, Trolling und Cybermobbing wird Flaming auch im Kontext von Cyberkriminalität diskutiert, was in dieser Arbeit jedoch unberücksichtigt bleiben soll. Hier wird auf entsprechende Arbeiten verwiesen, z. B. Jewkes (2005).

‘flaming’ refers to the expression of strong and inflammatory opinions to others electronically (Siegel et al., 1986, S. 161, Herv. i. O.)

Während der Kommunikationsinhalt extensiv erörtert wird, klammern gängige Definitionen in der Forschungsliteratur die Benennung möglicher *Adressatengruppen* vollständig aus<sup>18</sup>. Das bedeutet jedoch nicht, dass Adressierte in der theoretischen Auseinandersetzung und empirischen Umsetzung unberücksichtigt bleiben. Die zugrundeliegende Annahme besteht darin, dass Flaming im Gruppenkontext auftreten kann. Dies geschieht jedoch nicht übergreifend in allen Gruppen, sondern etabliert sich als Kommunikationskultur in bestimmten Gruppen (Kayany, 1998; Postmes et al., 2000) und ist abhängig von Kontextfaktoren (Hutchens et al., 2014; McKee, 2002; O'Sullivan & Flanagin, 2003)<sup>19</sup>. Damit widersprechen sie frühen technologiedeterministischen Ansätzen von Sproull, Kiesler und Kollegen, die davon ausgehen, dass Flaming in allen cvK-Kontexten auftritt (Kiesler & Sproull, 1992; Siegel et al., 1986; Sproull & Kiesler, 1986). Kayany (1998) unterscheidet konzeptionell zwischen Flaming und Venting, je nachdem ob Mitglieder der Ingroup (Flaming) oder Mitglieder der Outgroup (Venting) von Flaming betroffen sind. In seiner Untersuchung scheitert er aber aus methodischen Gründen am Nachweis eines Unterschieds zwischen Ingroup und Outgroup (Kayany, 1998). Eine differenzierte Betrachtung der Adressatengruppen wurde für Flaming bislang nicht beschrieben, lässt sich aber konzeptionell skizzieren: Generell richtet sich Flaming gegen Gruppenmitglieder, die von den eigenen Meinungen und Überzeugungen abweichen (Hutchens et al., 2014; McKee, 2002), was diese zu primären Adressaten macht. Als sekundäre Adressat:innen sind in diesem Sinne ebenfalls Sympathisierende des Flamers sowie unbeteiligte Dritte zu nennen. O'Sullivan und Flanagin (2003) unterscheiden zwischen Verfasser:innen und Adressat:innen von Flaming sowie unbeteiligten Dritten ohne jedoch diese näher zu charakterisieren.

In der Forschung wird die *kommunikative Wirkung* von Flaming zumeist als problematisch dargestellt, wenn auch positive Folgeeffekte diskutiert werden. So ist Flaming vielfach ein Mittel, um eine Gruppenidentität zu entwickeln und zu stärken (Lea et al., 1992; Vrooman,

---

<sup>18</sup> Auch Jenny Arendholz (2013) verzichtet auf die Benennung von Adressat:innen. Sie unterstreicht aber die Qualität von Flaming als persönliche Verbalattacken: „my reading of flaming is restricted to *ad hominem* attacks only. Whether repetitious punctuation, constant capitalization, swearing and other expressive and uninhibited verbal means count as flaming, needs to be decided by interlocutors, making flaming another emergent and evaluative notion” (Arendholz, 2013, S. 100, Herv. i. O.).

<sup>19</sup> So unterscheiden Hutchens und Kollegen (2014) situative und individuelle Faktoren. Zu den situativen Faktoren zählen sie die Anonymität von cvK sowie die reaktive Aggressivität, also aggressive Reaktionen auf ein provokatives Verhalten. Eng damit verbunden ist die verbale Aggressivität als individueller Faktor, also die Tendenz von Individuen das Selbstkonzept anderer zu attackieren (ebd.). Als weitere individuelle Faktoren nennen die Autoren u. a. die Nutzung von Onlinemedien sowie die Akzeptanz von Flaming (Hutchens et al., 2014).

2002). Dies zeigen Postmes, Spears und Lea (2000) in ihrer Untersuchung von studentischen Kurzmails im Rahmen eines Online-Kurses. Die Autoren weisen Gruppenbildungsprozesse nach, welche sich als Unterschiede in den Kommunikationsnormen zwischen den Gruppen etablieren. Diese Kommunikationsnormen beinhalten Flaming mit zumeist eindeutigen Hinweisen auf Humor oder Ironie, was Postmes und seine Kollegen als Akzeptanz solchen sozialen Verhaltens interpretieren: „if flaming is normative to interactants themselves, then an exchange of insults could create a bond rather than a conflict. Flaming could thus be a defining aspect of group identity or a characteristic of the group“ (Postmes et al., 2000, S. 357). Was die problematischen Wirkungen im Zusammenhang mit Flaming betrifft, konzeptualisieren Friedman und Currall (2003) die Eskalation von Konflikten im E-Mail-Kontext im Zeitverlauf, was jedoch entgegen ihrer technologiedeterministischen Annahme nicht ausschließlich mit technischen Merkmalen als vielmehr mit sozialen, sprachlichen und emotionalen Faktoren zusammenhängt (Brett et al., 2007; Friedman et al., 2004; N. A. Johnson et al., 2009). Darüber hinaus löst Flaming bei manchen Menschen eine Beeinträchtigung des psychischen Wohlbefindens aus, wozu Angstzustände oder Depressionen gehören (Baruch, 2005; Jane, 2015). Dies führt zu Silencing, d. h. zu kommunikativem Schweigen, was sich in gehemmten Online-Aktivitäten äußert, wie z. B. die Abkehr von Video-Uploads bei YouTube (Lange, 2007; Moor et al., 2010).

Die Abgrenzung zu Hate Speech kann auf Basis dieser knappen Darstellung folglich eindeutig erfolgen. So sind zwar Unterschiede im Kommunikationsinhalt von Hate Speech und Flaming vernachlässigbar, denn für beide gilt, dass sie ihr Gegenüber mittels semantischer und rhetorischer Stilmittel sowie nonverbaler Kommunikationsformen attackieren. Deutliches Abgrenzungspotential der beiden problematischen Kommunikationsphänomene muss dafür den Adressierten eingeräumt werden. Während Hate Speech sozialhierarchisch schwache, marginalisierte Gruppen als primär Adressierte attackiert, richtet sich Flaming gegen potenziell gleichrangige Gruppenmitglieder. Der Gruppenbegriff bleibt jedoch fluide; als Gruppe werden die Mitglieder in einem Kommunikationszusammenhang bezeichnet. Wenn es auch Forschungsbestrebungen gibt, entsprechende Adressatenkreise zu definieren, ist hier im Vergleich zur Hate Speech-Forschung eine mangelnde Zuwendung zu diesem Aspekt erkennbar.

In der Hate Speech-Forschung wird die feindselige Intention von Hate Speech a priori angenommen und bewirkt je nach Adressatengruppe Diskriminierung, Empörung, aber auch Mobilisierung sowie Stärkung der sozialen Identität. Eine eindeutig feindselige Intention wird für

Flaming in Frage gestellt<sup>20</sup> und die zugrunde liegende Intention wird bei den wenigen Arbeiten zu Flaming-Effekten ohnehin nicht berücksichtigt. Insgesamt auffällig ist die Abwesenheit von kommunikativen Wirkungen auf gesellschaftlicher Ebene, was sich auf die primären Adressaten zurückführen lässt. Mittels Flaming werden nicht Minderheiten attackiert, somit entfallen auf den ersten Blick Empörungsreaktionen. Flaming ist darüber hinaus weniger sichtbar, da es vor allem in nicht-öffentlichen Kontexten verbreitet ist. Wird jedoch Flaming in der medialen Berichterstattung thematisiert (z. B. als Exempel dysfunktionaler Online-Kommunikation), tritt in der Folge Empörung ein. Auf individueller und auf Gruppenebene werden in der Flaming-Literatur zu Hate Speech vergleichbare kommunikative Wirkungen diskutiert (also Beeinträchtigung des psychischen Wohlbefindens, Stärkung der Gruppenidentität etc.).

Im nächsten Kapitel folgt die begriffliche Abgrenzung von Trolling zu Hate Speech auf Basis der erarbeiteten Kernmerkmale. Auch hier zeigen sich konzeptionelle Übereinstimmungen aber auch deutliche Abweichungen zwischen den Phänomenen.

### 2.3.2 Trolling

Mit der mythologischen Gestalt des Trolls hat Trolling wenig gemein. Der Begriff ist der Anglersprache entlehnt und meint dort das Auswerfen eines Köders aus einem Boot heraus. Der Köder ist an einer Angelschnur befestigt und wird durch das Wasser gezogen, bis ein Fisch angebissen hat (Donath, 1999). Analog dazu fungiert Trolling im Kontext computervermittelter Kommunikation. Der *Kommunikationsinhalt* wird einem kommunikativen Köder gleichsam in Form von Schimpfworten und Beleidigungen, aber vor allem „meaningless, irrelevant, or repetitive posts aimed at attention-seeking or response-generating“ (Hardaker, 2010, S. 232) in Kommentarspalten oder Diskussionsforen ausgelegt. Diskussionsteilnehmende, die auf Äußerungen des Trolls eingehen, quasi den ‚Köder schlucken‘, geraten sodann in lange und nicht endende Diskussionen mit ihm.

---

<sup>20</sup> O'Sullivan und Flanagan (2003) modellieren einen Analyserahmen und berücksichtigen für die Wahrnehmung von Flaming ein Perspektivendreieck aus Sprechenden, primären Adressat:innen und unbeteiligten Dritten. Demnach kann man nur dann von echtem Flaming sprechen, wenn kommunizierte Inhalte und Interaktionen vom Sprechenden als Flaming intendiert sind und gleichzeitig sowohl von primären Adressat:innen als auch unbeteiligten Dritten als intendiert wahrgenommen werden (O'Sullivan & Flanagan, 2003). „The *process* of ‚flaming‘ thus includes the creation, transmission, and interpretation of a message that is perceived from multiple perspectives as violating norms, and which can itself be the beginning of the norm negotiation, evolution, and realignment processes.“ (O'Sullivan & Flanagan, 2003, S. 85, Herv. i. O.). Werden kommunizierte Inhalte oder Interaktionen aus mindestens einer der drei Perspektiven nicht als Normverletzung wahrgenommen, handelt es sich O'Sullivan und Flanagan (2003) zufolge um Subformen der Fehlkommunikation (vgl. auch Moor et al. 2010).

In ihrer Untersuchung identifiziert Claire Hardaker (2013) sechs Strategien mittels derer Trolling gelingt, welche in der Praxis häufig als Mischformen auftauchen. Alle Strategien zielen auf eine Störung des Diskussionsverlaufs ab, um ein sadistisches Bedürfnis des Trolls nach Belustigung zu befriedigen (Buckels et al., 2014; Buckels et al., 2019; Hardaker, 2010; Sest & March, 2017). Internet-Trolle verfolgen darüber hinaus politische Ziele und setzen Trolling koordiniert ein, um Desinformationskampagnen zu betreiben oder die öffentliche Meinung zu beeinflussen, was als Propaganda mit Trolling-Mitteln verstanden wird (Bradshaw & Howard, 2017; Quandt, 2018). Die Analyse von Trolling-Strategien verdeutlicht die Kontextabhängigkeit von Äußerungen, da ähnlich zu Flaming die Intention des Sprechenden unbekannt bleibt<sup>21</sup>.

Hardaker (2013) führt als erste Trolling-Strategie an, dass der Troll thematisch *abschweift* und damit den Diskussionsverlauf aus Sicht anderer Nutzenden stört. Herring und Kollegen (Herring et al., 2002) zeigen, wie ein solches Abschweifen durch einen männlichen Nutzer, Kent, in einem feministischen Internet-Forum große Verärgerung unter den zumeist weiblichen Nutzenden auslöst. Kent gibt sich als aufrichtig am Thema interessiert aus, provoziert aber die Nutzenden mit Beleidigungen, antagonistischen Positionen und versucht sie gegeneinander aufzubringen. Insgesamt acht Tage dauert der Streit, bis es den Nutzenden gelingt, Kent aus ihrem Forum zu bannen und neue Diskussionsregeln auszuhandeln.

Auch indem sich Trolle als eifrige und *kleinliche Kritiker* vernachlässigbarer Tipp-, Rechtschreib- oder Interpunktionsfehler zeigen (Hardaker, 2013), verstricken sie die anderen Nutzenden in sinnentleerte Diskussionen. Oder sie agieren ähnlich wie Kent in der Untersuchung von Herring und Kollegen (Herring et al., 2002) und säen Zwietracht (Hardaker, 2013 spricht von *Antipathie*), indem sie versuchen die Gruppe mit kontroversen oder pseudo-naiven Positionen zu spalten. Eine vierte Strategie von Trollen ist die *Gefährdung* der Diskussionsteilnehmenden, indem schlechte oder schädliche Ratschläge erteilt werden, häufig in pseudo-naiver Manier vorgetragen (auch in Donath, 1999). Aussagen, die die Diskussionsteilnehmenden *schockieren* sollen, sind laut Hardaker (2013) ebenfalls eine vielbenutzte Trolling-Strategie. Dazu werden provokante Haltungen zu heiklen, Tabu behafteten Themen (Tod, Religion, Sexualität) geäußert. Als letzte Strategie nennt Hardaker (2013) verbale *Aggressionen*, welche so unbegründet wie verletzend sind und eingesetzt werden, um das Gegenüber zu

---

<sup>21</sup> Die frühere Trolling-Forschung nähert sich inhaltsanalytisch an den Untersuchungsgegenstand, weswegen Äußerungen in Mailinglisten, Kommentarspalten und Diskussionsforen von Forscher:innen und Diskussionsteilnehmenden nur als Trolling *interpretiert* werden (Donath, 1999; Hardaker, 2010; 2013, Herring et al., 2002). Neuere Untersuchungen hingegen richten ihr Augenmerk auf den Kommunikator, also den Troll und seine Motive selbst (z. B. in Buckels et al., 2014; Quandt, 2018).

einer emotionalen Reaktion zu provozieren. Diese Strategie deckt sich mit Janes (2015) Verständnis von Trolling: “‘Trolling’ [...] is also used more specifically to describe the online posting of deliberately inflammatory or off-topic material with the aim of provoking textual responses and/or emotional reaction” (Jane, 2015, S. 66).

Die *Adressat:innen* von Trolling sind prinzipiell alle aktiven (Diskussionsteilnehmenden) wie passiven (stille Leser:innen, also Lurker) Nutzenden einer Online-Community, wobei vor allem unerfahrene, naive Nutzende den Köder des Internet-Trolls „schlucken“ (A. J. Kim, 2000). Durchschauen sie nicht den Trolling-Mechanismus, erwachsen daraus die oben erwähnten langwierigen und fruchtlosen Diskussionen zwischen aktiven Nutzenden und dem Troll. Mit Empörung und Verärgerung reagieren beide Gruppen, also die aktiven und die passiven Nutzenden. Abhängig von seiner Raffinesse durchschauen erfahrenere Nutzende den Troll und warnen vor weiteren Diskussionen mit den Worten „Don’t feed the troll“. Dass Trolle selbst zu Adressaten von Gegentrolling werden, ist ebenfalls vor allem erfahrenen Nutzenden zu verdanken. Coles und West (2016) setzen sich mit Nutzerkommentaren zur Studie von Buckels und seinen Kollegen (Buckels et al., 2014) auseinander und zeigen, dass insbesondere engagierte und erfahrene Nutzende Internet-Trolle mit ihren Handlungen mittels Gegentrolling konfrontieren. Kritisch anzumerken ist, dass es sich hierbei um Selbstzuschreibungen handelt und sich Nutzende möglicherweise als erfahrener und offensiver darstellen. Da jedoch Beispiele für Gegentrolling existieren (Golf-Papez & Veer, 2017), genügt die Feststellung an dieser Stelle.

Die *kommunikative Wirkung* von Trolling findet sich auf allen drei Analyseebenen wieder. Auf der Makroebene sind vor allem die Destabilisierung der öffentlichen Meinung zu nennen. Wie Gunitsky (2015) darlegt, setzen illiberale Staaten wie Russland oder China (Mounk, 2018) Social Media dazu ein, um die öffentliche Meinung nicht nur im eigenen Land sondern weltweit zu beeinflussen. Dazu werden „Trollarmeen“ eingesetzt, staatlich gelenkte Nutzende, die nach Vorgabe in Online-Diskussionen Zwist und Streit provozieren oder Desinformation säen (Bradshaw & Howard, 2017; Kreißel et al., 2018). Auch im Fall von nicht-politischem Trolling lösen die bereits oben erwähnten endlosen und absurden Diskussionen z. T. heftige Emotionen aus. Neben Empörung, Wut und Verärgerung handelt es sich aber auch um durchaus positive Emotionen (Freude, Vergnügen, Stolz) v. a. wenn es Diskussionsmitgliedern gelingt, den Troll zu enttarnen (Coles & West, 2016) oder wenn die Gruppennormen Trolling akzeptieren und dies zu Unterhaltungszwecken von den Diskussionsteilnehmenden genutzt wird (March & Marrington, 2019). Es sind bislang keine Untersuchungen bekannt, die sich

mit langfristigen Folgen von Trolling bei Rezipierenden auseinandersetzen, vergleichbar der Beeinträchtigung des psychischen Wohlbefindens aufgrund von Hate Speech.

Die Gegenüberstellung von Trolling und Hate Speech zeigt, dass auch hier die drei analytischen Kernmerkmale nützlich bei der Abgrenzung der Online-Phänomene sind. Trolling beansprucht neben aggressiven vor allem solche Kommunikationsinhalte für sich, die dazu geeignet sind, Diskussionsverläufe im Internet störend zu beeinflussen. Im Unterschied zu Hate Speech ist Trolling aber nicht allein auf der semantischen, sondern insbesondere auf der pragmatischen Ebene angesiedelt. Äußerungen von Internet-Trollen ergeben wörtlich genommen keinen Sinn (z. B. im Zusammenhang mit den Trolling-Strategien „kleinliches Kritisieren“, „Gefährden“ oder „Schockieren“) und gelten als mutwillige Störung von Diskussionsverläufen. Gelingt es den Kommunikationsteilnehmenden diese Äußerungen als Trolling zu identifizieren, entfaltet sich die nicht-wörtliche Bedeutung (Golf-Papez & Veer, 2017).

Auch für Trolling besteht ein eindeutiges Abgrenzungspotential hinsichtlich der Adressat:innen. Diese sind Diskussionsteilnehmende in einem bestimmten Kontext, z. B. innerhalb eines Diskussionsthreads in einem Diskussionsforum. In der Literatur wird v. a. die Unerfahrenheit von Diskutierenden als Adressatenmerkmal hervorgehoben. Es sind also nicht solche Merkmale, die eine soziale Gruppe charakterisieren und mehrheitlich unveränderbar sind, wie sie als Basis für Hate Speech dienen. Wer aufgrund seiner Unerfahrenheit auf einen Internet-Troll hereinfällt und sich von ihm in sinnlose Gespräche und Argumentationen verwickeln lässt, kann prinzipiell daraus lernen und solche Situationen vermeiden.

Die kommunikative Wirkung als Abgrenzungskriterium ist hingegen weniger aussagekräftig, denn analog zu Hate Speech entfaltet Trolling sowohl gesellschaftliche wie individuelle Wirkungen. Während jedoch Hate Speech auf der Makroebene Vorurteile und Stereotype gegen sozialhierarchisch niedrige Gruppierungen zementiert, sorgt Trolling für eine Destabilisierung der öffentlichen Meinung (Kreißel et al., 2018). Auf der Mikroebene nivellieren sich die Unterschiede zwischen Trolling und Hate Speech: Sichtbare Folgen sind eher kurzfristige Empörung und Verärgerung, aber auch eine gruppenbezogene Stärkung der sozialen Identität.

Ähnlich zu Flaming und Trolling agiert Cybermobbing ebenfalls in einem begrenzten sozialen Kontext. Wenn sich auch ein Großteil der Forschung dem Cybermobbing zwischen Kindern und Jugendlichen widmet, ist Cybermobbing auch unter Erwachsenen (z. B. im beruflichen Kontext) verbreitet. Wie unterscheidet sich der kommunikative Inhalt von Cybermobbing und Hate Speech? Welche Wirkungen werden in der Literatur diskutiert? Der Beantwortung dieser Fragen widmet sich der nächste Abschnitt.

### 2.3.3 Cybermobbing

Cybermobbing<sup>22</sup> wird häufig im schulischen Kontext untersucht (Festl, 2015; United Nations, 2016) und gilt als „peer harassment“ (Li, 2007, S. 1780). Dabei drangsalieren und schikanieren junge Menschen Gleichaltrige mittels digitaler Technologien: „aggression expressed by peers through digital (online or mobile) technologies targeting a child victim“ (Livingstone et al., 2016, S. 116). „Cyberbullying is any behavior performed through electronic or digital media by individuals or groups that repeatedly communicates hostile or aggressive messages intended to inflict harm or discomfort on others.“ (Tokunaga, 2010, S. 278). Manche Autoren gehen davon aus, dass sich Cybermobbing und traditionelles Mobbing nicht wesentlich voneinander unterscheiden (Kowalski & Limber, 2013), da das Attackenrepertoire offline durch Online-Handlungen ausgetauscht wird. Eine differenziertere Betrachtung zeigt jedoch, dass die Potenziale digitaler Technologien um ein Vielfaches schädlicher sind: Unter Ausnutzung der Anonymität im Internet können Täter ihre Opfer auch abseits des Schulhofs rund um die Uhr drangsalieren (Festl, 2015; Livingstone et al., 2016). Damit lösen sich nicht nur räumliche und zeitliche Rückzugsräume auf, auch der Zuschauerkreis erweitert sich durch einfach zu vervielfältigendes und zu distribuierendes Mobbingmaterial, z. B. in Form von Bildern, Filme (Kowalski et al., 2012).

Gängige Definitionen von Cybermobbing benennen nicht unbedingt explizit junge Menschen als Adressaten: „An aggressive, intentional act carried out by a group or individual, using electronic forms of contact, repeatedly and over time against a victim who cannot easily defend him or herself“ (P. K. Smith et al., 2008, S. 376). Tatsächlich untersucht jedoch ein Großteil der Forschung Cybermobbing bei Kindern und Jugendlichen (Baek & Bullock, 2014; G. Brewer & Kerlake, 2015; Festl, 2015; Hinduja & Patchin, 2019; Juvonen & Gross, 2008; Olweus, 2012; Tokunaga, 2010). Betroffen sind auch Studierende (Kowalski et al., 2016; Savage & Tokunaga, 2017), z. T. solche mit gesundheitlichen Einschränkungen (Kowalski et al., 2012) und sogar Berufstätige am Arbeitsplatz (Farley et al., 2015; Gardner et al., 2016; Privitera & Campbell, 2009), wenn es auch in diesem Bereich nicht annähernd so viel Forschung wie bei Cybermobbing von Kindern und Jugendlichen gibt. Aus diesem Grund beziehen sich die weiteren Ausführungen hauptsächlich auf die jüngere Altersgruppe.

Festl, Scharnow und Quandt (2015) betonen, dass Cybermobbing in einem sozialen Gefüge stattfindet, in dem Täter und Opfer einander aus der Offlinewelt kennen. Besonders

---

<sup>22</sup> Der Begriff Cyberbullying wird in der internationalen Forschung sowie im angelsächsischen Sprachraum verwendet. Im deutschen Sprachraum hat sich dagegen der Begriff Cybermobbing durchgesetzt (Livingstone et al., 2016; Nocentini et al., 2010). In dieser Arbeit werden die beiden Begriffe synonym verwendet.

persistent ist Cybermobbing dann, wenn gruppeninterne Normen Cybermobbing tolerieren (Festl et al., 2015; Festl, 2016), wenn also im Klassenverband ein Mobbing-Klima vorherrscht. Dabei nutzen die Täter:innen die Anonymität des Internets aus, so dass den Opfern von Cybermobbing nicht oder nur selten bekannt ist, welche:r Mitschüler:in tatsächlich hinter den Angriffen steckt (Barlett, 2015). Generell teilen Adressat:innen und Täter:innen eine Reihe von Merkmalen: So wurden Jugendliche, die andere im Internet mobben, in der Vergangenheit selbst gemobbt (Li, 2007). Beide Gruppen haben ein geringes Selbstwertgefühl (G. Brewer & Kerlake, 2015) und leiden unter sozialen Ängsten (Kowalski et al., 2012), was durch das Cybermobbing noch verstärkt wird. Auch wenn der Geschlechtseffekt nicht eindeutig ist (Barlett & Coyne, 2014; A. Sorrentino et al., 2019), sprechen Befunde dafür, dass eher Mädchen Opfer von Cybermobbing sind (Kowalski et al., 2012, diskutieren die Thematik eingehender). Auch die Mediennutzung hat einen Effekt auf die Häufigkeit von Cybermobbing: Opfer nutzen Internettechnologien häufiger und intensiver (Festl et al., 2015; Juvonen & Gross, 2008; Ybarra & Mitchell, 2004).

Dabei wird die ganze Bandbreite an digitalen Technologien genutzt, etwa geschriebener Text, gesprochene Sprache, (manipulierte) Bilder und Video (Kowalski et al., 2012; Slonje & Smith, 2008; P. K. Smith et al., 2006). "Cyberbullying also takes various forms and electronic communication tools – from email, listserve, cell phone, to websites" (Li, 2007, S. 1779). In einer Studie vergleichen die Autoren sieben verschiedene Kommunikationskanäle und stellen fest, dass Cybermobbing per Telefon oder Instant Messaging am häufigsten ist (Slonje & Smith, 2008). Befunde zur Nutzung bestimmter Plattformen spiegeln deren Popularität zum jeweiligen Zeitpunkt wider (z. B. MySpace in den frühen 2000er Jahren, TikTok seit 2016), und so widmen sich aktuelle Forschungsarbeiten Cybermobbing in jenen SNS zu, die von Jugendlichen verstärkt genutzt werden (Barlett et al., 2018).

Gerade weil sich Cybermobbing im sozialen Umfeld des Opfers vollzieht, wird eine starke Beeinträchtigung des Wohlbefindens dokumentiert (G. Brewer & Kerlake, 2015). Dies geht mit Verhaltensänderungen bzw. Vermeidungshandlungen einher: Schulleistungen verschlechtern sich, was auf das Fernbleiben vom Unterricht zurückzuführen ist (Kowalski & Limber, 2013). Körperliche Auswirkungen sind v. a. bei jüngeren Kindern nachgewiesen worden, wozu Kopfschmerzen, Appetitverlust oder Bettnässen zählen (Fekkes et al., 2006; Kowalski & Limber, 2013). Das psychische Wohlbefinden verschlechtert sich zusätzlich, und zieht ein verringertes Selbstwertgefühl, schwere Depressionen bis hin zu häufigen Selbstmordgedanken nach sich (Hinduja & Patchin, 2010, 2019; Kowalski et al., 2012; Olweus, 2012; Patchin & Hinduja, 2010; Tokunaga, 2010). Sabella, Patchin und Hinduja (2013) mahnen an, diese

korrelationale Befunde keineswegs kausal zu interpretieren. So werden in der Literatur Depressionen und ein geringes Selbstwertgefühl sowohl als Ursache als auch als Wirkung von Cybermobbing diskutiert (Fekkes et al., 2006), welche gleichermaßen Täter wie Opfer betreffen (G. Brewer & Kerlake, 2015). Und schließlich bewirkt Cybermobbing (wie traditionelles Mobbing auch), dass Opfer zu Tätern werden und selbst Gleichaltrige drangsalieren und schikanieren (Li, 2007). Die Wirkung auf Peers, etwa im Klassenverband, ist geprägt von der Erleichterung, nicht selbst Opfer zu sein oder Genugtuung darüber, dass das Opfer drangsaliert wird, denn „schließlich hat er/sie es verdient“ (Zych et al., 2019). Nur wenige Zuschauer werden aktiv und helfen dem Opfer (Koehler & Weber, 2018).

Was die Adressat:innen betrifft, unterscheiden sich Cybermobbing und Hate Speech in einigen Aspekten voneinander. Häufig befinden sich sowohl Täter:innen als auch Opfer von Cybermobbing in einem bestimmten sozialen Kontext und kennen einander aus der Offlinewelt (Olweus & Limber, 2018). Sie teilen bestimmte soziodemographische Merkmale, insbesondere das Alter bzw. den Jahrgang (Schule, Studium) oder das Geschlecht. Innerhalb ihres Peer-Netzwerks besteht ein Machtungleichgewicht zuungunsten der Opfer (Olweus, 1993), was zusätzlich davon befeuert wird, dass Cyberbullying-Adressat:innen Angehörige sozialhierarchisch niedriger Gruppen sein können, z. B. Menschen mit Behinderungen (Didden et al., 2009), ethnische Minderheiten (Llorent et al., 2016). Ob explizit Hate Speech im Rahmen von Cyberbullying eine Rolle spielt, ist jedoch nach Kenntnis der Autorin bislang nicht erforscht.

Die kommunikativen Inhalte ähneln einander, so werden beim Cybermobbing wie beim Hate Speech sämtliche verfügbaren Kommunikationsformen genutzt. Während jedoch Hate Speech primär auf Öffentlichkeit angelegt ist, vollziehen sich Cybermobbing-Attacken sowohl öffentlich als auch im Verborgenen. Beispiele für verborgenes Cybermobbing sind das Eindringen in E-Mail- oder Social Media-Konten mit dem Ziel, Konversationen und Interaktionen zu manipulieren oder Informationen zur Bloßstellung des Opfers zu gewinnen. So können (manipulierte) Bilder oder Videos den Radius der Peer-Gruppe verlassen und sich rasend schnell verbreiten. Der Öffentlichkeitsgrad der Cybermobbing-Handlungen wird um eine zeitliche Dimension ergänzt, die für Hate Speech in der bisherigen Forschung nicht diskutiert wird. Vor allem Tokunaga (2010) plädiert dafür, Cybermobbing als langfristig angelegte Aggressionen zwischen Opfern und Täter:innen zu definieren.

Die Cybermobbing-Forschung ist von einer medienpsychologischen, -pädagogischen und damit individualistischen Herangehensweise geprägt, auch wenn neuere Arbeiten den sozialen Kontext berücksichtigen (Festl, 2015). Ein Großteil der Untersuchungen fokussiert die Auswirkung auf das psychische Wohlbefinden der Opfer und die Motivation der unbeteiligten

Zuschauer helfend einzugreifen. Unerforscht bleiben bislang gesellschaftliche Wirkungen von Cyberbullying. In dieser Hinsicht ist die Hate Speech-Forschung differenzierter in der Bearbeitung von Fragestellungen, die die Folgen von Hate Speech für Individuen, Gruppen und Gesellschaft nachzeichnen.

Im nächsten Abschnitt werden die Erkenntnisse aus den vorherigen Unterkapiteln zusammengetragen, um die begriffliche Abgrenzung in einem Analyseraster zu veranschaulichen, was der Eingrenzung des Phänomens Hate Speech dient.

### 2.3.4 Resümee

Die begriffliche Abgrenzung von Hate Speech und verwandten Online-Phänomenen dysfunktionaler Kommunikation anhand der drei erarbeiteten Kernmerkmale erweist sich als fruchtbar, wie die Übersicht in Tabelle 2 zeigt. Als eindeutiges Abgrenzungsmerkmal kristallisieren sich insbesondere die *primären Adressat:innen* heraus. Dabei erweisen sich Hate Speech und Cybermobbing als konzeptionell nächstehend, da die hierarchische Positionierung innerhalb der Gesellschaft (Hate Speech) bzw. der Peer-Group (Cybermobbing) die Wahrscheinlichkeit beeinflusst, Opfer solcher Attacken zu werden. Kinder und Jugendliche werden häufig Opfer von Cybermobbing mit Hate Speech-Charakter, beispielsweise Angehörige von ethnischen oder religiösen Minderheiten (Pells et al., 2016). Primäre Adressat:innen von Flaming und Trolling sind dagegen Onlinenutzende, die entweder naiv und unerfahren (Trolling) sind oder die im Diskussionverlauf abweichende Ansichten (Flaming) vertreten. Flaming und Trolling kann im weiteren Diskussionsverlauf durchaus Verbalattacken (Beschimpfungen, auch Drohungen) enthalten, welche jedoch nur dann als Hate Speech verstanden werden, wenn sie sich gegen gesellschaftlich marginalisierte Gruppen richten. Als *sekundäre Adressat:innen* werden für alle vier betrachteten Kommunikationsphänomene unbeteiligte Dritte, Sympathisant:innen und Anhänger:innen festgelegt. Das Potenzial sekundäre Adressat:innen zu aktivieren und/oder zu mobilisieren, ist abhängig vom hierarchischen Status der primären Adressat:innen. Da sich sekundäre Adressat:innen nicht grundlegend unterscheiden und ihnen über alle betrachteten Kommunikationsphänomene hinweg die gleiche Funktion zukommt, spielen diese bei der begrifflichen Abgrenzung eine nachgeordnete Rolle.

Die Relevanz des *kommunikativen Inhalts* ist ungleich größer. Wenn auch übergreifend über alle Kommunikationsphänomene Beschimpfungen und Beleidigungen zentral und Unterschiede allenfalls graduell sind, ist die Abwertung und Verunglimpfung von primären Adressat:innen im Fall von Hate Speech ein deutlich hervorstechendes Merkmal. Eng damit verbunden ist die *kommunikative Wirkung* von Hate Speech welche sich auf gesellschaftlicher

Ebene als diskriminierende Praktiken und Stereotype gegenüber der adressierten Gruppe äußert. Auch gesellschaftlich destabilisierende Wirkungen von Hate Speech werden aktuell diskutiert (Pérez-Escobar & Noguera-Vivo, 2021), wie sie bereits für politisches Trolling dokumentiert wurden (Aro, 2016). Eine Makro-Wirkung für Flaming und Cybermobbing erfolgt medial vermittelt und äußert sich in Empörung der Mehrheitsgesellschaft gegenüber diesen Phänomenen. Wirkungen auf Gruppen- und Individualebene ähneln sich hingegen und sind für die Gruppenkohäsion bzw. das beeinträchtigte psychische Wohlbefinden der Angegriffenen von Relevanz.

Im nächsten Abschnitt wird in einem Exkurs die Kommunikatorperspektive von Hate Speech eingehend diskutiert. Diese ist zwar nicht zentral für die definitorische Eingrenzung des Phänomens; für die kommunikationswissenschaftliche Ausrichtung der Arbeit dennoch nicht vernachlässigbar.

**Tabelle 2:** Gegenüberstellung von Hate Speech und verwandten Online-Phänomenen auf Basis der drei Kernmerkmale. Der graue Druck weist auf vermutete, jedoch nicht ausreichend belegte Annahmen

		<b>Hate Speech</b>	<b>Flaming</b>	<b>Trolling</b>	<b>Cybermobbing</b>
<b>Adressat:innen</b>	<b>Primär</b>	Marginalisierte niedrige Gruppen und ihre Vertreter:innen	Diskussionsteilnehmende in Online-Foren u. ä. mit abweichenden Meinungen	v. a. unerfahrene Diskussionsteilnehmende in Online-Foren	Gruppenhierarchisch niedrige Individuen (v. a. Kinder und Jugendliche in ihren Peer-Groups) in frequentierten Online-Kanälen
	<b>Sekundär</b>	Unbeteiligte Dritte, Sympathisant:innen u. Anhänger:innen	Unbeteiligte Dritte, Sympathisant:innen u. Anhänger:innen	Unbeteiligte Dritte, Sympathisant:innen u. Anhänger:innen	Unbeteiligte Dritte, Sympathisant:innen u. Anhänger:innen
<b>Kommunikativer Inhalt</b>		Alle Kommunikationsformen zur Abwertung und Verunglimpfung primärer Adressat:innen	Alle Kommunikationsformen zur Beschimpfung und Beleidigung primärer Adressat:innen	Alle Kommunikationsformen, mit dem Ziel Diskussionen und Interaktionen zu stören	Alle Kommunikationsformen zur Abwertung und Verleumdung primärer Adressat:innen
<b>Kommunikative Wirkung</b>	<b>Individuelle Ebene</b>	Beeinträchtigung des psychischen Wohlbefindens	Beeinträchtigung des psychischen Wohlbefindens	Beeinträchtigung des psychischen Wohlbefindens	Beeinträchtigung des körperlichen und psychischen Wohlbefindens
	<b>Gruppenebene</b>	Stärkung der sozialen Identität, Mobilisierung	Stärkung der sozialen Identität	Stärkung der sozialen Identität	Stärkung der sozialen Identität
	<b>Gesellschaftliche Ebene</b>	Empörung, Diskriminierung	Empörung	Destabilisierung der öffentlichen Meinung (politisches Trolling)	Empörung

(Quelle: Eigene Darstellung)

## 2.4 Exkurs: Befunde zur psychosozialen Persönlichkeit von Hate Speech-Produzierenden

Einschlägige Kommunikationstheorien heben die Bedeutung von Kommunikator:innen im Kommunikationsprozess hervor (eine Übersicht in Burkart, 2003), welche je nach Modell als Sender<sup>23</sup>, Sprecher, Kommunikator oder Kommunikationspartner bezeichnet werden. In den bisherigen definitorischen Ausführungen in dieser Arbeit ist der Blick auf Kommunikator:innen ausgeklammert worden, da nicht die Kommunikator:innenebene ausschlaggebend für das Verständnis des Hate Speech-Phänomens ist, sondern vielmehr die feindseligen und diskriminierenden Kommunikationsinhalte, welche sich auf Adressatengruppen-/Rezipient:innenebene niederschlagen. Sicherlich profitiert eine kommunikationswissenschaftliche Perspektive auf Hate Speech jedoch von einem umfassenden Bild auf diesen Abwertungsmechanismus. Denn Hate Speech entfaltet sich in einem bestimmten psychosozialen Kontext, welcher einem Zusammenspiel individueller, sozialer wie struktureller Faktoren unterliegt. Im Folgenden wird der Begriff Produzent:in<sup>24</sup> von Hate Speech eingeführt, da diese:r im Kommunikationsprozess eine aktive Rolle einnimmt, indem er/sie Hassbotschaften generiert, gestaltet und verbreitet.

In einer international vergleichenden Studie liegt der Anteil an Hate Speech-Produzent:innen durchschnittlich bei 3%, in Deutschland bekennen sich gerade einmal 0,9 Prozent der Befragten dazu, Hate Speech Online zu produzieren (Kaakinen, Räsänen et al., 2018), was sich auch auf ein verzerrtes Antwortverhalten zurückführen ließe, da es sich hierbei um sozial unerwünschtes, delinquentes Verhalten handelt. Die Zahlen schwanken zwar, befinden sich aber im niedrigen einstelligen Prozentbereich, wie auch eine inhaltsanalytische Erfassung von Social Media Accounts zeigt, die sich an Hate Speech beteiligen (Kreißel et al., 2018). Dabei handelt es sich mehrheitlich um Männer (Costello & Hawdon, 2018; Crown Prosecution

<sup>23</sup> An dieser Stelle wird auf die gegenderte Schreibweise verzichtet, da es sich um die Bezeichnungen aus den Modellen handelt, welche zum jeweiligen Zeitpunkt ihrer Veröffentlichung das generische Maskulinum verwenden.

<sup>24</sup> Axel Bruns (2008) entwickelt den Begriff des „Producers“, ein Portmanteau, welches auf der Beobachtung des aktiven Online-Nutzenden fußt. Producers nutzen demnach nicht nur intensiv das Netz, sondern tragen auch aktiv an der Herstellung von Inhalten bei, ohne sich in starren Hierarchiekonstellationen zu verfangen. Sie arbeiten kollaborativ mit anderen Producers zusammen, wobei sie nicht den Abschluss ihres Produktionsprojekts anvisieren, sondern engagieren sich vielmehr kontinuierlich darin. Die kollaborativ generierten Inhalte werden als Allgemeingut von Producers und Rezipient:innen wahrgenommen, wie etwa Wikipedia.

Wenn auch (organisierte) Produzierende von Hate Speech kollaborativ ihre Hassbotschaften ins Netz stellen, scheint der Begriff des Producers hier ungeeignet. So trifft auf sie nicht immer das Prinzip der flachen Hierarchien, gerade wenn sie (parteilich) organisiert sind und ihre Botschaften werden in der breiten Öffentlichkeit auch nicht als Allgemeingut wahrgenommen.

Service, 2018) oder männliche Jugendliche (Blaya & Audrin, 2019), wobei einschlägige soziodemografische und -ökonomische Merkmale, wie Alter, Ethnie, Religionszugehörigkeit, Bildung oder Einkommen keinen Einfluss darauf haben, ob jemand Hate Speech produziert und verbreitet (Costello & Hawdon, 2018; Vorländer et al., 2016). Ein wesentlicher Einflussfaktor für die Produktion von Hate Speech besteht vor allem in einer überdurchschnittlichen Internetnutzung (Costello & Hawdon, 2018). Hier nennen die Autoren die Nutzung einzelner SNS wie *Reddit*, wobei sich insbesondere eine enge Bindung an eine Social Media Plattform als signifikant erweist, d. h. die intensive Nutzung bevorzugter SNS. Hate Speech ist in reichweitenstarken SNS wie *Facebook* oder *YouTube* weit verbreitet (Ben-David & Matamoros-Fernández, 2016), aber das Ausmaß von Hate Speech auf *Reddit* wird wegen unzureichender Regulierungsaktivitäten als gravierender wahrgenommen (Mohan et al., 2017). Als weiterer wichtiger Faktor erweist sich der Zusammenhang zwischen eigener Viktimisierung und Täterschaft, denn wer in der Vergangenheit Ziel von Hate Speech Attacken war, greift selbst andere Nutzende mit aggressiven Hassbotschaften an (Blaya & Audrin, 2019; Costello & Hawdon, 2018). Der Zusammenhang von Viktimisierung und Täterschaft wurde bereits für Cybermobbing gezeigt (Festl, 2015, 2016)<sup>25</sup>.

Soziale Einstellungen erweisen sich als starke Prädiktoren für die Produktion und Verbreitung von Hate Speech (Craig, 2012; Erjavec & Kovacic, 2012). Adorno und seine Forschungsgruppe wollten nach 1945 den Faschismus und Antisemitismus während des Dritten Reichs erklären und gingen davon aus, dass eine strenge und autoritäre elterliche Erziehung die Persönlichkeit von Menschen dahingehend beeinflusst, sich gesellschaftlichen Autoritätsfiguren zu unterwerfen und die eigene Ohnmacht auf Statusniedrige (sowohl Individuen als auch soziale Gruppen) zu projizieren (Adorno et al., 1969). Trotz umfassender methodologischer und inhaltlicher Kritik an dem Autoritarismus-Konstrukt existiert heute eine Vielzahl von Modifikationen der ursprünglichen Autoritarismus-Skala (Petzel, 2009). Als besonders einflussreich hat sich die *Right-Wing Authoritarianism-Skala* (RWA) (Altemeyer, 1981) erwiesen. Ursprünglich als Persönlichkeitsmerkmal konzipiert (Adorno et al., 1969; Altemeyer, 1981), diskutieren aktuellere Forschungsarbeiten Autoritarismus als Kombination psychologischer Persönlichkeitsmerkmale und sozialpsychologischer Interaktionen (Craig, 2012; Duckitt & Sibley, 2010). Autoritäre Individuen vertreten konventionelle Einstellungen und lehnen moderne Lebensstile ab, fügen sich Autoritäten unter und sind feindselig und aggressiv gegenüber

---

<sup>25</sup> Neuere Forschungsarbeiten zeigen nicht nur eine konzeptuelle Nähe zwischen Hate Speech und Cybermobbing, sondern auch einen empirisch messbaren Zusammenhang zwischen den beiden Phänomenen. Wer andere im Online-Kontext drangsaliert und mobbt, wendet auch Hate Speech an. Moderiert wird der Zusammenhang durch toxische Online-Enthemmung, vgl. Wachs et al. (2019)

Minderheiten, sofern sich diese nicht an die Normen und Werte der Mehrheitskultur anpassen (Altemeyer, 1988)

Die *Social Dominance Orientation* (SDO) (Pratto et al., 1994) wird ebenfalls als soziale Einstellung im Zusammenhang mit der Produktion und Verbreitung von Hate Speech diskutiert (Thomsen et al., 2008). Ausgangspunkt der SDO ist eine fundamental kompetitive Sicht auf die Welt, und zwar im Sinne eines „ruthlessly competitive jungle in which might is right, the strong win, and the weak lose“ (Duckitt & Sibley, 2010, S. 1869). Befürwortet wird eine hierarchische Struktur der Gesellschaft, in der eine Statusdifferenz aus der Dominanz statushoher sozialer Gruppen über statusniedrige soziale Gruppen besteht und in der vorhandene Ungleichheiten zwischen sozialen Gruppen aufrecht erhalten werden.

Sowohl RWA als auch SDO haben sich als Erklärungsfaktoren für Vorurteile gegenüber statusniedrigen sozialen Gruppen sowie für die Befürwortung, Produktion und Verbreitung von Hate Speech erwiesen, wenn sich auch die zugrundeliegenden Motivationen bei beiden Konstrukten unterscheiden (Thomsen et al., 2008). In ihrer Untersuchung weisen Thomsen und Kolleg:innen nach, dass SDO im Zusammenhang mit einer starken Ablehnung von anpassungswilligen Migrant:innen in den USA und in der Schweiz steht (ebd.). Begründet wird der Befund mit einem wahrgenommenen Verstoß gegen die ordnende Kraft der gesellschaftlichen Hierarchie. Durch ihre Anpassung an gesellschaftliche Normen und Werte lösen Migrant:innen ihren niedrigen sozialen Status auf, der für SDO-Individuen jedoch von zentraler Bedeutung ist. Werden Migrant:innen als anpassungsunwillig dargestellt, werden sie von Befragten mit einem hohen RWA-Score abgelehnt. Ablehnung, Aggression und Verurteilung von Individuen und sozialen Gruppen, die gegen konventionelle gesellschaftliche Normen und Werte verstoßen, sind charakteristisch für RWA (Thomsen et al., 2008). Diese Ablehnung eines Verstoßes gegen soziale Normen hat zur Konsequenz, dass RWA im Gegensatz zu SDO sogar mit der Befürwortung eines Hate Speech-Verbots in Beziehung steht. Diesen überraschenden Befund liefern Bilewicz und Kolleg:innen und begründen das Ergebnis damit, dass Hate Speech gegen konventionelle Regeln des Anstands und der Moral verstößt (Bilewicz et al., 2017). Damit erweist sich RWA als facettenreiches Konstrukt, welches Anknüpfungspunkte in der Hate Speech-Forschung offenbart.

Mit dem Konzept der *Dark Triad* bzw. *Dark Tetrad* wenden sich aktuelle Forschungsarbeiten der Untersuchung von subklinischen, also nicht pathologischen Persönlichkeitsmerkmalen von Hate Speech-Produzierenden, Trollen oder Cybermobbern zu (Buckels et al., 2019; Craker & March, 2016; Lopes & Yu, 2017; March et al., 2017; Sest & March, 2017). Zentral ist die Annahme, dass antisoziales (also disruptives, manipulatives oder feindseliges) Verhalten

im Internet von Menschen ausgeübt wird, die über eine entsprechende Disposition – bezogen auf ihre Persönlichkeit – verfügen (Fox & Rooney, 2015). Die Dark Triad umfasst hohe Anteile an Narzissmus, Macchiavellismus und Psychopathie (Paulhus & Williams, 2002), während die Dark Tetrad zusätzlich noch Sadismus als Bestandteil einer „dunklen“ Persönlichkeit integriert (Buckels et al., 2013; Buckels et al., 2019). Ausgangspunkt für die Konzeption der Dark Triad ist Paulhus' und Williams' Feststellung, wonach die drei Persönlichkeitsmerkmale sich auf das Verhalten von Individuen auswirken: „[...] all three entail a socially malevolent character with behavior tendencies toward self-promotion, emotional coldness, duplicity, and aggressiveness“ (Paulhus & Williams, 2002, S. 557). *Narzissmus* in seiner subklinischen Ausprägung ist definiert als Grandiosität, Anspruchsdenken, einem Hang zur Dominanz und einem Überlegenheitsgefühl (Paulhus & Williams, 2002). Vom florentinischen Philosophen Niccolò Machiavelli (1469-1527) abgeleitet ist der *Machiavellismus* und bezeichnet manipulative Persönlichkeiten mit einer geringen affektiven Beteiligung in interpersonellen Beziehungen, Ablehnung konventioneller Moralvorstellungen und Ideologien sowie einem ausgeprägten Realitätsbezug (Christie & Geis, 1970; Cloetta, 1983). Als subklinische *Psychopathie* wird ein Persönlichkeitsmerkmal bezeichnet, welches sich durch hohe Impulsivität einerseits und geringe Empathie und Ängstlichkeit andererseits auszeichnet. *Sadismus* als vierter Bestandteil der Dark Tetrad ist definiert als Neigung zu grausamem Verhalten (Buckels et al., 2013; Paulhus & Dutton, 2016).

Die *politischen Einstellungen* und *politischen Ideologien* von Hate Speech-Produzierenden dienen ebenfalls der Erklärung und Charakterisierung dieser. Täter:innen fürchten insbesondere sozialen Wandel (z. B. bezogen auf Einwanderung, Mischehen oder moderne Geschlechterrollen) und gehen aktiv gegen diesen vor, entweder durch Hass motivierte Gewalttaten oder durch Proteste (Glaser et al., 2002; Green et al., 1999; Vorländer et al., 2016). Die Längsschnittuntersuchung von Green und Kollegen widerlegt die These, wonach vor allem die Konkurrenz um knappe Ressourcen ursächlich für die vertretenen feindseligen Einstellungen ist (Green et al., 1998). Sie zeigen, dass nicht Not, sondern die Angst vor Neuem bzw. die Unsicherheit vor Unbekanntem Menschen feindselig agieren lässt.

Aufgefangen wird diese Angst von politischen Bewegungen (z. B. *PEGIDA* in Deutschland, *Tea Party Movement* in den USA) oder politischen Parteien (z. B. *UKIP* in Großbritannien, *Front National* in Frankreich) populistischer Couleur (Loch, 2019)<sup>26</sup>. Diese greifen gesellschaftliche

---

<sup>26</sup> In seinem 1993 erschienen Buch „Die Erfindung des Politischen“ arbeitet Ulrich Beck seine Theorie der reflexiven Modernisierung aus. Er vertritt die These von der Ablösung der industriellen Moderne durch eine reflexive Moderne. Letztere ist von tiefen gesellschaftlichen Umbrüchen geprägt, welche sich aber nicht zielgerichtet, sondern vielmehr ungeplant ereignen und unbeabsichtigte

Misstimmungen auf (Selektion), komprimieren sie auf wenige Streitpunkte (Aggregation) und stellen Öffentlichkeit für ihre Standpunkte her (Präsentation). Einzelpersonen, die personell oder ideologisch mit einer solchen Organisation verflochten sind, greifen zu Hate Speech, um im Sinne ihrer Organisation zu sprechen (Erjavec & Kovacic, 2012). Die Mobilisierung ihrer Mitglieder und Sympathisant:innen ist eminentester Bestandteil dieses politischen Stils (Loch, 2019; Mounk, 2018). Hate Speech und Trolling gelten auch als Propagandainstrument, mittels derer der öffentliche Diskurs gestört wird (Aro, 2016). Als Öffentlichkeitsarena werden häufig soziale Netzwerke gewählt (Ben-David & Matamoros-Fernández, 2016; Blakemore, 2016), denn diese versprechen eine kostengünstige und reichweitenstarke Ansprache ihrer Zielgruppen. Des Weiteren spielt auch die Vernetzung der Gruppierungen untereinander eine herausragende Rolle, und zwar sowohl was die Koordination von Themen als auch die Planung von Hate Speech-Attacken angeht (Burriss et al., 2000; Costello & Hawdon, 2018). Ihre Bindung zu einer Organisation macht diese Personengruppen zu Partei-soldat:innen, wie es Erjavec und Kovacics Typ des „soldiers“ nahelegt (Erjavec & Kovacic, 2012, S. 909f.). Die Mission des „soldiers“ besteht darin, die Organisationsinteressen durchzusetzen, ein Kulturkampf um die Deutungshoheit mit Hate Speech-Mitteln (ebd.). Iganski (2008) mahnt an, dass ein Großteil der Täter:innen lebensweltlich verankert sind und ohne Organisationsbindung, quasi im Alleingang, Hate Speech produzieren und verbreiten. Der Typ des „believers“ verfolgt ideologische Ziele, „players“ erfreuen sich an der puren Provokation und sind mit Internet-Trollen vergleichbar (eingehend dazu Kapitel 2.3.2 Trolling) und „watch-dogs“ sehen sich als Kämpfer gegen soziale Missstände und nutzen Hate Speech als rhetorisches Stilmittel, um öffentlichkeitswirksam Aufmerksamkeit zu generieren (Erjavec & Kovacic, 2012).

Die angewandten Strategien zur Produktion und Verbreitung von Hate Speech lassen sich aus der *Social Identity Theory* (SIT) von Henry Tajfel (1981; Tajfel & Turner, 1979) ableiten. Demnach sind Menschen grundsätzlich bestrebt, ein positives Selbstkonzept zu erlangen. Teil des Selbstkonzepts ist die soziale Identität, welche sich aus dem Wissen um die Zugehörigkeit zu

---

politische Konsequenzen nach sich ziehen. Vor allem vor dem Hintergrund des Falls des Eisernen Vorhangs 1989 und in den Folgejahren sieht Beck eine Abkehr von Politikherstellung im Rahmen klassischer Institutionen: Neue politische Akteure etablieren sich und erarbeiten eine regelverändernde Politik, denn „[w]ir befinden uns in immer mehr Fragen in Situationen, die von den gängigen Institutionen und Begriffen, Konzeptionen des Politischen weder erfaßt noch angemessen beantwortet werden können.“ (Beck, 1993, S. 18). Beck stellt den Modernisierungsprozessen der reflexiven Moderne, welche geprägt sind von Fortschritt im Sinne einer Verbesserung der individuellen und gesellschaftlichen Lebensumstände, eine „Gegenmoderne“ gegenüber, welche sich als „Modernisierung der Barbarei“ (Beck, 1993, S. 16) geriert. Gegenmoderne Tendenzen politischer Akteure zeigen sich in entsprechenden Partei- und Wahlprogrammen (rechts-)populistischer Parteien (Loch, 2019).

einer sozialen Gruppe, der bereits erwähnten Ingroup, ihrer Bewertung und einer damit verbundenen emotionalen Komponente, speist (Tajfel, 1981). Die Zugehörigkeit zur Ingroup ist internalisiert, d. h. die relevanten Kriterien werden vom Individuum anerkannt, sind zusätzlich sozial geteilt, und zwar sowohl von Mitgliedern der Ingroup als auch von Mitgliedern der Outgroup (z. B. Geschlecht<sup>27</sup>, ethnische Zugehörigkeit, Religion, Nationalität). Wird die Gruppenzugehörigkeit in einer gegebenen Situation salient, können psychologische Prozesse ausgelöst werden, die sowohl das Wissen um die Gruppenzugehörigkeit als auch deren kognitive und emotionale Bewertung in Relation zur Outgroup beinhalten. Fällt der soziale Vergleich negativ aus, stehen zwei Strategien<sup>28</sup> zur Verfügung, um das Selbstkonzept positiv zu beeinflussen: (1) soziale Kreativität und (2) sozialer Wettbewerb (Tajfel, 1981; Tajfel & Turner, 1979). *Soziale Kreativität* meint das Anwenden unkonventioneller Methoden, um das Selbstkonzept der Mitglieder der Ingroup zu verbessern. Dazu gehört das Ändern der Vergleichsdimensionen (z. B. in einer Diskussion zur Geschlechtergerechtigkeit im beruflichen Kontext werden Argumente gegen den Gender Pay Gap mit Gerechtigkeitsaspekten von Quotenregelungen beantwortet), der Vergleichsgruppe (z. B. „Den Frauen in Deutschland geht es deutlich besser als iranischen Frauen.“) oder eine geänderte Interpretation der Vergleichsattribute (z. B. „Das niedrigere Einkommen von Frauen hängt nicht mit dem Gender Pay Gap zusammen, sondern lässt sich auf ihr geringes Verhandlungsgeschick zurückführen.“). Die gewählten Beispiele können in Hate Speech umschlagen, wenn sie in feindseliger und abwertender Weise vorgetragen werden. Plausibel erscheint in diesem Zusammenhang auch die Neu-Interpretation von Ereignissen und der resultierenden Medienberichterstattung oder die pauschalisierende Medienkritik („Lügenpresse“, „Systemmedien“) unter den Begriff der sozialen Kreativität zu fassen, wie von Vorländer und Kollegen (2016) für PEGIDA-Demonstrierende gezeigt. Befürwortungen oder Drohungen von Gewalt bis hin zu tätlichen Angriffen und Konflikten werden als *sozialer Wettbewerb* bezeichnet. Die Konfrontation der Ingroup mit der Outgroup wird genutzt, um die soziale Position der Gruppen zu kalibrieren. Karen

---

<sup>27</sup> Die Zugehörigkeit zu einem von zwei Geschlechtern und damit verbundenen Geschlechterrollen gilt als internalisiert (Steffens & Ebert, 2016), wenn auch zu diesem Verständnis eine diversere Lesart von Geschlechtsidentität hinzugekommen ist, wie Küppers (2012) darlegt.

<sup>28</sup> Tajfel (1981) zufolge können Individuen eine positive Selbsteinschätzung auch durch *individuelle soziale Mobilität* gewinnen. Gemeint ist das Bestreben von Individuen eine andere (zumeist höher-rangige) soziale Position innerhalb der Outgroup zu erreichen, was vor allem dann eintritt, wenn die Ingroup keinen weiteren sozialen Aufstieg bietet. Für die hier vertretene Argumentation ist diese Strategie nicht relevant, da es gerade nicht die Hate Speech-Produzierenden sind, die soziale Mobilität zu erreichen, sondern zu verhindern suchen. Sie wenden sich vornehmlich gegen das Eindringen von gruppenfremden Personen, was sich in einer Überbetonung der Statusunterschiede äußert, um die Grenzziehung zwischen Ingroup und Outgroup als unüberwindbar darzustellen. Statusniedrige Gruppen werden als minderwertig und unterlegen herabgesetzt, statushohe Gruppen als elitär und unnahbar kritisiert.

Douglas und ihre Kolleg:innen (2005) können zeigen, dass „White Supremacist“-Gruppen im Internet vor allem Strategien der sozialen Kreativität anwenden, indem sie den Argumentationskontext verschieben (z. B. „Keine Rasse oder Kultur ist einer anderen überlegen. Integration gefährdet alle Völker.“)

Straftaten im Zusammenhang mit Hasskriminalität<sup>29</sup> werden in der US-amerikanischen Polizeiarbeit einer von drei (J. Levin & McDevitt, 1993) bzw. einer von vier (McDevitt et al., 2002) Gruppen der Tätermotivation zugeordnet. Diese forensische Täterkategorisierung ist auch in der akademischen Forschung einflussreich (Coester, 2007) und baut auf Strafakten sowie Interviews mit Polizeibeamt:innen, Opfern von Hasskriminalität und Täter:innen auf. Levin und McDevitt (1993) zeigen, dass Täter:innen in idealtypischer Weise

- (1) das Gefühl von Nervenkitzel und Aufstachelung reizt (*thrill hate crimes*), ausgelöst durch Gruppendynamiken,
- (2) einen physischen oder sozialen Raum vor dem Eindringen von Outgroups verteidigen (*defensive hate crimes*) oder
- (3) sich berufen fühlen, die Eigengruppe von „allem Bösen in der Welt“ zu befreien (*mission hate crimes*).

Die Autoren erweitern 2002 die Typologie um das Motiv der Vergeltung gegen vermeintliches oder tatsächliches Unrecht (*retaliatory hate crimes*) (McDevitt et al., 2002, S. 305f.). Dazu unterziehen sie 169 Strafakten, die bereits 1993 für die Täter:innentypologie herangezogen wurden, einer erneuten Auswertung. Die Autoren stellen fest, dass allen Täter:innenmotiven eine von blindem Eifer getriebene Geisteshaltung gemein ist, wenn auch die psychosozialen Auslöser variieren (ebd.). Wer den Nervenkitzel und die Aufregung erfahren möchte, wünscht sich durch seine Tat Macht und Anerkennung innerhalb der Ingroup. Täter:innen, die „ihren“ Raum zu verteidigen suchen, tun dies aus einem ausgeprägten Schutzbedürfnis gegenüber der Ingroup heraus. Sie handeln häufig proaktiv, während diejenigen, die vom Wunsch nach Vergeltung getrieben sind, sich selbst als „reaktive Kreuzritter“ wahrnehmen, d. h. idealistisch motivierte Beweggründe angeben. Blinder Fanatismus kennzeichnet die Täter:innengruppe der *mission hate crimes*, die es sich zur Lebensaufgabe gemacht hat, als Einzeltäter:in oder als Mitglied einer Hate Group verhasste Outgroups zu bekämpfen (McDevitt et al., 2002). In Anlehnung an McDevitt und Kollegen, schlagen Jacks und Adler (2015) eine

---

<sup>29</sup> Hate Crimes ist ein Sammelbegriff worunter sämtliche Formen der Ausgrenzung, Abwertung bis hin zu Gewalttaten gegenüber Mitgliedern bestimmter sozialer Gruppen fallen (Coester, 2007; Craig, 2012; McDevitt & Levin, 2002).

Täter:innentypologie vor, die im Onlinekontext die Motivation der Hate Speech-Produzierenden sowie deren Einfluss auf die Online-Community in einem Londoner Stadtbezirk zu ergründen versucht. Deren Typologie umfasst Browser, Commentators, Activists und Leaders. *Browser* sind Nutzende, die zwar Hate Speech rezipieren, jedoch nicht damit interagieren. *Commentators* äußern sich in SNS, in Kommentarspalten der Onlineausgaben von Lokalzeitungen, aber auch in Chatgruppen von Hate Sites. *Activists* sind in extremistischen Organisationen eingebunden und sind an der Produktion und Verbreitung von Hate Speech beteiligt. *Leaders* kommt eine Schlüsselrolle zu, da ihr Bestreben ihre Ideologie zu verbreiten, als besonders substantiell von den Autoren charakterisiert wird (ebd.). Exemplarisch wird der „Stormfront“-Gründer Don Black als Leader bezeichnet, der gut vernetzt und finanziell unterstützt seit den 1990er Jahren seine Hate Site betreibt (ebd.).

Online agierende Täter werden laut Quandt (2018) aus drei Beweggründen zu ihren Taten verleitet: Boshaftigkeit sowie taktische und strategische Motive. Die Beweggründe lassen sich in einem zweidimensionalen Raum spannen: Planung/Intentionalität und Zeit. Boshaftigkeit als Motiv ist vergleichbar mit Levin und McDevitts Tätermotiv der *defensive hate crimes*. „These crimes are in their intended effect very much like acts of terrorism, meant to send a signal by means of fear and horror.“ (J. Levin & McDevitt, 2002, S. 79). Ob Boshaftigkeit auch für die Sensationslüsternen aus der Levin- und McDevitt-Typologie eine Rolle spielt, kann abschließend nicht eindeutig festgestellt werden, da die Täter:innen eher situationsgetrieben vorgehen. *Mission hate crimes* deuten darauf hin, dass sie eine strategische (langfristige) Planung beinhalten, wenn auch die Täter:innenmotivation in der gesellschaftlichen „Befreiung“ besteht: „the perpetrator seeks to rid the world of evil rather than to respond to any specific event that threatens him“ (McDevitt et al., 2002, S. 309). Taktische Motive lassen sich in der konkreten Umsetzung erkennen, z.B. indem Täter:innen in das Wohngebiet des Opfers eindringen (*thrill hate crimes*) oder indem die Opfer belästigt oder eingeschüchtert werden, wenn sie in die Täter:innengegend ziehen. Weiterer Forschungsbedarf ist notwendig, um die psychosozialen Beweggründe von Onlinetäter:innen zu analysieren, da hier ein Zugang zu Präventionsmaßnahmen identifiziert werden kann.

Befunde zur konstanten oder episodenhaften Verbreitung von Hate Speech sind uneindeutig. Haben die Anschläge auf das World Trade Center in New York am 11. September 2001, am 13. November 2015 in Paris oder am 19. Dezember 2016 auf den Weihnachtsmarkt am Berliner Breitscheidplatz vermehrt Hate Speech-Attacken und physische Gewalt gegen Muslime hervorgerufen? Die Attentate wurden jeweils von islamistischen Tätern verübt und im Nachgang der Ereignisse kam es immer wieder zu Gewalttätigkeiten gegen Muslime oder

muslimisch aussehenden Menschen (Human Rights Watch, 2002; Seibert, 2016). Aber abgesehen von solchen anekdotischen Evidenzen sind wissenschaftliche Auseinandersetzungen mit der Thematik selten. Die Arbeiten von Disha und Kollegen (2011), King und Sutton (2013) oder Hanes und Machin (2014) sind in diesem Kontext zu nennen, da sie den Zusammenhang zwischen gesellschaftlichen Schockereignissen und einem signifikanten Anstieg von Hate Crimes nachweisen. Nach besonders Aufsehen erregenden Ereignissen, die eine umfängliche massenmediale Berichterstattung erfahren haben (Kepplinger, 2001), nehmen Diskriminierungs- und Abwertungspraktiken bis hin zu offener Gewalt zu, insbesondere dann, wenn potentielle Opfer als Vertreter:innen der angefeindeten Gruppe wahrnehmbar sind (King & Sutton, 2013). Der Ereigniseffekt zeigt sich bereits auf Einstellungsebene: Joscha Legewie zeigt in einer quasi-experimentellen Untersuchung, dass sich die Einstellungen in der Bevölkerung aufgrund von islamistischen Terrorangriffen zuungunsten von Migrant:innen und insbesondere zuungunsten von Muslim:innen verändern (Legewie, 2013). Die Nachwirkungen von Triggerereignissen sind auch in der Polizeiarbeit beobachtbar und sprechen für die Existenz eines Ereigniseffekts: Sind die mutmaßlichen Täter in einem Mordfall Afroamerikaner, werden in der Folgezeit verstärkt afroamerikanische Männer im öffentlichen Raum angehalten und kontrolliert (Legewie, 2016). Für weiße oder hispanoamerikanische Verdächtige konnte der Effekt nicht nachgewiesen werden, was auf Diskriminierung durch Racial Profiling hindeutet (ebd.).

Im nächsten Kapitel geht es darum, den Blick auf Ursachen und Wirkungen von Hate Speech gegen Frauen als soziale Gruppe zu lenken. Auf den ersten Blick erscheint es paradox, Frauen als sozialhierarchisch untergeordnete Gruppe aufzufassen – zumal in einer aufgeklärten und gleichberechtigten Gesellschaft. In der sozialpsychologischen Vorurteilsforschung sind diese Phänomene gut belegt, wie im Folgenden eingehend aufgearbeitet wird.

**Tabelle 3: Zusammenfassende Übersicht über psychosoziale Eigenschaften von Hate Speech-Produzierenden**

<p><b>Soziodemografische Merkmale</b></p>	<ul style="list-style-type: none"> <li>• (Junge) männliche Erwachsene</li> <li>• Überdurchschnittliche Internetnutzung</li> <li>• Zusammenhang zwischen Viktimisierung und Täterschaft</li> </ul>
<p><b>Persönlichkeitsmerkmale</b></p>	<ul style="list-style-type: none"> <li>• Autoritäre Persönlichkeit             <ul style="list-style-type: none"> <li>○ Anpassung an Einstellungen und Handlungen von Höherrangigen</li> <li>○ Konformistisch gegenüber Ingroup, feindselige Ablehnung der Outgroup</li> <li>○ Vorrang von Gruppeninteressen, strenge Moralvorstellungen</li> </ul> </li> <li>• Dark Triad/Dark Tetrad             <ul style="list-style-type: none"> <li>○ Narzissmus (Überlegenheitsgefühl)</li> <li>○ Machiavellismus (manipulative Persönlichkeit)</li> <li>○ Psychopathie (hohe Impulsivität, geringe Empathie)</li> <li>○ Sadismus (Neigung zu grausamem Verhalten)</li> </ul> </li> </ul>
<p><b>Politische Einstellungen</b></p>	<ul style="list-style-type: none"> <li>• Ablehnung von sozialem Wandel</li> </ul>
<p><b>Organisationsbindung</b></p>	<ul style="list-style-type: none"> <li>• Hauptsächlich Einzelpersonen ohne Organisationsbindung, aber mit Sympathie/ideologischer Bindung zu populistischen Parteien, Hate Groups</li> <li>• Vereinzelt Mitglieder populistischer Parteien bzw. extremistischer Hate Groups</li> </ul>
<p><b>Strategien</b></p>	<ul style="list-style-type: none"> <li>• Soziale Kreativität (Anwenden unkonventioneller Methoden)             <ul style="list-style-type: none"> <li>○ Vergleichsdimensionen</li> <li>○ Vergleichsgruppe</li> <li>○ Interpretation der Vergleichsattribute</li> </ul> </li> <li>• Sozialer Wettbewerb (verbale und tätliche Auseinandersetzungen)</li> </ul>
<p><b>Motive</b></p>	<ul style="list-style-type: none"> <li>• Forensische Forschung:             <ul style="list-style-type: none"> <li>○ Nervenkitzel</li> <li>○ Verteidigung der eigenen Domäne vor Outgroups</li> <li>○ Berufung, Hass als Lebensaufgabe</li> <li>○ Vergeltung</li> </ul> </li> <li>• Kommunikationswissenschaftliche Forschung:             <ul style="list-style-type: none"> <li>○ Boshaftigkeit</li> <li>○ Strategie</li> <li>○ Taktik</li> </ul> </li> </ul>
<p><b>Zeitliche Komponente (Triggerereignisse)</b></p>	<ul style="list-style-type: none"> <li>• Verbale und tätliche Anfeindungen</li> <li>• Negative Einstellungsänderungen</li> </ul>

(Quelle: Eigene Darstellung)

### 3 Sexismus als Einstellungskonstrukt

Warum herrschen Vorbehalte und Vorurteile gegen Frauen vor, die sich für eine Karriere in der IT-Branche entscheiden? Wie äußern sich diese Vorbehalte und Vorurteile und wie beeinflussen sie ihre Berufswahl? Warum kommt es in spezifischen Kontexten immer wieder zu Hate Speech gegen Frauen? Diesen Fragen geht das folgende Kapitel nach. Zunächst werden in Kapitel 3.1 verwandte Phänomene voneinander abgegrenzt, um schließlich Sexismus als theoretisches Konstrukt zu fokussieren. Es wird gezeigt, dass Sexismus in der sozialpsychologischen Forschung als paradoxes Einstellungskonstrukt konzipiert wird, was in dem Teilkapitel 3.2 eingehend dargestellt wird. Der Abschnitt 3.3 folgt mit einer Analyse von Faktoren, die den Einstieg und Verbleib von Frauen in der IT-Branche, aber auch den Ausstieg aus derselben erklären. Das Kapitel endet mit einer Zusammenfassung der Erkenntnisse, welche für die eigene Erhebung zentral sind.

#### 3.1 Begriffsbestimmung und Abgrenzung von verwandten Phänomenen

„Geschlecht ist in unserer Kultur einer der wesentlichen Gesichtspunkte, nach denen Personen kategorisiert werden“ (Steffens & Ebert, 2016, S. 13). Menschen nehmen aber nicht nur das Geschlecht ihres Gegenübers wahr. Sie beurteilen ihr Gegenüber auch auf dieser Basis (Stangor et al., 1992) und sehen sich darüber hinaus mit geschlechtsspezifischen Vorurteilen konfrontiert (Barrett & Bliss-Moreau, 2009), was weitreichende Folgen, individueller wie gesellschaftlicher Art, nach sich zieht (Barrett & Bliss-Moreau, 2009; M. J. Brandt, 2011; Caprariello et al., 2009; Fiske, 2000; Poteat & Birkett, 2016; Sidanius & Pratto, 1999). Aus dieser Gemengelage können feindselige Begegnungen zwischen den Geschlechtern entstehen<sup>30</sup>, die eine Begriffsinflation auslösen, welche im folgenden überblicksartig dargestellt wird. Das verfolgte Ziel besteht darin, terminologische Klarheit für den weiteren Verlauf der Arbeit zu gewinnen. Wie unterscheiden sich Antifeminismus und Antigenderismus von männlichem Chauvinismus, dieser wiederum von Frauenfeindlichkeit und Sexismus (Schmincke, 2018)? Warum besteht ein deutlicher Geschlechterbias in den Termini?

*Antifeminismus* versteht sich als ideologischer Antagonist zu feministischen Positionen und Forderungen. Zu Beginn des 20. Jahrhunderts formieren sich in Deutschland erste Frauenvereine und Organisationen, die sich für die gesellschaftliche Teilhabe der Frauen einsetzen, was

---

<sup>30</sup> Sicherlich sind nicht alle Begegnungen ausnahmslos feindseliger Natur, dagegen sprechen allein die täglichen Begegnungen im Alltag, sei es im beruflichen oder privaten Kontext. Diese sind üblicherweise geprägt von einer kooperativen Grundhaltung (Eagly, 1987).

jedoch rasch für Widerstand und Ablehnung in (groß-)bürgerlichen und bäuerlichen Kreisen sorgt (S. Maurer, 2018). Um ihre Ablehnung gegen weitere Modernisierungsschübe nachhaltig zu verbreiten und diese sogar zu verhindern, schließen sich antifeministische Stimmen infolgedessen in Vereinen und Bündnissen zusammen, wie Ute Planert (2000) aus historischer Perspektive darstellt (Schmincke, 2018). Diese antifeministischen Gruppierungen delegitimieren Errungenschaften, z. B. politische Teilhabe in Parteien und Vereinen, aber auch weitergehende Forderungen von Frauenrechtlerinnen<sup>31</sup> in der Öffentlichkeit mittels Broschüren, Karikaturen, Zeitungsartikeln und Büchern (Planert, 2000).

Dass Antifeminismus nicht der Vergangenheit angehört, lässt sich am gegenwärtigen antifeministischen Diskurs als Reaktion auf die Durchsetzung von Gleichstellungsthematiken erkennen (Klaus & Lünenborg, 2013; Möller, 1999; Rauchut, 2018; Schmincke, 2018). Neben konservativen Journalisten, christlich-fundamentalistischen Gruppierungen und rechtspopulistischen Parteien (Schmincke, 2018) treten mit sogenannten Maskulisten neue Akteure auf, die um die Deutungshoheit über die Geschlechterverhältnisse ringen und sich gleichzeitig als Opfer feministischer Forderungen sowie Errungenschaften und als Frauen überlegen stilisieren. „Der Maskulismus lehnt feministische Theorien ab, thematisiert männliche Unterdrückung und fokussiert ein positives, männliches Selbstbewusstsein.“ (Claus, 2014, S. 17). Darüber hinaus ist „eine *Verschiebung* von antifeministischen zu sogenannten ‚antigenderistischen‘ Diskursen“ (Maihofer & Schutzbach, 2015, S. 202, Herv. i. O.) beobachtbar, die auch – aber nicht nur – von Maskulisten betrieben wird.

*Antigenderismus* meint „die Ablehnung verschiedener Formen institutionalisierter Geschlechterpolitik“ (Ganz & Meßmer, 2015, S. 60). Andere Autorinnen unterstreichen die Kritik an und die Infragestellung der akademischen Geschlechterforschung, was letztlich in einem Kulturkampf um gesellschaftspolitische Deutungsmacht mündet (Maihofer & Schutzbach, 2015). Damit ist Antigenderismus als gesellschaftliches Phänomen zu deuten, das vordergründig nicht in erster Linie Frauen an sich ablehnt, sondern vielmehr die institutionalisierten Formen von Frauen- und Geschlechterpolitik (Ganz & Meßmer, 2015). Die Positionen in diesem Deutungskampf sind verhärtet, zumal von antigenderistischer Seite die Politikinhalt diskreditiert, verhöhnt und verunglimpft werden (Schmincke, 2018).

*Männlicher Chauvinismus* geht von einer grundsätzlichen Überlegenheit von Männern bzw. einer Unterlegenheit von Frauen aus und hat damit ebenfalls politische Implikationen und zwar insofern als dass ein Ungleichgewicht der Kräfte zwischen den Geschlechtern

---

<sup>31</sup> Das Frauenwahlrecht wurde in Deutschland erst 1918 eingeführt.

politischen Ausdruck findet (Gilmore, 2001; Mansbridge & Flaster, 2007). Ursächlich für männlichen Chauvinismus werden psychologische Faktoren genannt, wie z. B. Ängste und verminderte Selbstwertgefühle (Woods, 1976).

*Frauenfeindlichkeit* ist die ungerichtete Abwertung von Frauen, die im Englischen häufig als *misogyny* (dt. Misogynie, also Frauenhass) bezeichnet wird (Schmincke, 2018) und mit der vorwiegend online Frauen konfrontiert werden (Jane, 2014, 2017; Kahn et al., 2013). Diese Feindseligkeit, häufig als Furcht interpretiert, ist weit verbreitet, findet Eingang in „cultural institutions, in writings, in rituals, or in other observable activity. Misogyny, then, is a sexual prejudice that is symbolically exchanged (shared) among men, attaining praxis“ (Gilmore, 2001, S. 9). Der Topos der „naturegegebene[n] Minderwertigkeit der Frauen“ (Holland, 2007, S. 97) ist in der abendländischen Geschichte, Religion und Kultur ubiquitär wie kaum ein anderer, wie vor allem literaturwissenschaftliche Untersuchungen zeigen (Bloch, 1987, 1991; Millett, 2016 [1969]; Sielke, 2009). Holland (2007) und auch Gilmore (2001) tragen eine umfangreiche Materialsammlung zusammen, um zu zeigen, dass Frauenfeindlichkeit sich nicht auf westliche Gesellschaften beschränkt, sondern kultur- und zeitübergreifend ist. Misogynie zielt darauf ab, „Frauen auf eine Rolle, die im höchsten Maße restriktiv ist, [zu reduzieren] und [dafür zu sorgen], dass dieser Status quo erhalten bleibt“ (Holland, 2007, S. 301). Erklärungsversuche für Frauenfeindlichkeit lassen sich hinsichtlich der eingenommenen Perspektive in psychologische/psychoanalytische, soziostrukturelle und historisch-materialistische Ansätze unterscheiden (Gilmore, 2001), wobei Gilmore Frauenfeindlichkeit ausschließlich als psychologischen internen Konflikt von Männern definiert (ebd.).

In der sozialpsychologischen Forschung, und dort insbesondere in der Vorurteilsforschung, ist *Sexismus* als Konzept mit einem besonders regen Forschungsengagement etabliert. Seinen Ursprung findet das Konzept des Sexismus in wissenschaftlichen Publikationen der 1970er Jahre in Anlehnung an den Rassismus-Begriff, um strukturelle Ungleichheit zwischen den Geschlechtern zu erforschen (J. C. Becker, 2014; Glick & Rudman, 2010; Schmincke, 2018; Swim et al., 1995). Im Gegensatz zu den oben erläuterten Begriffen beschränkt sich Sexismus jedoch nicht auf ein Geschlecht als soziale Kategorie der Ungleichbehandlung, also nicht auf Frauen als Objekt sexistischer Aussagen und Handlungen, denn Männer sind ebenso davon betroffen. Darüber hinaus können beide Geschlechter sexistische Aussagen verbreiten und damit an der (Re-)Produktion von Sexismus beteiligt sein (ebd.). Meist wird Sexismus als individuelle Einstellung definiert und operationalisiert (Eckes, 2010), aber der Begriff wird von einigen Autoren zusätzlich auf gesellschaftlicher Ebene verortet, etwa als „institutionelle und kulturelle Praktiken“ (J. C. Becker, 2014, S. 29).

Dominierten anfangs eindimensionale Perspektiven auf Sexismus, sind mittlerweile deutlich ausdifferenzierte und komplexe Sichtweisen im Forschungsdiskurs maßgeblich. Insgesamt werden vier Erscheinungsformen unterschieden: traditioneller bzw. offener Sexismus (Spence & Helmreich, 1972), Moderner Sexismus (Swim et al., 1995; Swim & Cohen, 1997), Neosexismus (Tougas et al., 1995) und Ambivalenter Sexismus (Glick et al., 1997; Glick & Fiske, 1996, 1999, 2001b).

Die Ansicht von Allport (1954), wonach Vorurteile vor allem negative, unwillkommene Eigenschaften, Merkmale und Zuschreibungen beinhalten, regt in den 1970er Jahren erste Forschungsarbeiten zu Sexismus an. Spence und Helmreich (1972) entwickeln mit Kolleg:innen (Spence et al., 1973) die Attitudes-toward-Women-Scale (AWS), eine der meistgenutzten Skalen zur Erfassung sexistischer Einstellungen gegenüber Frauen (Glick & Rudman, 2010). Sie konzipieren in Allports Tradition Sexismus als Feindseligkeit gegenüber Frauen, welche als *offener Sexismus* zu Tage tritt. Die Items enthalten präskriptive Aussagen darüber, welche Erwartungen an Frauen gerichtet werden, etwa die traditionelle Rolle der Frau als tugendhafte Mutter und Hausfrau („Women should be concerned with their duties of childbearing and house tending, rather than with desires for professional and business careers.“) und Männern untergeordnet („On the average, women should be regarded as less capable of contributing to economic production than are men.“) (Spence et al., 1973). Ein Einstellungswandel gegenüber Frauen macht sich jedoch spätestens in den 1990er Jahren bemerkbar, als deutlich wird, dass den offen diskriminierenden Aussagen, wie sie in der AWS formuliert sind, immer seltener zugestimmt wird (Spence & Hahn, 1997; Twenge, 1997). 1972 waren die Skalenwerte noch annähernd normalverteilt, aber nur zwanzig Jahre später zeigt sich eine ausgeprägte rechtsschiefe Verteilung der Antworten, was auf deutlich egalitäre Einstellungen hinweist (Spence & Hahn, 1997). Der Toleranzbereich für unverhohlenen frauendiskriminierende Aussagen hat sich mit der Zeit immer weiter verengt. Damit sind die Grenzen der AWS klar erreicht, da diese Skala nicht mehr ausreichend differenziert die Einstellungen der Befragten erfasst (Eckes, 2010; Eckes & Six-Materna, 1998).

Sexistische Einstellungen gegenüber Frauen sind damit jedoch mitnichten verschwunden, wie die Konzepte des Modernen Sexismus bzw. Neosexismus verdeutlichen. Diese Konzepte sehen Sexismus nicht als modern im Sinne von fortschrittlich an, sondern sind in Abgrenzung zu traditionellem („old-fashioned“) Sexismus benannt worden und fassen Einstellungen zusammen, die sich gegen zeitgenössische Errungenschaften wenden (J. E. Cameron, 2001). *Moderner Sexismus* versteht sich als explizite Analogie zu modernem Rassismus und ist durch die Leugnung fortgesetzter Diskriminierung gekennzeichnet (Eckes & Six-Materna, 1998;

Swim et al., 1995). Sowohl Frauen als auch ethnische Minderheiten haben in den Bürger- und Frauenrechtsbewegungen ihre jeweilige Ungleichbehandlung angeprangert und schließlich erfolgreich abgeschafft, indem sie die Durchsetzung egalitärer Rechte errungen haben. Darüber hinaus hat sich die Äußerung progressiver Einstellungen in breiten Gesellschaftsschichten durchgesetzt. Diese Argumente werden angeführt, wenn weitergehende Forderungen als obsolet abgetan werden, denn „Frauen und Männer haben in der heutigen Gesellschaft die gleichen Chancen, etwas zu erreichen.“<sup>32</sup> (Eckes & Six-Materna, 1998). Damit reiht sich das Konzept des Modernen Sexismus in Arbeiten der Vorurteilsforschung ein, die implizit vorhandene Vorurteile nachweisen (Dovidio & Gaertner, 1986b; Eckes & Six-Materna, 1998; Greenwald et al., 1998). Ähnlich argumentieren Tougas und Kolleg:innen (Tougas et al., 1995), die zeitgleich zum Modernen Sexismus von Swim und ihren Ko-Autor:innen das Konzept des *Neosexismus* vorstellen. Die Operationalisierung der beiden Skalen enthält als gemeinsamen Faktor die Leugnung fortgesetzter Diskriminierung (Eckes & Six-Materna, 1998). Zudem verknüpft Neosexismus mit dem Widerstand gegen vermeintliche Privilegien von Frauen und der Ablehnung weiterer Forderungen zwei weitere Faktoren (Tougas et al., 1995), die an Sears' (1988) Kennzeichen für symbolischen Rassismus angelehnt sind (Campbell et al., 1997; Eckes, 2010; Eckes & Six-Materna, 1998). Dieser nennt drei Merkmale als charakteristisch für symbolischen (bzw. modernen) Rassismus: (1) Leugnung fortgesetzter Diskriminierung, (2) Widerstand gegenüber weitergehenden Forderungen und (3) Abneigung gegen besondere Förderung (Sears, 1988).

Modern- und traditionell-sexistische Einstellungen bilden, so gesehen, Komponenten eines Einstellungssystems [...] Die moderne Komponente hat die traditionelle nicht abgelöst, vielmehr ist davon auszugehen, daß sie nebeneinander bestehen und ihre jeweiligen Anwendungsbereiche bzw. Anwendungssituationen haben. Traditioneller Sexismus [...] ist keineswegs überholt. Viele seiner Elemente beschreiben auch heute noch weit verbreitete Einstellungen gegenüber der Rolle von Mann und Frau in der Gesellschaft. Nur entziehen sich diese Einstellungen mehr und mehr dem differenzierenden Zugriff durch Fragebogen. (Eckes & Six-Materna, 1998, S. 235)

Diese Differenzierung sexistischer Einstellungen ist Glick und Fiske (1996, 1999, 2001b) mit ihrem Konzept des Ambivalenten Sexismus gelungen. *Ambivalenter Sexismus* ist demnach ein mehrdimensionales Konstrukt, welches hostile (feindselige) und benevolente (wohlwollende) Einstellungen gegenüber Frauen (Eckes & Six-Materna, 1999; Glick & Fiske, 1996) und Männern gleichermaßen (Glick & Fiske, 1999) erfasst. Hostile Einstellungen sind solche, die der Herabsetzung des eigenen oder anderen Geschlechts<sup>33</sup> dienen („Most women interpret

---

<sup>32</sup> Item aus der deutschen Skala zur Erfassung des modernen Sexismus von Eckes und Six-Materna (1998)

<sup>33</sup> Sowohl Frauen als auch Männer zeigen ambivalent sexistische Einstellungen gegenüber Angehörigen des eigenen Geschlechts (Becker, 2010; Glick et al., 2000).

innocent remarks or acts as being sexist.“<sup>34</sup> „Men would be lost without women to guide them.“<sup>35</sup>), wohingegen Benevolenz vordergründig positive Haltungen zum Ausdruck bringt, die jedoch dem Einhalten traditioneller Geschlechterrollen dienen („Men should be willing to sacrifice their own well being in order to provide financially for the women in their lives.“<sup>36</sup>). Zwischen den beiden Dimensionen gibt es einen stark positiven Zusammenhang (Glick et al., 2004; Glick & Fiske, 2001b). Individuen mit benevolent sexistischen Einstellungen tendieren auch zu hostilem sexistischen Aussagen. Die Autor:innen begründen diese widerstreitenden, ambivalenten Haltungen mit einem andauernden Geschlechterkonflikt sowohl auf gesellschaftlicher als auch auf individueller Ebene, den sie als „curious combination of power difference and intimate interdependence“ entschlüsseln (Glick & Fiske, 2001b, S. 115). Denn einerseits sorgt das Ungleichgewicht von Macht zwischen den Geschlechtern (Männer als gesellschaftlich dominante Gruppe, vgl. auch Richard et al., 2003) für Feindseligkeit, andererseits befinden sich Männer und Frauen aufgrund ihrer intimen heterosexuellen Liebesbeziehungen in einem Abhängigkeitsverhältnis zueinander (Glick & Fiske, 2001b, 2001a).

Das Kapitel 3.2.2 behandelt das Konstrukt des Ambivalenten Sexismus ausführlich, weswegen an dieser Stelle eine begriffliche Abgrenzung bzw. eine Einführung in die verschiedenen Konstrukte erfolgt ist. Zunächst wird in Kapitel 3.2 der sozialpsychologische Ansatz vertieft und Sexismus als Einstellung mit seinen drei Komponenten (Affekt, Kognition sowie Verhalten) eingehend dargestellt. Kapitel 3.3 arbeitet am Beispiel sexistischer Einstellungen gegenüber Frauen in der IT-Branche die Relevanz für die empirische Untersuchung in dieser Arbeit heraus. Das Kapitel endet mit einer Zusammenfassung der Erkenntnisse in Abschnitt 3.4.

## 3.2 Forschungsstand zu Sexismus als Einstellungskonstrukt

Folgt man Eagly und Chaikens Definition von Einstellungen als „*a psychological tendency that is expressed by evaluating a particular entity with some degree of favor or disfavor.*“ (Eagly & Chaiken, 1993, S. 1, Herv. i. O.), stellt das Geschlecht des Gegenübers ein Einstellungsobjekt dar, welches ein evaluatives Repertoire erfährt. Dieses wird als *Multikomponentenmodell* der Einstellung bezeichnet, da es kognitive, affektive oder verhaltensmäßige Anteile enthält (Eagly & Chaiken, 1993; Maio & Haddock, 2010), was in Kapitel 3.2.1 mit besonderem Fokus auf sexistische Einstellungen vertieft wird. Darüber hinaus wird in der Einstellungsforschung die *Struktur von Einstellungen* unterschieden, also die Annahme ein- oder zweidimensionaler mentaler Repräsentationen und darauf folgender Messungen (Eagly & Chaiken, 1993). So

---

<sup>34</sup> Item aus der Ambivalent Sexism Inventory (ASI) von Glick und Fiske (1996)

<sup>35</sup> Item aus der Ambivalence toward Men Inventory von Glick und Fiske (1999)

<sup>36</sup> Item aus der Ambivalent Sexism Inventory (ASI) von Glick und Fiske (1996)

verstehen ältere Operationalisierungen Sexismus als eindimensionales Maß wie z. B. die AWS (Spence & Helmreich, 1972), wohingegen mit dem Ambivalenten Sexismus (Connor et al., 2016; Glick & Fiske, 1996) eine zweidimensionale Skala etabliert wurde. Die Struktur ambivalent sexistischer Einstellungen wird in Kapitel 3.2.2 umfänglich dargestellt. Eher selten behandelt die Einstellungsforschung die *Funktion von Einstellungen*, meist wird diese implizit angenommen (Eagly & Chaiken, 1993). Die funktionale Sicht auf Einstellungen ergründet deren Ursachen, Ziele und Zwecke und hat damit weitreichende Folgen: „knowledge of the motivational basis for an attitude is key to understanding how it can be changed“ (Eagly & Chaiken, 1993, S. 479). Um Redundanzen zu vermeiden, wird die funktionale Sicht auf Sexismus nicht gesondert dargestellt, sondern ist in Kapitel 3.2.2 integriert.

#### 3.2.1 Multikomponentenmodell

Das Multikomponentenmodell geht von drei trennscharfen Komponenten einer Einstellung aus, nämlich Kognitionen, Affekte und Verhalten (Eagly & Chaiken, 1993; Maio & Haddock, 2010). Diese sind primär als Heuristik zu verstehen, um eine analytische Konzeption zu ermöglichen, da die Befunde bezüglich der Dimensionalität widersprüchlich sind (Breckler & Wiggins, 1991; Eagly & Chaiken, 1993). In den folgenden Abschnitten werden die Einstellungskomponenten hinsichtlich ihrer Bedeutung in der Vorurteilsforschung mit besonderem Fokus auf die Sexismusforschung dargestellt.

##### 3.2.1.1 *Stereotype als kognitive Komponente*

Der kognitive Anteil von Einstellungen betrifft ganz allgemein alle mentalen Prozesse, also den Erwerb, das Speichern und das Anwenden von Wissensstrukturen, in Form von Attributionen, Gedanken, Meinungen oder Überzeugungen zu einem bestimmten Einstellungsobjekt (Eagly & Chaiken, 1993; Wegener & Carlston, 2005). Individuen verfügen also über *einstellungsrelevante Wissensstrukturen* (z. B. „Um ein bestimmtes Studienfach zu studieren, ist das Interesse am Fach und nicht das Geschlecht ausschlaggebend.“) und können diese bei Aktivierung als *kognitive Antwort* abrufen (z. B. „Informatik ist ein Studienfach, das gleichermaßen für Frauen und Männer geeignet ist.“).

*Stereotype* und im Speziellen Geschlechterstereotype werden als solche einstellungsrelevante Wissensstrukturen verstanden (Ashmore & Del Boca, 1979; Locke & Johnston, 2001); bei der *Stereotypisierung* handelt es sich entsprechend um die Anwendung dieser Wissensstrukturen im Sinne einer kognitiven Antwort (Eckes, 2010). Stereotype werden in sozialen Kontexten erlernt und kulturell geteilt (Ashmore & Del Boca, 1979; Eckes, 2010; Fiske, 2000; Wood & Eagly, 2002). Sie stellen kognitive Abkürzungen und Vereinfachungen dar, die eine

schnelle Einordnung des Gegenübers erlauben. Gemeint ist eine Bandbreite an Attributen, wie z. B. Persönlichkeitseigenschaften, Aussehen oder Verhaltensweisen, mit denen sowohl die Mitglieder als auch die soziale Gruppe selbst charakterisiert werden (Dovidio et al., 2010; Locke & Johnston, 2001). Dabei ist es unerheblich, ob die Attribute individuell oder kollektiv vorliegen. Darüber hinaus sind Stereotype nicht an faktische Richtigkeit gebunden (Deaux & Kite, 1993). Sie weisen deskriptive und präskriptive Komponenten auf (Dovidio et al., 2010; Locke & Johnston, 2001): Die *deskriptive* Komponente von Stereotypen beinhaltet Zuschreibungen, wie die Mitglieder einer sozialen Gruppe sind, während die *präskriptive* Komponente Erwartungen an diese formuliert.

In ihrem Überblick zeigt Heilman (2001) Konsequenzen, die aus einem Verstoß gegen deskriptive und präskriptive geschlechtsstereotype Erwartungen an weibliche Karrierepfade resultieren. Deskriptive geschlechtsstereotype Eigenschaften („Frauen sind liebevoll und fürsorglich.“, „Frauen sind nicht kompetente Informatikerinnen.“) haben laut Heilman zur Folge, dass die Leistung von Frauen abgewertet und/oder ihnen Anerkennung für Erfolge vorenthalten wird. Gerade ein Verstoß gegen präskriptive geschlechtsstereotype Eigenschaften („Frauen müssen liebevoll und fürsorglich sein.“, „Frauen sollten keinen Männerberuf ergreifen.“) löst offene negative Reaktionen aus, und zwar in Form von persönlicher Missbilligung, Herabwürdigung oder Ablehnung (s. a. Kapitel 3.2.1.3). Ähnliche Befunde wurden immer wieder repliziert (Brescoll & Uhlmann, 2008; Eagly et al., 1992; Heilman et al., 2004; Heilman & Okimoto, 2007; Rudman et al., 2012; Rudman & Fairchild, 2004).

Fiske und Russell (2010) stellen fest, dass seit den 1980er Jahren Erklärungsansätze für die Entstehung von Stereotypen auf kognitionspsychologische Erkenntnisse verweisen. Noch immer aktuell bleibt demnach Taylors (1981) metaphorische Darstellung des „kognitiven Geizhalses“ (engl. *cognitive miser*), der sich bemüht, seine kognitiven Kapazitäten zu schonen. Taylor zufolge geschieht dies, da Individuen konsistente Wahrnehmungen und Erwartungen bevorzugen, hingegen inkonsistente, dissonante Inhalte meiden bzw. nur wichtige und relevante Informationen hinzuziehen (Stangor & Ford, 1992). Dass der Mensch im sozialen Alltag nach Komplexitätsreduktion strebt und sich standardmäßig auf einfache soziale Kategorisierungen verlässt, wurde mehrfach belegt (Bargh, 1999; Tajfel, 1981; Turner et al., 1987).

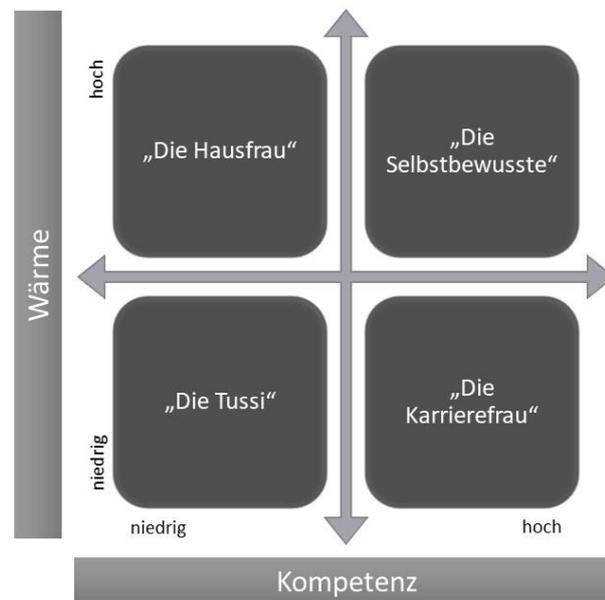
Die aktuelle Forschung zu Stereotypen widmet sich eher subtilen, uneindeutigen Ausprägungen, die entweder unbewusst ablaufen oder sich mehrdeutig und sogar ambivalent äußern (Fiske & Russell, 2010). Unbewusst oder unkontrolliert aktivierte Stereotype werden als *automatisch* bezeichnet. Das Geschlecht als soziale Kategorie ist besonders häufig von automatisch ablaufenden Stereotypen betroffen. In diesem Zusammenhang sind Verwechslungen

innerhalb einer Kategorie verbreitet. Androgyne Gesichter werden beispielsweise häufiger als männlich wahrgenommen, wenn die Hautfarbe der Stimulusperson dunkel ist und als eher weiblich, wenn der Stimulus asiatisch wirkt (K. L. Johnson et al., 2012). *Mehrdeutige*, ambige Stereotype sind solche, die nicht auf den ersten Blick erkennbar sind, sondern eher Raum für Interpretationen lassen. Emotionale Gesichtsausdrücke bei Frauen werden etwa als typisch weibliche Eigenschaften bewertet, während sie bei Männern als situativ bedingt angesehen werden (Barrett & Bliss-Moreau, 2009). Biernat und Vescio (2002) konnten zeigen, dass athletische Männer im Vergleich zu ähnlich athletischen Frauen als bessere Softball-Spieler wahrgenommen wurden. Wenn sowohl positive als auch negative Attribute zu einem Stereotyp verschmelzen, spricht man von *ambivalenten* Stereotypen, wie das Konzept des Ambivalenten Sexismus zeigt (Glick & Fiske, 1996, 1999) .

Das *Stereotypinhaltmodell* (*stereotype content model*, SCM) von Susan Fiske (Cuddy et al., 2008; Fiske et al., 2002; Fiske, 2015) erklärt auf der Basis zweier universeller Dimensionen, nämlich Wärme/Kommunalität und Kompetenz/Instrumentalität, wie sowohl Individuen als auch Gruppen wahrgenommen werden. Als „warm“ gelten Persönlichkeitseigenschaften wie gutmütig, vertrauenswürdig, fürsorglich. Unter Kompetenz werden Eigenschaften wie fähig, erfahren, intelligent etc. versammelt. Anwendung findet das SCM auch in der Sexismusforschung, gerade weil sich Wärme und Kompetenz als weibliche bzw. männliche Charakteristika attribuieren lassen (Abele, 2003; J. B. Miller, 2012). Setzt man nun die beiden Dimensionen mit den Polen hoch/niedrig in Beziehung zueinander, resultiert dies in vier Substereotypen innerhalb der Vierer-Matrix des Stereotypinhaltmodells. Diese Vierer-Matrix enthält je zwei ambivalente und eindeutige Stereotype: *Ambivalente Stereotype* sind solche, deren Dimensionspole in die entgegengesetzte Richtung zeigen. Die Dimensionspole eindeutiger Stereotype zeigen hingegen in die gleiche Richtung. Jeder Quadrant der Vierer-Matrix entspricht einem Substereotyp (Eckes, 2002): Exemplarisch steht „die Hausfrau“ als ambivalentes Stereotyp, denn sie wird als warm(herzig), aber nicht als kompetent angesehen. Ebenfalls ambivalent ist das Stereotyp „die Karrierefrau“, welche zwar als kompetent, dafür eher kühl wahrgenommen wird. „Die Tussi“ als Stereotyp ist sowohl kühl als auch inkompetent, hier zeigen die beiden Dimensionspole jeweils einen niedrigen Wert an. Der vierte Quadrant ist für „die Selbstbewusste“ aber auch die Ingroup reserviert, die sowohl als warm im Sinne von vertrauenswürdig als auch als kompetent wahrgenommen wird (Caprariello et al., 2009; Cuddy et al., 2007, 2008). Männer und Frauen erfahren gleichermaßen stereotype Bewertungen (Eckes, 2002), und zwar zeitüberdauernd (Twenge 1997) und über kulturelle Grenzen hinweg (J. E. Williams & Best, 1990). Im Einzelnen ist der Stereotypinhalt

kulturell und zeitlich nicht universell, dafür aber die Existenz ambivalenter Stereotype (Cuddy et al., 2008; Fiske & Russell, 2010).

**Abbildung 5:** Exemplarische weibliche Substereotype im Stereotypinhaltsmodell



(Eigene Darstellung in Anlehnung an Eckes, 2002, 2010; Fiske, 2015)

Stereotype als kognitive Komponente von Einstellungen werden methodisch auf vielfältige Weise erfasst. Dies erfolgt zum einen explizit, also *direkt*, z. B. mittels direkter Fragen („Was halten Sie davon, dass Frauen in Informatikstudiengängen unterrepräsentiert sind?“) oder mittels Skalen, wenn es darum geht, latente Einstellungsstrukturen zu messen (z. B. AWS, s. o.). Messartefakte treten jedoch nicht selten auf, v. a. bei heiklen Themen sind verzerrte Antworten erwartbar, aufgrund der Neigung sozial erwünschte Antworten zu geben. Der Implicit Association Test (IAT) ist zum anderen eine bewährte Methode, um unbewusste Einstellungen *indirekt* zu messen (Greenwald et al., 1998; Greenwald & Banaji, 1995). Dabei werden die interessierenden Einstellungsobjekte (z. B. soziale Kategorien wie Geschlecht, Ethnie, Religion) gemeinsam mit positiven oder negativen Begriffen gezeigt. Die Versuchsteilnehmenden durchlaufen in der klassischen Experimentalanlage von Greenwald und Kolleg:innen (Greenwald et al., 1998; Greenwald & Banaji, 1995) insgesamt fünf Aufgabenstellungen, in denen sie anfangs die Einstellungsobjekte als disjunkte binäre Elemente durch Handheben zu unterscheiden lernen (linke Hand – männliche Vornamen und rechte Hand – weibliche Vornamen). Im nächsten Schritt werden die präsentierten Begriffe (z. B. Freude, Gift) einer von zwei Kategorien (positiv/negativ) zugeordnet. Die verbleibenden

Aufgaben werden schrittweise kognitiv anspruchsvoller, bis die Probanden zuletzt dazu angehalten werden, die Einstellungsobjekte mit den dargebotenen Begriffen in Kombination zu bewerten. Die Menge der resultierenden Fehler (bzw. die Bearbeitungsdauer in computerunterstützten Messungen) wird als Indikator für vorhandene Vorurteile gegenüber einer der beiden Einstellungsobjekte gedeutet. Kritik an der Validität des Verfahrens ist vielfach geäußert worden (Gawronski et al., 2006; Karpinski & Hilton, 2001; Wilson et al., 2000), wenngleich metaanalytische Verfahren die Aussagekraft des IAT stützen (Hofmann et al., 2005).

Stereotype sind also einfache kognitive Schemata, die standardmäßig aktiviert werden und häufig automatisch, also unbewusst, geäußert werden. Stereotype Aussagen können durchaus bewusst und absichtlich geäußert werden, um das Gegenüber verbal wie nonverbal zu attackieren. Im Fall von sexistischen Stereotypen geschieht dies in der Folge von Verstößen gegen geschlechtsstereotype Attribuierungen. Nicht selten werden Stereotype jedoch eher kaschiert und erscheinen uneindeutig oder ambivalent. Der nächste Abschnitt widmet sich Emotionen als affektive Komponente (sexistischer) Einstellungen. Dabei soll es nicht um Geschlechterstereotype von Emotionen gehen, im Sinne stereotyp weiblicher Labilität und männlicher Gefühlskälte, sondern um emotionale Reaktionen auf Geschlecht als Einstellungsobjekt.

#### *3.2.1.2 Emotionen als affektive Komponente*

Affektive Reaktionen auf Angehörige von Minderheiten, Outgroups oder schlicht verachteten Gruppen sind in der Vorurteilsforschung vielfach belegt: So reagierten etwa Individuen mit Entrüstung und Wut auf die Übergriffe von Migranten auf (deutsche) Frauen in der Kölner Silvesternacht 2015 (Arendt et al., 2017), obdachlose Menschen erfahren, dass Passanten auf sie mit Abscheu und Ekel reagieren (Scheufele & Schieb, 2014) und Menschen mit Behinderung werden mitleidig behandelt (Hughes, 2019), um nur einige Beispiele zu nennen. Frauen und Männer stellen zwar keine Minderheit dar und in der Regel sind deren Interaktionen kooperativ und nicht von negativen Emotionen durchsetzt. Sie interagieren auf professionelle Weise im beruflichen Kontext und führen intime (heterosexuelle) Liebesbeziehungen miteinander; Männern und Frauen schlagen m. a. W. nicht ununterbrochen negative Emotionen entgegen. Dennoch lassen sich Erkenntnisse der Vorurteilsforschung in die Sexismusforschung übertragen (Connor et al., 2016). Begreift man nämlich ganz allgemein Sexismus als „*bias based on gender categorization*“ (Glick & Rudman, 2010, S. 331, Herv. i. O.), wird deutlich, dass diese geschlechterstereotype Kategorisierung durchaus emotionale Reaktionen bewirkt. Dies geschieht etwa dann, wenn Männer und Frauen nicht dem jeweiligen Geschlechterstereotyp entsprechen: Sogenannte „Karierefrauen“ werden beispielsweise

benedet, einfühlsame Männer als „Pantoffelhelden“ belächelt oder bemitleidet (Cuddy et al., 2007).

Smith und Mackie (2010) nennen drei Gebiete in der Vorurteilsforschung, die die emotionale Komponente von Einstellungen betrachten: (1) Emotionen beeinflussen kognitive Prozesse, (2) Emotionen werden bei Begegnungen zwischen Gruppen ausgelöst und schließlich (3) Emotionen resultieren aus der Identifikation mit der Ingroup. Der erste Forschungsstrang sieht die *Valenz von Emotionen* (d. h. negative Emotionen wie Angst oder positive Emotionen wie Glücksempfinden) als ursächlich für eine beeinträchtigte Art der kognitiven Verarbeitung an, da diese die Fähigkeit und Motivation hierzu herabsetzen (Bodenhausen, 1993; Bodenhausen et al., 1994). Aufgrund widersprüchlicher Befunde werden in neueren Untersuchungen zum Einfluss von Emotionen Bewertungstheorien herangezogen: Nicht die Valenz, sondern die kognitiv-bewertende Dimension einer Emotion beeinflusst, wie Menschen ein Einstellungsobjekt (z. B. Individuen, Gruppen, Objekte, Produkte etc.) beurteilen (Lerner & Keltner, 2000). Angst und Wut sind etwa zwei Emotionen mit der gleichen Valenz, da es sich bei beiden um negative Emotionen handelt. Allerdings wirken sich Angst und Wut unterschiedlich in kognitiv-bewertender Hinsicht aus. Angst dient als Interpretationsschablone für Unsicherheit und Risikovermeidung, während Wut die Verantwortung bei der Outgroup sucht und Handlungsbereitschaft auslöst (Lerner & Keltner, 2001). So zeigte sich bei der US-Präsidentenwahl 2016, dass Versuchspersonen, die als wütende Sexist:innen identifiziert wurden, häufiger als ängstliche Sexist:innen an die Wahlurnen gingen und sie stimmten häufiger für Trump (Valentino et al., 2018).

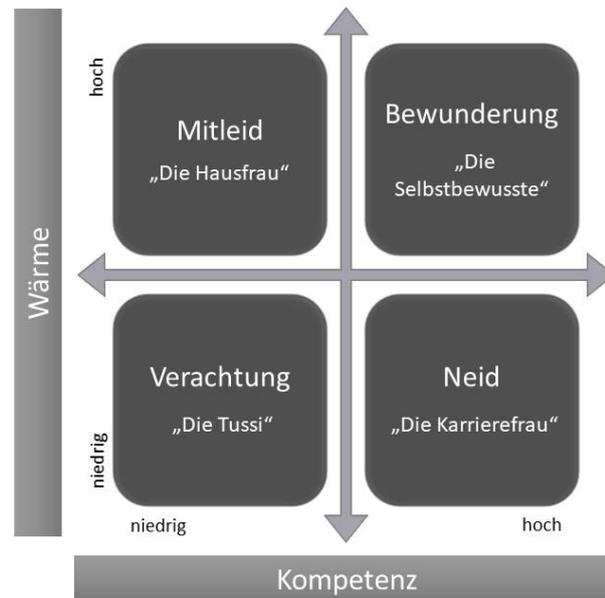
Der zweite Forschungsstrang gehört in der Vorurteilsforschung zwar zu den zentralen Ansätzen, ist aber nur bedingt auf Geschlechterstereotype anwendbar. Dessen Kernthese ist in der *Intergroup Anxiety* konzeptualisiert (W. G. Stephan, 2014; W. G. Stephan & Stephan, 1985) und geht von emotionalen Reaktionen aus, die aus Interaktionen zwischen Angehörigen verschiedener Gruppen (z. B. Ethnien, soziale Schichten etc.) resultieren (E. R. Smith & Mackie, 2010). Da Interaktionen zwischen Männern und Frauen andauernd (zeitliche Dimension) und dabei mehrheitlich harmonischer Natur (qualitative Dimension) sind<sup>37</sup>, erscheinen Untersuchungen zu Intergruppenängsten in diesem Kontext ungeeignet. Intergroup Anxiety ließe sich aber aus Perspektive des Ambivalenten Sexismus diskutieren, da Männer wie Frauen

---

<sup>37</sup> Sicherlich sind an dieser Stelle häusliche Gewalt (Bundeskriminalamt, 2020), sexuelle Übergriffe am Arbeitsplatz, z. B. in Cassino & Besen-Cassino (2019) als exemplarisch für gewalttätige Interaktionen zwischen den Geschlechtern zu nennen.

durchaus negative Emotionen, Ängste und Vorbehalte gegenüber dem anderen Geschlecht zeigen (C. W. Stephan et al., 2000).

**Abbildung 6:** Emotionen gegenüber exemplarischen weiblichen Substereotypen im SCM



(Eigene Darstellung in Anlehnung an Eckes, 2002, 2010; Fiske, 2015)

Der dritte Forschungsstrang fokussiert noch stärker die Zugehörigkeit zu einer sozialen Kategorie, die sich in Form von sozialer Identifikation äußert. Der *Intergroup Emotions Theory* (IET) zufolge lösen Ereignisse, die die Ingroup betreffen, sowohl kollektiv als auch individuell erfahrbare Emotionen aus (Mackie et al., 2000). Dies lässt sich in unterschiedlichen Kontexten nachweisen, etwa in Form von Freude, Stolz oder Enttäuschung der Fans bei Sportereignissen, wenn die „eigene“ Mannschaft einen Sieg erringt bzw. das Spiel verliert (Crisp et al., 2007), oder auf Muslim:innen gerichteter Hass und Abscheu nach den Anschlägen des 11. September 2001 (King & Sutton, 2013; Lickel et al., 2006). Der IET zufolge wird die Outgroup immer dann hinsichtlich ihrer Bedeutung für die Ingroup bewertet, wenn die eigene Zugehörigkeit zu einer sozialen Gruppe salient ist (E. R. Smith & Mackie, 2010). Als Bewertungsmaßstab für die Outgroup dient deren Bedrohungspotential ganz allgemein, mögliche Normverletzungen im Besonderen, aber auch das Ausmaß an überzogenen Forderungen an die Ingroup. Emotionale Reaktionen unterscheiden sich respektive (Cottrell & Neuberg, 2005; Neuberg & Cottrell, 2002). In ihrer Studie konnten Cottrell und Neuberg (2005) zeigen, dass feministische Aktivist:innen negative Emotionen in Form von Vorurteilen bei weißen US-amerikanischen Studierenden auslösen, während amerikanische Ureinwohner:innen Mitleid

erzeugen. Das deutet darauf hin, dass emotionale Reaktionen nicht als uniform zu verstehen sind, sondern sich von Outgroup zu Outgroup unterscheiden.

Das SCM, welches in Kapitel 3.2.1.1 eingeführt wurde, erklärt, warum unterschiedliche emotionale Reaktionen auf Outgroups gezeigt werden. Demzufolge hängt diese von deren wahrgenommener Position in der Wärme-/Kompetenz-Vierermatrix ab. Sowohl korrelative als auch experimentelle Befunde konnten vier Emotionen nachweisen (vgl. Abbildung 6): Bewunderung, Verachtung, Neid und Mitleid (Cuddy et al., 2004, 2007, 2008; Eckes, 2002; Fiske et al., 2002; Fiske, 2015). Wer eine Outgroup als kühl und wenig kompetent wahrnimmt, empfindet *Verachtung* für sie. Solche Outgroups werden als unmoralisch (z. B. Promiskuität) und ihre Situation als selbstverschuldet (z. B. Schulabbruch) angesehen (Cuddy et al., 2008). Verachtung trifft das Stereotyp der „Tussi“ oder des „Proleten“ (Eckes, 2002), um zwei Geschlechterstereotype exemplarisch zu nennen. Vertreter:innen eines Stereotyps, die als warm und kompetent wahrgenommen werden, lösen hingegen *Bewunderung* aus. Das trifft häufig auf dominante Mitglieder der Ingroup (z. B. „der/die Selbstbewusste“) oder Referenzgruppen (z. B. Mittelschicht als Gruppe, zu der man gehören will) zu. Bewundert wird, wer positive Werte (z. B. Erfolg) vertritt, die als erstrebenswert gelten (Cuddy et al., 2008). Im Unterschied zum bewunderten Stereotyp ruft eine Outgroup dann *Neid* hervor, wenn sie zwar als sehr kompetent, aber kühl wahrgenommen wird. „Die Karrierefrau“ als exemplarisches Stereotyp kann hier angeführt werden, da ihre Kompetenz zu stark vom weiblichen Globalstereotyp (geringe Kompetenz, hoher Wärmegrad) abweicht. Darüber hinaus wird ihr aufgrund der wahrgenommenen fehlenden Wärme Feindseligkeit unterstellt, die mit einer Bedrohung von Normen und Werten der Ingroup einhergeht (Cuddy et al., 2008). *Mitleid* erfahren Outgroups, die als warm, aber inkompetent wahrgenommen werden, zumal sie ihre Situation nicht schuldhaft herbeigeführt haben. Cuddy und ihre Kollegen (2008) bezeichnen dieses Mitleid als paternalistisch, da diese Emotion in erster Linie statusniedrigeren Outgroups (z. B. „die Hausfrau“, „der Softie“) zuteil wird.

Aus den Ausführungen in diesem Kapitel lässt sich schließen, dass die affektive Komponente eine herausragende Rolle bei der Bewertung eines Einstellungsobjekts spielt. Smith und Mackies (2010) Übersichtsartikel wurde ausgewählt, um die drei Forschungsstränge zu Emotionen in der Vorurteilsforschung darzulegen und die gewonnenen Erkenntnisse auf die Sexismusforschung zu übertragen. Dazu gehört zum einen der Einfluss von Emotionen auf Kognitionen, zum anderen Emotionen bei Intergruppeninteraktionen und schließlich Emotionen und Ingroupidentifikation. Es ließ sich zeigen, dass von den drei Forschungssträngen insbesondere der dritte für die Sexismusforschung am fruchtbarsten ist, wie mittels SCM von

Fiske und Kollegen vermittelt wurde. Im nächsten Kapitel wird die Diskriminierung als Verhaltenskomponente in Hinblick auf ihre Bedeutung für die Sexismusforschung vorgestellt.

### 3.2.1.3 *Diskriminierung als Verhaltenskomponente*

Im Rahmen dieses Kapitels soll der Frage nachgegangen werden, wie sich Diskriminierungsverhalten aus einer sozialpsychologischen Perspektive äußert und welche Ursachen dafür genannt werden. Verhalten nimmt einen zentralen Stellenwert in der Einstellungsforschung ein und wird bisweilen als Determinante, als Ergebnis oder als Indikator von Einstellungen verstanden (Jaccard & Blanton, 2005). Leon Festingers *Theorie der kognitiven Dissonanz* (1957) steht exemplarisch für Einstellungen, die konsistent mit früherem Verhalten sind. In diesem Zusammenhang lässt sich Verhalten als Determinante von Einstellung konzipieren. Die *Theory of Planned Behavior* (Ajzen, 1991) geht hingegen davon aus, dass Einstellungen zwar nicht das Verhalten direkt, sondern *Verhaltensintentionen* beeinflussen. Die dritte Sicht auf die Verhaltenskomponente betrachtet *Verhalten als Manifestation* des hypothetischen Konstrukts Einstellung. Krosnick und Kollegen warnen jedoch davor, Verhalten mit Einstellungen gleichzusetzen: „[...] the attitude is *not* the numerical summary or the behavioral response that our measurement procedure produces as a product“ (Krosnick et al., 2005, S. 24). Gerade bei der indirekten Einstellungsmessung müssen sich Forschende bewusst machen, dass sie vom beobachteten Verhalten auf die Einstellung nur schließen (ebd.).

In der Diskriminierungsforschung wird häufig vom beobachteten Verhalten auf etwaige sexistische, rassistische oder andere Einstellungen geschlossen. Dieser Forschungszweig ist fest in der Vorurteilsforschung verankert, integriert aber in neueren Arbeiten neben historischen auch institutionelle und organisatorische Faktoren der Diskriminierung (Scherr et al., 2017). In gängigen Definitionen von Diskriminierung werden darunter soziale Kategorisierungsprozesse und damit einhergehende Verhaltensweisen verstanden, die ausgrenzen oder benachteiligen (J. Correll et al., 2010; Dovidio et al., 2010; Dovidio & Gaertner, 1986a; Scherr, 2011; Zick, 2017), wie es bspw. in der Definition von Correll und anderen anklingt: „Discrimination involves behavior directed toward members of a social category that is consequential for their outcomes and that is directed toward them simply because of their category membership“ (J. Correll et al., 2010, S. 55). Diese Definition lässt eine sehr breite Auslegung von Diskriminierung zu, denn nicht alleine eine Benachteiligung der Outgroup, sondern auch eine Bevorzugung der Ingroup ist diskriminierend (ebd., A. S. Haslam & Dovidio, 2010). In einem Klassifikationsvorschlag zählt Zick (2017, S. 74) bis zu 18 Ausdrucksformen der Diskriminierung auf, die sich wiederum auf individueller, kollektiver bzw. institutioneller Ebene unterscheiden. Für die sozialpsychologische Forschung lässt sich ableiten, dass v. a. stabile,

einstellungsbezogene Formen der Diskriminierung von Individuen und Kollektiven relevant sind. Wie im nächsten Kapitel eingehend dargelegt wird, trifft diese Anforderung auf Ambivalenten Sexismus als mehrdimensionales Einstellungskonstrukt zu.

Zentrale theoretische Zugänge, um die Ursachen von Diskriminierung zu erklären, sind: (1) Theorie des realistischen Gruppenkonflikts, (2) Soziale Kategorisierung, (3) Social Identity Theory und (4) Social Dominance Theory (Dovidio et al., 2010; A. S. Haslam & Dovidio, 2010; Zick, 2017). Die *Theorie des realistischen Gruppenkonflikts* führt Diskriminierung auf kompetitive Mechanismen zwischen sozialen Gruppen bei (wahrgenommener) Ressourcenknappheit zurück (Sherif & Sherif, 1969). Exemplarisch sind Ergebnisse von Zárate und Kollegen, die nachweisen, dass mexikanische Migrant:innen in der US-Grenzstadt El Paso hinsichtlich beruflicher Kompetenzen negativ bewertet werden, gerade weil die Einwohner:innen der Stadt mit den Migrant:innen um berufliche Perspektiven konkurrieren (Zárate et al., 2004). Da Geschlechterbeziehungen primär nicht auf Wettbewerb, sondern auf Interdependenz beruhen, hat der Ansatz bislang keine Anwendung für solche Fragestellungen gefunden, ist aber anschlussfähig an politische und kulturelle Aushandlungsprozesse, wie sie im Antigenderismus abgebildet werden.

*Soziale Kategorisierungsprozesse* beinhalten eine klar kognitive Komponente (vgl. Kapitel 3.2.1.1), aber sie beeinflussen ebenso das Verhalten, wenn diesbezüglich deutlich zwischen Mitgliedern der Ingroup und der Outgroup unterschieden wird. Gerade wenn eine soziale Kategorie salient wird, ist die Bevorzugung der Ingroup evident. Ganz egal ob es sich dabei z. B. um Fußball- oder Parteianhänger:innen (Weisel & Böhm, 2015) oder Mitarbeitende handelt (Cunningham, 2005): Menschen neigen dazu, sich einer Gruppe zugehörig zu fühlen. Wie stark sich Individuen als Mitglieder einer sozialen Gruppe verstehen und welche Implikationen damit verbunden sind, geht aus der *Social Identity Theory* (SIT) hervor (Tajfel & Turner, 1979). Demnach beeinflusst die Gruppenzugehörigkeit anteilig das Selbstwertgefühl von Individuen, d. h. das individuelle Selbstwertgefühl steigt bzw. sinkt mit dem Status dieser Gruppe. „Groups are real and have psychological significance for members“ (Reynolds & Turner, 2001, S. 178). Wer sich stark mit der sozialen Gruppe identifiziert, ist umso motivierter Unterschiede zwischen Ingroup und Outgroups zu maximieren und den Status der Ingroup zu erhöhen. Damit ebnet die soziale Identität den Weg für Diskriminierung. „Diskriminierung ist die Herstellung der Vormachtstellung der Ingroup und der Differenz zur Outgroup“ (Zick, 2017, S. 68). Folgen einer solchen Diskriminierung auf Basis der sozialen Identität zeigen sich beispielsweise in Form von geringeren kognitiven Leistungen. Kinder von chinesischen Wanderarbeiter:innen zeigen substantiell schlechtere schulische Leistungen, wenn ihre Herkunft

salient wird (Afridi et al., 2015). Frauen schneiden in Mathematik-Tests schlechter ab, wenn sie mit dem weiblichen Stereotyp der „Mathe-Null“ konfrontiert werden (Schmader, 2002; Spencer et al., 1999; Steele, 1997). Die Lesekompetenz von Jungen ist geringer, wenn Lehrer:innen Jungen für schlechtere Leser halten (Retelsdorf et al., 2015).

Während die SIT soziale Hierarchien zwischen Gruppen nur marginal thematisiert, ist die Annahme einer gruppenbasierten sozialen Hierarchie im Rahmen der *Social Dominance Theory* (SDT) zentral (Sidanius & Pratto, 1999). Die Autor:innen nehmen an, dass sich SDT auf individueller Ebene in der Dominanzorientierung, der *Social Dominance Orientation* (SDO), niederschlägt, nämlich als „[...] degree to which individuals desire and support group-based hierarchy and the domination of ‚inferior‘ groups by ‚superior‘ groups“ (Sidanius & Pratto, 1999, S. 48). Die Zugehörigkeit und Identifikation mit dominanten sozialen Gruppen, Sozialisation, Persönlichkeit und das Geschlecht beeinflussen die Dominanzorientierung von Individuen. Vereinfacht gesagt: Weiße, ältere Männer aus der Mittelschicht mit niedrigerem Bildungshintergrund zeigen eine höhere Dominanzorientierung als junge, liberal eingestellte Frauen (Pratto et al., 1994). Diskriminierung als Verhaltenskomponente von Einstellungen führt unter der Fragestellung der vorliegenden Arbeit dazu, sexistisches Hate Speech gegen Frauen in der IT-Branche als Folge von sozialen Kategorisierungs- und Identitätsprozessen sowie einer Dominanzorientierung zu konzeptualisieren. In Situationen, in denen bestimmte soziale Kategorien salient werden (z. B. Geschlecht, Ethnie, Bildung), erfolgt der Abruf von Stereotypen instantan (Fiske, 2000). Übertragen auf das Geschlecht als soziale Kategorie lässt sich folgern: Schneidet im sozialen Vergleich das eigene Geschlecht schlechter als die gegengeschlechtliche Gruppe ab, werden Strategien zur Aufwertung des Selbstkonzepts angewandt (s. a. Kapitel 2.4). Es muss sich dabei nicht um eine faktisch zutreffende Schlechterstellung handeln, sondern auch schon deren subjektive Wahrnehmung führt zur genannten Aufwertungsstrategie. Sexistisches Hate Speech ist in dem Zusammenhang mit einer ausgeprägten Dominanzorientierung ein Mechanismus der Abwertung mit dem Ziel das individuelle Selbstkonzept aufzuwerten.

„Für das Selbstverständnis moderner Gesellschaften ist der Anspruch konstitutiv, dass prinzipiell alle Gesellschaftsangehörigen als Gleichberechtigte um Positionen im sozialen Gefüge konkurrieren können“ (Scherr, 2010, S. 41), was sich in einem Diskriminierungsverbot mit rechtlich bindendem Charakter (z. B. Art. 3 Grundgesetz) äußert (Bielefeldt, 2010). Geschlecht als Diskriminierungskategorie ist trotz seiner grundgesetzlich verankerten Unzulässigkeit jedoch soziale Praxis in der Arbeitswelt. Geschlechtsbezogene Diskriminierung ist dabei nicht immer ostentativ (etwa als Lohnunterschied), sondern zeigt sich nur in bestimmten

Zusammenhängen (Olian et al., 1988). Frauen bevorzugen etwa Bewerbungen weiblicher Kandidatinnen (Foschi & Valenzuela, 2012), Managerinnen werden von Aktionär:innen stärker kontrolliert (Gupta et al., 2018). Männer werden in geschlechtsinkongruenten Arbeitsverhältnissen (weibliche Vorgesetzte in einer Männerdomäne oder männlicher Vorgesetzter in einer Frauendomäne) schlechter bezahlt und sind weniger angesehen (Brescoll et al., 2012), dafür werden männliche Bewerber im naturwissenschaftlichen Kontext als kompetenter bewertet (Moss-Racusin et al., 2012).

In diesem Kapitel wurde eingangs die unklare Rolle der Verhaltenskomponente in der Einstellungsforschung angesprochen, um Verhalten als Manifestation einer Einstellung zu identifizieren und im Verlauf des Kapitels das Hauptaugenmerk auf Diskriminierung als Ausprägung der Verhaltenskomponente zu legen. Will man die Ursachen für Diskriminierung erklären, sind vier theoretische Ansätze zentral. Aus der Theorie des realistischen Gruppenkonflikts geht hervor, dass v. a. Ressourcenknappheit zu diskriminierenden Praktiken führt. Soziale Kategorisierungsprozesse unterscheiden Ingroup und Outgroups, was im Rahmen der SIT und SDT Bedeutung erlangt. Die SIT betont Identifikationsprozesse mit der Ingroup und zielt auf eine Erhöhung des Selbstwerts vermittelt durch eine Stuserhöhung der eigenen Gruppe. Eine explizite Hierarchie zwischen Gruppen nimmt dagegen die SDT an, die sich in einer individuellen Dominanzorientierung niederschlägt. Forschungsergebnisse zu geschlechtsbezogener Diskriminierung zeigen, dass sowohl Männer als auch Frauen von Diskriminierung im Berufskontext betroffen sind. Damit ist die Darstellung des Multikomponentenmodells von Einstellungen abgeschlossen. Das folgende Kapitel widmet sich dem Ansatz des Ambivalenten Sexismus.

#### **3.2.2 Eine integrative Einstellungsstruktur – Ambivalenter Sexismus als paradoxes Vorurteil**

In der Sexismusforschung wird der Ansatz des Ambivalenten Sexismus von Glick und Fiske (1996, 1999, 2001b, 2001a) als Meilenstein bewertet (J. C. Becker, 2014; Eckes, 2010; Eckes & Six-Materna, 1999). Bahnbrechend an ihrem Ansatz ist die Abkehr von der Annahme, Sexismus äußere sich als undifferenzierte negative Einstellung, wie etwa in der *Attitudes Toward Women Scale* (AWS) operationalisiert (Spence et al., 1973; Spence & Helmreich, 1972). Spätestens als die AWS keine validen Ergebnisse mehr lieferte (Spence & Hahn, 1997), wurde deutlich, dass Allports dominantes Paradigma des „prejudice as an antipathy“ (Allport, 1954, S. 9) nicht mehr aufrecht erhalten werden kann. Ambivalenter Sexismus umfasst zwei Dimensionen, die als zusammengehörig betrachtet werden: hostiler und benevolenter Sexismus (Glick & Fiske, 1996). Negative Haltungen und Verhaltensweisen, die Frauen abwerten,

werden als feindseliger oder *hostiler Sexismus* (HSe) bezeichnet. Als *benevolent Sexismus* (BS) bezeichnen die Autoren vermeintlich positive Einstellungen, die subjektiv als ‚Ritterlichkeit‘ (Schmerl & Steinbach, 1973) getarnt sind (Glick & Fiske, 1996). Aber die Überlegenheit des Ambivalenten Sexismus zeigt sich nicht nur in seiner mehrdimensionalen Konzeptualisierung, sondern auch in der Grundannahme, dass sich Sexismus sowohl gegen Frauen (Glick & Fiske, 1996) als auch gegen Männer (Glick & Fiske, 1999)<sup>38</sup> als Adressatengruppen richten kann. Männer wie Frauen sind ambivalent sexistisch gegenüber dem eigenen und dem anderen Geschlecht eingestellt (Glick & Fiske, 2001b).

Glick und Fiske (1996) bauen ihre Argumentation auf dem Konzept der *Social Dominance Orientation* (SDO) auf (Sidanius et al., 1994). Sie identifizieren Männer als dominantes Geschlecht in der Geschlechterhierarchie, da sie über einen höheren Status, mehr Macht und einen besseren Zugang zu Ressourcen verfügen (Glick & Fiske, 1996). Um den Status quo der Geschlechterungleichheit zu rechtfertigen und aufrechtzuerhalten, dienen sowohl HSe als auch BS als ‚legitimizing ideologies‘ (vgl. Glick & Fiske, 2001a, S. 110f.) und werden als zwei Seiten „ein und derselben ‚sexistischen Münze‘“ (Eckes, 2010, S. 184) verstanden. Beide Dimensionen zeigen einen mittleren bis stark positiven statistischen Zusammenhang (Glick et al., 2000; Glick & Fiske, 2001b). Vordergründig erscheint es paradox, die beiden Dimensionen als funktionell zusammengehörig zu betrachten. Während HSe noch eindeutig als schädlich identifiziert und von Frauen als Adressatinnen im Wesentlichen abgelehnt wird, erscheint das Wohlwollen des BS harmlos und erst auf den zweiten Blick problematisch (Barreto & Ellemers, 2005). Generell dient BS dazu, den Status quo in der Geschlechterhierarchie zu legitimieren (Jost & Kay, 2005) und ist mit einer höheren Lebenszufriedenheit verbunden (Connelly & Heesacker, 2012). BS verspricht Frauen, die sich an der traditionellen Geschlechterrolle orientieren, heterosexuelle Intimität und Schutz von Männern („Women should be cherished and protected by men.“)<sup>39</sup> und hindert sie zugleich daran, einen eigenen unabhängigen Status zu erreichen (Glick & Fiske, 2001a). BS motiviert Frauen Intimität im Austausch für Schutz und Versorgung zu geben. Damit besiegelt BS die Interdependenz zwischen den Geschlechtern. HSe droht hingegen mit strafenden, abwertenden Einstellungen bei Abweichen von der Geschlechterrolle (Connor et al., 2016; Glick & Fiske, 2001a). Als ‚legitimizing ideology‘ dient dabei die traditionelle Geschlechterrolle, die mittels „iron hand in a velvet

---

<sup>38</sup> In der weiteren Darstellung des Ambivalenten Sexismus erfolgt eine Beschränkung auf Frauen als Adressatinnen, was zum einen der Thematik der Arbeit und zum anderen der Tatsache geschuldet ist, dass Frauen Männern hinsichtlich gesellschaftlichem Status, Macht und Ressourcenzugang unterlegen sind (Sidanius, Pratto & Bobo, 1994).

<sup>39</sup> Item aus der Ambivalent Sexism Inventory (ASI) von Glick und Fiske (1996)

glove“ (Jackman, 1994) durchgesetzt wird: Schutz und Intimität erfolgt im Austausch für Wohlverhalten (BS), Strafe dagegen bei Abweichen (HSe).

Als Ausgangspunkt für ambivalent sexistische Einstellungen gilt das Zusammenspiel von drei wesentlichen, wenn auch nicht disjunkten Faktoren: patriarchale Strukturen, ausdifferenzierte Geschlechterrollen und zwischengeschlechtliche Interdependenz (Glick & Fiske, 1996, 2001b, 2001a). Männliche Dominanz, die sich in *patriarchalen Strukturen* niederschlägt, findet sich in vielen Gesellschaften wieder und zeigt sowohl eine feindselige als auch eine wohlwollende Ausprägung (und korrespondiert somit mit HSe und BS). Es wird evolutionspsychologisch mit der sexuellen Selektion argumentiert, die Frauen dazu veranlasst, Männer mit einer guten genetischen und materiellen Ausstattung auszuwählen, um ihren Nachwuchs zu versorgen und beschützen (BS). Männer wiederum erhöhen ihre Fitness, indem sie sich mit möglichst vielen Frauen paaren, was wiederum einen Konkurrenzkampf mit Geschlechtsgegnossen auslöst. Männliche Aggression und Dominanz richtet sich infolgedessen nicht nur gegen Konkurrenten, sondern auch gegen Partnerinnen (HSe), um sicherzustellen, dass diese ihre eigenen Nachkommen aufziehen (Archer, 1996). Sozialpsychologische Ansätze verstehen patriarchale Strukturen als verfestigt. Sie zeigen sich in der fortgesetzten geschlechterspezifischen Arbeitsteilung, obwohl geschlechtlicher Dimorphismus (Unterschiede zwischen Männern und Frauen hinsichtlich der Körpergröße, Muskelmasse, sekundäre Geschlechtsmerkmale etc.) für die Arbeitsteilung in der postindustriellen Dienstleistungsgesellschaft größtenteils unbedeutend geworden ist (Eagly & Wood, 2012; Wood & Eagly, 2012). Frauen sind z. B. in der IT-Branche als Führungskräfte deutlich unterrepräsentiert (vertikale Segregation) und sie übernehmen auch andere Berufsfelder als Männer (horizontale Segregation), beispielsweise im Marketing, Backoffice oder Design statt in den besser bezahlten und statushöheren Bereichen Management, Programmierung oder etwa Netzwerkadministration (Brasse, 2003).

Der Einfluss *ausdifferenzierter Geschlechterrollen* ist in Alice Eaglys *Theorie der Sozialen Rollen* (Social Role Theory) zentral (Eagly, 1987). Ihre Argumentation beginnt bei der geschlechtsspezifischen Verteilung von Männern und Frauen in berufliche und häusliche Rollen (Eagly, 1987; Eagly & Steffen, 1984). Das dabei beobachtete Rollenverhalten von Männern und Frauen bewirkt ein Verfestigen der Überzeugungen hinsichtlich der jeweiligen Geschlechterrollen (gender role beliefs) und die Attribution von entsprechendem Verhalten auf geschlechtsstereotype Persönlichkeitseigenschaften (Eagly & Wood, 2012; Wood & Eagly, 2009, 2012). Wenn Frauen im beruflichen Kontext überdurchschnittlich häufig Menschen beraten, betreuen, erziehen, unterrichten oder pflegen, werden sie von anderen als besonders

kooperativ, empathisch und fürsorglich empfunden und entsprechen der traditionellen Geschlechterrolle (BS). Männer hingegen werden als leistungsfähig, durchsetzungsstark und selbstbewusst wahrgenommen, weil sie überdurchschnittlich häufig Tätigkeiten ausüben, die entweder körperlich anstrengend sind oder analytische und/oder Führungsaufgaben beinhalten (Eagly & Wood, 2012; Wood & Eagly, 2009, 2012). In einer Untersuchung konnte gezeigt werden, dass fiktive Praktikantinnen hinsichtlich ihrer Eignung einen MINT-Beruf zu ergreifen von ambivalent sexistischen Proband:innen als weniger geeignet im Vergleich zu fiktiven Praktikanten bewertet wurden (Reilly et al., 2017).

Ungeachtet der Diversität an Familien- und Lebensformen sind heterosexuelle Partnerschaften gegenwärtig noch immer die maßgeblich bedeutendste Beziehungsart (Lenz, 2006). Auch die empirische „Instabilität von Zweierbeziehungen“ (Lenz, 2006, S. 20) mit seriellen Beziehungen und einer relativ hohen Scheidungsrate, beeinflusst kaum den Wunsch in einer intimen Partnerschaft zu leben (Meier & Allen, 2009). Heterosexuelle Partnerschaften sind darüber hinaus (pop-)kulturell geradezu überhöht und werden als heteronormatives Ideal stilisiert (J. Brandt, 2014). Im traditionellen Sinne sind diese Beziehungen geprägt von *zweigeschlechtlicher Interdependenz*. Diese Interdependenz löst komplexe Interaktionen aus, denn sie bewirkt und verhindert gleichermaßen Ambivalenzen zwischen den Partnern (Glick & Fiske, 2001b; Hammond & Overall, 2016). Dabei zeigt sich die wohlwollende Komponente (BS), wenn Frauen sich exklusiv an einen Mann binden und im Gegenzug für sexuellen Austausch und Sicherheit der Partnerschaft Schutz und Versorgung erhalten (Glick & Fiske, 2001b). Diese gegengeschlechtlichen Motive weisen jedoch in postmodernen Gesellschaften „große milieuspezifische Unterschiede“ (Lenz, 2006, S. 94) auf und haben wegen des gesellschaftlichen Wandels an Relevanz eingebüßt. Gerade die Berufstätigkeit von Frauen oder Fortschritte in der Fortpflanzungsmedizin sind hierfür wichtige Faktoren. Intime Partnerschaften können aber auch eine feindselige Komponente (HSe) enthalten, welche in Form von psychischer, physischer und sexueller Gewalt zwischen den Partnern sichtbar wird (Garcia-Moreno et al., 2005; Lenz, 2006; World Health Organization, 2013).

Die Auswirkungen von Ambivalentem Sexismus finden sich in denjenigen Bereichen wieder, in denen Männer vermeintlich oder tatsächlich über mehr Kraft und Macht verfügen. Dazu zählt der potentiell zwar offene Zugang, aber die dennoch erschwerte Teilhabe an ökonomischen und politischen Machtressourcen (Connor et al., 2016). Frauen erleben feindseligen Sexismus (HSe) immer dann, wenn sie nicht dem traditionellen Globalstereotyp (hohes Maß an Wärme, niedrige Kompetenz) entsprechen oder die traditionelle Geschlechterrolle in Frage stellen. Weibliche Karrierepfade können als Verstoß gegen deskriptive und präskriptive

geschlechtsstereotype Erwartungen gewertet werden und resultieren in fehlender Anerkennung für Leistungen bis hin zu Missbilligung und Ablehnung (Heilman, 2001, 2012; Heilman & Okimoto, 2007). Eine solche negative Evaluation von dominanten Frauen im Arbeitskontext wird auch als „Backlash“ bezeichnet (Rudman, 1998; Rudman et al., 2012; Rudman & Glick, 1999, 2001) und dieser Effekt konnte in einer Meta-Studie bestätigt werden (M. J. Williams & Tiedens, 2016). Der weibliche Körper gilt ebenfalls als Bereich, der ambivalent sexistisch kontrolliert wird; ist etwa Gegenstand herablassender Bemerkungen und Verhaltensweisen, der Bewertung oder gar Bedrohung, wie unten gezeigt wird (Connor et al., 2016).

Eagly und Karau (2002) kommen in ihrer *Rollenkongruenztheorie* (Role Congruity Theory) zu dem Schluss, dass Frauen seltener Positionen als Führungskräfte besetzen, weil die unterstellten weiblichen Persönlichkeitseigenschaften (kommunal expressiv) nicht mit den instrumentellen Erfordernissen an eine (männliche) Führungsrolle kompatibel sind. Als Konsequenz aus der wahrgenommenen Inkongruenz werden weibliche Führungskräfte als unsympathischer wahrgenommen. Zudem zeigt sich ein Bias hinsichtlich ihrer Führungsleistung, die in männlich dominierten Abteilungen als niedriger als die von Männern wahrgenommen wird (Eagly et al., 1995; Eagly & Karau, 2002; Paustian-Underdahl et al., 2014). Handelt es sich um geschlechtsneutrale oder geschlechtsspezifisch weibliche Abteilungen, ist der Effekt nicht signifikant (ebd.). Eine unterstellte mangelnde Kongruenz konnte ebenfalls für die Geschlechter- und Wissenschaftler:innenrolle (Carli et al., 2016) oder für Politikerinnen nachgewiesen werden (S. J. Gervais & Hillard, 2011; Okimoto & Brescoll, 2010). Sowohl Sarah Palin als auch Hillary Clinton weichen als US-amerikanische Politikerinnen vom weiblichen Globalstereotyp ab (Schneider & Bos, 2014) und sind Vorurteilen ausgesetzt. Während jedoch Palin (inkompetent, warm) die Führungsrolle verletzt, weicht Clinton (kompetent, kühl) von der Geschlechterrolle ab (S. J. Gervais & Hillard, 2011). Bei der US-Präsidentenwahl 2016 stimmten demnach hostile Sexist:innen eher für Trump als für Clinton (Ratliff et al., 2017).

Generell gilt für Politikerinnen, wie Hillary Clinton (Brescoll et al., 2018), Angela Merkel (Scholz, 2007), Julia Gillard (Worth et al., 2016), und viele andere Frauen, die von der traditionellen Geschlechterrolle abweichen, dass sie offenen Anfeindungen und feindseligem Sexismus (HSe) ausgesetzt sind. Besonders gravierend sind solche Angriffe, weil sie nicht einfach etwaiges sachbezogenes Fehlverhalten thematisieren, sondern darauf abzielen, den weiblichen Körper zu attackieren (Connor et al., 2016). Gerade im Online-Kontext flammt Hate Speech in Form von Verbalattacken immer wieder auf, das den weiblichen Körper als unrein oder krankhaft verunglimpft und mit sexualisierter Gewalt droht (Citron, 2014; Megarry, 2014). So erfährt die kanadische Bloggerin Anita Sarkeesian 2012 massiven Hass und

Gewaltandrohungen online in Form von Vergewaltigungs- und Morddrohungen, nachdem sie weibliche Stereotype in Computerspielen in einer Videoserie thematisiert hat (Lavigne, 2013). Im Zuge der Gamergate-Kontroverse im Internet 2014 wird Sarkeesian neben Spieleentwicklerinnen und Wissenschaftlerinnen erneut Ziel sexualisierter Gewaltandrohungen (Mortensen, 2018). Die Hetzkampagne gegen die österreichische Autorin Stefanie Sargnagel erreicht 2017 zwar keine globale Aufmerksamkeit, ist aber in seinen Ausmaßen angedrohter Gewalt ebenfalls enorm (Mesch, 2017). Losgetreten wird die Hetze von der Wiener Boulevardzeitung „Krone“, die einen satirischen Reisebericht von Sargnagel benutzt, um die „Krone“-Leser:innen gegen die Autorin aufzuwiegeln (Schmitt, 2017). In ihrem teilweise fiktiven Reisebericht provoziert die Autorin mit Promiskuität, Drogenkonsum und Tierquälerei und verletzt damit gleich drei Geltungsbereiche weiblicher Geschlechtnormen („Frauen sind enthaltsam, gesittet und nicht aggressiv.“). Die resultierende Aggression der „Krone“-Leser:innen ist im Sinne hostiler Sexisten legitim, da Sargnagel (wie auch andere Abweichlerinnen) für ihre Normverletzung bestraft werden müssen.

Das Konzept des Ambivalenten Sexismus hat in den vergangenen 20 Jahren seine Überlegenheit gegenüber anderen Sexismus-Konzepten bewiesen, da es wohlbegründet im sozialpsychologischen Theoriekanon liegt. Es fokussiert geschlechtsbezogene Einstellungen aus einer mehrdimensionalen Perspektive. Neben hostilem (HSe), also feindseligem Sexismus, der im Sinne Allports (1954) oder der AWS von Spence und Helmreich (1972) konzeptualisiert ist, findet sich benevolenter Sexismus (BS), der in älteren Untersuchungen auch als Ritterlichkeit Eingang in die Literatur (Schmerl & Steinbach, 1973) findet. Beide Dimensionen im Ambivalenten Sexismus tragen dazu bei, die traditionelle Geschlechterhierarchie mit ihren Geschlechterrollen aufrechtzuerhalten. Während BS geschlechterrollenkonformes Verhalten positiv sanktioniert bzw. als Erwartungsrahmen definiert, sanktioniert HSe jede Abweichung davon. Dies äußert sich in drei wesentlichen Bereichen, wie Connor und Kolleg:innen (2016) feststellen: Ökonomie und Politik sowie dem weiblichen Körper. Der weibliche Körper ist auffallend häufig Gegenstand von Verbalattacken, auch wenn das unterstellte Fehlverhalten einem der beiden anderen Bereiche zuzurechnen ist. Dieser Aspekt wurde bei der Auswahl der Stimuli im Rahmen dieser Arbeit aufgegriffen und integriert. Im nächsten Teilkapitel werden weibliche Karrierepfade in der IT-Branche und Einstellungen dazu dargestellt, um nachvollziehbar zu machen, dass dieser als männlich dominierter Industriezweig als umkämpfter Schauplatz gilt.

### 3.3 Sexistische Einstellungen zu Frauen in der IT-Branche

Postindustrielle Gesellschaften sind heute geprägt von einem tiefgreifenden Wandel auf gesellschaftlicher, kultureller und wirtschaftlicher Ebene. Dessen Auswirkungen zeigen sich etwa in der Auflösung traditioneller Lebensläufe und „Ungleichheit im Status und den Lebenschancen [werden] dem Einzelnen als Folge individueller und biografischer Entscheidungen“ (Geissler & Oechsle, 2000) zugerechnet. Konnten Wood und Eagly (2002) in ihrer Untersuchung vorindustrieller Gesellschaften kulturübergreifende Muster identifizieren, die die Arbeitsteilung zwischen Männern und Frauen erklären, sind gegenwärtige Ungleichgewichte zwischen den Geschlechtern uneindeutig. Sie lassen sich nämlich nicht nur am Beschäftigungsgrad (Hausfrau vs. Berufstätige) identifizieren (Archer, 1996), denn Frauen sind heute mehrheitlich berufstätig, wenn auch in Teilzeit (Mischke & Wingerter, 2012), v.a. während der Familienphase (Webber & Williams, 2008). Darüber hinaus üben Frauen andere – wenn auch geschlechtsstereotype – Berufe aus (Eagly, 1997), die nicht nur schlechter bezahlt sind (Abele & Spurk, 2011; Valcour & Tolbert, 2003), weniger Aufstiegschancen haben (Ng et al., 2005), oder für die sie überqualifiziert sind (Abele, 2000).

Vor der Familienphase ist das Selbstverständnis junger Frauen geprägt von einer wahrgenommenen Handlungsfreiheit und vielzähligen Entscheidungsalternativen (Geissler & Oechsle, 2000). So sind Bildungsabschlüsse junger Männer und Frauen sehr ähnlich, wobei Schülerinnen sogar höhere Schulabschlüsse erlangen, wie der Bildungsbericht aus dem Jahr 2020 zeigt. Entscheiden sich Schulabgänger zu einem Studium, liegt der Anteil von Frauen an den Studienanfängern bei 52 Prozent. Der Anteil an Absolventinnen zu Absolventen ist dann etwa gleich hoch (ebd.). Die Wahl der Studienfächer fällt geschlechtsspezifisch aus: Frauen entscheiden sich überproportional für Veterinär- und Humanmedizin, Kunst-, Sprach- und Kulturwissenschaft, in MINT-Fächern (z.B. Mathematik, Ingenieurwissenschaften) hingegen sind z. T. weniger als ein Viertel der Absolventen Frauen (Statistisches Bundesamt, 2014). In einer Befragung von US-Studierenden fand Beyer (1999) heraus, dass Studienfächer von Studierenden als geschlechtsstereotyp männlich oder weiblich wahrgenommen werden, indem der wahrgenommene Frauenanteil im Vergleich zum tatsächlichen Anteil über- bzw. unterschätzt wird. MINT-Fächer werden als überwiegend männlich dominiert und anspruchsvoll, geistes- und sozialwissenschaftliche Fächer (z.B. Kunst, Psychologie, Geschichte) als vorherrschend weiblich und einfach bewertet. Darüber hinaus unterschätzen sowohl weibliche als auch männliche Befragte den Notendurchschnitt weiblicher Studierender deutlich. Ähnliche Ergebnisse, die Frauen als stereotyp inkompetent im mathematisch-naturwissenschaftlichen und gerade im Informatik-Bereich charakterisieren, wurden oft repliziert (Cheryan et al.,

2009; Cheryan et al., 2013; Diekman et al., 2010; Margolis & Fisher, 2001; Spencer et al., 1999). Die Folgen, die daraus erwachsen, sind dramatisch: Frauen studieren deutlich seltener MINT-Fächer und lassen sich damit Karrierechancen in innovativen und wachstumsstarken Branchen entgehen (Diekman et al., 2010).

Entscheiden sich Frauen für ein MINT-Studium, bringen sie einen bestimmten Dispositionsmix mit, wie Janshen und Rudolph (1987) für Ingenieurinnen zeigen und was sich auf andere MINT-Fächer übertragen lässt. Dazu gehören sowohl männliche Peers als auch eine enge Bindung an den Vater, ein bildungsnahe familiärer Hintergrund, sowie ein positives Selbstwertgefühl als psychische Disposition. Des Weiteren sind entsprechende Begabungen, Neigungen und Interessen mit ausschlaggebend (Janshen & Rudolph, 1987; auch Stewart, 2003). Extrinsische Motive für die Studienwahl bestehen in der Erwartung hinsichtlich des Einkommens und des beruflichen Status' (Stewart, 2003). Als Barrieren für ein MINT-Studium nennt Gräßle (2009) die emotional herausfordernde Unsicherheit beim Übergang von der Schule zur Hochschule, was sich aber generell als Krise junger Schulabgänger:innen enttarnen lässt. Spezifischer erscheinen verbreitete geschlechterstereotype Bilder vom Frausein und von technischen Studienfächern, die junge Frauen die Entscheidung für ein MINT-Fach erschweren (ebd.). Wenn sie sich für Informatik als Studienfach entscheiden, wählen sie eher ein Fach aus dem Bereich der angewandten Informatik (Ahuja et al., 2006). In Informatik-Studiengängen sind im Wintersemester 2020/21 gerade einmal 19 Prozent der Studierenden Frauen (Statistisches Bundesamt, 2021). Der Anteil an Informatik-Hochschulabsolventinnen liegt 2020 bei 16 Prozent (Statista, 2021b). Die Zahlen hängen nicht nur mit den üblichen Ursachen für Studienabbrüche (problematische Studienbedingungen, mangelnde Eignung, Betreuung etc.), sondern auch mit einer erschwerten sozialen Integration zusammen, da Frauen deutlich unterrepräsentiert sind (Gensch & Kliegl, 2011).

Nach Studium und Ausbildung beträgt der Anteil an weiblichen IT-Fachkräften weniger als ein Fünftel (Bitkom, 2018; Honeypot, 2018), was die geschlechterbezogene Arbeitsteilung in IT-Berufen weiter evident macht (Sax et al., 2017). Um Ursachen für die Zugänge und vor allem den Verbleib in einer männerdominierten Branche wie der IT zu erklären, sind die Arbeiten von Trauth und ihren Kolleginnen vielversprechend (Quesenberry & Trauth, 2012; Trauth et al., 2008, 2009; Trauth, 2017). Trauths *Theorie der individuellen Unterschiede in Geschlecht und IT (Individual differences theory of gender and IT)* verweist deutlich auf kulturelle Erklärungsfaktoren (z.B. persönliche Identifikationsfiguren, Einstellungen gegenüber Frauen in der IT), aber sie berücksichtigt auch eine Vielzahl struktureller Elemente, etwa die Größe der Branche, die Position im Unternehmen etc. (Trauth, 2017). Die Autorin

interpretiert Berufswahl und Karriereentscheidungen als Ergebnis eines Komplexes aus interindividuellen Eigenschaften und strukturellen Bedingungen:

„[...] the theory examines the individual variations across genders as a result of both personal characteristics and environmental influences in order to understand the participation of women in the IT profession. Hence, the individual differences theory of gender and IT focuses on the differences *within* rather than *between* the genders through the understanding of specific influencing factors.“ (Trauth et al., 2008, S. 8)

Drei Konstrukte sind in ihrer Theorie zentral und setzen sich aus der individuellen Identität, dem individuellen Einfluss und schließlich Umwelteinflüssen zusammen (vgl. Abbildung 7, Trauth, 2017). Die *individuelle Identität* umfasst strukturelle Variablen, nämlich soziodemografische und berufsbezogene Größen, wie etwa der IT-Bereich, in dem jemand tätig ist. Faktoren, die es Frauen ermöglichen sich in der Branche durchzusetzen, werden als *individueller Einfluss* bezeichnet. Dazu gehören neben den erworbenen Fertigkeiten und Fähigkeiten in der Ausbildung auch Mentor:innen und Vorbilder. Bei den *Umwelteinflüssen* handelt es sich um Makroeinflüsse, wie politische, ökonomische oder soziokulturelle Rahmenbedingungen (vgl. Trauth, 2017, S. 12). Das Zusammenspiel dieser drei Konstrukte in unterschiedlicher Konstellation bringt in zahlreichen Forschungsarbeiten neue Erkenntnisse darüber zu Stande, wie Frauen in diese männerdominierte Branche Zugang finden, aber auch welche Faktoren ihren Verbleib langfristig beeinflussen (Trauth et al., 2009).

**Abbildung 7: Konstrukte der Theorie der individuellen Unterschiede in Geschlecht und IT**

#### Individuelle Identität

- Soziodemografie (Alter, Geschlecht, Ethnie, soziale Schicht, Elternschaft etc.)
- Beruf (Branche, IT-Bereich)

#### Individueller Einfluss

- Individuelle Eigenschaften (Bildung, Ausbildung, Fähigkeiten, Persönlichkeitseigenschaften)
- Persönlicher Einfluss (z.B. Mentoren, Vorbilder, einschneidende Lebenserfahrungen)

#### Umwelteinflüsse

- Kultur (z.B. Normen und Werte, Einstellungen ggü. IT/Frauen in der IT)
- Infrastruktur (ÖPNV, institutionelle Kinderbetreuung)
- Ökonomie (Größe der Branche, Gehälter etc.)
- Politik (Gesetzgebung, auch Unternehmenspolitik)

(Eigene Darstellung in Anlehnung an Trauth, 2017, S. 12)

Studien zeigen beispielsweise den Umgang von Frauen mit beruflichen Männernetzwerken (Morgan et al., 2004), die Herausforderungen, denen Frauen sich stellen, um Familie und Beruf miteinander zu vereinbaren (Quesenberry et al., 2006), kulturelle Unterschiede hinsichtlich der wahrgenommenen Rolle, die Frauen in einer Gesellschaft einnehmen (Trauth et al., 2008) oder organisationale Einflussvariablen, wie Mentor:innen oder das Betriebsklima (Trauth et al., 2009). Dass der Einfluss dieser Faktoren nicht statisch ist, sondern sich im Zeitverlauf wandeln kann, zeigen Trauth und ihre Kolleginnen (2008). Sie führten Leitfadenterviews mit weiblichen IT-Fachkräften aus vier westlichen Ländern, darunter Irland. Der Datensatz beinhaltet Interviews mit Irinnen aus den Jahren 1990 und 2003. Sagten irische Frauen 1990 noch, dass in ihrer traditionellen Gesellschaft die Akzeptanz arbeitender Mütter gering sei, wird nur 13 Jahre später das gesellschaftliche Klima diesbezüglich als fortschrittlicher wahrgenommen. Insgesamt werden die Karrierechancen von Frauen allerdings trotz dieser Verbesserungen als schlecht bewertet, gerade wenn sie Mütter sind bzw. werden (ähnliche Ergebnisse in Quesenberry et al., 2006). In einer Untersuchung von Cuddy und Kolleg:innen werden berufstätige Mütter als wärmer, aber weniger kompetent im Vergleich zu arbeitenden Frauen ohne Kinder empfunden; berufstätige Männer werden hingegen uniform als kompetent wahrgenommen – unabhängig davon, ob sie Väter sind oder nicht (Cuddy et al., 2004). Arbeitende Mütter erleben nicht nur Gehaltseinbußen, sondern werden auch an höheren beruflichen Standards als Männer gemessen, beispielsweise was ihre Leistung und ihre Pünktlichkeit betrifft (S. J. Correll et al., 2007).

Familienpolitische Reformansätze, wie das deutsche Bundeselterngeld- und Elternzeitgesetz oder Initiativen zur Förderung von Schülerinnen (z.B. *Girls' Day* in Deutschland, *Technology Takes You Anywhere* in Australien), sind zwar vielversprechend, haben aber ein zu geringes bzw. ein eher langfristig angelegtes Wirkungspotential. Wie wichtig die Bedeutung der Kultur und Sozialisation ist, zeigen Trauth und Kolleginnen (Trauth, 2002; Trauth et al., 2008): Sowohl weibliche Geschlechterrollen als auch das Bild von Frauen in IT-Berufen ist gesellschaftlich konstruiert. Frauen mit südostasiatischen Wurzeln (China, Indien, Taiwan) berichten von Erwartungen im familiären Umfeld, etwa einen prestigereichen Beruf in der IT-Branche zu ergreifen<sup>40</sup>. Auch Schulen und Lehrer nehmen eine wichtige Funktion ein, Schülerinnen zu einem Informatikstudium zu ermutigen. Das fällt jungen asiatischen Frauen auch deshalb deutlich leichter, da es nicht als Beleidigung, sondern als Kompliment gilt, als Geek – also Computerfreak – bezeichnet zu werden (ebd.). Unterschiede zwischen Asiatinnen und Europäerinnen bzw. Nordamerikanerinnen lassen sich immer wieder nachweisen und belegen

---

<sup>40</sup> Für japanische Frauen gilt das nicht in gleichem Maße, da von ihnen die traditionelle Rolle als Hausfrau und Mutter erwartet wird, vgl. in Trauth (2002); Trauth et al. (2008).

soziokulturelle Mechanismen, wie Bildungswege oder familiäre Einflüsse (Adya, 2008). So sind für Südostasiatinnen IT-Berufe nicht geschlechtsstereotyp männlich, wohl aber Ingenieursberufe (Adya, 2008; Trauth et al., 2008). Befragt nach Diskriminierungserfahrungen am Arbeitsplatz, verneinen Südostasiatinnen solche Erfahrungen, was mit einer größeren Akzeptanz von Statusunterschieden, einer langfristigen Karriereplanung, die Rückschläge akzeptiert, sowie dem attribuierten positiven Stereotyp der „tüchtigen Asiatin“ zusammenhängt (Adya, 2008).

Aus diesen Ausführungen geht hervor, dass Karrierepfade in die IT für Frauen mit einigen Herausforderungen verbunden sind. Zwar erwerben Frauen heute z. T. qualifiziertere Bildungsabschlüsse als Männer, aber sie entscheiden sich dennoch für geschlechtsspezifische Berufe im Verwaltungs- und Dienstleistungssektor. Diejenigen Studienanfängerinnen, die sich für ein Studienfach im MINT-Bereich entscheiden, haben trotz vorhandener Interessen und Begabungen Schwierigkeiten mit der Diskrepanz zwischen weiblichem Selbstbild und dem geschlechtsstereotyp männlichen Bild des Studienfachs. Trauths *Theorie der individuellen Unterschiede in Geschlecht und IT* vermittelt zwischen individuellen und gesellschaftlichen Ursachen für den niedrigen Anteil von Frauen in der IT-Branche. Dabei verdeutlicht sie zum einen die Bedeutung von Makrofaktoren (Kultur, Wirtschaft, Politik) und interindividuellen biografischen Faktoren.

#### 3.4 Zusammenfassung

Die eingangs gestellte Frage nach den Ursachen für Vorbehalte und Vorurteile gegenüber Frauen in der IT-Branche konnte in den vorhergehenden Kapiteln auf der Basis von Erkenntnissen aus der sozialpsychologischen Vorurteilsforschung beantwortet werden. Vor diesem theoretischen Hintergrund ist Sexismus als Einstellungsstruktur konzipiert worden, das anhand des Multikomponentenmodells von Einstellungen hinsichtlich seiner kognitiven, affektiven und Verhaltensdimension dargelegt wurde. Als hilfreicher theoretischer Ansatz hat sich das Stereotypinhaltsmodell erwiesen, welches nicht nur kognitive Wissensstrukturen in Form von Stereotypen von Individuen oder sozialen Gruppen anhand der Wärme-/Kompetenz-Dimensionen, sondern analog auch affektive Bewertungen modelliert. Diskriminierung als Verhaltenskomponente ist innerhalb der Sexismus-Forschung hoch relevant, ist sie doch ein Indikator für sexistische Einstellungen, welche nicht nur individuell auftreten, sondern auch auf Mesebene innerhalb von Unternehmen und Organisationen bis hin zur gesellschaftlichen Ebene angesiedelt sind.

Die Verankerung von Sexismus in gesellschaftlichen Normen und Werten findet seinen Niederschlag im Konstrukt des Ambivalenten Sexismus, das Peter Glick und Susan Fiske zufolge dazu dienen soll, die traditionelle Geschlechterhierarchie zu rechtfertigen und aufrecht zu erhalten. Ambivalenter Sexismus ist zweidimensional konzipiert mit benevolentem Sexismus, welcher geschlechtsrollenkonformes Verhalten belohnt bzw. als Erwartungsrahmen definiert und hostilem Sexismus, der Abweichungen von diesem Erwartungsrahmen scharf sanktioniert. Hate Speech ist in diesem Zusammenhang ein Mechanismus, um Verstößen gegen die traditionelle Geschlechterrolle (also auf der HS-Dimension) Ausdruck zu verleihen. Weibliche Karrierepfade in der IT-Branche werden exemplarisch herangezogen, um zu illustrieren, dass es sich hierbei um Abweichungen von weiblichen Berufsentscheidungen handelt, weswegen die Thematik der Frauen in der IT-Branche als besonders vulnerabel bewertet wird, Adressatinnen von Hate Speech zu werden.

## 4 Das Elaboration Likelihood Model als Ansatz zur Erklärung von Einstellungsänderungen

Wenn es um die theoretische Einbettung von Persuasionseffekten geht, gehört das *Elaboration Likelihood Model* (ELM) von Petty und Cacioppo (1986a) zu einem der am häufigsten herangezogenen Ansätze (Klimmt, 2011; Knoll, 2015; Petty & Briñol, 2012; Wirth & Kühne, 2013; Wood, 2000). Ähnlich dem *Heuristic Systematic Model* von Shelly Chaiken (1980) handelt es sich bei dem ELM um ein Zwei-Prozess-Modell der Informationsverarbeitung oder „Dual Process Model“, welches die Verarbeitung von Medieninhalten, Werbebotschaften oder anderweitig persuasiven Botschaften (z. B. gesundheitliche Aufklärungskampagnen) auf unterschiedlichen kognitiven Pfaden modelliert. Die beiden Pfade werden als zentrale und periphere Route der Informationsverarbeitung bezeichnet und ermöglichen eine Integration der divergierenden Erkenntnisse der Einstellungsforschung bis Ende der 1970er Jahre. Die *zentrale Route* beschreibt eine vertiefte kognitive Verarbeitung oder gedankliche Auseinandersetzung, während die *periphere Route* eine oberflächliche und flüchtige Informationsverarbeitung impliziert. „A general framework for understanding attitude change must consider that in some situations people are avid seekers and manipulators of information, whereas at other times people are best described as "cognitive misers" who eschew any difficult information-processing activity“ (Petty & Cacioppo, 1984a, S. 79). Ihren Ursprung hat die Konzeptionalisierung in einer Reihe von Theorieansätzen, die eine sorgfältige gedankliche Auseinandersetzung eintreffender Informationen postulieren (Petty et al., 1983), wozu u. a. die Theorie der kognitiven Dissonanz (Festinger, 1957) oder der Cognitive Response-Ansatz (Greenwald, 1968) gehören.

Ziel des folgenden Kapitels ist, das ELM als Analyseperspektive im Hate Speech-Kontext fruchtbar zu machen, da hierzu nur wenig Forschung existiert (E. Lee & Leets, 2002; Schieb & Preuss, 2018). Zunächst wird in Abschnitt 4.1 der Persuasionsbegriff definiert und der persuasive Charakter von Hate Speech mit bereits gewonnenen Erkenntnissen aus Kapitel 2 verknüpft. Abschnitt 4.2 gibt einen knappen Überblick zu den klassischen Annahmen im Persuasionsprozess und Abschnitt 4.3 schließt daran an und vertieft die Erkenntnisse. Mit Abschnitt 4.4 erfolgt eine integrative Sicht auf Hate Speech und seine Persuasivität im Rahmen computervermittelter Kommunikation. Das Kapitel schließt mit einer Zusammenführung der gewonnenen Einsichten aus den drei Theoriekapiteln (Abschnitt 4.5).

## 4.1 Einführung in den Persuasionsbegriff

Persuasion ist ein zentraler Begriff innerhalb des theoretischen ELM-Gebäudes und wird häufig als Überreden oder Überzeugen des Gegenübers mit dem Ziel der Einstellungsänderung (dazu später) definiert (Petty & Cacioppo, 1986b). Aus linguistischer Perspektive zielt Persuasion darauf ab, „gewünschte Wirkungen in kooperativer Übereinstimmung/Auseinandersetzung mit dem Gesprächspartner herbeiführen zu wollen“ (Ortak, 2004, S. 6). Gerald R. Millers kommunikationswissenschaftlicher Definition zufolge beinhaltet Persuasion drei charakteristische Eigenschaften, um sie von anderen kommunikativen Handlungen zu unterscheiden: „‘being persuaded’ applies to *situations where behavior has been modified by symbolic transactions (messages) that are sometimes, but not always, linked with coercive force (indirectly coercive) and that appeal to the reason and emotions of the person(s) being persuaded*“ (G. R. Miller, 2002, S. 6, Herv. i. O.).

Persuasion bedarf also erstens entsprechender *Botschaften*, die in Kombination mit nonverbalem Handeln ihre Wirkung verstärken. Persuasive Botschaften werben als Medieninhalte, Produktwerbung, Wahlwerbespots und dergleichen um die Aufmerksamkeit der Rezipienten. *Zwang* als zweites Charakteristikum von Persuasion aufzufassen, erscheint auf den ersten Blick abwegig, da der Rezipient durch die persuasive Botschaft und nicht durch Druckausübung, gerade nicht durch „ein Angebot, das man nicht ablehnen kann“<sup>41</sup>, überzeugt oder überredet werden soll. Perloff verweist auf „an atmosphere of free choice“ (Perloff, 2003, S. 8) in der sich Persuasion vollzieht und hebt die grundsätzliche Entscheidungsfreiheit hervor. Bei Ortak (2004) klingt der Zwangsaspekt im Begriff der Auseinandersetzung als Gegensatz zur kooperativen Übereinstimmung an. Die sprechende Person kann nicht davon ausgehen, dass ihre persuasive Botschaft auf aufgeschlossene und empfangsbereite Rezipierende trifft. Sie muss damit rechnen, dass diese ihr ablehnend, skeptisch oder mindestens desinteressiert begegnen. Um nun indirekten Zwang im Sinne Millers auszuüben, stehen dem Sprecher oder der Sprecherin unterschiedliche Strategien und Stilmittel zur Verfügung, wie das Polarisieren, Motivieren und Plausibilisieren, das Aufzeigen der Relevanz für die Rezipierenden oder die Kompetenz des Sprechers bzw. der Sprecherin (ausführlich bei Ortak, 2004). Dem dritten Charakteristikum von Persuasion kommt ein besonderer Stellenwert zu, da sich die Wirkung auf *Emotionen und Kognitionen* der Rezipierenden in der Dualität zwischen Überreden und Überzeugen widerspiegelt. Während die emotionale Wirkung mit einem

<sup>41</sup> Zitat aus Francis Ford Coppolas Verfilmung „Der Pate“ von 1972, das die Machtausübung eines fiktiven New Yorker Mafiabosses prägnant zusammenfasst, der weder vor Erpressung, tätlichen Auseinandersetzungen bis hin zu Mord zurückschreckt, um somit unter Zwang seine Interessen durchzusetzen.

Überreden in Zusammenhang gebracht wird, gilt das Überzeugen als Königsdisziplin, wenn die Rezipierenden die Argumentation in der persuasiven Botschaft als logisch und nachvollziehbar anerkennen (G. R. Miller, 2002). In der Persuasionsforschung werden stets beide Aspekte berücksichtigt und als zwei Seiten einer Medaille betrachtet (Wood, 2000).

Die kommunikationswissenschaftliche Auseinandersetzung mit dem Persuasionsbegriff setzt an „Änderungen des mentalen Zustands“ (Wirth & Kühne, 2013, S. 314) an und fasst darunter die Beeinflussung von Einstellungen, aber auch von Meinungen, Überzeugungen und Werten. Dass sich Persuasionsforschung primär *Einstellungen* widmet, ist forschungsökonomischen Zwängen geschuldet. Einstellungen lassen sich mittels Fragebögen, Skalen und mehr oder minder aufwendigen Messverfahren verhältnismäßig einfach ermitteln (s. Kapitel 3.2). Eine wichtige Rolle spielt dabei ihre kognitive Zugänglichkeit (*accessibility*) (Fazio, 1990; Rhodes, 2012). Kognitiv zugängliche Einstellungen lassen sich leichter abrufen (ebd.; Fazio et al., 2000) und beeinflussen das Verhalten in einer gegebenen Situation (Fazio, 1986). Insgesamt gestaltet sich die Messung von *Verhalten* im Forschungsprozess deutlich aufwendiger, weswegen persuasive Effekte seltener direkt auf Verhaltensebene nachgewiesen werden (Rhodes, 2012). Stattdessen dienen Verhaltensintentionen oder aggregierte Verhaltensweisen als angemessene Prädiktoren für tatsächliches Verhalten (Ajzen, 1991). Darüber hinaus wird Nancy Rhodes (2012) zufolge in persuasiven Botschaften an *soziale Normen* appelliert, also all jene Regeln, die in sozialen Interaktionen entstehen und verfestigt werden (Cialdini & Trost, 1998). Auch für Normen gilt ein gut belegter Einfluss auf Verhalten (Ajzen & Fishbein, 2005). Gesundheitskampagnen etwa setzen darauf, ein gesundheitsförderliches Verhalten („Blutspender retten Leben“) zu etablieren oder ein gesundheitsschädliches Verhalten zu bekämpfen („Julia hört auf. Das schaffst auch Du“), indem soziale Normen als Orientierungsmaßstab propagiert werden (Flynn et al., 2011).

Die genannten Auswirkungen persuasiver Botschaften auf Einstellungen, Verhalten und soziale Normen äußern sich nicht nur als Änderung (*response changing*), sondern auch als Ausformung (*response shaping*) und Verstärkung (*response reinforcing*) auf ebendiesen Wirkungsebenen (G. R. Miller, 2002). Holbert zeigt, dass Änderungen, Ausformungen und Verstärkungen von Einstellungen, Verhaltensweisen und sozialen Normen immer wieder nachgewiesen werden (Holbert, 2012). Tatsächlich gestaltet sich jedoch eine Änderung als ungleich schwieriger, aufwendiger und unsicherer als die Verstärkung und Bestätigung vorhandener Einstellungen, Verhaltensweisen oder sozialer Normen (ebd.). In der Erie County Studie wurde erstmals belegt, dass (Medien-)Inhalte vor allem in Übereinstimmung mit bestehenden Einstellungen selektiv rezipiert werden (Lazarsfeld et al., 1944). Klappers

Verstärkerthese baut auf diesen Befunden auf und hält fest, dass Massenkommunikation zur Verstärkung bestehender Einstellungen beiträgt (Klapper, 1960).

Wie in Kapitel 2.1.4.3 deutlich wurde, muss Hate Speech ein persuasiver Charakter im Sinne Millers (2002) zugeschrieben werden, da es – verbal und nonverbal geäußert – durch indirekten Zwang (aufgrund der außerordentlichen Polarisierungskraft) kognitive und emotionale Reaktionen bei Rezipierenden hervorruft. Hate Speech ist nicht frei von Zwang, da primäre Adressatengruppen nicht wählen können, ob sie sich von feindseligen und abwertenden Kommunikationsinhalten diskriminiert fühlen. Sekundären Adressatengruppen kann es dagegen deutlich einfacher gelingen, sich von dem indirekten Zwang der Wirkung abzugrenzen, wenn auch bestimmte Kontexte einen Wirkungsverlauf verstärken können, da Anhänger:innen mit Argumenten versorgt und emotional mobilisiert werden, während Gegner:innen nach Gegenargumenten suchen und mit Empörung reagieren.

Persuasion wird im ELM-Kontext als komplexer Wirkungsprozess verstanden, der von den beiden Hauptautoren Richard E. Petty und John T. Cacioppo in ihren Arbeiten expliziert und mit empirischen Befunden unterlegt wurde. Im nächsten Abschnitt werden die sieben Postulate des ELM dargestellt, um die Grundzüge des Modells darzustellen.

## 4.2 Klassische Modellannahmen

1986 erscheinen von Petty und Cacioppo zwei Hauptwerke, welche dazu dienen, das Theoriegebäude des ELM anhand von sieben Postulaten zu errichten. In der Monografie *Communication and persuasion: Central and peripheral routes to attitude change* geben die Autoren zunächst einen Überblick über ihre Annahmen, und belegen anhand von Studien deren Gültigkeit (Petty & Cacioppo, 1986a). Deutlich ausführlicher diskutieren die beiden die Annahmen in dem Zeitschriftenbeitrag *The Elaboration Likelihood Model of Persuasion* (Petty & Cacioppo, 1986b). Die Beschreibung und Begründung der Postulate dient der Darstellung der komplexen Beziehungen der einzelnen Faktoren im Persuasionsprozess, welche in den nachfolgenden Jahren noch weiter ausgefeilt wurde (Petty & Wegener, 1999).

Das *erste Postulat* bezeichnen Petty und Cacioppo mit Verweis auf Festinger (1950) als „important guiding principle“ (Petty & Cacioppo, 1986b, S. 127) da es das Fundament des ELM darstellt: Sie gehen davon aus, dass Menschen bestrebt sind, stimmige Einstellungen zu erlangen und zu vertreten, was in konvergentem Verhalten mündet (Festinger, 1954). Die gedankliche Auseinandersetzung mit einer persuasiven Botschaft wird in einem Elaborationskontinuum modelliert (*zweites Postulat*) welches von individuellen und situativen Faktoren beeinflusst wird. Im *dritten Postulat* weisen Petty und Cacioppo darauf hin, dass

Einstellungsänderungen sich auf eine Vielzahl von Variablen zurückführen lassen, welche entweder als Argumente oder als periphere Hinweisreize verstanden werden. Zusätzlich beeinflussen den Autoren zufolge bestimmte Variablen die Elaborationsstärke und die Elaborationsrichtung, was im *vierten Postulat* als objektive Elaboration und im *fünften Postulat* als verzerrte Elaboration<sup>42</sup> näher erläutert wird. Das sechste Postulat hebt die steigende Bedeutung peripherer Hinweisreize bei niedriger Elaboration hervor. Die Konsequenzen, die sich aus der gedanklichen Auseinandersetzung im Persuasionsversuch ergeben, werden im siebten Postulat diskutiert.

Auf dem behandelten Überblick aufbauend werden im nächsten Abschnitt die zentralen Elemente im ELM und ihre Beziehungen untereinander vertieft.

**Tabelle 4:** *Postulate des ELM*

1. Postulat: Stimmigkeit der Einstellungen	Menschen sind motiviert stimmige Einstellungen zu haben.
2. Postulat: Elaborationskontinuum	Obwohl Menschen stimmige Einstellungen zeigen möchten, variieren Ausmaß und Art der sachbezogenen Elaboration, zu der sie bereit oder fähig sind in Abhängigkeit von individuellen und situativen Faktoren.
3. Postulat: Multiple Rollen von Variablen	Variablen beeinflussen Ausmaß und Richtung der Einstellungsänderung indem sie: <ul style="list-style-type: none"> <li>• als persuasive Argumente dienen</li> <li>• als periphere Hinweisreize dienen und/oder</li> <li>• Ausmaß und Richtung der Themen- und Argumentationsverarbeitung beeinflussen</li> </ul>
4. Postulat: Objektive Elaboration	Variablen, die die Motivation und/oder die Fähigkeit zur objektiven Verarbeitung einer Botschaft beeinflussen, tun dies, indem sie die Auseinandersetzung mit den Argumenten verstärken oder verringern.
5. Postulat: Verzerrte Elaboration	Variablen, die die Verarbeitung einer Botschaft eher verzerren, können die Motivation und/oder Fähigkeit einstellungsbezogene Gedanken zu entwickeln entweder in eine positive (befürwortende) oder eine negative (ablehnende) Richtung verzerren.
6. Postulat: Ausgleich	Ist die Motivation und/oder Fähigkeit zur Verarbeitung von Argumenten gering, üben periphere Hinweisreize einen relativ großen Einfluss auf die Persuasion aus.
7. Postulat: Einstellungsstärke	Einstellungsänderungen, die hauptsächlich aus der Verarbeitung sachbezogener Argumente (zentrale Route) resultieren, sind zeitlich persistenter, prognostizieren Verhalten genauer und sind resistenter gegenüber Gegenargumenten im Vergleich zu Einstellungsänderungen, die aus der Verarbeitung peripherer Hinweisreize resultieren.

(Quelle: Eigene Darstellung in Anlehnung an Kitchen, Kerr, Schultz, McColl, & Pals, 2014; Petty & Cacioppo, 1986; Petty & Wegener, 1999)

<sup>42</sup> Petty und Wegener (1999) vertauschen die Reihenfolge der Postulate und besprechen die verzerrte Elaboration als fünftes Postulat. Ursprünglich wird diese Elaborationsart als sechstes Postulat besprochen (Petty & Cacioppo, 1986b).

## 4.3 Informationsverarbeitung innerhalb des Elaborationskontinuums

### 4.3.1 Einflussgrößen auf das Elaborationskontinuum

Die gedankliche Auseinandersetzung wird von Petty und Cacioppo als Kontinuum mit zwei idealtypischen Ausprägungen an den äußeren Enden modelliert, der zentralen (s. Kapitel 4.3.2.1) und der peripheren Route der Informationsverarbeitung (s. Kapitel 4.3.2.2).

„One can view the extent of elaboration received by a message as a continuum going from no thought about the issue-relevant information presented to complete elaboration of every argument, and complete integration of these elaborations into the person's attitude schema.“ (Petty & Cacioppo, 1986b, S. 129)

Sind die situativen wie individuellen Bedingungen günstig, steigt die Elaborationswahrscheinlichkeit, d. h. die Wahrscheinlichkeit sich gedanklich mit der persuasiven Botschaft auseinander zu setzen (Petty et al., 1993). Unter situativen Bedingungen verstehen die Autoren solche Faktoren, die die *Fähigkeit* zur kognitiven Verarbeitung beeinflussen. Individuelle Faktoren wirken sich auf die *Motivation* zur Elaboration aus.

#### 4.3.1.1 Fähigkeit zur Verarbeitung

Im Vordergrund stehen zwei wesentliche Faktoren, die die Fähigkeit zur Auseinandersetzung mit einer persuasiven Botschaft beeinflussen: Ablenkung und Wiederholung. *Ablenkung* setzt die Verarbeitungskapazität herunter (Kahneman, 1973), so dass unter erschwerten Bedingungen dieser Art kaum Gegenargumente formuliert werden können (Festinger & Maccoby, 1964). Es zeigt sich, dass vor allem schwache Argumente bei Ablenkung persuasiv wirken (Petty et al., 1976). Im heutigen Rezeptionsprozess ist „Media Multitasking“ verbreitet und äußert sich als Kombination von audiovisuellen (z. B. Musikstreams, Radio oder TV) und Internetinhalten. Die ablenkende Wirkung bei paralleler Medienrezeption setzt signifikant die Fähigkeit herab, Gegenargumente vorzubringen (Jeong & Hwang, 2012).

Bei *Wiederholung* werden vermehrt kognitive Auseinandersetzungen mit dem Persuasionsversuch registriert, was sich zunächst in Zustimmung äußert, um dann in zunehmende Ablehnung, eine steigende Anzahl an Gegenargumenten und irrelevante Gedanken zu kippen (Cacioppo & Petty, 1979). Starke Argumente profitieren zunächst von jeder Wiederholung und gewinnen an Überzeugungskraft. Mit zunehmender Wiederholung treten jedoch Sättigungseffekte auf und Reaktanzeffekte werden wahrscheinlicher (Brehm, 1966); vor allem schwache Argumente verlieren rasch an Überzeugungskraft, was sich als Ablehnung der persuasiven Botschaft äußert (Cacioppo & Petty, 1979; Petty & Cacioppo, 1986b).

Des Weiteren konnte gezeigt werden, dass *Vorwissen* die Fähigkeit zur Elaboration der persuasiven Botschaft erhöht. Je ausgeprägter das Vorwissen, desto stärker werden dargebotene Argumente einer persuasiven Botschaft reflektiert, durchdacht und bei widersprechenden Positionen abgelehnt. Dagegen übernehmen Probanden ohne Vorwissen die Argumente der Persuasionsbotschaft, selbst wenn diese ihren Einstellungen widerspricht (Wood, 1982). Der Umfang von Vorwissen wird durch Indexbildung aus der Anzahl richtiger Antworten (z. B. Al-Ansari et al., 2003) oder per Selbsteinschätzung („Wie gut fühlen Sie sich zum Thema informiert?“) und der Anzahl genutzter Informationsquellen in einem definierten Zeitraum (z. B. Wood, 1982) operationalisiert. Ein direkter Einfluss von vorhandenem Wissen auf Einstellungen und Verhalten bzw. Verhaltensintentionen ist jedoch nicht haltbar, wie Ajzen und Kolleg:innen zeigen (Ajzen et al., 2011).

#### 4.3.1.2 *Motivation zur Verarbeitung*

Das Involvement, das Kognitionsbedürfnis und die persönliche Verantwortung sind individuelle Faktoren, die sich auf die Motivation zur Elaboration auswirken. Unter *Involvement* verstehen Petty und Kollegen (Petty et al., 1983; Petty & Cacioppo, 1981) die persönliche Relevanz (outcome-relevant involvement bei B. T. Johnson & Eagly, 1989) oder die wahrgenommenen Konsequenzen des Einstellungsthemas (*issue involvement*) bzw. -objekts (*product involvement*). In einem Persuasionsversuch<sup>43</sup> wird die persönliche Relevanz z. B. durch zeitliche, lokale/regionale oder situative Dimensionen aktiviert und bewirkt, dass die Bereitschaft sich mit der Botschaft auseinanderzusetzen, steigt: „[...] people are more motivated to devote the cognitive effort required to evaluate the true merits of an issue or product when involvement is high rather than low“ (Petty et al., 1983, S. 137). Diese kognitive Komponente wird mit vorhandenem Wissen noch verknüpft (Wirth, 2006), was nicht nur die Motivation erhöht, sich mit der persuasiven Botschaft auseinanderzusetzen, sondern sogar die Fähigkeit, da die Versuchsperson in die Lage versetzt wird, bei disparaten Einstellungen Gegenargumente zu formulieren bzw. bei konsistenten Einstellungen die Argumentstärke zu prüfen (Petty & Cacioppo, 1986b).

Das *Kognitionsbedürfnis* (*Need for Cognition, NFC*) wird definiert als „[...] an individual's tendency to engage in and enjoy effortful cognitive endeavors“ (Cacioppo et al., 1984, S. 306) und ist damit als zeitlich überdauerndes interindividuelles Persönlichkeitsmerkmal konzipiert (Cacioppo & Petty, 1982). Die Autoren weisen nach, dass das NFC-Konstrukt nicht die generelle kognitive Fähigkeit (im Sinne des Intelligenzquotienten o. ä.) misst, sondern die

---

<sup>43</sup> Schönbach (2009) zeigt, dass nicht jede Persuasionsbotschaft die erwünschte Wirkung erzielt, sondern lediglich die Wahrscheinlichkeit dafür erhöht.

Motivation und das Bedürfnis zur vertieften Informationsverarbeitung (Cacioppo & Petty, 1982). Wer gerne und ausdauernd kognitiv anspruchsvolle Denkanstrengungen unternimmt, Argumente abwägt und zu einer begründeten Entscheidung kommt, hat ein ausgewiesenes Kognitionsbedürfnis. Wer hingegen wenig motiviert ist, sich eher auf einfache Heuristiken verlässt oder sich nur unter bestimmten Voraussetzungen kognitiven Herausforderungen widmet, weist einen niedrigen NFC-Wert auf (Cacioppo & Petty, 1982; Petty et al., 2009). Aus den skizzierten Eigenschaften des NFC-Konstruktes lässt sich die Relevanz innerhalb des ELM-Theoriegebäudes klar ableiten. Im Gegensatz zu einem niedrigen Kognitionsbedürfnis, führt ein hohes Kognitionsbedürfnis zu einer eindeutigeren Unterscheidung zwischen überzeugenden und schwachen Argumenten, einer verbesserten Erinnerungsleistung und einer intensiveren gedanklichen Auseinandersetzung mit der persuasiven Botschaft (Cacioppo et al., 1983).

Ganz allgemein wirkt sich die Wahrnehmung *persönlicher Verantwortung* auf die Motivation des Individuums im Gruppenkontext Leistungen zu erbringen aus. Der Begriff entstammt dem sozialpsychologischen Phänomen des Social Loafings (*soziales Faulenzen*) und wird definiert als "tendency to reduce one's effort when working collectively compared with coactively on the same task" (Karau & Williams, 1993, S. 683). Bei kollektiver Zusammenarbeit wird das Ergebnis der Gruppe bewertet, während ko-aktive Zusammenarbeit die Einzelleistungen der Gruppenmitglieder berücksichtigt (Feuchter & Funke, 2004). Eine Vielzahl von Variablen beeinflusst die wahrgenommene persönliche Verantwortung und in Folge dessen die erbrachte Leistung, wozu u. a. die Schwierigkeit der Aufgabenstellung, die wahrgenommene Bedeutsamkeit der Aufgabe, die wahrgenommene Gruppengröße etc. gehören (Harkins & Petty, 1982; Karau & Williams, 1993). So liegen die individuellen Leistungen bei einfachen Aufgaben und wahrgenommener Arbeitsteilung unter denen von Versuchspersonen, die schwierige Aufgaben alleine erfüllen müssen (Harkins & Petty, 1982). Relevant ist die wahrgenommene persönliche Verantwortung auch im Persuasionskontext und wirkt sich auf die Motivation zur Elaboration von Argumenten im Persuasionsversuch aus. In einem Experiment von Petty und Kollegen sollen Versuchspersonen einen von drei Redaktionsbeiträgen (starke, schwache und sehr schwache Argumente) entweder als Teil einer Gruppenleistung oder in alleiniger Verantwortung bewerten (Petty et al., 1980). Die Autoren zeigen, dass die individuelle kognitive Auseinandersetzung mit starken Argumenten signifikant stärker und deren Bewertung positiver ist als in den beiden anderen Bedingungen.

Die Fähigkeit und Motivation zur gedanklichen Auseinandersetzung bestimmen maßgeblich die eingeschlagene Route der Informationsverarbeitung. Sind Fähigkeit und/oder Motivation

hoch, setzen sich Menschen intensiv mit der Persuasionsbotschaft auseinander und verarbeiten diese auf der zentralen Elaborationsroute. Sind hingegen Fähigkeit und/oder Motivation niedrig, spielen eher periphere Hinweisreize bei der Informationsverarbeitung eine Rolle (Petty & Cacioppo, 1986a).

### 4.3.2 Zwei Verarbeitungspfade

Die beiden Pfade der Informationsverarbeitung unterscheiden sich hinsichtlich der Tiefe der kognitiven Auseinandersetzung und der Berücksichtigung von peripheren Hinweisreizen und damit einhergehend der Konsequenzen, die sich aus dem Persuasionsprozess ergeben. Dabei kommt es zu einer gegenläufigen Austauschbeziehung zwischen vertiefter Elaboration und unberücksichtigten Hinweisreizen einerseits beziehungsweise oberflächlicher Elaboration und einer gestiegenen Berücksichtigung von Hinweisreizen andererseits.

#### 4.3.2.1 Zentrale Route

Als zentrale Route der Informationsverarbeitung wird der äußere Pol des Elaborationskontinuums bezeichnet, welcher eine erhöhte Wahrscheinlichkeit einer umfangreichen kognitiven Auseinandersetzung mit einem Persuasionsversuch nach sich zieht: „[...] that which likely resulted from a person’s careful and thoughtful consideration of the true merits of the information presented in support of an advocacy (Petty & Cacioppo, 1986b, S. 125). Als notwendige Bedingungen zur zentralen Verarbeitung werden die Fähigkeit und/oder Motivation zur gründlichen Informationsverarbeitung gesehen. Während Ablenkung und die Wiederholung (schwacher Argumente) den kognitiven Prozess stören, sind hohe Ausprägungen von Involvement, Kognitionsbedürfnis und persönlicher Verantwortung förderliche Faktoren um die Elaborationswahrscheinlichkeit zu erhöhen (Petty & Briñol, 2012; Petty & Cacioppo, 1986b).

Fallen die Bedingungen günstig aus, sind also Motivation und Fähigkeit zur Elaboration vorhanden, werden große Anstrengungen bei der Auseinandersetzung mit einstellungsrelevanten Argumenten unternommen („Die IT-Branche benötigt eine Frauenquote, um wettbewerbsfähig zu werden.“), indem Argumente mit vorhandenem Wissen und Einstellungssets abgeglichen („Frauen sind gut qualifiziert und sollten Karriere machen.“), auf Logik und Stichhaltigkeit geprüft („Studienergebnisse zeigen, dass diverse Teams zum Unternehmenserfolg führen.“) und bei positiver Begutachtung („Die Frauenquote ist sinnvoll.“) schließlich in das Einstellungsschema integriert werden (Petty & Briñol, 2012). In der ELM-Logik lässt sich folgern, dass Persuasionsversuche, die negative Gedanken auslösen („Eine Frauenquote ist nicht sinnvoll, weil nicht das Geschlecht, sondern Qualifikation und berufliche Kompetenz für

die Stellenbesetzung ausschlaggebend sind.“), Einstellungen kaum oder nicht ändern bzw. beeinflussen können.

Die Valenz der Rezeptionsgedanken lässt sich also auf zwei Einflussfaktoren zurückführen, und zwar auf (1) einstellungsrelevante Argumente und (2) deren Argumentstärke, also deren Qualität im Sinne der Überzeugungskraft (O'Keefe, 2008, 2012; Petty & Cacioppo, 1986b). Enthält die Persuasionsbotschaft *einstellungsrelevante Argumente*, die passend in die kognitive Struktur integriert werden können, werden positive Gedanken hervorgerufen. Entsprechend resultieren Argumente, die nicht einstellungskonform sind, in negativen Rezeptionsgedanken, sodass der Persuasionsversuch misslingen kann (Petty & Cacioppo, 1986b). Zum jetzigen Stand der Forschung fehlen geprüfte Erkenntnisse zur Valenz der Gedanken nach der Rezeption von (sexistischem) Hate Speech bei Menschen mit stark bzw. schwach ausgeprägten dazugehörigen Einstellungen. Überspitzt formuliert: Sind Sexist:innen positiv gestimmt, wenn sie sexistisches Hate Speech rezipieren?

Die *Argumentstärke* ist bei Petty und Cacioppo als empirische Größe konzipiert und nicht theoretisch fundiert (Petty & Cacioppo, 1986b, S. 133). Dazu generieren sie zunächst eine möglichst hohe Anzahl an überzeugenden wie fadenscheinigen Argumenten zu einem Einstellungsobjekt und lassen sie in einem Pretest hinsichtlich ihrer Überzeugungskraft bewerten. Auf dieser Bewertungsgrundlage werden starke und schwache Argumente unter Experimentalbedingungen getestet. Die empirische Messung der kognitiven Antwort erfolgt mittels Thought-Listing-Technik, bei der die Versuchspersonen aufgefordert werden, alle Rezeptionsgedanken mündlich oder schriftlich zu äußern (Cacioppo & Petty, 1981). Sowohl starke wie auch schwache Argumente werden auf der zentralen Route gründlich durchdacht, aber während starke Argumente überzeugen und positive Gedanken bewirken, sind Einwände die Konsequenz von schwachen Argumenten und münden in negativen Gedanken. Positive Gedanken werden als geglückte Persuasion bewertet, da die Versuchsperson den Aussagen der Persuasionsbotschaft zustimmt. Bei negativen Gedanken ist hingegen keine Einstellungsänderung zu erwarten. Darüber hinaus gilt die Anzahl der Rezeptionsgedanken als, wenn auch nur korrelativer, Gradmesser für die Verarbeitungstiefe (Petty & Cacioppo, 1986b).

Während die vorhergehenden Abschnitte eine objektive Verarbeitung, d. h. im Elaborationskontinuum entlang der Argumentationslinie der Persuasionsbotschaft skizziert haben, soll es im Folgenden um die verzerrte Verarbeitung im Persuasionsversuch gehen. Die ersten vier Postulate gehen von einer idealtypischen Wirkungsweise im Persuasionsprozess aus, bei der starke oder schwache Argumente eine hohe bzw. niedrige Elaboration zur Folge haben (Klimmt, 2011; Petty & Cacioppo, 1986b). Abseits von Laborbedingungen sind solche

schematische Ursache-Wirkungsbeziehungen jedoch selten anzutreffen, nicht nur weil die Argumentstärke selten homogen ist (Klimmt, 2011), sondern weil Moderatorvariablen in den Verarbeitungsprozess eingreifen und in der Folge die Persuasionswirkung verzerren, also abweichend von der Argumentationslinie der Persuasionsbotschaft wirken. Vorwissen und Voreinstellungen gehören zu den rezipientenbezogenen Moderatorvariablen, die die Richtung der Verarbeitung (im Sinne einer Weichenstellung für die Valenz der Gedanken) beeinflussen. Diese können durch Vorwarnungen bzw. Vorabinformationen zum Botschaftsinhalt aktiviert werden und führen dazu, dass im Persuasionsversuch Argumente, die bestehenden Einstellungen widersprechen, besonders kritisch geprüft werden (Petty & Cacioppo, 1986b). Liberman und Chaikens Studie weist diesen Effekt nach: Weibliche Studierende bewerteten Gesundheitsrisiken für Frauen im Zusammenhang mit Kaffeegenuss deutlich defensiver, wenn sie selbst Kaffeetrinkerinnen waren (Liberman & Chaiken, 1992). Obwohl sie sich intensiv mit der persuasiven Botschaft während des Experiments auseinandersetzten, verarbeiteten sie die Argumente nicht objektiv, in dem Sinne, dass sie sich von ihnen überzeugen ließen (ebd.). Auch einfache Hinweisreize (z. B. Enthusiasmus der Publikumsreaktion) wirken sich im Persuasionsprozess verzerrend aus (Axson et al., 1987; Erb et al., 2007). Vergleichbare Befunde hinsichtlich einer ausgeprägten verzerrten Informationsverarbeitung liefern verwandte theoretische Zugänge, die auch in der Kommunikationswissenschaft hohe Relevanz besitzen, wie z. B. Effekte, die als Confirmation Bias (Nickerson, 1998) bezeichnet werden, wie die Theorie kognitiver Dissonanz (Festinger, 1957) oder die Selective-Exposure-Hypothese (Stroud, 2008).

An Petty und Cacioppo's Konzeption der Argumentqualität wurde in der wissenschaftlichen Community viel Kritik geübt. Insbesondere Daniel O'Keefe hat sich ausführlich (O'Keefe, 1999, 2003, 2004, 2008) mit der Stärke von Argumenten als Variable innerhalb des ELMs beschäftigt und kommt zu dem Schluss: „In ELM research, the nature of argument quality (or argument strength) has not been a focus of explicit attention, because argument quality variations have been defined in terms of persuasive effects.“ (O'Keefe, 2012, S. 142). Folgt man O'Keefe's Auslegung, ist die Argumentqualität einem Zirkelschluss gleich definiert, in dem Sinne, dass „ein Argument stark ist, weil es persuasiv ist und es persuasiv ist, weil es stark ist.“ (vgl. ebd.). Petty und Cacioppo's Operationalisierung eines Prädiktors an seiner Wirkung (O'Keefe, 2003, S. 267) wird als nicht überzeugend angesehen, weswegen eine theoriegeleitete Herangehensweise zwingend erscheint. In der vorliegenden Arbeit wurde weiter oben Hate Speech eine persuasive Kraft zugesprochen, weswegen sich nicht nur die normative Frage stellt, ob man im Kontext von Hate Speech von starken bzw. schwachen Argumenten sprechen kann. Zur Erinnerung: Starke Argumente sind solche, die eine Vielzahl an (1)

zusätzlichen und (2) befürwortenden Gedanken zu denen im Persuasionsversuch genannten, auslösen (Petty & Cacioppo, 1986b, S. 133). Ausführlich wurde in Kapitel 2.1.4.3 begründet, dass Hate Speech bei sekundären Adressaten verzerrte Reaktionen hervorruft: Während Anhänger:innen und Sympathisant:innen Hate Speech als einstellungskongruent und wünschenswert wahrnehmen und somit den Kommunikationsinhalt befürworten, reagieren unbeteiligte Rezipierende weniger eindeutig. Dieses Publikum zeigt zwar mehrheitlich Ablehnung und Empörung gegenüber Hate Speech, so dass der Persuasionsversuch im Sinne des ELM als gescheitert bewertet würde, andererseits konnten mögliche Persuasionseffekte durch inzivile Kommunikation (Anderson et al., 2014; Ziegele et al., 2018) für diese Rezipierendengruppe nachgewiesen werden. Es erscheint angemessen, Widerspruch erregende Botschaften wie Hate Speech, nicht im Sinne starker/schwacher Argumente zu konzipieren, sondern für die empirische Untersuchung eine Variation auf Basis inhaltlicher Erwägungen vorzunehmen (vgl. Kapitel 5).

Wie wirken sich persuasive Botschaften aus, sofern sie objektiv (d. h. entlang der Argumentationslinie) verarbeitet werden? Petty und Kollegen greifen 1995 die Frage auf, um die Konsequenzen des Persuasionsprozesses auf die Einstellungsstärke darzulegen (Petty et al., 1995). Demnach wird die Einstellungsstärke hinsichtlich ihrer Persistenz und Resistanz sowie daraus abgeleiteten Verhaltensprognosen bewertet. Die *Persistenz* der Einstellung greift den zeitlichen Aspekt auf, im Sinne einer temporalen Beständigkeit: „*persistence* refers to the extent to which a newly changed attitude endures over time even if it is never attacked directly“ (Petty et al., 1995, S. 100). Persistente Einstellungen bleiben also im Zeitverlauf bestehen und unterliegen nicht kognitiven Verlusten (z. B. Vergessen). Die *Resistenz* akzentuiert die Widerstandsfähigkeit gegenüber Persuasionsversuchen „*resistance* refers to the ability of an attitude to hold up to an explicit attack“ (ebd.). Persistenz und Resistanz werden als disjunkte Effekte verstanden, welche nicht notwendigerweise gemeinsam auftreten, d. h. Einstellungen können resistent gegenüber Gegenargumenten sein, aber nach einiger Zeit verblassen. Zur Veranschaulichung dienen folgende Beispiele: Einstellungen gegenüber Bildungschancen für Schülerinnen und Schüler sind in Deutschland sowohl persistent als auch resistent (Stocké, 2005). Einstellungen (insbesondere deren kognitive Aspekte, also Wissen, Vorstellungen Bewertungen etc.) gegenüber gesunde Ernährung sind persistent, aber nicht resistent gegenüber fett- und zuckerreichen Verlockungen: Dass Hamburger, Pizza und Kekse ungesund sind, wissen die meisten Befragten, aber situative und interindividuelle Faktoren führen zur Wahl diätetisch fragwürdige Lebensmittel (Cason & Wenrich, 2002).

Als weitere Folge einer vertieften kognitiven Auseinandersetzung mit der Persuasionsbotschaft betrachten Petty und Kollegen Verhaltensprognosen, die konsistent mit der Botschaft sind (Petty et al., 1995). Diese wird als Elaboration-Konsistenz-Hypothese bezeichnet und gilt sowohl für bestehende Einstellungen als auch für neue Einstellungen. Damit knüpfen Petty und Kollegen an eine reiche Forschungstradition, wie beispielsweise Ajzens und Fishbeins *Theory of Reasoned Action* (Ajzen & Fishbein, 1980) und Fazios *Einstellungs-Verhaltens-Modell* (Fazio, 1986; Fazio & Roskos-Ewoldsen, 2005), die sich extensiv der Frage widmen, wie Einstellungen Verhalten determinieren. Einstellungen, die auf der zentralen Route gebildet wurden, prognostizieren Verhalten zuverlässig, sind konsistent zum vorhandenen Einstellungsschema, persistent (in zeitlicher Hinsicht) und dazu resistent gegenüber anderslautenden persuasiven Botschaften (Haugtvedt & Petty, 1992; Petty et al., 1995). Für die periphere Route gilt dies nicht, wie im nächsten Abschnitt dargelegt wird.

#### 4.3.2.2 *Periphere Route*

Auf der peripheren Route der Informationsverarbeitung werden persuasive Botschaften weniger gründlich, sondern eher oberflächlich, heuristisch verarbeitet. Das bedeutet, dass nur eine geringe Auseinandersetzung mit Argumenten stattfindet und zwar in dem Sinne, dass nicht das Für und Wider abgewogen, sondern wenn überhaupt, dann eher die Anzahl der Argumente denn deren inhaltliche Aussage wahrgenommen wird (Petty et al., 1983; Petty & Cacioppo, 1984a). Ursächlich werden eine geringe Motivation oder Fähigkeit zur Verarbeitung einstellungsrelevanter Argumente angeführt (Petty et al., 1997; Petty & Cacioppo, 1986b). Zu den Elaborationshemmnissen zählen Ablenkung als motivationaler Faktor (Petty et al., 1976) und eine geringe persönliche Relevanz des propagierten Einstellungsobjekts (Petty et al., 1983), ein eher geringes Kognitionsbedürfnis (Cacioppo et al., 1983) sowie keine oder nur geringe wahrgenommene persönliche Verantwortung (Harkins & Petty, 1982) als Faktoren, die die Fähigkeit zur Elaboration beeinflussen (s. ausführlich Kapitel 4.3.1.). Wenn also Motivation und Fähigkeit zur Elaboration fehlen, werden eher kontextuelle Faktoren, sogenannte Hinweisreize, verarbeitet. Statt sich also mit den propagierten Argumenten der Persuasionsbotschaft (z. B. „Warum wir einen höheren Frauenanteil in der IT-Branche benötigen“) auseinanderzusetzen und eigene Argumentationsketten zu aktivieren, kommen mentale Abkürzungen (z. B. in Form von geschlechterstereotypen Vorurteilen „Frauen bringen Wärme und Empathie zu den gefühlskalten Nerds.“) oder einfache Entscheidungsregeln (z. B. „Wenn so viele Gründe genannt werden, ist das Einstellungsthema bedeutsam/richtig/gut.“) zum Tragen. Auch können Emotionen, die während der Rezeption entstehen, Einfluss auf die Persuasionssituation nehmen (Petty & Briñol, 2015), indem die Persuasionsbotschaft weniger sorgfältig verarbeitet wird (Tiedens & Linton, 2001).

Aktiviert werden die Heuristiken von salienten peripheren Hinweisreizen (Chaiken & Eagly, 1983), welche im Kontext des Persuasionsversuches auftreten und sich zwei unterschiedlichen Klassen zuordnen lassen, namentlich Merkmalen der Botschaft (*message cues*) und Merkmalen des Kommunikators oder der Kommunikatorin (*source cues*). Im Falle einer geringen Elaborationswahrscheinlichkeit konnten Petty und Cacioppo zeigen, dass die Anzahl von Argumenten als *botschaftsimmanenter Hinweisreiz* wirkt, und zwar unabhängig von der Argumentqualität (Petty & Cacioppo, 1984a). Die klassischen Experimentalbedingungen sind derart aufgebaut, dass Studierende drei bzw. neun Argumente unterschiedlicher Qualität (starke vs. schwache Argumente) lesen, die die Einführung von Prüfungsänderungen befürworten. Um die Elaborationswahrscheinlichkeit zu variieren, wird der zeitliche Rahmen der diskutierten Prüfungsänderungen in die nahe Zukunft (hohes Involvement) oder ferne Zukunft (niedriges Involvement) versetzt. Gering involvierte Studierende lassen sich danach stärker von vielen Argumenten überzeugen und stimmen ihnen zu, ohne auf die Argumentqualität zu achten (Petty & Cacioppo, 1984a).

Persuasive Effekte, die aus der peripheren Verarbeitung von Hinweisreizen, wie Glaubwürdigkeit, Attraktivität, Sympathie oder Expertise des Kommunikators resultieren (Axson et al., 1987; Chaiken & Stangor, 1987; Eagly et al., 1978; Heinbach et al., 2018; Petty et al., 1981; Petty et al., 1983; Petty et al., 1997; Petty & Briñol, 2012; Petty & Cacioppo, 1984b; Priester & Petty, 1995; Winter & Krämer, 2014), gehören zu den intensiv untersuchten *Merkmalen des Kommunikators* respektive *der Kommunikatorin*. Deren Zugehörigkeit zu einer (wahrgenommenen) Minderheit bzw. Mehrheit wird zwar seltener als peripherer Hinweisreiz diskutiert, aber die Meta-Analyse von Wendy Wood und ihren Kolleginnen zeigt, dass persuasive Effekte von Minderheiten bzw. Mehrheiten in bestimmten Situationen mit erhöhter Wahrscheinlichkeit auftreten (Wood et al., 1994). So treten Minderheiteneffekte in privaten Kontexten ohne direkten Bezug zum Einstellungsobjekt verstärkt auf, d. h. Versuchspersonen nehmen die Position eines Minderheitenkommunikators eher an oder stimmten ihr zu, wenn sie dies nicht öffentlich bekunden (müssen) und wenn die kommunizierte Position (z. B. Karrierechancen für Frauen in der IT-Branche) nur indirekt mit dem Einstellungsobjekt (hier: Sexismus) verwandt ist.

Die klassischen Versuchsbedingungen sehen die binäre Manipulation der Hinweisreize in Verbindung mit interessierenden weiteren Einflussgrößen (z. B. Involvement, Argumentqualität) vor, um den persuasiven Effekt peripherer Hinweisreize zu messen (Petty et al., 1995). Eine niedrige Elaborationswahrscheinlichkeit begünstigt die Verarbeitung peripherer Hinweisreize und lässt die Argumentstärke unberücksichtigt (Priester & Petty, 1995). In einer Studie

zeigen Johnston und Coolen, dass Versuchspersonen mit geringem Involvement (womit eine geringe Elaborationswahrscheinlichkeit verbunden ist) stereotyp-widerlegende Aussagen über Altenheim-Bewohnerinnen (z. B. „Seniorinnen kleiden sich stets modisch.“) eher übernehmen, wenn der Kommunikator bzw. die Kommunikatorin als glaubwürdige Quelle wahrgenommen werden (Johnston & Coolen, 1995). Hoch Involvierte berücksichtigen die Qualität der Argumente und lassen sich von der Quelle nicht beeinflussen (ebd.).

Petty und Kollegen zufolge sind Einstellungsänderungen, die auf der peripheren Route erzielt werden, also durchaus möglich, auch wenn sie sich deutlich von zentral verarbeiteten Einstellungen unterscheiden (Petty et al., 1997; Petty & Cacioppo, 1986b). Was nämlich Persistenz und Resistanz betrifft, geht eine niedrige Elaborationswahrscheinlichkeit mit einem potentiellen Verlust der gewonnenen Einstellungen einher, d. h. bereits nach kurzer Zeit ist der Persuasionseffekt abgeklungen bzw. beeinflussbar durch Gegenargumente (Haugtvedt & Petty, 1992; Petty et al., 1995). Ähnliches gilt für Verhaltensprognosen, da aufgrund der peripheren Verarbeitungsweise keine konsistenten Persuasion-Elaboration-Verhaltenskette erzeugt werden (Cacioppo et al., 1986; Petty et al., 1995).

Zusammenfassend sind periphere Verarbeitungsprozesse immer dann erwartbar, wenn Motivation und Fähigkeit zur Elaboration fehlen bzw. wenn Hinweisreize salient sind. Auch wenn Persuasionseffekte instabiler sind im Vergleich zu solchen, die auf der zentralen Route erzielt wurden, dient geringe Elaboration als wertvolle „Entlastungsstrategie“ (Klimmt, 2011, S. 44), um die geistigen Ressourcen zu schonen und den Fokus auf relevante Themen, Inhalte und Situationen zu richten.

### 4.3.3 Multiple Rollen von Variablen

“Although the theory is probably best known for its central and peripheral routes to persuasion [...] its most powerful aspect is the specification of a small number of mechanisms by which any given variable can produce attitude change” (Briñol & Petty, 2012, S. 297). Wenn jede Variable Einstellungsänderungen zu bewirken vermag, können dann einzelne Variablen der zentralen oder der peripheren Verarbeitungsrouten zugeordnet werden? Diesen Wunsch nach Eindeutigkeit lehnen Petty und Cacioppo ab und führen stattdessen das dritte Postulat ins Feld, welches *multiple Rollen von Variablen* diskutiert (Petty & Cacioppo, 1986b). Ihrer Argumentation zufolge ist nicht nur der Wesenseigenschaft des Elaborationskontinuums<sup>44</sup> zu verdanken, dass keine Liste peripher oder zentral verarbeiteter Variablen existiert, sondern

---

<sup>44</sup> Die periphere und die zentrale Route stellen die idealtypischen Extrema des *Elaborationskontinuums* dar.

auch der kontextabhängigen Persuasionssituation (Petty et al., 1993; Petty & Cacioppo, 1986b).

Die Autoren gehen davon aus, dass Variablen im Persuasionsprozess drei unterschiedliche Rollen annehmen können, und zwar indem sie (1) als Argument oder (2) als peripherer Hinweisreiz dienen, oder (3) indem sie das Ausmaß (i. S. der Verarbeitungstiefe) und die Richtung (objektiv vs. verzerrt) der Elaboration beeinflussen (Petty & Cacioppo, 1986b). Anhand eines Beispiels verdeutlichen die Autoren ihre These: Eine schöne Landschaft kann in einem Werbespot für eine Automarke als peripherer Hinweisreiz dienen, da sie nicht im Zusammenhang mit Eigenschaften des Fahrzeugs steht. In einem Reisekatalog hingegen dient die Schönheit der Landschaft interessierten Urlaubswilligen als Argument, die beworbene Reise zu buchen (Petty et al., 1993). Das angeführte Beispiel lässt sich noch weiter ausführen: Ist die abgebildete Landschaft im Reisekatalog persönlich relevant („Urlaub in der Toskana“), dürfte sich die kognitive Auseinandersetzung mit den Botschaftsinhalten zusätzlich erhöhen (objektive Verarbeitung auf der zentralen Route). Eine verzerrte Verarbeitung ist dagegen zu erwarten, wenn die Reise von einem persönlich nicht sympathisch wahrgenommenen Testimonial beworben wird.

Die intendierte Wirkung, die von einer persuasiven Botschaft ausgehen soll, kann unter Maßgabe des Zusammenwirkens aller Variablen (Merkmale der Botschaft, der Rezipierenden, des Kanals etc.) erreicht werden. Absendende der persuasiven Botschaft können die Wahrscheinlichkeit für das Eintreten der intendierten Wirkung erhöhen, indem sie ein erfolgsversprechendes Variablenset (überzeugende Argumente, Statistiken, optisch ansprechende Gestaltung, sympathische:r Expert:in als Sprecher:in etc.) zusammenstellen (Felser, 2015; Schenk, 2007). Neue Kommunikationsumgebungen, wie z. B. SNS ermöglichen die Prüfung, Neubewertung sowie ggf. Anpassung der bisherigen theoretischen Annahmen und Befunde des ELM, wie im folgenden Kapitel diskutiert wird.

Vielfach wurde Kritik an der Annahme geübt, wonach Variablen im Persuasionsprozess unterschiedliche Rollen annehmen, da hier eine Beliebigkeit in der Interpretation der Ergebnisse zu befürchten ist, die Operationalisierung der Argumentqualität fragwürdig erscheint bzw. nur solche Ergebnisse veröffentlicht werden, die dem Rahmenmodell entsprechen (Bitner & Obermiller, 1985; B. T. Johnson & Eagly, 1989; Kitchen et al., 2014; Mongeau & Stiff, 1993; O'Keefe, 2012; Stiff & Boster, 1987). Christoph Klimmt bemängelt Petty und Cacioppo's Umgang mit der akademischen Kritik: „Die Autoren sehen in den (unvermeidlichen) Kritikpunkten anderer Wissenschaftler keine fundamentalen Probleme und haben die inhaltliche Substanz des Modells bis heute nicht angetastet“ (Klimmt, 2011, S. 95).

#### 4.4 Persuasion von Hate Speech in sozialen Netzwerkseiten

Die theoretische Einordnung des Hate Speech-Begriffs in Kapitel 2 hat gezeigt, dass ein bestimmter Kommunikationsraum nicht konstitutiv für das Verständnis dieser Kommunikationsform ist und Hate Speech damit sowohl in Online- als auch in Offline-Räumen stattfindet. Im folgenden Abschnitt soll die Persuasivität von Hate Speech (vgl. auch Kapitel 2.1.4.3) aus Perspektive der Forschung zu computervermittelter Kommunikation (cvK) erörtert werden, um eine integrative Sicht für den Fortgang der Arbeit zu erzielen.

SNS sind seit den späten 1990er Jahren weit verbreitet (Boyd & Ellison, 2007). Wurden sie früher als Nischenphänomen für Jugendliche bewertet, sind derzeit weltweit 3,8 Milliarden Menschen aktive Nutzende. Allein in Deutschland nutzen im Januar 2020 rund 45 Prozent der Bevölkerung SNS aktiv, rund jede:r zweite Deutsche ist Mitglied mindestens einer SNS, also insgesamt 42,8 Millionen Nutzende (Statista, 2020). Facebook ist die meistgenutzte Plattform und wird in Deutschland von 31 Prozent in der Bevölkerung mindestens einmal wöchentlich genutzt (Beisch et al., 2019; Frees & Koch, 2018). Niemann (2016) unterscheidet zwischen freizeit-orientierten (Facebook, Twitter, Snapchat) und beruflich-orientierten (LinkedIn, Researchgate) SNS, wobei es vielfältige Überschneidungen gibt, da Nutzende plattformübergreifend ihre Inhalte teilen, oder die Nutzungsweisen sich sowohl der Freizeit wie auch dem Beruf zurechnen lassen (z. B. Facebooknutzung durch Wissenschaftler:innen kann der Prokrastination wie auch der beruflichen Vernetzung dienen).

SNS lassen sich der Klasse der beziehungs-orientierten Social Media-Plattformen zuordnen (Springer et al., 2014). Eigenschaften, die als konstituierend für SNS festzulegen sind, stellen jedoch eine Herausforderung dar, denn SNS unterliegen einem steten soziotechnologischen Wandel von Nutzungspraktiken sowie aktualisierter technologischer Implementierungen, was eine Abgrenzung zu anderen Social Media-Anwendungen (z. B. zu inhalts-orientierten Plattformen wie YouTube) erschwert und Übergänge fließend erscheinen (Ellison & Boyd, 2013). Generell lässt sich sagen, dass SNS Plattformen sind, die ihren Nutzenden die Abbildung ihrer bestehenden Netzwerke (Freundschafts- oder Berufsnetzwerke etc.) ermöglichen (sofern jene Nutzende der SNS sind), ergänzt um neue Kontakte (Privatpersonen oder professionelle Kommunikator:innen in ihren beruflichen Rollen, also Politiker:innen, Journalist:innen; des Weiteren Unternehmen, Marken, Organisationen etc.) (Boyd & Ellison, 2007). „The desire to communicate and share content is a primary driver of SNS use“ (Ellison & Boyd, 2013, S. 159). Wichtiger Treiber für die Nutzungsweise sind individuelle Impression Management-Strategien (Marder et al., 2016). Entsprechend erstellen die Nutzenden selbst Inhalte (Postings) oder teilen Inhalte anderer Nutzer:innen (z. B. Fotos, Videos), sie nutzen Inhalte

externer Content-Provider (z. B. journalistische Beiträge, um sich zu einem Thema zu äußern) oder kreieren durch Mash-up-Techniken neue Inhalte (z. B. Memes) aus bereits produziertem Material (Springer et al., 2014). Um die Inhalte auffindbar und kategorisierbar zu machen, können diese zusätzlich mit Stichworten (*hashtags*) verschlagwortet (getaggt) werden (ebd.). Die Gesamtheit der Inhalte aller Kontakte von Nutzenden wird im Newsfeed veröffentlicht und erlaubt eine Vielzahl von Interaktionen mit den Inhalten, in Form von Kommentaren, Zustimmungsfeatures (z. B. Gefällt-mir-Markierungen) oder Funktionen, die das Teilen von Inhalten anderer Nutzer:innen ermöglichen. Dazu wird eine Vielzahl von Kommunikationsarten und -modi eingesetzt:

- privat in Form von Direct Messages (private Nachrichten zwischen Nutzenden),
- öffentlich als Kommentar oder Reply unter einem Posting, was je nach Reichweite und Privatsphäreneinstellungen halböffentlich (innerhalb des persönlichen Netzwerks oder einer Teilmenge dessen) oder öffentlich und damit sichtbar für eine (potenziell) große Anzahl von Nutzenden ist,
- synchron oder asynchron
- wechselseitig (dialogisch) oder einseitig (one-to-many), v.a. bei professionellen Kommunikator:innen (Niemann, 2016).

Generell lassen sich sämtliche oben beschriebenen Merkmale, Modi und Funktionen von SNS für die Produktion von Hate Speech ebendort ausnutzen. Nutzerprofile sind schnell erstellt oder lassen sich als automatisierte Programme in Form von Social Bots einsetzen (Frischlich et al., 2017). Unabhängig davon, ob es sich um menschliche oder automatisierte Nutzerprofile handelt, können sämtliche Interaktionen von Hate Speech durchdrungen sein: das Erstellen und Teilen von Postings, Kommentaren, Privatnachrichten, Hashtags oder die Nutzung von Zustimmungsfeatures (Döring & Mohseni, 2020; Jane, 2014; Kreißel et al., 2018). Konzentrierte Hate Speech-Aktionen von menschlichen Nutzenden und Social Bots setzen auf Wiederholung, Kontinuität und Konsonanz herabwürdigender Kommunikationsinhalte, um eine möglichst weitreichende Wirkung bei primären wie sekundären Adressat:innen zu erreichen (Frischlich et al., 2019). Angesichts der Auslegung der Programmarchitektur von SNS auf einen hohen Emotionalisierungsgrad, um durch Interaktion mit emotionalisierenden Inhalten die Reichweite zu maximieren und das Reaktionsvolumen zu vergrößern, bedeuten dies günstige Bedingungen für die Verbreitung von Hate Speech als hoch emotionalisierendes Kommunikationsphänomen.

Im Rahmen dieser Arbeit wird für Hate Speech in SNS ein großes Persuasionspotential angenommen, da sowohl die Botschaft, die Interaktion durch die Programmarchitektur als auch

bestimmte Nutzungspraktiken dies begünstigen. Dies gilt insbesondere für diejenigen Nutzer:innen, die die Ansichten der Hate Speech Produzent:innen teilen und Hate Speech als geeignete Kommunikationsform ansehen oder zumindest tolerieren. Wie wirkt sich das Persuasionspotential von Hate Speech auf diejenige Nutzer:innengruppe aus, die als sekundäre Rezipient:innen keine Sympathien für Hate Speech und die vertretenen Positionen hegen?

Mit bis zu 90 Prozent ist der Anteil an Lurkern, d. h. Nutzer:innen, die auf eine aktive SNS-Nutzung im oben beschriebenen Sinne verzichten, sehr hoch (van Mierlo, 2014). Diskutiert wird ein Ursachenbündel, das sich individuellen und rezeptionssituativen Faktoren zuordnen lässt, wie z. B. soziodemografische Merkmale (Alter, Geschlecht), ein selbst zugeschriebener fehlender Expert:innenstatus, ein ausreichend befriedigtes Informationsbedürfnis, negative Erfahrungen mit anderen Nutzer:innen, technische Probleme etc. (Preece et al., 2004; Springer, 2014; Sun et al., 2014). Kann Lurking als Sonderform des Social Loafings (s. Kapitel 4.3.1.2) im Kontext von SNS-Communities verstanden werden (Shiue et al., 2010)? Dies hätte Implikationen für die Wahrnehmung, Verbreitung und Bekämpfung von Hate Speech, da Social Loafing die Motivation zur Elaboration von Argumenten hemmt. Aus der Social Loafing-Perspektive würden Lurker auf Hate Speech seltener mit Gegenrede bzw. Befürwortung antworten, weil sie ihre persönliche Verantwortung als gering einschätzen. Karau und Williams (1993) identifizieren in einer meta-analytischen Untersuchung mehrere Antezedenzen für Social Loafing. Dazu zählen eine diffundierende Verantwortlichkeit in einer potenziell großen Nutzergruppe („Eine andere kann sich darum kümmern.“), das Unterlassen einer eindeutigen Positionierung („Wegen diesem unbedeutenden Kommentar lasse ich mich nicht auf eine Diskussion ein.“) oder geringe Erwartungen an eine engagierte Nutzerkultur („Hier engagiert sich niemand gegen Hate Speech.“). Der Unterschied zwischen Lurking und Social Loafing ist offenbar, auch wenn beide Phänomene eine Zurückhaltung bei der aktiven Online-Nutzung unterstellen: Während Lurker sich auf ein zielgerichtetes Rezeptionsverhalten zur Informationsgewinnung beschränken (Han et al., 2014; Preece et al., 2004), beschreibt Social Loafing einen bewussten Verzicht Verantwortung zu üben (Karau & Williams, 1993). Um Konsequenzen dieser Art zu vermeiden, sind Buddy-Systeme<sup>45</sup> im Gespräch, welche die persönliche Verantwortung der Nutzer:innen zur aktiven Teilnahme aktivieren (Bishop, 2007; H.-S. Kim & Sundar, 2014). Dazu gehören Interventionen, wie das Melden von Postings oder der Einsatz von Gegenrede (Benesch, 2014; Schieb & Preuss, 2018), aber es ist bislang ungeklärt, ob die

---

<sup>45</sup> Dabei kann es sich sowohl um Nutzende handeln, die eine Moderatorenrolle einnehmen oder indem Interaktionen zwischen Nutzenden angeregt werden, um an deren soziale Verantwortung innerhalb der SNS zu appellieren.

Zuschreibung persönlicher Verantwortung Nutzer:innen dazu veranlasst, sich gegen Hate Speech zu positionieren.

Theoretische Zugänge zur Kommunikation in SNS sind so zahlreich wie widersprüchlich. Joseph B. Walther (2011) unterscheidet drei Theoriestränge, darunter sogenannte Cues-Filtered-Out-Theorien und Ansätze, die die Wahrnehmung oder Nutzung von cvK im engeren Sinne behandeln, letztere werden jedoch an dieser Stelle ausgeklammert, da sie nicht Wirkungen im engeren Sinne berücksichtigen. *Cues-Filtered-Out Theorien* gehen von Rezeptionswirkungen aus, welche auf das Fehlen von Hinweisreizen in der cvK zurückzuführen und als Theoriegerüst in der Hate Speech-Forschung verbreitet sind (Fox et al., 2015; Kaspar, 2017; McCluskey & Hmielowski, 2011; Reader, 2012; Santana, 2014). Ausgangspunkt ist ein vorherrschender Technikdeterminismus, der das Fehlen von Wärme und emotionaler Nähe, dafür enthemmte und angriffslustige Nutzer begünstigt, wie von der Hypothese des *Lack of Social Context Cues* postuliert (Siegel et al., 1986). Auch wenn SNS durchaus ein reichhaltiges Spektrum an sozialen Hinweisreizen ermöglichen, indem sie zur Verwendung eines Profilbilds und zur Preisgabe persönlicher Informationen (Klarname, Geschlecht, Alter, Bildung, Wohnort etc.) und persönlicher Vorlieben und Interessen anregen (Niemann, 2016), ist Hate Speech in SNS verbreitet (Keipi et al., 2017). Also ist die Idee eines verbesserten, da verständigungsorientierten Kommunikationsoutcomes je mehr Sinnesmodalitäten angesprochen werden gemäß der *Media Richness Theorie* (Daft & Lengel, 1986) eingeschränkt haltbar. Das liegt zum einen daran, dass Nutzende ein Profil mit falschen persönlichen Informationen erstellen (L. R. Smith et al., 2017) oder strenge Privatsphäreneinstellungen ihres echten Profils vornehmen (Christofides et al., 2009), wenn sie mit Nutzenden außerhalb ihres Netzwerks interagieren und zum anderen sind bestimmte Effekte auf psychologische und medienspezifische Faktoren zurückzuführen. Hinweise hierzu lassen sich Sulers *Online Disinhibition Effect* (Suler, 2004) entnehmen, der sechs Faktoren für enthemmtes Online-Verhalten nennt: Anonymität, fehlende Sichtbarkeit des Gegenübers, asynchrone Kommunikationsmodi, solipsistische Introjektion (d. h. Merkmale der anonymen und nicht unmittelbar sichtbaren Online-Diskussionsteilnehmenden werden in das eigene psychische Erleben integriert), dissoziative Vorstellungskraft (Online- und Offline-Welt werden als verschiedene Entitäten wahrgenommen) und schließlich fehlende Autorität. Lapidot-Lefler und Barak (2012) zeigen, dass insbesondere der fehlende Blickkontakt dazu führt, dass Diskussionsteilnehmende auf toxische Weise miteinander kommunizieren.

Bedeutet fehlende Hinweisreize für die Persuasionsforschung, dass eine periphere Verarbeitung ausgeschlossen oder zumindest reduziert ist? Bereits 1983 konnten Shelly Chaiken und

Alice Eagly zeigen, dass einzelne Kommunikationsmodi (text-basiert, audiovisuell und auditiv) einen unterschiedlichen Einfluss auf Persuasion nehmen und textbasierte Botschaften mit höherer Wahrscheinlichkeit zentral verarbeitet werden (Chaiken & Eagly, 1983). Hinweisreize scheinen dagegen nur im Rahmen von Kommunikationsmodi Bedeutung zu gewinnen, in denen sie wahrnehmbar und damit salient werden (ebd.). Forschungsergebnisse von Guadagno und Cialdini (2002, 2007) oder Di Blasio und Milani (2008) deuten ebenfalls darauf hin. In einer Studie diskutieren 108 Studierende in Kleingruppen von bis zu fünf Personen Face-to-Face oder im Chat über den Ausbau der Mensa ihrer Universität (Di Blasio & Milani, 2008). Gegenübergestellt werden zwei Optionen, die Argumente für den Ausbau mit unterschiedlichen Konsequenzen skizzieren. Um einen Einfluss auf die Meinung zu provozieren, präsentieren die Forschenden die gegensätzliche Option zur präferierten Mehrheitsmeinung innerhalb der Gruppe als Beschlussvorlage der Universität. Nur fünf Prozent der Teilnehmer im Chat gegenüber 38 Prozent der Diskussionsteilnehmenden lassen sich von der Verwaltungsentscheidung umstimmen, was die Autor:innen darauf zurückführen, dass soziale Einflussprozesse in Kontexten mit einer Vielzahl von Sinnesmodalitäten stärker sind (Di Blasio & Milani, 2008). Die Befunde lassen jedoch unberücksichtigt, dass es sich hierbei um kurzfristige Effekte handelt (Walther, 2011).

Einen differenzierenden Blick auf Anonymität und fehlende Identifizierbarkeit der Nutzenden nehmen Spears und Lea ein (Spears & Lea, 1994). Das von ihnen konzipierte und im weiteren Verlauf mit Kolleg:innen ausgearbeitete *Social Identity Modell of Deindividuation Effects (SIDE)* macht deutlich, dass diese beiden Faktoren dazu führen, dass sich Menschen im Kontext von cvK an salienten sozialen Kategorien einer Gruppe und den dort vorherrschenden Normen orientieren (Postmes et al., 1998; Postmes et al., 2001; Reicher et al., 1995). Da der wahrgenommene Anonymitätsgrad verhältnismäßig hoch ist, werden soziale Kategorien bedeutsam, welche die Forenmitglieder/SNS-Kommentierenden vereint, d. h. also, dass Fans polnischer Trickfilme, Bahnliebhaber etc. sich selbst und ihr Gegenüber im Trickfilm- oder Bahnliebhaber-Forum als Mitglieder einer sozialen Gruppe mit ähnlichen Interessen wahrnehmen und einander prinzipiell positiv zugetan sind. Erwartbar ist z. B., dass junge Mütter in einem Online-Forum oder in einer Facebook-Gruppe sich unterstützen und hilfsbereit interagieren, da sie sich der sozialen Gruppe der Mütter zugehörig fühlen. Die Verbundenheit kann jedoch situationsbedingt umschlagen (Postmes et al., 1998), wenn divergierende Themen aufgenommen werden, wozu z. B. die Säuglingsernährung zählt. Säuglingsernährung stellt gesellschaftsweit, aber v. a. unter jungen Müttern einen kontinuierlichen Konfliktherd dar (Murphy, 1999). Polarisierende Auseinandersetzungen in solchen Gruppen können dann gemäß SIDE auftreten, wenn die Zugehörigkeit zur Gruppe der stillenden Mütter im Vergleich

zu Müttern, die ihre Säuglinge mit Babynahrung zufüttern, salient wird (Regan & Brown, 2019). Generell gilt: Nutzende, die sich als zugehörig zu einer situativ erzeugten sozialen Gruppe und damit als Ingroup verstehen, nehmen Nutzer:innen der Gruppe als Outgroup wahr, wenn eine Aufspaltung der Gruppe aufgrund salienter Merkmale erfolgt (Postmes et al., 1998)<sup>46</sup>. Für die Hate Speech-Forschung in SNS ist das SIDE dann relevant, wenn soziale Kategorien wie Religion, Hautfarbe etc. Salienz erlangen, also solche, die gemeinhin mit primären Adressatengruppen von Hate Speech assoziiert werden (Döring & Mohseni, 2020). Es wird angenommen, dass die salienten sozialen Kategorien primärer Adressatengruppen von Hate Speech als soziale Hinweisreize fungieren und aufgrund der Annahme der multiplen Rolle von Variablen sowohl zentral als auch peripher verarbeitet werden können.

Wie in den vorhergehenden Abschnitten erläutert, finden sich vielfach Anschlussstellen Hate Speech als Persuasionsbotschaft zu verstehen. Vorteilhaft erweist sich dabei das ELM, weil es von Wirkmechanismen ausgeht, die den Persuasionsprozess zunächst in einzelne Versatzstücke auflöst, um sodann einen integrativen Rahmen zu spannen, welcher persönlich-individuelle, botschaftsbezogene wie rezeptions-situative Variablen umfasst. Auf diese Weise lässt sich nachvollziehen, welche Faktoren den Rezeptions- und Persuasionsprozess auf welche Weise beeinflussen.

#### **4.5 Zwischenfazit: Auf dem Weg zu einer Modellierung der Persuasionswirkung von sexistischem Hate Speech anhand des Elaboration Likelihood Models**

Mit der vorliegenden Arbeit wird das Ziel verfolgt, Erkenntnisse der sozialpsychologischen Sexismusforschung mit einer kommunikationswissenschaftlichen Perspektive auf Hate Speech zu verknüpfen und mit Hilfe des Elaboration Likelihood Models eine systematische Untersuchung des Persuasionsprozesses von sexistischem Hate Speech anzustreben. Dieses Unterkapitel fasst die bisher zusammengetragenen Erkenntnisse zusammen.

Wie in dieser Arbeit hergeleitet, dient Hate Speech als Kommunikationsform der gruppenbezogenen Abwertung, Verunglimpfung und Diskriminierung, was sowohl auf Kommunikationsinhalts-, Kommunikator- als auch auf Wirkungs- und Rezipierendenebene (primäre Adressatengruppen) in den unterschiedlichsten Fachdisziplinen (Informatik, Kriminalpsychologie, Rechtswissenschaft, Sozialpsychologie und Psychologie etc.) untersucht wurde (Blaya & Audrin, 2019; Costello & Hawdon, 2018; Crowley, 2014; T. Davidson et al., 2017; Glaser et al.,

---

<sup>46</sup> In Kapitel 3 werden die Konsequenzen aus ausgeprägten Ingroup-/Outgroup-Kategorisierungen ausführlich dargestellt.

2002; Kahn et al., 2013; Leets, 2002; McDevitt et al., 2002; Sponholz, 2018). Da der Hate Speech-Begriff unscharf verwendet wird, kommt es zu terminologischen Überlappungen mit dem Begriff der Inzivilität (Muddiman & Stroud, 2017) oder Dangerous Speech (Fink, 2018). Die Autorin der vorliegenden Arbeit stellte eingehend dar, dass drei Kernmerkmale konstituierend für Hate Speech sind, welche einen definitorischen Rahmen aufspannen. Als erstes Kernmerkmal wurden primäre und sekundäre Adressat:innen identifiziert. Primäre Adressat:innen sind marginalisierte, sozialhierarchisch niedrig gestellte Kollektive bzw. deren (wahrgenommene) Vertreter:innen, welche auf der Basis von sozialen Kategorien (Geschlecht, Religion, Ethnie etc.) als Outgroup und damit als disjunkt von der Ingroup eingestuft werden. Kennzeichnend ist, dass die Ingroup in der Regel bevorzugt wird und Outgroups bei wahrgenommener Bedrohung im sozialen Wettbewerb Abwertung erfahren, was als Intergruppenbias bezeichnet wird (M. B. Brewer, 2001). Als sekundäre Adressat:innen werden Angehörige im Publikum bezeichnet, welche sich in Anhänger:innen und Sympathisant:innen bzw. in unbeteiligte Dritte unterscheiden lassen. Sie selbst sind nicht Ziel von Hate Speech, zumindest nicht auf der Basis der jeweils salienten sozialen Kategorie.

Der Kommunikationsinhalt ist das zweite Kernmerkmal von Hate Speech, worunter sämtliche verbale und nonverbale Hate Speech-Äußerungen fallen, mittels derer Abwertung und Feindseligkeit ausgedrückt und deren Aussagekraft verstärkt wird. Als drittes Kernmerkmal wurde die Wirkungsdimension diskutiert, da Hate Speech als sinnhaftes soziales (auf andere ausgerichtetes) Handeln aufgefasst wird. Die Sinnhaftigkeit ist subjektiv aus Perspektive der Kommunizierenden zu verstehen, die Hate Speech dazu nutzen, primäre Adressatengruppen zu verunglimpfen, zu diffamieren und zu diskriminieren und sekundäre Adressat:innen zu mobilisieren bzw. zu empören. Effekte auf sekundäre Adressatengruppen (Sympathisant:innen oder unbeteiligte Dritte) konnten in Form von Mobilisierungswirkungen (Disha et al., 2011; King & Sutton, 2013; Oberschall, 2000; Straus, 2007) oder in Form von negativen Affekten und Kognitionen in Richtung der primären Adressatengruppen nachgewiesen werden (Leader et al., 2009; Soral et al., 2018). Ferner wurde deutlich, dass die Mobilisierungswirkung auf sekundäre Adressat:innen vornehmlich einen Ereignisbezug aufweist oder mit politischen Instabilitätslagen zusammengebracht wird, während ablehnende Affekte und Kognitionen im Zusammenhang mit verbreiteten Stereotypen gegenüber Outgroups sichtbar werden (ebd.). Mobilisierungswirkungen werden in diesen Arbeiten zumeist in Form von Hate Crimes nachgewiesen, was sich auf forschungspraktische Gründe zurückführen lässt, denn Hate Crimes sind sichtbarer, sowohl hinsichtlich ihrer kriminologischen Erfassung wie auch deren Folgen, insbesondere im Vergleich zu den flüchtigen mündlichen Ausprägungen

von Hate Speech<sup>47</sup>. Die vorliegende Arbeit widmet sich explizit der Untersuchung von Hate Speech-Effekten auf Handlungsebene wie auf Ebene der Einstellungskomponenten.

Die Herausarbeitung der drei Dimensionen erweist sich von grundlegender Bedeutung, um Hate Speech von verwandten Phänomenen – gerade im Online-Bereich – abzugrenzen. Zudem wird hervorgehoben, dass Hate Speech keineswegs ausschließlich online stattfindet, sondern seinen Ursprung in der Offline-Lebenswelt hat und mit der Verbreitung der Internetnutzung dorthin gelangt ist und Aufmerksamkeit in breiten Gesellschaftsschichten erlangt hat. Die kommunikationswissenschaftliche Forschung profitiert von der konzeptionellen Klärung des Begriffs in besonderem Maße, da die theoriebasierte Auseinandersetzung mit dem Hate Speech-Begriff dazu beiträgt, ein umfassendes Verständnis der Zusammenhänge von Adressatengruppen, kommunikativem Inhalt und Wirkungsdimension zu erreichen, ohne sich auf einen Kommunikationsraum zu beschränken. Aus dem definitorischen Rahmen wurde die Auseinandersetzung mit den Kommunikator:innen ausgeklammert, da diese das Hate Speech-Konzept nur mittelbar betreffen. Um jedoch den Entstehungskontext aus Kommunikatorsicht zu berücksichtigen, wurden die psychosozialen Eigenschaften der Hate Speech-Produzent:innen herausgearbeitet. Somit deckt Kapitel 2 alle kommunikationswissenschaftlichen Forschungsfelder ab, um Hate Speech in dieser Disziplin theoretisch zu verorten.

Um das Verständnis von Hate Speech im Zusammenhang mit der Geschlechtszugehörigkeit zu vertiefen, wurde der Forschungsstand zu Sexismus als Einstellungskonstrukt aus der Perspektive sozialpsychologischer Vorurteilsforschung mit einem besonderen Fokus auf Ambivalenten Sexismus zusammengefasst. Der oben beschriebene Intergruppenbias kann bei Individuen mit einer Dominanzorientierung in Aggressionen und Hate Speech münden, und zwar wenn soziale Wandlungsprozesse den sozialen Wettbewerb gesellschaftlicher Gruppen verschärfen bzw. wenn die soziale Position der eigenen Gruppe dadurch als gefährdet wahrgenommen wird. Vor diesem Hintergrund lässt sich auch ein Zusammenhang zwischen der Geschlechterkategorie und Hate Speech nachvollziehen. Geschlecht als soziale Kategorie unterliegt seit Jahrzehnten einer weitreichenden Auflösung geschlechtsbedingter Ungleichheit (Zugang zu Bildungs- und wirtschaftlichen Ressourcen etc.), welche aber nicht vollständig erfolgt ist. Von dem Wandlungsprozess sind auch traditionelle Geschlechterrollen betroffen, die einen persistenten Norm- und Erwartungsrahmen vorgeben, der überspitzt ausgedrückt submissive Frauen und Männer in der Versorger- und Beschützerrolle vorsieht. Es wird

---

<sup>47</sup> Die Dokumentation von Hate Speech ist ebenfalls möglich, z. B. durch Screenshots von Postings in SNS oder audiovisuelle Aufnahmen in Form von Filmen, Audiomitschnitten oder Fotografien von Hate Speech offline.

angenommen, dass Situationen, die den Gegensatz zwischen traditionellen und progressiven Charakteristika salient erscheinen lassen, bei vorhandener Dominanzorientierung Hate Speech befördern.

Das Konzept des Ambivalenten Sexismus erlaubt ein differenziertes Bild geschlechtsstereotyper Einstellungen, da es von benevolenten und hostilen Anteilen in der Beurteilung des Gegenübers ausgeht. Benevolente Einstellungen würdigen die Ausrichtung des Verhaltens an traditionellen Geschlechterrollen. Ein Abweichen von der traditionellen Geschlechterrolle aktiviert hingegen hostile sexistische Einstellungen, und zwar unabhängig vom Geschlecht des Einstellungsobjekts. Damit tragen beide Dimensionen dazu bei, die traditionelle Geschlechterhierarchie aufrechtzuerhalten. Sichtbar wird der Zusammenhang der beiden Dimensionen messmethodisch, da beide signifikant miteinander korrelieren. Ungeklärt ist, ob ambivalent sexistische Einstellungen mit einer höheren Akzeptanz von sexistischem Hate Speech einhergehen, sichtbar daran, dass ihnen nicht widersprochen wird oder sie befürwortet werden und ob zudem eine erhöhte Bereitschaft zu Hate Speech-Äußerungen erkennbar wird.

Herausgearbeitet wurde das Berufsfeld als bedeutsamer Bereich, in dem Geschlechtsstereotype zum Tragen kommen und die IT-Branche exemplarisch als deutlich männlich dominierter Sektor, der davon besonders betroffen ist (vgl. Kapitel 3.3). Frauen, die eine Karriere in der IT-Branche anstreben, verstoßen gegen präskriptive Anteile des weiblichen Geschlechtsstereotyps, d. h. traditionellen Annahmen und Erwartungen darüber, welche Berufe Frauen ergreifen *sollen* (vgl. Eckes, 2010, S. 178). Abweichungen dieser Art erfahren Abwertungen, die verdeckt oder offen vorgetragen werden und die sich in Form von Hate Speech äußern können (vgl. Kapitel 3.2.1.1). Zugleich treiben weibliche Karrieren aber auch die Auflösung geschlechtsbedingter Ungleichheit voran, indem Frauen Zugang zu ökonomischen Ressourcen erreichen, was den oben beschriebenen Intergruppenbias vergrößert. Damit sind zwar generelle Ursachen für Abwertungen mittels sexistischem Hate Speech begründbar, aber nicht die spezifischen Umstände für die Bereitschaft zu Hate Speech-Äußerungen. So ist bislang unzureichend geklärt, welche Faktoren Hate Speech begünstigen, insbesondere, was die Mobilisierung von unbeteiligten Rezipient:innen selbst Hate Speech zu produzieren, betrifft. Oder anders gefragt: Wie persuasiv ist sexistisches Hate Speech tatsächlich? Zu prüfen ist ein Geschlechtereffekt, also ob Nutzerinnen, die sexistischen Hate Speech verbreiten und damit gegen das weibliche Globalstereotyp der empfindsamen, aber weniger kompetenten und durchsetzungsstarken Frau verstoßen, andere Effekte auf Rezipientinnen und Rezipienten auslösen als männliche Nutzer.

Um die Persuasivität von sexistischem Hate Speech zu ermitteln, wurde in der vorliegenden Arbeit das ELM von Petty und Cacioppo (1986a, 1986b) als relevantes Rahmenwerk zur Analyse gewählt, da es den Persuasionsprozess als Interaktion von botschaftsimmanenten Faktoren (Kommunikator:in, Botschaft, Situation) und Merkmalen der rezipierenden Person modelliert. Die Anwendung des ELM in diesem Kontext trägt dazu bei, den Wirkungsprozess auf sekundäre Adressat:innen, v. a. auf das unbeteiligte Publikum, nachzuvollziehen. Das Modell postuliert eine duale Informationsverarbeitung, bei der die persuasiven Argumente einer Botschaft auf einem Elaborationskontinuum von kognitiv vertieft (zentral) bis oberflächlich (peripher) verarbeitet werden. Als notwendige Bedingung für die zentrale Verarbeitung werden ein hohes Maß an Fähigkeit und Motivation zur Verarbeitung vorausgesetzt. Die Fähigkeit und Motivation zur Verarbeitung sind latente Konstrukte, welche mittels situativer (Ablenkung und Wiederholung im Persuasionsprozess) und individueller Variablen (Vorwissen, Involvement, Kognitionsbedürfnis, wahrgenommene persönliche Verantwortung) gemessen werden. Sind Fähigkeit und/oder Motivation gering, wird eine periphere Verarbeitung der Persuasionsbotschaft angenommen.

Das ELM hat seinen Ursprung in der Konsumentenforschung (Petty et al., 1983), weswegen die Übertragung des Modells auf ein als gesellschaftlich relevant betrachtetes Problemfeld wie Hate Speech mit einigen Herausforderungen verbunden ist. Bisherige Anwendungen lassen sich im Bereich von Aggressionen am Arbeitsplatz (S. C. Douglas et al., 2008), in der Gesundheitskommunikation (Angst & Agarwald, 2009; Dinoff & Kowalski, 1999), im Kontext anti-islamischer Propaganda (Müller et al., 2014), vereinzelt im Zusammenhang mit Counterspeech (Schieb & Preuss, 2018) oder Hate Speech (E. Lee & Leets, 2002) finden. Damit wird deutlich, dass potenziell jeder Kommunikationsinhalt mit einer persuasiven Absicht für die Rezeption mittels ELM auf seine Persuasionwirkung hin überprüft werden kann. Klar einsichtig ist das für persuasive Kommunikationsinhalte, die ihre Botschaften mittels *starker* (im Sinne begründeter, stichhaltiger) Argumente transportieren, wie das z. B. für Gesundheitskampagnen oder Counterspeech der Fall ist. Ungeklärt ist bislang, ob Hate Speech ebenfalls als starkes (nicht im normativen Sinne), also persuasives Argument gelten kann. Entgegen Petty und Cacioppo's Konzeptionalisierung einer empirisch bestimmten Argumentstärke, die Ursachen- und Wirkungszuschreibungen konfundieren (vgl. Kapitel 4.3.2.1) soll im Rahmen dieser Arbeit eine aus Kapitel 2 theoretisch abgeleitete Operationalisierung der Intensität von Hate Speech verwendet werden. Vermutet wird eine Verarbeitung von Hate Speech auf der zentralen Route, weil Hate Speech bei primären wie sekundären Adressat:innen den Umfang und die Qualität der kognitiven Auseinandersetzung stimuliert, aber auch emotionale Reaktionen provoziert. Ob die Emotionen reflektiert und entsprechend kognitiv eingeordnet

werden, bleibt eine empirische Frage, die nicht Gegenstand dieser Arbeit ist. Denkbar ist schließlich auch, dass die aus der Rezeption resultierende Emotion als einfacher Hinweisreiz (periphere Verarbeitung) fungiert und die Valenz der Emotion sich auf das Einstellungsobjekt überträgt (Petty & Briñol, 2015).

Ob das ursprüngliche Modell auf den Online-Kontext in unveränderter Form übertragen werden kann, wird von einigen Autoren bezweifelt. Gayle Kerr und ihre Kollegen (2015) fordern Anpassungen an veränderte Mediennutzungsgewohnheiten, weil z. B. fehlende periphere Hinweisreize den Persuasionsprozess beeinflussen (Chaiken & Eagly, 1983; Guadagno & Cialdini, 2007). Dass solche Anpassungen an den Online-Kontext sinnvoll sind (z. B. mittels Operationalisierung der Variablen, Ergänzungen um andere theoretische Zugänge) und durchaus Persuasionseffekte belegt werden können, zeigen ELM-Studien der letzten Jahre (Knoll, 2015; SanJosé-Cabezudo et al., 2009; Wang et al., 2019). In klassischen Studien der ELM-Forschung aus den 1980er Jahren wurde gezeigt, dass sich das Konstrukt der persönlichen Verantwortung auf die Motivation zur Elaboration im Persuasionsversuch auswirkt (Harkins & Petty, 1982). Übertragen auf die Rezeptionssituation online wird die persönliche Verantwortung von Lurkern als gering wahrgenommen, da sie ihren Einfluss in der großen Nutzermenge innerhalb des SNS als gering wahrnehmen. Bislang ist noch ungeklärt, ob die Wahrnehmung von persönlicher Verantwortung Nutzende im Umgang mit Hate Speech beeinflusst.

Die vorliegende Arbeit hat fachübergreifende Erkenntnisse zusammengetragen, indem Bezüge zwischen der Kommunikationswissenschaft und Sozialpsychologie hergestellt wurden, mit dem Ziel, den Persuasionsprozess von sexistischem Hate Speech in Abgrenzung zu Inzivilität (verstanden als Unhöflichkeit) und Dangerous Speech (verstanden als Hate Speech mit Gewaltpotenzial) erklären zu können. Für die empirische Überprüfung der Zusammenhänge wurde das ELM herangezogen. Die Autorin dieser Arbeit verfolgt das Ziel, die Übertragbarkeit des Modells auf Hate Speech in SNS zu testen und auftretende Rezeptions- und Persuasionseffekte aufgrund von Botschafts- oder Rezipierendenmerkmalen zu analysieren. In den nächsten Kapiteln wird die empirische Überprüfung der theoretischen Überlegungen vorgestellt, zunächst die Ableitung der Hypothesen im nächsten Abschnitt (Kapitel 5) und sodann in Kapitel 6 die Vorgehensweise bei der Messung und die Ergebnisse aus zwei experimentellen Studien und schließlich das Fazit in Kapitel 7, in dem die Befunde zusammengefasst, diskutiert und eingeordnet werden.

## 5 Ableitung der Hypothesen

Das folgende Kapitel widmet sich der Ableitung der zu prüfenden Hypothesen, mit dem Ziel, die bisher gewonnenen Erkenntnisse der empirischen Überprüfung zuzuführen. Dazu werden drei Themenblöcke vorgestellt, welche das Ausmaß der Elaboration, die Auswirkungen der Elaboration auf Einstellungen und die Auswirkungen der Elaboration auf Verhaltensintentionen zur Diskussion stellen.

### 5.1 Elaboration von Hate Speech

Das Persuasionspotential von Hate Speech gilt als beträchtlich, wenn auch zumeist quasi-experimentelle Evidenzen oder normative Erwartungen angeführt werden, um persuasive Effekte bei sekundären Adressat:innen von Hate Speech zu erörtern (Delgado & Stefancic, 2014; Legewie, 2013; Maynard & Benesch, 2016; Parekh, 2006; Sorial, 2015) bzw. gar nicht Hate Speech, sondern verwandte Phänomene (Inzivilität, Cybermobbing etc.) Gegenstand der Untersuchung sind (B. T. Gervais, 2015; Rösner et al., 2016). Mobilisierungswirkungen oder die Aktivierung von Stereotypen und Vorurteilen marginalisierter sozialer Gruppen bei sekundären Adressat:innen wurde bereits im Kontext von politischen Konfliktlagen belegt (Kimotho & Nyaga, 2016; Leets & Bowers, 1999; Straus, 2007).

In Kapitel 4 wurde Elaboration als kognitive Auseinandersetzung mit persuasiven Inhalten vorgestellt: „By *elaboration* in a persuasion context, we mean the extent to which a person thinks about the issue-relevant arguments contained in a message.“ (Petty & Cacioppo, 1986b, S. 128, Herv. i. O.). Unter dem Ausmaß an Gedanken ist eine quantitative Messgröße nach Rezeption der Persuasionsbotschaft zu verstehen. Messmethodisch wird die Elaboration mittels Thought-Listing-Technik ermittelt, was der Erfassung der Rezeptionsgedanken in mündlicher oder schriftlicher Form entspricht und eine vertiefende Analyse erlaubt (Cacioppo & Petty, 1981). Die mengenmäßige Erfassung allein ist jedoch nur ein korrelativer Indikator für die Elaborationstiefe. Darüber hinaus ermöglicht die Thought-Listing-Technik die Analyse der Gedankenvalenz. Gerade die Gedankenvalenz wird im Rahmen des ELM als Indikator für eine stattgefundene Persuasionwirkung bewertet: Positive Gedanken gelten als erfolgreiche Persuasionwirkung, negativen Gedanken wird eine solche abgesprochen. Ob sich diese Annahme aus dem ELM im Hate Speech-Kontext widerspiegelt, muss jedoch angezweifelt werden und es deuten sich Inkompatibilitäten der theoretischen Rahmenwerke an (s. Kapitel 4.4). Die Thought-Listing-Technik erlaubt darüber hinaus die inhaltliche Auswertung der Rezeptionsgedanken, bei der je nach Erkenntnisinteresse kaum Grenzen bestehen. Die ELM-Autoren schlagen die Auswertung der gedanklichen Ausgangsbasis vor, die als

Gradmesser für die Verarbeitungstiefe verstanden wird. Die quantitativ-qualitative Analysestrategie erlaubt ein umfassendes Bild der Elaborationstiefe und -weise. Folglich werden unter Elaboration stets die drei Elaborationsdimensionen verstanden: Gedankenanzahl, Gedankenvalenz und gedankliche Ausgangsbasis.

Aus den klassischen ELM-Annahmen geht die Funktion von Motivation und Fähigkeit für die Informationsverarbeitung hervor. Vorwissen hat einen positiven Einfluss auf die Fähigkeit zur Elaboration. Die Motivation sich kognitiv mit der Persuasionsbotschaft auseinanderzusetzen, steigt mit der Höhe des Kognitionsbedürfnisses und des persönlichen Involvements. Daraus folgt, dass die Auseinandersetzung mit der Persuasionsbotschaft kognitiv stärker, d. h. auf der zentralen Route verarbeitet wird, wenn Motivation und Fähigkeit vorhanden sind. Die zentrale Verarbeitung würde sich demnach sowohl auf die Anzahl (im Sinne eines Anstiegs) als auch auf die Gedankenvalenz (im Sinne vorwiegend positiver Gedanken) sowie die inhaltlich-kognitive Ausrichtung der Rezeptionsgedanken auswirken. Diese idealtypische Verarbeitungsrichtung bei sekundären Adressat:innen von Hate Speech darf angezweifelt werden. Zwar ist davon auszugehen, dass unter der Bedingung hoher Fähigkeit und Motivation zur Verarbeitung die Gedankenanzahl ansteigt, allerdings muss davon ausgegangen werden, dass es sich dabei mehrheitlich um negative Rezeptionsgedanken handelt. Dafür sprechen Befunde aus der Hate Speech-Forschung, die zeigen, dass sekundäre Adressat:innen mit starker Ablehnung und Entrüstung auf die Rezeption von Hate Speech reagieren (s. Kapitel 2.1.4.3). Andererseits werden Persuasionseffekte von inzivilen Äußerungen diskutiert (Anderson et al., 2014). Für die gedankliche Ausgangsbasis als Elaborationsdimension wird vermutet, dass die Motivation und Fähigkeit zur Verarbeitung kognitiv-inhaltliche Rezeptionsgedanken stimuliert. Andererseits müssen auch Reaktanzeffekte berücksichtigt werden, welche trotz vorhandener Elaborationsbereitschaft diese verweigert.

Der persuasive Erfolg einer Botschaft hängt maßgeblich von ihrer Aufbereitung ab, wozu die Argumentstärke und periphere Hinweisreize (z. B. ästhetische Gestaltungsvarianten, Eigenschaften der Sprecher:innen) zählen. Wie ausführlich im Abschnitt 4.3.2.1 erörtert, muss die Argumentstärke im Fall von Hate Speech durch inhaltliche Erwägungen ersetzt werden. Zum jetzigen Zeitpunkt fehlt eine Gegenüberstellung von Hate Speech zu Inzivilität und Dangerous Speech, um mögliche Effekte näher zu betrachten. Denn Inzivilität adressiert keine bestimmte gesellschaftliche Gruppierung, beinhaltet jedoch respektlose, unhöfliche und vulgäre Kommunikationsinhalte (Anderson et al., 2014; Coe et al., 2014; Ziegele & Jost, 2016). Dangerous Speech ist Hate Speech im definierten Sinne und enthält zusätzlich eine Gewaltkomponente, indem zu Gewaltakten gegen primäre Adressatengruppen aufgerufen bzw.

damit gedroht wird. Ungeklärt ist bislang die Frage, ob und wie sich die Elaboration von inzi-  
vilen Äußerungen im Vergleich zu Hate Speech oder sogar Dangerous Speech (Aufrufe zu o-  
der Androhungen von Gewalt in Verbindung mit Hate Speech), d. h. also hinsichtlich der *Hate*  
*Intensität* unterscheidet. Zu prüfen ist, ob die Hate Intensität die Anzahl der Gedanken erhöht  
oder ob Reaktanzeffekte auftreten, welche die Bereitschaft sich zu äußern, herabsetzen. Aus  
diesem Grund wird Hypothese 1a offen formuliert:

*H1a: Die Hate Intensität beeinflusst die Anzahl der Rezeptionsgedanken. Dieser Zu-  
sammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

Abgeleitet aus den Erkenntnissen aus Kapitel 2.1.4.3, ist davon auszugehen, dass mit steigen-  
der Hate Intensität die Gedankenvalenz negativer wird.

*H1b: Je höher die Hate Intensität ist, desto negativer sind die Rezeptionsgedanken.  
Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration mode-  
riert.*

Die inhaltliche Dimension der Rezeptionsgedanken lässt sich sowohl als zentrale als auch als  
periphere Verarbeitung vermuten. Bei zentraler Verarbeitung der Hate Intensität würden  
Versuchsteilnehmende die Positionen diskutieren und die Persuasionsbotschaft stützende  
Argumente oder Gegenargumente vorbringen, je nachdem ob objektiv oder verzerrt verar-  
beitet wird. Bei peripherer Verarbeitung würden vornehmlich affektiv gefärbte oder irrele-  
vante Rezeptionsgedanken artikuliert werden. Aus diesem Grund wird ebenfalls eine offene  
Formulierung für Hypothese 1c gewählt:

*H1c: Die Hate Intensität beeinflusst die gedankliche Ausgangsbasis der Rezeptionsge-  
danken. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elabora-  
tion moderiert.*

Quellenmerkmale, also Merkmale der sprechenden Person, werden im ELM-Kontext häufig  
peripher verarbeitet, beispielsweise die Attraktivität des Testimonials. Aber wie bereits in  
Kapitel 4.3.3 erörtert, können Variablen im Persuasionskontext unterschiedliche Rollen an-  
nehmen, indem sie bei geringer Fähigkeit und Motivation zur Elaboration peripher, hingegen  
bei vorhandener Motivation und Fähigkeit als Argument oder verzerrt verarbeitet werden.  
Im Zusammenhang mit sexistischem Hate Speech ist das Geschlecht der kommentierenden  
Person zunächst als ein solcher peripherer Hinweisreiz zu bewerten. Denkbar ist aber auch,  
dass das Geschlecht zentral verarbeitet wird und im Sinne eines Bewertungsmaßstabs als  
Argument fungiert, wie Befunde aus der Sexismusforschung nahelegen (Chang & Hitchon,

2004). Es wird erwartet, dass Hate Speech von männlichen wie von weiblichen Produzent:innen als Abweichung von der normierten Geschlechterrolle wahrgenommen wird. Abgeleitet aus der Wärme-/Kompetenz-Wahrnehmung von Individuen und Gruppen lässt sich die Wahrnehmung von Hate Speech bzw. deren Produzent:innen als kalt und inkompetent charakterisieren. Damit verstößt Hate Speech sowohl gegen das männliche (kühl, kompetent) als auch gegen das weibliche (warm, inkompetent) Globalstereotyp. Analog zu den Hypothesen 1a-c wird deshalb postuliert:

*H2a: Die Geschlechterkategorie von Hate Speech-Produzent:innen beeinflusst die Anzahl der Rezeptionsgedanken. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

Für die Gedankenvalenz kann keine einseitige Hypothese formuliert werden, da hier Effekte sowohl für Männer als auch für Frauen denkbar sind.

*H2b: Die Geschlechterkategorie von Hate Speech-Produzent:innen beeinflusst die Gedankenvalenz. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

*H2c: Die Geschlechterkategorie von Hate Speech-Produzent:innen beeinflusst die gedankliche Ausgangsbasis der Rezeptionsgedanken. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

Die Interaktion der Nutzer:innen mit Inhalten auf SNS reicht von eher niedrighschwelligen Funktionen (Klick auf Like- und Share-Funktionen) bis hin zu aufwändigeren Beteiligungsformen, wie dem Verfassen eines Kommentars oder Postings etc. Der Zusammenhang zwischen dem Emotionalisierungsgrad der geteilten Inhalte und dem Reaktionsvolumen konnte wiederholt nachgewiesen werden (Stieglitz & Dang-Xuan, 2013; Tellis et al., 2019). Dabei ist das Reaktionsvolumen nicht als deskriptive, rein technologische Interaktion misszuverstehen, sondern bildet die sozialen Dynamiken zwischen Nutzer:innen ab und stellt eine Kommunikationspraxis mit eigenen Normen dar (Ellison & Boyd, 2013; Light & Cassidy, 2014). Die Anzahl von Likes, Shares und Retweets gibt Nutzenden Aufschluss über den Grad der Informationsdiffusion, die Relevanz der Inhalte oder das Meinungsklima zu einem Thema (John, 2013; Kwak et al., 2010; E.-J. Lee & Jang, 2010). Die derart wahrgenommene soziale Umwelt im SNS-Kontext beeinflusst die Nutzung der Interaktionsfeatures und führt zur Verstetigung oder dem Abbruch der Interaktion (Bishop, 2007). Die Befunde sind widersprüchlich, denn sowohl ein hohes wie auch ein niedriges Reaktionsvolumen beeinflussen Einstellungen

(Alhabash et al., 2013) oder Verhaltensintentionen von Nutzer:innen (Alhabash, Baek et al., 2015; Alhabash, McAlister et al., 2015) bzw. zeigen keine Effekte (Winter et al., 2015).

Die Wahrnehmung des Reaktionsvolumens im Kontext von Hate Speech in SNS wird hier als situatives Merkmal im Persuasionsprozess aufgefasst, was den klassischen Konzeptionierungen von Petty und Cacioppo (1986b) entspricht. Das Reaktionsvolumen ist ein Indikator für die persönliche Verantwortung und es wird ein Einfluss auf die Elaboration von Hate Speech angenommen. Ungeklärt ist jedoch, ob das Reaktionsvolumen zentral als Argument oder peripher als Hinweisreiz verarbeitet wird. Folglich lassen sich folgende Hypothesen aus den bisherigen Überlegungen wie folgt ableiten:

*H3a: Das Reaktionsvolumen beeinflusst die Anzahl der Rezeptionsgedanken. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

*H3b: Das Reaktionsvolumen beeinflusst die Gedankenvalenz. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

*H3c: Das Reaktionsvolumen beeinflusst die gedankliche Ausgangsbasis der Rezeptionsgedanken. Dieser Zusammenhang wird von der Fähigkeit und Motivation zur Elaboration moderiert.*

## **5.2 Elaboration und wahrgenommene Relevanz**

Eine zentrale Variable innerhalb des ELM ist die persönliche Relevanz einer Botschaft. Die persönliche Relevanz prognostiziert eine ausführliche und kognitiv vertiefte Informationsverarbeitung (Petty & Cacioppo, 1986b) und wurde vor allem in der Konsumentenforschung repliziert (C.-H. Cho, 1999; Knoll, 2015). Abweichend davon sind Befunde aus der Gesundheitskommunikation, die eine kognitiv defensive Verarbeitung von Furchtappellen belegen, bei der die persönliche Relevanz der Botschaft abgelehnt wird (Liberman & Chaiken, 1992). Die Auseinandersetzung mit den Argumenten einer Botschaft findet selektiv statt, indem bedrohliche Argumente zwar intensiv, aber vergleichsweise kritischer als beschwichtigende Argumente verarbeitet werden.

Für die Rezeption von Hate Speech bei sekundären Adressat:innen (unbeteiligte Dritte) wird ein ähnlicher Distanzierungseffekt erwartet. Begründen lässt sich die Distanzierung mit der Social Judgement-Theorie, der zufolge Botschaften abgelehnt werden, die inkongruent mit individuellen Einstellungen sind (Sarup et al., 1991). Gilt andererseits, dass Einstellungen, die als kongruent zur rezipierten Botschaft wahrgenommen werden, die Distanzierung reduzieren? Argumentiert wird, dass Ambivalenter Sexismus als einstellungskongruent zur

Rezeption von Hate Speech wahrgenommen werden kann, da die Thematik „Frauenkarrieren in der IT-Branche“ die Salienz der Geschlechterkategorie als einstellungsrelevanten Bewertungsmaßstab hervorhebt. Die Verfügbarkeit der Geschlechterkategorie wird in der Rezeptionssituation auch anhand des Geschlechts der Hate Speech-Produzent:innen erhöht. Unter der Bedingung ambivalent sexistischer Einstellungen kann vermutet werden, dass die Rezeption von Inzivilität, Hate Speech oder Dangerous Speech als Abweichung von den prototypischen Geschlechterrollen wahrgenommen wird (Wärme-Kompetenz-Dimensionen der Bewertung von Männern und Frauen) und die Distanzierung verstärken. Zu prüfen ist also, ob sich unter der Bedingung ambivalent sexistischer Einstellungen ein negativer Einfluss des Geschlechts der Hate Speech-Produzent:innen sowie der Hate Speech-Intensität auf die wahrgenommene persönliche Relevanz nachweisen lässt.

*H4a: Die Hate Intensität beeinflusst die wahrgenommene persönliche Relevanz. Dieser Effekt wird von ambivalent sexistischen Einstellungen moderiert.*

*H4b: Das Geschlecht der Kommentierenden beeinflusst die wahrgenommene persönliche Relevanz. Dieser Effekt wird von ambivalent sexistischen Einstellungen moderiert.*

Der Zusammenhang für den situativen Einfluss durch das wahrgenommene Reaktionsvolumen auf die persönliche Relevanz ist uneindeutig. Einerseits sind Distanzierungseffekte denkbar, die bei einem hohen Reaktionsvolumen eine geringere persönliche Relevanz bewirken, andererseits können Bandwaggon-Effekte auftreten, so dass die Wahrnehmung eines hohen Reaktionsvolumens die persönliche Relevanz verstärkt. Dieser Effekt würde dann durch ambivalent sexistische Einstellungen verstärkt. Entsprechend wird folgende Hypothese aufgestellt:

*H4c: Das Reaktionsvolumen beeinflusst die wahrgenommene Relevanz. Dieser Effekt wird von ambivalent sexistischen Einstellungen moderiert.*

### **5.3 Einstellungen zur Botschaft und Einstellungen zu Frauenkarrieren**

In der Persuasionsforschung gilt die Änderung und Neubildung von Einstellungen als unverhältnismäßig schwieriger als die Festigung bestehender Einstellungen. Dabei erweisen sich starke (bestehende) Einstellungen als resistenter gegenüber Persuasionsversuchen, insbesondere wenn sie als wichtig, verfügbar und strukturell konsistent wahrgenommen werden (Haugtvedt & Petty, 1992). Im Kontext der vorliegenden Arbeit befinden sich Einstellungen

zu Hate Speech, Ambivalentem Sexismus und modernen Frauenkarrieren im Mittelpunkt, und es ist nach derzeitigem Erkenntnisstand noch ungeklärt, welche Beziehungen zwischen den Einstellungsvariablen und der Rezeption von Hate Speech bestehen.

Die kognitive Bewertung von Hate Speech entsteht adhoc in der Rezeptionssituation. Dagegen spricht aber, dass Nutzer:innen bereits Hate Speech in sozialen Netzwerken rezipiert haben und diesbezüglich schon starke Einstellungen gebildet haben (Santos et al., 2020).

*H5: Die Geschlechterkategorie der Produzent:innen beeinflusst die kognitive Bewertung von Hate Speech. Dieser Zusammenhang wird von der wahrgenommenen Relevanz vermittelt.*

Ambivalenter Sexismus kann als bestehende Einstellung verstanden werden, weil es sich dabei um eine interindividuelle Bewertungstendenz zu geschlechtsrollenspezifischen Verhaltensweisen handelt, welche unabhängig von der Rezeptionssituation ist. Es wurde dargelegt, dass berufliche Domänen u. a. stark von Geschlechterzugehörigkeiten geprägt sind und Angehörige von Berufsgruppen auf der Basis von Geschlechterrollen wahrgenommen werden. Die Geschlechtszugehörigkeit in einem Berufszweig, der als nicht geschlechtsrollenkonform wahrgenommen wird, würde folglich Einstellungen zu Berufsentscheidungen (hier: Frauenkarrieren in der IT-Branche) aktivieren. Diese Einstellungen können als einstellungsverstärkend verstanden werden, da sie auf ambivalent sexistischen Bewertungen basieren. Wenn in der Rezeptionssituation die Geschlechterkategorie der Hate Speech-Produzent:innen salient wird, kann bei objektiver Verarbeitung eine Verstärkung des Zusammenhangs von ambivalent sexistischen Einstellungen auf die Stärke der Einstellungen erwartet werden. Denkbar ist aber auch eine verzerrte Verarbeitung, bei der Ambivalenter Sexismus als Grundlage für die Bewertung der Produzent:innen herangezogen wird.

*H6a: Die Beziehung zwischen Ambivalentem Sexismus und Einstellungen zu Frauenkarrieren wird von der Geschlechterkategorie der Produzent:innen beeinflusst.*

In der Rezeptionssituation wird die individuelle Geschlechtszugehörigkeit salient und deswegen wird die nächste Hypothese folgendermaßen formuliert:

*H6b: Die Beziehung zwischen Ambivalentem Sexismus und Einstellungen zu Frauenkarrieren ist abhängig von dem Geschlecht der Versuchspersonen.*

Da darüber hinaus denkbar ist, dass der Einfluss von ambivalent sexistischen Einstellungen auf Einstellungen zu Frauenkarrieren sich je nach Rezeptionssituation und nach Geschlecht

der Versuchspersonen unterscheidet, wird also eine moderierte Moderation des Zusammenhangs vermutet:

*H6c: Der Einfluss von Merkmalen der Produzent:innen auf die Beziehung zwischen Ambivalentem Sexismus und Einstellungen zu Frauenkarrieren wird vom Geschlecht der Versuchspersonen moderiert.*

## 5.4 Verhaltensintentionen

Die Elaboration-Konsistenz-Hypothese (EKH) ermöglicht die Anschlussfähigkeit des ELM an Einstellungs-Verhaltens-Modellen wie der Theory of Planned Behavior von Ajzen und Fishbein (Ajzen, 1991; Ajzen & Fishbein, 1980). Zunächst gilt es zu prüfen, welche Verhaltensintentionen nach Rezeption von Kommentar-, Produzent:innen- sowie situativen Merkmalen als wahrscheinlich wahrgenommen werden. Hinsichtlich der Verhaltensintention können die beiden zweipoligen Dimensionen Aktivitätsgrad (aktiv-passiv) und Akzeptanzrichtung (zustimmend-ablehnend) unterschieden werden. So ermöglichen auf technischer Seite die implementierten Interaktionsfeatures der Plattformarchitekturen sowohl stärker involvierende, also *aktive* (z. B. Nutzung der Kommentierfunktion, auch das Melden unangemessener Kommunikationsinhalte) als auch eher *passive* bzw. weniger involvierende Reaktionsformen (z. B. Share- und Like-Interaktionen). In diesem Sinne stellt das Ignorieren bzw. genauer das sogenannte „Durchscrollen“ des Social Media-Feeds sicherlich die Extremform passiven Nutzungsverhaltens dar, ist dabei aber weder dem zustimmenden noch dem ablehnenden Nutzungsverhalten zuzurechnen. Die *Akzeptanzrichtung* aktiver Reaktionsformen reicht von unterstützenden, bestätigenden oder befürwortenden bis hin zu kritisierenden oder ablehnenden Verhaltensweisen (Ellison & Boyd, 2013; John, 2013). Was eher passive zustimmende Interaktionsmöglichkeiten betrifft, so bieten die meisten Social Media-Plattformen in erster Linie Zustimmungsfeatures an (z. B. „Gefällt mir“-Markierungen in Herzform bei Twitter oder Instagram). Facebook stellt den Nutzenden eine differenzierte Reaktionspalette zur Verfügung, welche den Ausdruck von Ablehnung ermöglicht (z. B. trauriges oder wütendes Emoji als Reaktionsbutton) oder zu kontextabhängigen Nutzungspraktiken (z. B. kann ein lachendes „Haha“-Emoji unter Schadensmeldungen als Häme interpretiert werden) und damit zu Mehrdeutigkeit führt. Die Interaktion mit Inhalten in sozialen Netzwerken ist darüber hinaus von soziodemografischen Merkmalen abhängig (Pempek et al., 2009; Santos et al., 2020).

Aus den bisherigen Argumentationen folgt daher die Aufstellung folgender Hypothesen:

*H7a: Je stärker die Hate Intensität ist, als desto wahrscheinlicher werden ablehnende und aktive Verhaltensintentionen genannt. Soziodemografische Faktoren verstärken den Effekt.*

*H7b: Die Geschlechterkategorie beeinflusst die artikulierte Wahrscheinlichkeit des Aktivitätsgrads und der Akzeptanzrichtung von Verhaltensintentionen. Soziodemografische Faktoren verstärken den Effekt.*

*H7c: Das Reaktionsvolumen beeinflusst die artikulierte Wahrscheinlichkeit des Aktivitätsgrads und der Akzeptanzrichtung von Verhaltensintentionen. Soziodemografische Faktoren verstärken den Effekt.*

Petty und Kollegen (1995) diskutieren Evidenzen für die EKH, bei der Verhalten sowohl auf Basis bestehender als auch auf Basis neu gebildeter Einstellungen prognostiziert wird. Bestehende Einstellungen würden etwa die Verhaltensintention bei Hervorheben der thematischen Salienz beeinflussen (ebd). Werden Einstellungen während der Rezeptionssituation neu gebildet, können sie ebenfalls Verhaltensintentionen prognostizieren, sofern der Inhalt der Persuasionsbotschaft als persönlich relevant wahrgenommen wird (ebd.).

Auf Grundlage der bisherigen Argumentation wird daher vermutet:

*H8: Ambivalent sexistische Einstellungen erhöhen die artikulierte Wahrscheinlichkeit zustimmender Verhaltensintentionen. Die Hate Intensität, weibliche Kommentierende und ein höheres Reaktionsvolumen verstärken diesen Effekt.*

*H9: Der Zusammenhang zwischen ambivalent sexistischen Einstellungen und der artikulierten Wahrscheinlichkeit zustimmender Verhaltensintention wird von Einstellungen zu Frauenkarrieren moderiert.*

*H10: Je höher die persönliche Relevanz ist, desto stärker ist die artikulierte Wahrscheinlichkeit von ablehnenden Verhaltensintentionen. Dieser Zusammenhang wird von Einstellungen zu Frauenkarrieren beeinflusst.*

## 6 Empirische Umsetzung

Präsentiert werden im Folgenden Befunde aus zwei experimentellen Studien, die im Rahmen dieser Arbeit durchgeführt wurden. Kapitel 6.1 behandelt die erste und Kapitel 6.2 entsprechend die zweite Studie. Der Aufbau der Kapitel gleicht sich: Die jeweils ersten Abschnitte (Kapitel 6.1.1 und 6.2.1) widmen sich der Darstellung und Begründung des gewählten methodischen Vorgehens. Darauf aufbauend werden die Studienergebnisse vorgestellt (Kapitel 6.1.2 und 6.2.2) und schließlich in den Kontext des Forschungsinteresses eingeordnet (Kapitel 6.1.3 und 6.2.3).

### 6.1 Studie I – Elaboration und Relevanzwahrnehmung von Hate Speech

#### 6.1.1 Forschungsdesign und Methodik

##### 6.1.1.1 Studiendesign und Umsetzung

Ziel der Studie ist die Untersuchung der Elaboration von Hate Speech. Hate Speech als gruppenbasierte Abwertung wurde in Kapitel 2.2 definiert und kann aber im Hinblick auf die Gestaltung des Kommunikationsinhalts in seiner Intensität graduell variieren. Inzivilen Äußerungen lassen sich als Offensive Speech dementsprechend von Hate Speech unterscheiden und diese von Dangerous Speech, da sie zusätzliche eine Gewaltkomponente enthalten (Benesch, 2013; Fink, 2018; Gagliardone et al., 2016; Maynard & Benesch, 2016). Mit der Unterscheidung der Intensitätsstufen sind unterschiedliche Wirkungsantizipationen verbunden (Maynard & Benesch, 2016).

Der Nachweis eines Kausalzusammenhangs zwischen verschiedenen Intensitätsstufen von Hate Speech und ihren antizipierten persuasiven Wirkungen lässt sich mittels wissenschaftlicher Experimente erbringen. Experimente sind nicht als Datenerhebungsmethode misszuverstehen, sondern stellen eine bestimmte Form der Untersuchungsanlage dar (Brosius et al., 2016, S. 217). Drei Kriterien müssen Andreas Diekmann (2007) zufolge erfüllt sein, um von wissenschaftlichen Experimenten zu sprechen: (1) Auf mindestens zwei Experimentalgruppen werden (2) die Versuchspersonen randomisiert diesen Gruppen zugewiesen. Und (3) die unabhängige Variable wird von der Untersuchungsleitung manipuliert, d. h. unter kontrollierten Bedingungen systematisch variiert (Diekmann, 2007, S. 337). Abgeleitet sind diese aus den drei Logikthesen zur Inferenz kausaler Beziehungen von John Stuart Mill: (1) Ursache geht Wirkung zeitlich voraus, (2) Ursache und Wirkung kovariieren und (3) Alternativerklärungen

für den Zusammenhang zwischen Ursache und Wirkung werden ausgeschlossen (Eifler & Leitgöb, 2019).

Vorab wurde die Befragung auf Verständnisprobleme, Skalenlänge sowie technische Mängel (Filterführung, Funktionalität der Zufallsauswahl des Stimulusmaterials) von 26 Personen aus dem persönlichen und akademischen Netzwerk der Autorin geprüft. Die gewonnenen Daten aus diesem Pretest wurden nicht statistisch ausgewertet. Die hier vorgestellte Untersuchung wurde als teil-standardisierte Online-Befragung mit einem experimentellen Between-Subjects-Design konzipiert und zwischen dem 15. Februar und dem 03. Mai 2018 realisiert. Rekrutiert wurden die Versuchspersonen aus dem deutschen SoSci-Panel (s. ausführlich Kapitel 6.1.1.4). Es handelt sich um ein 4 x 2 x 2-faktorielles Design, bei dem Botschaftsmerkmale, und zwar die *Intensität von Hate Speech* (sachliche Kommentare, Offensive Speech, Hate Speech und Dangerous Speech) sowie das *Geschlecht der Produzent:innen* (weiblich, männlich) und situative Merkmale<sup>48</sup>, nämlich das *Reaktionsvolumen* auf die Kommentare (hohe Likeanzahl, niedrige Likeanzahl), variiert wurden. Die vier Gruppen, die sachliche Kommentare gelesen haben, wurden als Kontroll- bzw. Vergleichsgruppen zu den Experimentalgruppen aufgestellt. Die sachlichen Kommentare sind dennoch insofern keine „echte“ Kontrollgruppe, da sie ebenfalls ein Treatment darstellen, bei dem die Intensität von Hate Speech bei Null liegt, da keinerlei abwertende, aggressive und diskriminierende Aussagen geäußert werden. Aus diesem Grund wird in den folgenden Ausführungen die Bedingung „sachliche Kommentare“ als Experimentalgruppe ohne Hate Speech verstanden.

Die Zuweisung zu einer der Experimentalbedingungen erfolgte randomisiert und erlaubt damit die Neutralisierung von unbekanntem Drittvariablen (Diekmann, 2007, S. 339). Mögliche Störeinflüsse durch unbekanntem Drittvariablen fallen dadurch nicht mehr ins Gewicht und mögliche Unterschiede in der gemessenen abhängigen Variablen können ausschließlich auf das experimentelle Treatment zurückgeführt werden (ebd.).

Die große Beliebtheit von Online-Experimenten ist auf ihre Einfachheit bei der Rekrutierung einer großen Zahl an Versuchspersonen und in der Durchführung zu sehen (Anwyl-Irvine et al., 2020). Wesentlicher Kritikpunkt ist bis heute die fehlende Kontrolle während der Durchführung der Untersuchung (Heerwegh, 2009; Reips, 2000). So könnten Versuchspersonen durch die Anwesenheit weiterer Personen im Raum abgelenkt sein oder verzerrt antworten, da ihre Eingaben beobachtet oder kommentiert werden. Verzerrungseffekte sind auch zu erwarten, wenn Müdigkeit oder Langeweile eine ernsthafte Beantwortung der Fragen

---

<sup>48</sup> Begründet wird die Auffassung des Reaktionsvolumens als situatives Merkmal in Kapitel 4.3.1 und in Kapitel 5.2.

gefährden oder in Item-Nonresponse münden. Darüber hinaus können externe Situationen, die zum Abbruch führen, nicht kontrolliert werden. Aber auch mehrfache Teilnahmen von besonders motivierten Versuchspersonen sind problematisch, v. a. bei großzügigen materiellen und immateriellen („der Meinung Gehör verschaffen“) Incentivierungen. Dazu stellt Selbstselektion im Zusammenhang mit empirischer Sozialforschung ein Problem im Rekrutierungsstadium dar. Personen, die freiwillig an wissenschaftlichen Untersuchungen teilnehmen, können sich hinsichtlich wichtiger Merkmale (Einstellungen, Soziodemographie, Sozioökonomie) von Nicht-Teilnehmer:innen unterscheiden. Selbstselektion bei der Zuordnung zu einer Experimental- oder Kontrollgruppe wird in experimentellen Untersuchungen hingegen durch Einsatz von Randomisierungsverfahren ausgeschlossen.

**Tabelle 5:** Schematische Notation der Versuchsanordnung

	O <sub>1</sub>	X <sub>A1B1C1</sub>	O <sub>17</sub>
	O <sub>2</sub>	X <sub>A1B2C1</sub>	O <sub>18</sub>
	O <sub>3</sub>	X <sub>A1B1C2</sub>	O <sub>19</sub>
R	O <sub>4</sub>	X <sub>A1B2C2</sub>	O <sub>20</sub>
	.	.	.
	.	.	.
	.	.	.
	O <sub>16</sub>	X <sub>A4B2C2</sub>	O <sub>32</sub>
R	Randomisierung		
O <sub>1-16</sub>	Vorher-Messung (Einstellungen, SNS-Nutzung)		
O <sub>17-32</sub>	Nachher-Messung (Elaboration, Relevanz, Bewertung)		
X <sub>A1B1C1</sub> bis X <sub>A4B2C2</sub>	Treatment in den Experimentalgruppen, mit A = Intensität (aufsteigend von <i>sachlich</i> bis <i>dangerous</i> ), B = Geschlecht (weibliche Kommentarschreiberinnen, männliche Kommentarschreiber), C = Reaktionsvolumen (viele und wenige Likes)		

(Quelle: Eigene Darstellung, in Anlehnung an Diekmann, 2007, S. 338)

Befunde aus Laborexperimenten sind zwar weitgehend frei von Störfaktoren, da die Versuchssituation stark kontrolliert ist (*interne Validität*), lassen sich aber nur eingeschränkt auf realweltliche soziale Situationen übertragen. Online durchgeführte experimentelle Untersuchungen finden hingegen in der natürlichen Umgebung der Versuchspersonen statt und zeichnen sich dadurch durch eine hohe *externe Validität* aus (Brosius et al., 2016, S. 240). Ein weiterer forschungspragmatischer Grund spricht für die Durchführung der Studie als Online-Experiment: Untersucht werden soll die Wirkung sexistischer Hasskommentare in dem

reichweitenstärksten SNS in Deutschland zum Zeitpunkt der Erhebung, nämlich Facebook (Kemp, 2018). Die Versuchspersonen können in ihrer persönlichen Nutzungsumgebung, an dem Gerät, das sie auch zur Facebook-Nutzung verwenden, an der Studie teilnehmen.

Die gewonnenen Ergebnisse aus einem Experiment lassen keine Generalisierbarkeit auf Populationsebene zu, es sei denn die Stichprobe wäre repräsentativ zur Grundgesamtheit (Brosius et al., 2016). Wohl können aber Aussagen getroffen werden, die den Einfluss der unabhängigen Variable auf die abhängige Variable („Alkohol am Steuer beeinträchtigt die Fahrtüchtigkeit.“) nachweisen (ebd.).

#### 6.1.1.2 *Aufbau des Fragebogens*

Zunächst wurden die Versuchspersonen zu einer Studie zum Thema „Karrierechancen von Frauen in der IT-Branche“ begrüßt und es wurde eine informierte Einwilligung zur Teilnahme eingeholt, bei der die Freiwilligkeit, die Anonymität, der Erhebungszweck sowie mögliche Risiken der Teilnahme erläutert wurden. Bei Nicht-Einwilligung wurde die Befragung abgebrochen. Die Aufklärung über Risiken (*„Es ist nicht zu erwarten, dass Sie gesundheitlichen Schaden durch die Teilnahme an dieser Studie nehmen. Da es sich um eine medienpsychologische Studie handelt, kann es jedoch vorkommen, dass Ihre Stimmung kurzzeitig beeinflusst wird. Sie werden Facebook-Kommentare lesen, die für manche Menschen problematisch sind.“*) war eine von mehreren notwendig gewordenen Bedingungen der SoSci-Panel-Gutachter:innen, um forschungsethische Bedenken auszuräumen bzw. zu entkräften. Es ist nicht auszuschließen, dass Befragte bereits zu diesem frühen Zeitpunkt der Teilnahme entsprechend geprimt wurden, was sich auf die interne Validität der Daten auswirken kann. Auf diese Einführung folgte eine Filterfrage, um regelmäßige Facebook-Nutzer:innen als Zielgruppe der Erhebung zu identifizieren und für die weitere Befragung zu behalten. Auf diese Weise wurde die externe Validität erhöht, da anzunehmen war, dass Nutzer:innen mit der Nutzungsweise des SNS und der optischen Aufbereitung des Stimulus vertraut sind, was für Nicht-Nutzende problematisch wäre.

Im nächsten Schritt erfolgte die Abfrage der gewohnheitsmäßigen Nutzungsweise des SNS, um den Aktivitätsgrad der Nutzung zu identifizieren. Vermutet wurde, dass ein Zusammenhang zwischen der gewohnheitsmäßigen SNS-Nutzung und der Verhaltensintention nach der Rezeption des Stimulus besteht. Als nächstes wurden die Kurzsкала *Need for Cognition* zur Messung des Kognitionsbedürfnisses gemeinsam mit einer Kurzsкала erwünschten Verhaltens und im nächsten Schritt die Skala zu ambivalent sexistischen Einstellungen gegenüber Frauen abgefragt. Vor Präsentation des Stimulus wurden das Vorwissen und das persönliche Involvement erhoben. Bei Verneinung sich regelmäßig zu „Karrierechancen von Frauen“ zu

informieren, gelangten die befragten Personen direkt auf die Einleitungsseite vor Präsentation des Stimulus. Diese Einleitungsseite wurde zum einen von den SoSci-Panel-Gutachter:innen gefordert, um die Befragten auf die Stimulus-Präsentation vorzubereiten („*Unter dem Posting ist eine Diskussion zum Thema Karrierechancen von Frauen entbrannt.*“), zum anderen dient sie als Vorbereitung zur Elaboration, enthält sie doch die Aufforderung zur sorgfältigen Rezeption: „*Bitte lesen Sie sich sowohl das Posting als auch die Kommentarbeiträge der Nutzer aufmerksam durch. Für den weiteren Verlauf der Befragung ist es wichtig, dass Sie diese Seite nicht nur überfliegen.*“

Von dieser Hinweisseite aus gelangten die Versuchspersonen randomisiert zu einer der 16 Experimentalgruppen, welche im nächsten Abschnitt (Kapitel 6.1.1.3) detailliert dargestellt werden. Die Versuchspersonen konnten die Dauer und Intensität der Rezeption selbst bestimmen, um so eine möglichst natürliche Rezeptionssituation zu ermöglichen. Im Anschluss wurden die Versuchspersonen nach der persönlichen Relevanz der gelesenen Kommentare befragt und sollten im nächsten Schritt bis zu zehn Rezeptionsgedanken in freie Textfelder eintragen. Diese beiden Fragen sollten möglichst direkte Eindrücke nach der Stimulus-Rezeption abbilden, weswegen die Fragen zum Treatment-Check nicht unmittelbar im Anschluss gestellt wurden. Diese folgten erst nach Erhebung der Verhaltensintention, bei der die Befragten die Wahrscheinlichkeit der vorgelegten Reaktionsitems angeben sollten. Schließlich wurden die Befragten gebeten, soziodemografische Angaben (Geschlecht, Alter, Bildung) zu machen. Auf der letzten Seite folgte das Debriefing, bei dem das Erkenntnisinteresse und Hintergründe zur Studie erläutert wurden. Zum Schluss wurden Online-Hilfsangebote verlinkt, welche Aufklärung und Hilfestellungen für Betroffene von Hate Speech und Interessierte leisten.

### 6.1.1.3 *Stimulusmaterial*

Das verwendete Stimulusmaterial wurde einem Facebook-Posting des offiziellen *heise online*-Kanals mit einer Auswahl von fünf Kommentaren nachempfunden. *heise online* ist ein Online-Nachrichtenanbieter, der schwerpunktmäßig das Feld der Informations- und Kommunikationstechnologien (IKT) inklusive deren gesellschaftspolitischen Auswirkungen bedient. Damit ist *heise online* thematisch näher an der Ausrichtung der Arbeit als beispielsweise eine inhaltlich vielfältige Tageszeitung. Der verwendete Beitrag mit dem Titel „*Wo sind die Frauen? Die IT-Branche hat ein Problem*“ (Tobien, 2017) ist vom 01. März 2017, wurde jedoch nicht bei Facebook veröffentlicht. Für die vorliegende Studie wurde der Artikel als *heise online* Facebook-Posting nebst angeschlossenen Diskussionsthread mit schriftlicher Erlaubnis

manipuliert<sup>49</sup> und als Bild-Stimulus gezeigt. Damit ist die externe Validität zwar eingeschränkt, was forschungsethischen und forschungsökonomischen Aspekten geschuldet ist. Die Präsentation des Diskussionsthreads unter realen Bedingungen (d. h. innerhalb des Facebook-Nutzungsprofils der Teilnehmenden) wurde aus forschungsethischen Bedenken verworfen, da es die Belastung der Versuchspersonen in bedenklichem Maße erhöht hätte. Um die Belastung in Grenzen zu halten, wurde die Präsentation des Diskussionsthreads als Bild gewählt, da die Versuchspersonen auf diese Weise kognitiven und emotionalen Abstand zum Stimulusmaterial einnehmen konnten. Auch rein pragmatische Gründe sprachen für die Präsentation des Stimulus als Bild, da die Experimentalbedingungen konstant gehalten werden konnten, z. B. sind interaktive Elemente in einem Online-Experiment kaum kontrollierbar, selbst geringfügigere Störquellen wie die Zeitstempel bei der Veröffentlichung des Postings und der Kommentare sind problematisch.

Die Präsentation des Postings selbst wurde konstant gehalten, nur die Diskussionsumgebung wurde manipuliert. Der Beitrag entsprach einem üblichen Facebook-Posting von *heise online* mit zwei kurzen einleitenden Sätzen und einer Bildervorschau auf den Artikel, um interessierte Leser:innen für die Thematik zu gewinnen. Vermieden wurden Reizwörter wie „Gender“, „Sexismus“, „Geschlechterkluft“, um keinen wertenden Eingriff des Nachrichtenportals zu suggerieren. Des Weiteren wurden auf Beitragsebene die Zeitstempel der Veröffentlichung, die Anzahl an Likes sowie die sichtbaren Reaktionsemojis für das *heise-online*-Posting in allen Experimentalgruppen konstant gehalten.

Versuchspersonen sahen fünf Kommentare von *Frauen* bzw. von *Männern*. Indem das Geschlecht der kommentierenden Person innerhalb einer Stimulusdarbietung konstant gehalten wurde, sollte der Prime auf die Geschlechterkategorie aktiviert werden (s. Abbildung 8). Geschlechtszugehörigkeiten wurden mittels Verwendung üblicher deutscher Vornamen suggeriert (Steffi, Sabine oder Christian, Mark). Die Nachnamen wurden unkenntlich gemacht, ebenso die Profilbilder. Die Manipulation des Reaktionsvolumens bestand darin, die Anzahl der Likes auf die einzelnen Kommentare in *geringe* (niedrige einstellige Zahlen) und *hohe* Like-Zahlen (höhere zwei- bis niedrige dreistellige Zahlen) einzuordnen. Die Werte orientierten sich dabei an tatsächlichen Like-Zahlen auf dem Facebook-Kanal von *heise online* und weichen von tatsächlich als hoch einzustufenden Werten von tausenden Likes ab (Knoll, 2015).

---

<sup>49</sup> E-Mail vom 13.10.2017 mit dem stellvertretenden Chefredakteur von *heise online*, Jürgen Kuri, liegt vor.

**Abbildung 8:** Hate Speech Kommentar (Hate Intensität) von Frauen (Geschlecht der Kommentierenden) mit vielen Likes (Reaktionsvolumen)

**heise online**  
1 Std · 🌐

Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



**CRACKING THE CODE**  
GET 3X MORE WOMEN IN COMPUTING

Wo sind die Frauen? Die IT-Branche hat ein Problem  
Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.  
HEISE.DE

👍 Gefällt mir    💬 Kommentieren    ➦ Teilen

👍 😂 🤔 107    Top-Kommentare ▾

**Steffi** Das ist doch genetisch bedingt, dass eine Trulla besser mit Kindern umgehen kann. Wie kann man nur so dumm sein und es nicht einsehen...  
Gefällt mir · Antworten · 🗳️ 112 · 6 Min

**Mela** Jeden IT-Affinen den ich kenne würde vor Freude fast ausflippen wenn er eine träfe mit der er hemmungslos über Technik quatschen könnte. Statt Mathe und IT zu machen heulen die Dummschützen und fordern Sonderprivilegien.  
Gefällt mir · Antworten · 🗳️ 79 · 34 Min

**Sabine**



**FROLLEIN PROGRAMMIERERIN**  
**GEH ZURÜCK AN DEN HERD!**

Gefällt mir · Antworten · 🗳️ 130 · 27 Min

**Monika** Mit erzwungenen FeminiSStinnenquoten die besser gebildeten und wesentlich fähigeren männer in die Arbeitslosigkeit schicken. Tolle Geisteshaltung haben diese Bitterfregatten.  
Gefällt mir · Antworten · 🗳️ 35 · 1 Std

**Andrea** Eine feste Quote für Frauen in Chef-Positionen kann nur einem Fieberwahn entspringen. Lasst die unfähigen Personen ruhig in die Chef-Etagen...oh weh oh weh.  
Gefällt mir · Antworten · 🗳️ 18 · 58 Min

👤 Kommentieren ...    😊 📷 GIF 🗑️

Innerhalb des Diskussionsthreads wurde der dritte Kommentar konstant als Image Macro dargeboten, also einem Bild mit variierendem Text in einer bestimmten Schriftart in weißer Schriftfarbe. Abgebildet war in allen Experimentalgruppen eine Person mit lackierten Fingernägeln, die vor einem aufgeklappten Notebook ein Smartphone in der Hand hält. Das Gesicht war nicht Teil des Bildausschnitts. Es wurden vier unterschiedliche Versionen des Textes dargeboten, die aus den inhaltlichen Variationen erwachsen. In den Experimentalgruppen wurden die Kommentare auf Basis der von Gagliardone und Kollegen (2016) vorgeschlagenen Unterscheidung zwischen *Offensive* (OS), *Hate* (HS) und *Dangerous Speech* (DS) konstruiert. In der *Offensive*-Variation wurde das Konzept der Inzivilität (Muddiman & Stroud, 2017) operationalisiert, d. h. es wurden inzivile Äußerungen versammelt, die beleidigend und polemisch die Thematik kommentierten („*Das ist doch genetisch bedingt, dass Frauen besser mit Kindern umgehen können. Und außerdem hat die Floristik-Branche auch ein Gender-Problem. Keine Männer*“). Manipulationen in der *Hate*-Gruppe waren aggressiver in der Wortwahl, was sich in herabsetzenden, dehumanisierenden und deutlich diskriminierenden Aussagen zeigte („*Statt Mathe und IT zu machen, heulen die Dummschützen und fordern Sonderprivilegien.*“, „*Eine feste Quote für Frauen in Chef-Positionen kann nur einem Fieberwahn entspringen. Lasst die unfähigen Personen ruhig in die Chef-Etagen...*“). Die *Dangerous*-Variation ist forschungsethisch als besonders heikel zu bewerten, da Gewaltandrohungen oder -andeutungen geäußert wurden („*Schade, dass mir so eine nicht im Dunkeln begegnet.*“, „*Sowas entspringt auch nur einem kranken Asselhirn.*“). Forschungsethische Bedenken wurden daher über mehrere Maßnahmen angegangen, nämlich in Form von Hinweisen zu Beginn der Befragung (Freiwilligkeit und mögliche kurzzeitige Beeinträchtigungen der Stimmung), die Möglichkeit jederzeit die Befragung abzubrechen, vor allem mittels eines Exit-Buttons nach der Stimuluspräsentation und eines ausführlichen Debriefings inklusive der Nennung von Anlaufstellen für Hilfestellungen bei eigener Betroffenheit.

In der *Kontrollgruppe* wurden sachliche Kommentare (SK) zu dem gleichen Facebook-Posting verwendet („*Wenn ich Arbeitgeber wäre, wäre es mir wichtig, gute Leute zu bekommen, egal welchen Geschlechts.*“). Um das Stimulusmaterial zu konstruieren, wurden echte Kommentare und Versatzstücke daraus herangezogen, die in Kommentarspalten unterhalb von journalistischen Beiträgen, in SNS oder in Foren und Gruppen geäußert wurden und einer sprachlichen und inhaltlichen Anpassung durch die Autorin dieser Arbeit unterzogen. Da in SNS Fehler bezüglich Grammatik, Orthografie und Interpunktion verbreitet sind, wurden diese bei der Konstruktion des Stimulusmaterials ebenfalls berücksichtigt.

#### 6.1.1.4 Stichprobenbeschreibung

Die Versuchspersonen wurden aus dem SoSci-Panel rekrutiert, einem nicht-repräsentativen Convenience-Panel, das an dem Münchner Institut für Kommunikationswissenschaft und Medienforschung der Münchner Ludwig-Maximilians-Universität in Zusammenarbeit mit der Deutschen Gesellschaft für Publizistik und Kommunikationswissenschaft (DGPK) entstanden ist (Leiner, 2016). Die Nutzung des Panels versprach einen günstigen Zugang zu einer hohen Anzahl an teilnehmenden Personen, um die komplexe Experimentalanlage realisieren zu können.

Zwischen dem 15. Februar 2018 und dem 03. Mai 2018 wurde der Fragebogen insgesamt 1053-mal gestartet. 167 Personen, knapp 16 Prozent, wurden von der Befragung ausgeschlossen, da sie Facebook nicht regelmäßig nutzen, was als Teilnahmevoraussetzung definiert war. Weitere 28 Personen (rund 3 Prozent) haben auf dieser Seite die Befragung verlassen. Unmittelbar nach der Stimuluspräsentation gab es die meisten Abbrüche, was zum einen aufgrund von Belastungen, ausgelöst durch das Stimulusmaterial und damit verbunden den eingesetzten Abbruchoptionen bewirkt wurde, zum anderen aber auch aufgrund möglicher Reaktanzeffekte ausgelöst durch Fragen nach der persönlichen Relevanz, Rezeptionsgedanken und Verhaltensintentionen. Es handelt sich um 321 Personen (30,5%), die in diesem Zusammenhang die Befragung vorzeitig abgebrochen haben, was als sehr hoch bewertet werden muss. Insgesamt wurden 49,4 Prozent der begonnenen Befragungen abgebrochen. In der Bruttostichprobe verbleiben damit 537 Personen, die den Fragebogen abgeschlossen haben.

Um die Datenbereinigung weiter voranzutreiben, wurde die Bearbeitungszeit herangezogen, mit dem Ziel die Qualität der verbliebenen Antworten einzuschätzen. Im Durchschnitt benötigten die Versuchspersonen 11.46 Minuten ( $SD = 4.4$  Minuten). Da die Bearbeitung der Fragen mindestens 6 Minuten benötigt, wurden 34 Personen aus dem Datensatz entfernt, die schnellere Bearbeitungszeiten aufwiesen. Schließlich wurden zwei minderjährige Teilnehmende aus dem Datensatz entfernt. Der Datensatz musste ferner um 14 weitere Fälle bereinigt werden, da aus technischen Gründen (möglicherweise Inkompatibilität des Browsers; wurde jedoch nicht erhoben) bei diesen das Stimulusmaterial nicht angezeigt wurde. Zwar wurde vor der Stimuluspräsentation darauf hingewiesen, dass dieses Problem durch Neuladen der Seite im Browser behoben werden könne, was aber ganz offensichtlich übersehen wurde. Damit liegen in der vorläufigen Stichprobe 487 gültige Fälle für die Auswertung vor.

Die Befragten waren zwischen 18 und 79 Jahre alt und ihr Durchschnittsalter lag bei 33.89 Jahren ( $SD = 12.91$ ). Zum Zeitpunkt der Erhebung stellt die Altersklasse der 25-34-

Jährigen die größte unter den Facebook-Nutzer:innen in Deutschland dar (Kemp, 2018). Mehr als zwei Drittel der Befragten (68 Prozent) ordneten sich dem weiblichen Geschlecht zu, 155 Personen (32 Prozent) ordneten sich dem männlichen Geschlecht und 2 Personen (0,4 Prozent) einer anderen Geschlechtszugehörigkeit zu. Da der Anteil der beiden Personen, die sich weder dem weiblichen noch dem männlichen Geschlecht zuordnen, zu niedrig war, um weitergehende Berechnungen durchzuführen und die Geschlechtskategorie im Rahmen dieser Untersuchung jedoch zentral ist, wurden diese aus der vorläufigen Stichprobe entfernt, so dass die finale Stichprobengröße  $n = 485$  Personen beträgt.

**Tabelle 6:** Stichprobengröße vor und nach Datenbereinigung

Gruppe	Vorher	Nachher	Geschlecht der Versuchspersonen		Differenz
			Weiblich	Männlich	
SK1	50	30	23	7	-20
SK2	50	34	27	7	-16
SK3	48	27	16	11	-21
SK4	50	33	22	11	-17
OS1	49	31	24	7	-18
OS2	49	27	18	9	-22
OS3	50	29	18	11	-21
OS4	50	33	24	9	-17
HS1	50	29	20	9	-21
HS2	49	32	25	7	-17
HS3	48	34	25	9	-14
HS4	49	31	23	8	-18
DS1	51	24	11	13	-27
DS2	51	34	21	13	-17
DS3	50	23	15	8	-27
DS4	49	34	18	16	-15

*Anmerkungen:* **SK1** = Sachliche Kommentare, viele Likes, Frauen; **SK2** = Sachliche Kommentare, viele Likes, Männer; **SK3** = Sachliche Kommentare, wenig Likes, Frauen; **SK4** = Sachliche Kommentare, wenig Likes, Männer; **OS1** = Offensive Speech, viele Likes, Frauen; **OS2** = Offensive Speech, viele Likes, Männer; **OS3** = Offensive Speech, wenig Likes, Frauen; **OS4** = Offensive Speech, wenig Likes, Männer; **HS1** = Hate Speech, viele Likes, Frauen; **HS2** = Hate Speech viele Likes, Männer; **HS3** = Hate Speech, wenig Likes Frauen; **HS4** = Hate Speech, wenig Likes, Männer; **DS1** = Dangerous Speech, viele Likes, Frauen; **DS2** = Dangerous Speech, viele Likes, Männer; **DS3** = Dangerous Speech, wenig Likes, Frauen; **DS4** = Dangerous Speech, wenig Likes, Männer.

Angegeben sind absolute Häufigkeiten.

$N = 793$  vor der Datenbereinigung und  $N = 485$  nach der Datenbereinigung

Die Teilnehmenden waren überdurchschnittlich hoch gebildet: Mehr als die Hälfte (54,6 Prozent) verfügt über einen Hochschulabschluss, ein knappes Drittel (31,3 Prozent) gab das Fachabitur oder Abitur als höchsten Bildungsabschluss an. Knapp fünf Prozent haben eine abgeschlossene Lehre und die restlichen Werte verteilten sich auf niedrigere Abschlüsse, Schüler:innen oder sonstige Angaben. Hinsichtlich der Berufstätigkeit geht rund die Hälfte

der Befragten (49,1 Prozent) einer Beschäftigung nach, davon die meisten als Angestellte in Vollzeit (26,4 Prozent), in Teilzeit (11,8 Prozent), als Beamte (5,2) oder als Selbständige (5,6 Prozent). Mehr als ein Drittel (35,2 Prozent) studiert und 3,7 Prozent befinden sich in Ausbildung. Knapp jede:r Zehnte wählt die Kategorie „Sonstiges“, worunter sich mehrheitlich Rentner:innen, Mütter und Väter in Elternzeit oder Doktorand:innen einordnen. Mit drei Prozent ist der Anteil der Arbeitssuchenden in der Stichprobe sehr gering.

Von den ursprünglich jeweils rund 50 Personen je Experimentalgruppe, die sich als Facebook-Nutzer:innen vor der weitergehenden Datenbereinigung bekannten, brachen in den einzelnen Gruppen auffällig viele die Befragung ab (s. Tabelle 6). Da das Geschlecht erst nach Stimuluspräsentation abgefragt wurde, bleibt ungeklärt, ob eher Männer oder eher Frauen die Befragung abgebrochen haben. Auffällig ist jedoch, dass die Differenz bei den Verlusten durch Befragungsabbrüche in den Experimentalgruppen, in denen Frauen kommentieren, größer war. Der Phi-Koeffizient zeigte an, dass dieser Zusammenhang deutlich jedoch schwach signifikant mit  $\varphi = .5$ ,  $p < .046$  war. In Kapitel 3.2.2 wurde erläutert, dass Abweichungen vom globalen Geschlechtsstereotyp negative Gegenreaktionen nach sich ziehen. Ob abwertende und aggressive Aussagen von Frauen wie von Männern jeweils als eine solche Abweichung verstanden werden und welchen Einfluss sie zeigen, soll im Verlauf der Ergebnisdarstellung vertiefend analysiert werden. Die (sozial-)psychologische Experimentalforschung setzt zur Überprüfung der erfolgreichen Manipulation Zusatzfragen ein, um die intendierte Wirkung abzufragen (Perdue & Summers, 1986), was jedoch methodisch nicht unumstritten ist (Hauser et al., 2018). Dabei zeigte sich, dass die Manipulation der Likeanzahl nicht gelungen war, da nur 2,5 Prozent der Befragten das Treatment richtig wahrgenommen hatten. Auch wenn sich die Versuchspersonen nicht bewusst an die Anzahl der Likes für das Igemakro erinnern konnten, muss diese Experimentalbedingung dennoch aus zwei Gründen bei der Analyse berücksichtigt werden: (1) Methodische Gründe sprechen für die Berücksichtigung, da sichergestellt werden muss, dass alle Einflüsse kontrolliert werden und daraus abgeleitet sind (2) mit dem *Mere Exposure-Effekt* theoretische Gründe anzuführen. Damit lassen sich Einstellungsänderungen nachweisen, ohne dass sich Versuchspersonen an Stimulusdarbietungen erinnern (Janiszewski, 1993; Schemer et al., 2007). Das Geschlecht der kommentierenden Personen im Stimulusmaterial wurde von 340 Versuchspersonen (70,1 Prozent) richtig erkannt und diese Teilstichprobe wurde dann zur Überprüfung der Hypothesen herangezogen.

### 6.1.1.5 Operationalisierung der zentralen Konstrukte

Effekte von rezipierten Hate Speech-Aussagen sind auf kognitiver, affektiver wie auf Verhaltensebene zu erwarten. Im Folgenden werden die Variablen und Konstrukte vorgestellt, welche der Erfassung dieser Effekte dienen.

#### Messungen zur Erfassung kognitiver und affektiver Effekte

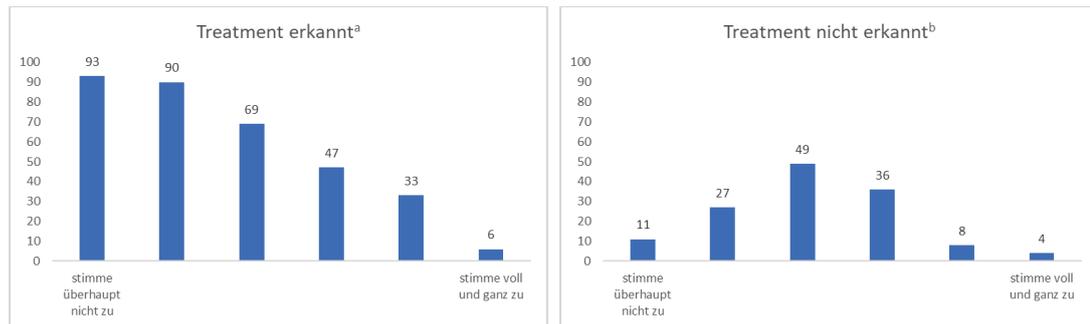
In der ELM-Forschung und hier insbesondere in der Konsumentenforschung konnte bereits ein direkter Zusammenhang zwischen der persönlichen Relevanz von Persuasionsbotschaften und Persuasionseffekten nachgewiesen werden. Die *persönliche Relevanz* wurde mit vier Items der Message-Relevance-Skala (Laczniak & Muehling, 1993) erfasst, die auch von Knoll (2015) eingesetzt wurde. Auf die Frage „Finden Sie die gelesenen Kommentare für sich selbst relevant?“ konnten die Versuchsteilnehmer:innen auf einer sechsstufigen Skala ihre Zustimmung von „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“ gegenüber den Items zum Ausdruck bringen. Die Items wurden zu einem Durchschnittsindex zusammengefasst mit einer ausgezeichneten Reliabilität der Skala mit  $\alpha = .88$  ( $M = 2.57$ ,  $SD = 1.37$ ). Die Antworten schöpften zwar die gesamte Skala ab, befanden sich aber mehrheitlich im niedrigen Zustimmungsbereich. Die wahrgenommene Relevanz fiel damit eher niedrig aus, was sich auch im Median von 2 niederschlug.

**Tabelle 7: Deskriptive Kennwerte für die Items für wahrgenommene persönliche Relevanz**

Item	M	SD	Schiefe	Wölbung
Die Kommentare sind wichtig für mich. ( $n = 338$ )	2.35	1.44	0.96	-0.67
Es lohnt sich, die Kommentare im Kopf zu behalten. ( $n = 339$ )	2.41	1.49	0.80	-0.50
Die Kommentare sind es wert, ihnen Aufmerksamkeit zu schenken. ( $n = 338$ )	2.51	1.57	0.72	-0.67
Die Kommentare sind interessant für mich. ( $n = 338$ )	2.62	1.59	0.66	-0.79
Index wahrgenommene persönliche Relevanz ( $n = 338$ )	2.57	1.37	0.56	-0.67

Die rechtsschiefe Verteilung der wahrgenommenen persönlichen Relevanz der Versuchspersonen, die die experimentelle Manipulation des Geschlechts nicht wahrgenommen hatte, lässt sich von der asymptotischen Normalverteilung in der Gruppe derjenigen, die das Treatment nicht wahrgenommen haben, in Abbildung 9 ablesen. Die Prüfung der statistischen Signifikanz bestätigte den visuellen Eindruck eines Zusammenhangs zwischen persönlicher Relevanz und Treatmentwahrnehmung mit  $(\chi^2(3) = 39.16, p < .001)$  und signifikanten Unterschieden in Kennwerten der Verteilung (Kolmogorov  $D(473) = 2.553, p < .001$ ).

**Abbildung 9:** *Wahrgenommene persönliche Relevanz in Abhängigkeit von der Treatment-Wahrnehmung*



<sup>a</sup>N = 338

<sup>b</sup>N = 135

Die *Elaboration* der dargebotenen Stimuli wurde anhand der *Thought-Listing-Technik* ermittelt. Hierbei handelt es sich um ein qualitatives Verfahren, bei dem die Rezeptionsgedanken als offene Texteingabe erfasst werden. Cacioppo und Petty (1981) betonen die Bedeutung einer sorgfältig formulierten Anweisung, um valide Gedanken zu generieren, die sich tatsächlich auf die Stimuluspräsentation beziehen, und um sprunghafte sonstige Gedanken auszuschließen. Abweichend von Caccioppo und Pettys Anweisung wurde darauf verzichtet, eine Reihenfolge der Gedanken einzufordern, um die Versuchspersonen keinem „Prüfungsstress“ auszusetzen. Verwendet wurde folgender Wortlaut: „*Beim Lesen der Kommentare haben Sie sich vermutlich einige Gedanken gemacht. In den Textfeldern haben Sie nun die Möglichkeit bis zu zehn Gedanken einzutragen. Die Reihenfolge der Gedanken spielt keine Rolle. Bitte tragen Sie in jedes Textfeld jeweils nur einen Gedanken ein. Es ist nicht erforderlich zehn Gedanken einzutragen, sondern nur so viele Ihnen einfallen. Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.*“

Die Rezeptionsgedanken wurden einer quantitativen Inhaltsanalyse zugeführt, bei der drei von vier Kategorien, wie von Cacioppo und Petty (1981) vorgeschlagen, codiert wurden: Summe, Valenz (*Polarity Dimension*) sowie Ausgangsbasis der Gedanken bzw. dem Gedankenursprung (*Origin Dimension*). Das Zielobjekt der Gedanken (*Target Dimension*) wurde nicht ausgewertet, da die Trennschärfe zum Gedankenursprung als nicht zufriedenstellend erachtet wird (z. B. ist der Stimulus sowohl Ausgangsbasis als auch Ziel der Gedanken). Die Auswertung der Ausgangsbasis der verschriftlichten Rezeptionsgedanken versprach die *Elaboration* der Befragten angemessen abzubilden. Eine studentische Hilfskraft<sup>50</sup> führte die Codierung blind durch, d. h. ohne Kenntnis der anderen erhobenen Daten; auch die Zuweisung zur Experimentalgruppe blieb der Codiererin und der Autorin dieser Arbeit unbekannt, um

<sup>50</sup> Besonderer Dank gebührt Natascha Löffler, deren Codierung sehr gewissenhaft erfolgte.

Verzerrungen bei der Codierung auszuschließen. In der Kategorie *Summe der Gedanken* wurden die Angaben je Textfeld gezählt, und zwar unabhängig davon, wie ausführlich der Rezeptionsgedanke je Textfeld ausfiel. Wurden also Gedanken in drei Textfeldern eingetragen, wurde dies mit der Ziffer 3 codiert. Die Anforderungen an die Codierung dieser Kategorie sind gering, so dass hier die Intercoder-Reliabilität zwischen der Autorin der vorliegenden Arbeit und der codierenden studentischen Hilfskraft im Pretest vollständig übereinstimmten und einen Holsti-Wert von 1 erreichten. Insgesamt wurden 1890 Rezeptionsgedanken geäußert ( $M = 4.68$ ,  $SD = 2.57$ ). Insgesamt 16,7 Prozent der Versuchspersonen nannten keinen Gedanken. Die Summe der Gedanken kann als korrelationaler Indikator für die Verarbeitungstiefe aufgefasst werden, d. h. je mehr Gedanken geäußert werden, umso höher ist die Verarbeitungstiefe (Cacioppo & Petty, 1981). Problematisch an dieser Überlegung ist jedoch, dass keinerlei Aussagen hinsichtlich der Gedankenqualität getroffen werden können.

Bei der Erfassung der *Valenz der Gedanken* wurde zwischen positiven, negativen und neutralen Gedanken unterschieden. Positive Gedanken wurden codiert, wenn positive Argumente, Eigenschaften oder Assoziationen im Zusammenhang mit dem Stimulus oder der Thematik genannt wurden, z. B. „*Der letzte Kommentar von Freddie... hat mir am besten gefallen und genau ausgedrückt, was ich denke: es geht um die Kompetenz der Leute und nicht um das Geschlecht!*“ Analog dazu wurden negative Gedanken codiert, z. B. „*Peinlich*“, „*Sehr unqualifizierte Beiträge*“. Sachliche Feststellungen und Fakten wurden als neutrale Gedanken codiert, z. B. „*Es handelt sich ausschließlich um Kommentare von Frauen*“. Auch wenn positive oder negative Aussagen durch eine Relativierung abgeschwächt wurden („*halbwegs themenrelevant*“, „*eigentlich ganz angenehm*“), erhielten sie diesen Code. Die Gedanken waren mehrheitlich negativ, denn durchschnittlich wurden 2.67 Gedanken mit negativem Tenor geäußert ( $SD = 2.54$ ), gefolgt von neutralen Gedanken ( $M = 1.77$ ,  $SD = 2.03$ ). Nur 12,9 Prozent der Gedanken waren positiv ( $M = .25$ ,  $SD = .78$ ). Die Valenz der Gedanken wurde aus dem Quotienten der Differenz aus positiven sowie neutralen und negativen Gedanken und der Summe aller Gedanken errechnet. Der Wertebereich lag zwischen 1 (ausschließlich positive und/oder neutrale Gedanken) und -1 (ausschließlich negative Gedanken). Der Valenzwert war deutlich negativ ( $M = -.75$ ,  $SD = 0.62$ ). Durch eine intensive Codierschulung konnte eine hohe Übereinstimmung bei der Codierung der Gedankenvalenz erzielt werden mit einem Holsti-CR von .85.

Der *Gedankenursprung* setzt an der Informationsverarbeitung der Teilnehmer:innen an. Unterschieden wurden Gedanken, die aus der kognitiven und affektiven Auseinandersetzung mit dem Stimulusmaterial resultieren, gewissermaßen „außerhalb“ der Versuchsperson

stehen. Cacioppo und Petty bezeichnen diese als „externally originated thoughts“ bzw. als „modified externally originated thoughts“ (Cacioppo & Petty, 1981, S. 320). Die von den Autoren als externen Ursprungs bezeichneten Gedanken werden im Verlauf der Arbeit als *kognitive Auseinandersetzung mit dem Stimulusmaterial* bezeichnet, denn sie waren durch die Experimentalsituation oder die Wahrnehmung des Stimulus induziert, thematisierten beispielsweise die Sachlichkeit oder Aggressivität der Kommentare, die Anzahl der Likes oder das Geschlecht der Kommentierenden („*Ich soll zehn Gedanken nennen.*“ „*Frauenstimmen hört man nicht.*“, „*Die haben Angst vor Frauen.*“). Rezeptionsgedanken, die bei Petty und Cacioppo umständlich als modifizierte externe Gedanken bezeichnet werden, wurden vorliegend als *affektive Reaktion auf das Stimulusmaterial* codiert, denn dabei handelte es sich um Bewertungen oder Gefühlsreaktionen, ausgelöst nach Rezeption des Stimulusmaterials („*Ich bin sauer, dass diese Männer so frauenfeindlich kommentieren.*“). Gedanken, die nicht direkt auf das Stimulusmaterial zurückgeführt werden konnten, sondern aus der kognitiven Verarbeitung durch einen Erfahrungs- und Wissensabgleich resultierten, wurden als *Vorwissen* („internally originated thoughts“) codiert. Vorwissen in den Rezeptionsgedanken ist dabei selbstredend nicht mit der Selbsteinschätzung bezüglich des Vorwissens der Proband:innen (s. u.) gleichzusetzen. Dabei wurden neue Aspekte der Thematik genannt, so dass eine besonders starke kognitive Auseinandersetzung erkennbar wurde („*Frauen, die in der IT arbeiten, sollen männlichen IT-Spezialisten die Arbeit wegnehmen? Es gibt ohnehin viel zu wenige Fachkräfte in dieser Branche, da nimmt niemand jemand anderem etwas weg.*“). Konnten die Gedanken keiner der drei Kategorien zugeordnet werden, wurden sie in eine „Sonstige“-Kategorie geschoben. Die Codierung des Gedankenursprungs erwies sich als anspruchsvoll, was sich im Reliabilitätsmaß mit dem Holsti-CR von .78 widerspiegelt. Die Kategorie ist inhaltlich komplex und die Ausprägungen ließen sich nur schwer voneinander abgrenzen, weswegen die erreichte Reliabilität als dennoch zufriedenstellend erachtet wurde (Früh, 2007, S. 194f.). Der Großteil der Rezeptionsgedanken entfiel auf erkennbare kognitive Verarbeitungen des Stimulus, nämlich 61,1 Prozent, gefolgt von Vorwissen mit 15,5 Prozent und affektiven Reaktionen mit 14,2 Prozent. Weniger als jeder zehnte Gedanke befand sich in der „Sonstige“-Kategorie (9,1 Prozent).

#### *Messung zur Erfassung der Verhaltensintention*

Um Effekte auf Verhaltensintentionen zu messen, wurden acht Items herausgearbeitet, welche sich an der unmittelbaren Reaktion im Facebook-Nutzungskontext nach Rezeption der Kommentare orientierten. Damit wurden Verhaltensweisen ausgeschlossen, die mit einem Medienbruch einhergehen, wie z. B. ein Anruf bei Freunden oder zwar online erfolgen, aber

dafür andere Nutzungsformen beinhalten, etwa eine weitergehende Online-Recherche o. ä. Die verwendeten Items wurden keiner validierten Skala entnommen, sondern eigens für den Studienzweck formuliert. Skalen zur Verhaltensintention nach Rezeption von Hate Speech existieren bislang nicht, weswegen weiterer Forschungsbedarf bezüglich einer Validierung von Verhaltensintentionen angenommen wird. Die Einschätzung der Verhaltensintention wurde über alle Items auf einer sechsstufigen endbepunkteten Skala von „*sehr unwahrscheinlich*“ bis „*sehr wahrscheinlich*“ gemessen. Damit wurde dem Umstand Rechnung getragen, dass mehrere Reaktionen möglich sind, z. B. die Vergabe eines Likes in Verbindung mit einem zustimmenden Kommentar. Die formulierten Items drückten ein vielfältiges Verhaltensrepertoire aus, das die Akzeptanzrichtung von Zustimmung „*Ich unterstütze meine Vorredner mit einem sachlichen Kommentar*“ über Ablehnung „*Ich widerspreche meinen Vorrednern entschieden mit einem sachlichen Kommentar*“ und Ignorieren „*Ich lese die Kommentare, ohne etwas zu tun*“ bis hin zu affektiven Verhaltensintentionen abbildet („*Ich reagiere mit einem missbilligenden Emoji (,Traurig‘ oder ,Wütend‘)*“). Dazu wurde auch der Aktivitätsgrad der Verhaltensintention berücksichtigt. Höchste Zustimmung erhielt das Ignorieren der Kommentare ( $M = 4.79$ ,  $SD = 1.51$ ). Auf die gelesenen Kommentare mit eigenen beleidigenden Kommentaren zu antworten, wurde als am wenigsten wahrscheinlich angegeben, weder als Widerspruch (VI5) ( $M = 1.27$ ,  $SD = .80$ ) und noch weniger als Unterstützung der Vorredner:innen (VI6) ( $M = 1.09$ ,  $SD = .34$ ). Weniger als zwei Prozent gaben an, dass sie es für wahrscheinlich bis sehr wahrscheinlich (Skalenstufen 5 und 6 von 6) ansehen, beleidigend zu widersprechen (VI5). Für VI6 sah es gerade einmal eine Versuchsperson als etwas wahrscheinlich (Skalenstufe 4 von 6) an, die gelesenen Kommentare zum Anlass zu nehmen und selbst beleidigend gegen die primären Adressat:innen aufzutreten. Für sich selbst schlossen die Befragten überwiegend den Gebrauch verbaler Aggression aus, was sich auch mit den Zahlen zur Prävalenz von Hate Speech-Produzent:innen deckt, denn nur ein kleiner Teil der Internetnutzer:innen ist für einen Großteil des Onlinehasses verantwortlich (Kreißel et al., 2018). Wenn auch die große Mehrheit der Befragten kaum zu einer aktiven Positionierung bereit war – sei es als Ablehnung oder als Befürwortung – gab immerhin jede:r Zehnte an, sachliche Gegenrede (VI4) als wahrscheinliche oder sehr wahrscheinliche Verhaltensintention (Skalenstufen 5 und 6 von 6) einsetzen zu wollen. Wie stark an dieser Stelle Effekte sozialer Erwünschtheit vorlagen, wird im Rahmen der Hypothesentests zu zeigen sein. Tabelle 8 fasst die Befunde zusammen. Sichtbar wird, dass die letzte Variable („Ignorieren“) stark linkschief und alle anderen stark rechtsschief verteilt waren.

**Tabelle 8:** Deskriptive Kennwerte zur Messung der Verhaltensintentionen

Item	M	SD	Schiefe	Wölbung
Ich reagiere mit einem Like (VI1).	1.91	1.50	1.42	0.61
Ich reagiere mit einem missbilligenden Emoji (VI2).	1.87	1.38	1.51	1.13
Ich unterstütze meine Vorredner mit einem sachlichen Kommentar (VI3).	1.74	1.29	1.72	1.97
Ich widerspreche meinen Vorrednern mit einem sachlichen Kommentar (VI4).	2.26	1.59	1.03	-0.19
Ich widerspreche meinen Vorrednern mit einem beleidigenden Kommentar (VI5).	1.27	0.80	3.47	12.70
Ich unterstütze meine Vorredner mit einem beleidigenden Kommentar (VI6).	1.09	0.34	4.72	26.38
Ich melde die Kommentare (VI7).	1.91	1.46	1.58	1.32
Ich lese die Kommentare, ohne etwas zu tun (VI8).	4.79	1.51	-1.11	0.08

Anmerkungen: N = 340

Die Korrelationsmatrix in Tabelle 9 weist einige mittlere bis starke Zusammenhänge auf, wobei aufgrund der Multidimensionalität der Items der Wahrnehmungsraum uneindeutig ist. Während die Korrelation mit einem Pearsons  $r = .37, p < .01$  zwischen VI5 (*beleidigender Widerspruch*) und VI6 (*beleidigende Unterstützung*) nachvollziehbar als *aktive Aggressivität* verstanden werden kann, erscheint die hohe negative Korrelation zwischen VI4 (*sachlicher Widerspruch*) und VI8 (*Ignorieren*) mit einem Pearsons  $r = -.47, p < .01$  auf den ersten Blick nicht vereinbar zu sein. Um den Wahrnehmungsraum der Versuchspersonen zu validieren, wurde eine Multidimensionale Skalierung mittels PROXSCAL-Verfahren durchgeführt (Backhaus et al., 2018). Als Distanzmaß wurde die Euklidische Distanz mit standardisierten Z-Scores gewählt (Rohrlack, 2009). Die Güte der Modellierung bewegt sich mit einem S-Stress-Wert von .06 und einer erklärten Streuung von .96 in einem akzeptablen Bereich (Borg, 2010). Das Ergebnis der multidimensionalen Skalierung ist in Abbildung 10 abgebildet.

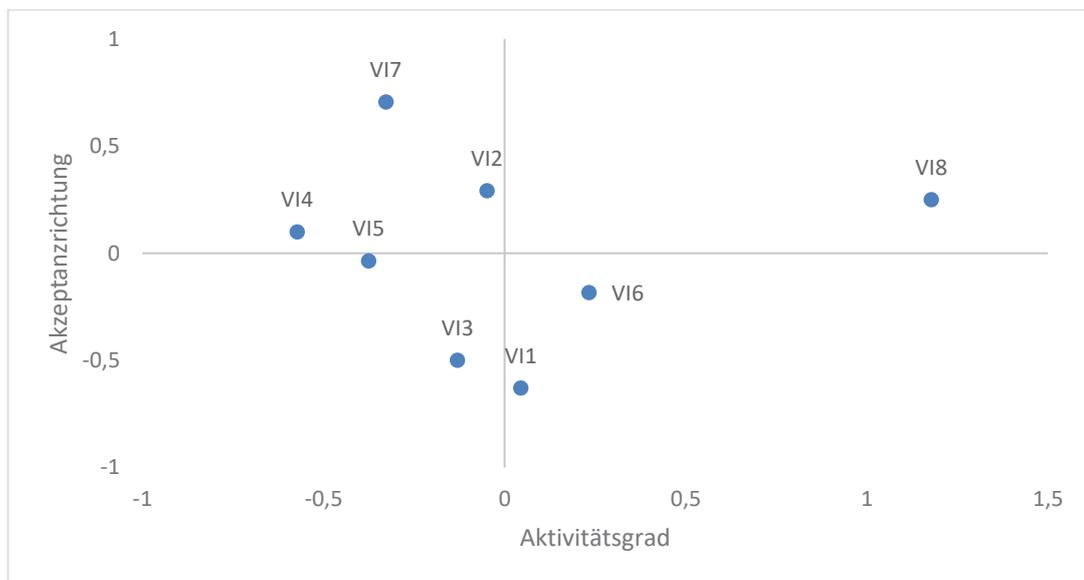
**Tabelle 9** Korrelationen der Items Verhaltensintention (VI)

	VI1	VI2	VI3	VI4	VI5	VI6	VI7	VI8
VI1	–							
VI2	.21**	–						
VI3	.57**	.23**	–					
VI4	.06	.34**	.35**	–				
VI5	.08	.21**	.20**	.38**	–			
VI6	.18**	.11	.26**	.13*	.37**	–		
VI7	-.12*	.18**	.02	.27**	.17**	.03	–	
VI8	-.10*	-.20**	-.26**	-.47**	-.22**	-.13*	-.25**	–

Anmerkungen: VI1 = Ich reagiere mit einem Like („Gefällt mir“) oder einem zustimmenden Emoji („Herz“); VI2 = Ich reagiere mit einem missbilligenden Emoji („Traurig“ oder „Wütend“); VI3 = Ich unterstütze meine Vorredner mit einem sachlichen Kommentar; VI4 = Ich widerspreche meinen Vorrednern entschieden mit einem sachlichen Kommentar; VI5 = Ich widerspreche meinen Vorrednern entschieden mit einem beleidigenden Kommentar; VI6 = Ich unterstütze meine Vorredner und schreibe einen beleidigenden Kommentar; VI7 = Ich melde Facebook die Kommentare; VI8 = Ich lese die Kommentare, ohne etwas zu tun.

N = 340; \* $p < .05$ . \*\* $p < .01$ .

**Abbildung 10** Zweidimensionale Anordnung der Verhaltensintention im Wahrnehmungsraum (Studie I)



Anmerkungen: VI1 = Ich reagiere mit einem Like („Gefällt mir“) oder einem zustimmenden Emoji („Herz“); VI2 = Ich reagiere mit einem missbilligenden Emoji („Traurig“ oder „Wütend“); VI3 = Ich unterstütze meine Vorredner mit einem sachlichen Kommentar; VI4 = Ich widerspreche meinen Vorrednern entschieden mit einem sachlichen Kommentar; VI5 = Ich widerspreche meinen Vorrednern entschieden mit einem beleidigenden Kommentar; VI6 = Ich unterstütze meine Vorredner und schreibe einen beleidigenden Kommentar; VI7 = Ich melde Facebook die Kommentare; VI8 = Ich lese die Kommentare, ohne etwas zu tun. (N = 340)

Die Ergebnisse der Multidimensionalen Skalierung knüpfen an die theoretischen Überlegungen an und bilden den Wahrnehmungsraum der Versuchspersonen zweidimensional ab. Auf der Ordinate ist die Akzeptanzrichtung mit den beiden Skalenendpunkten *Ablehnung* (VI2 = *Ablehnungsemoji*, VI4 = *sachlicher Widerspruch*, VI5 = *beleidigender Widerspruch* und VI7 = *Melden*) und *Befürwortung* (VI1 = *Like*, VI3 = *sachliche Unterstützung* und VI6 = *beleidigende Unterstützung*) dargestellt. Die Konfiguration auf der Abszisse ist auf den ersten Blick weniger eindeutig, kann jedoch um 30 Grad im Uhrzeigersinn gekippt die aktiveren Verhaltensintentionen von dem passiven Ignorieren identifizieren, das als VI8 deutlich abgesetzt ist. Auffällig ist, dass die Werte kontraintuitiv im positiven Bereich für *Passivität* bzw. *Ablehnung* und im negativen Wertebereich für *Aktivität* und *Akzeptanz* stehen, was der Frageformulierung geschuldet ist, die im Fragebogen verwendet wurde („Im Folgenden geht es darum, wie wahrscheinlich Sie diese Reaktionen auf die soeben gelesenen Kommentare zeigen würden.“). VI8 hat den höchsten Mittelwert und wurde somit als die wahrscheinlichste Verhaltensoption von den Versuchspersonen gewählt. Damit kann die Dimension *Aktivitätsgrad* auch als *Wahrscheinlichkeitsgrad* interpretiert werden, was Verhaltensweisen mit einer

aktiven Positionierung unwahrscheinlicher macht. Darüber hinaus lässt sich aus dem abgebildeten Wahrnehmungsraum ableiten, dass das Ignorieren als Verhaltensintention eher den ablehnenden Verhaltensweisen zuzurechnen ist.

#### *Messungen von Kovariaten und Moderatorvariablen*

Um ambivalent sexistische Einstellungen gegenüber Frauen zu messen, wurde das *Ambivalent Sexism Inventory* (ASI) von Glick und Fiske (1996) in der deutschen Übersetzung von Thomas Eckes und Iris Six-Materna (1999) verwendet. Die insgesamt 12 Items wurden auf einer sechsstufigen endbepunkteten Skala von „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“ gemessen. Die Reliabilität der Skala erwies sich mit  $\alpha = .87$  als sehr gut ( $M = 2.87, SD = 1.02$ ). Die Skala setzt sich zusammen aus der Subskala *Benevolenter Sexismus* mit  $\alpha = .77$  gerade noch zufriedenstellend reliabel ( $M = 2.75, SD = 1.09$ ) und der Subskala *Hostiler Sexismus* mit  $\alpha = .86$  ( $M = 2.51, SD = 1.12$ ). Die beiden Skalen korrelierten erwartungsgemäß stark mit  $r = .56$  ( $p < .001$ ), d. h. dass es einen positiven Zusammenhang zwischen der Zustimmung zu vordergründig positiven Aussagen wie „Frauen sollten von Männern umsorgt und geschützt werden“ und der Zustimmung zu feindseligen Aussagen wie „Frauen übertreiben Probleme, die sie am Arbeitsplatz haben“ gibt.

Das Kognitionsbedürfnis wurde mit der deutschsprachigen *Need for Cognition*-Kurzskala (NFC-K) von Hanna Beißert und Kolleginnen gemessen (Beißert et al., 2014). Die Autorinnen nehmen direkten Bezug zu den Arbeiten von Cacioppo und Kollegen (Cacioppo et al., 1983; Cacioppo et al., 1984) und legen ein vier Items umfassendes Konstrukt vor, das in der vorliegenden Arbeit mit einer sechsstufigen endbepunkteten Skala von „stimme gar nicht zu“ bis „stimme voll und ganz zu“ gemessen wurde. Die Reliabilität der Skala war zwar mit  $\alpha = .54$  nicht zufriedenstellend, da sich aber die Skalenlänge positiv auf die Reliabilität auswirkt, zweifeln Beißert und Kolleginnen (2014) die Eignung des Reliabilitätsmaßes Cronbachs Alpha zur Bestimmung der internen Konsistenz bei Kurzskalen an. Die Autorinnen weisen mittels Messwiederholungen eine zeitliche Stabilität bei bildungshomogenen Stichproben nach (ebd.). Da die Stichprobe in der vorliegenden Arbeit überdurchschnittlich hoch gebildet war mit mehr als vier von fünf Personen, die mindestens über ein Fachabitur oder die Allgemeine Hochschulreife verfügen, wurde auf dieser Basis eine hohe Stabilität angenommen. Das hohe Kognitionsbedürfnis in der Stichprobe äußerte sich in einem hohen gemessenen Mittelwert ( $M = 4.25, SD = .86$ ).

Das subjektiv wahrgenommene *Vorwissen* wurde auf der Basis eines Items („*Ich fühle mich kompetent an einer Diskussion zum Thema ‚Karrierechancen von Frauen‘ teilzunehmen.*“) gemessen ( $M = 3.64$ ,  $SD = 1.44$ ). Das *persönliche Involvement* (PI) basiert auf der Operationalisierung von Wendy Wood (1982) und wurde über eine Mehrfachauswahl von Informationsquellen abgefragt, nämlich von interpersonaler Kommunikation, über den Besuch von Vorträgen bis hin zu medialen Quellen reichend. Eine offene Antwortoption enthüllte die Bedeutung von beruflichen und studienbedingten Informationsquellen zum Thema. Die überwiegende Mehrzahl der Befragten, nämlich 81,5 Prozent, gab an, sich aus mindestens einer Quelle zu informieren ( $M = 3.05$ ,  $SD = 1.16$ ). Gespräche und mediale Quellen waren dabei die am häufigsten genannten Informationsquellen. Für die weiteren Berechnungen wurde ein einfacher Summenscore gebildet, um die Stärke des PI abzubilden (höhere Werte standen für ein höheres PI).

Um für die *habitualisierte Nutzung von Facebook* (HN) auf die Verhaltensintention zu kontrollieren, wurden 11 Items auf einer endbepunkteten sechsstufigen Skala von „nie“ bis „immer“ erhoben. Die verwendeten Items entstammten analog zu den gemessenen Verhaltensintentionenitems keiner validierten Skala. Auf eine Verbalisierung der Skala wurde verzichtet, um gleichmäßige Abstände zwischen den Skalenpunkten zu erzeugen und damit das Skalenniveau zu erhöhen (Porst, 2008). Wie jedoch Rohrmann (1978) zeigen konnte, werden Häufigkeitsskalen ausnahmslos als ordinalskaliert wahrgenommen. Demnach wären auf der sechsstufigen Skala zwischen 1 („nie“) und 6 („immer“) die Abstufungen dazwischen wie folgt zu interpretieren: „selten“ (2), „gelegentlich“ (3), „oft“ (4), „meistens“ (5). Unterschieden wurden aktive, Inhalte generierende („Ich teile eigene Inhalte in meiner Chronik.“, „Ich kommentiere Beiträge oder antworte auf Kommentare auf öffentlichen Seiten.“) von passiven Nutzungsweisen, die sich auf die Nutzung der bereitgestellten Interaktionsfeatures der Plattform beschränkten („Ich markiere Fotos von mir oder meinen Freunden.“, „Ich drücke den Gefällt-mir-Button oder Reaktionsbutton für Beiträge oder Kommentare.“). Das Ignorieren von Facebook-Inhalten wurde in Ignorieren von Beiträgen und Ignorieren von Kommentaren unterschieden. Um ein Priming der Befragten vor Stimuluspräsentation zu vermeiden, wurde auf die Abfrage der Zustimmung bzw. Ablehnung von Facebook-Inhalten verzichtet. Um von einer habitualisierten, gewohnheitsmäßigen Ausübung zu sprechen, muss diese hinreichend oft praktiziert werden. Um die Ordinalskala parametrischen Analyseverfahren zugänglich zu machen, wurde die sechsstufige Skala dichotomisiert. Daher wurden die Werte 1 bis 3 zum Wert 0 („keine HN“) und die Werte 4 bis 6 zum Wert 1 („HN“) zusammengefasst.

Nicht nur im Zusammenhang mit Hate Speech spielt die Selbst- und Fremdwahrnehmung eine bedeutende Rolle, bei der Versuchsteilnehmende bestrebt sind, ein möglichst positives Bild von sich zu präsentieren (Aronson et al., 2014). Befragte geben dann sozial erwünschte Antworten, antworten also besonders gefällig, tendieren dazu mittlere Antwortoptionen auszuwählen und extreme Antworten zu vermeiden, insbesondere wenn diese mit negativen, devianten Verhaltensweisen oder Einstellungen zusammenhängen (Stocké, 2004). Aus diesem Grund wurde die Kurzskala zur Erfassung von sozial erwünschten Antworttendenzen (SEA-K) bestehend aus zwei Items verwendet (Satow, 2012) auf einer sechsstufigen endbepunkteten Skala von 1 („stimme gar nicht zu“) bis 6 („stimme voll und ganz zu“) gemessen. Das erste Item *„Ich habe schon mal über andere gelästert oder schlecht über sie gedacht.“* wurde umcodiert ( $M = 2.30, SD = 1.31$ ) und die Korrelation mit dem zweiten Item *„Ich würde niemals schlecht über einen Kollegen oder meinen Arbeitgeber reden.“* ( $M = 3.18, SD = 1.38$ ) zeigt einen mittelstarken positiven Zusammenhang zwischen den beiden Items mit Pearsons Korrelationskoeffizient  $r = .35, p < .001$ . In dem gebildeten Summenindex wurden 10 Antwortpunkte (Wahl der beiden Antwortoptionen mindestens auf Stufe 5) als Schwellenwert definiert, ab dem sozial erwünschte Antworten zu erwarten sind ( $M = 5.48, SD = 2.21$ ).

## 6.1.2 Ergebnisse

In diesem Abschnitt werden die Ergebnisse der ersten Studie dargestellt. Die Reihenfolge der Bearbeitung orientiert sich dabei im Wesentlichen an der Reihenfolge der adressierten Hypothesen, um die Bedingungen unter denen persuasive Effekte von Hate Speech eintreten, zu untersuchen.

### 6.1.2.1 Prüfung der Zufallsaufteilung

Wie bereits oben beschrieben, ist die Zufallsaufteilung von Merkmalsausprägungen ein wesentlicher Vorteil von experimentellen Studien, da sich Effekte in den Experimentalgruppen auf die Stimuluspräsentation zurückführen lassen. Die Überprüfung der Randomisierung erfolgt bei metrischen Variablen mittels einfaktorieller Varianzanalyse und bei kategoriellen Variablen mittels  $\chi^2$ -Test (Backhaus et al., 2018). Das Ergebnis der Randomisierungsprüfung der metrischen Variablen über alle 16 Experimentalgruppen kann Tabelle 10 entnommen werden. Varianzhomogenität auf Basis des Medians konnte über alle Variablen hinweg angenommen werden.

**Tabelle 10: Prüfung auf Randomisierung metrischer Variablen**

Variable	F
Ambivalenter Sexismus	$F(15, 324) = 0.59, p = .881$
Kognitionsbedürfnis	$F(15, 324) = 1.04, p = .412$
Persönliches Involvement	$F(15, 261) = 1.05, p = .405$
Vorwissen	$F(15, 324) = 0.47, p = .955$
Soziale Erwünschtheit	$F(15, 324) = 1.11, p = .347$
Alter	$F(15, 324) = 1.61, p = .070$

In den 16 Experimentalgruppen waren *männliche wie weibliche Versuchspersonen* nicht überzufällig unterschiedlich aufgeteilt,  $\chi^2(15) = 14.15, p = .514$ . Bildungsunterschiede konnten nicht festgestellt werden,  $\chi^2(30) = 28.52$ , Monte-Carlo-Signifikanz  $p = .530$ , 99% CI [.517, .543]. Für die *habituelle Nutzung* von Facebook waren ebenfalls keine Unterschiede zu verzeichnen, bis auf die Variable, die das Ignorieren von Kommentaren betraf. Hier wurden signifikante Gruppenunterschiede identifiziert,  $\chi^2(15) = 27.08, p < .05$ .

### 6.1.2.2 Auswirkung von Hate Speech auf die Elaboration

Gegenstand der Ergebnisdarstellung sind im folgenden Abschnitt die erhobenen Rezeptionsgedanken der Versuchspersonen, die unmittelbar nach Präsentation der experimentellen Stimuli geäußert wurden. Unter der Bedingung hinreichender Motivation und Fähigkeit werden in der klassischen ELM-Theorie zumeist die quantitative Erfassung der Gedanken sowie deren Valenz, seltener (da aufwendiger) die inhaltliche Auseinandersetzung als Indikatoren für die Elaboration von Persuasionsbotschaften angesehen. Zunächst soll der Zusammenhang zwischen Kommentar- und situativen Merkmalen und resultierenden Rezeptionsgedanken analysiert werden, indem Ergebnisse von Verfahren aus der Familie der *generalisierten linearen Modelle* (GLM) dargestellt werden, darunter hierarchische Regressionsanalysen und Poissonregressionen für die Zählvariablen mit dummycodierten unabhängigen Variablen.

#### *Auswirkungen auf die Gedankenvalenz*

Mehrheitlich wiesen die Gedanken eine negative Valenz auf ( $M_{\text{positiv}} = 1.98$ ,  $SD_{\text{positiv}} = 1.12$ ;  $M_{\text{neutral}} = 2.41$ ,  $SD_{\text{neutral}} = 1.69$ ,  $M_{\text{negativ}} = 3.63$ ,  $SD_{\text{negativ}} = 2.29$ ) ohne signifikante Unterschiede des Anteils negativer Gedanken an allen Gedanken zwischen den Experimentalgruppen, in denen die Anzahl der Likes (Reaktionsvolumen) bzw. das Geschlecht der Kommentierenden manipuliert wurde. Einzig die Hate Intensität war statistisch hochsignifikant mit  $F(3, 310) = 38.91$ ,  $p < .001$ ,  $\eta_p^2 = .27$ ,  $\omega_p^2 = .27$ . Post-hoc Tests mit Bonferroni-Korrektur zeigten, dass dieser Effekt hauptsächlich von sachlichen Kommentaren ausging, welche einen geringeren Anteil an negativen Gedanken auslösten als Offensive Speech (0.46, 95%-CI[0.33, 0.59]), Hate Speech (0.46, 95%-CI[0.33, 0.59]) oder Dangerous Speech (0.56, 95%-CI[0.39, 0.66]).

Diese Befunde wurden in den folgenden Schritten hierarchischen Regressionsanalysen mit dummycodierten unabhängigen Variablen zugeführt, mit dem Ziel weitere Einflussquellen auf die Gedankenvalenz ausfindig zu machen. Diese Analyseverfahren ermöglicht eine größere Flexibilität im Vergleich zu Varianzanalysen durch das Hinzufügen einzelner Variablen oder von Variablensets in die Regressionsgleichung (Backhaus et al., 2018). Modell 1 beinhaltete wie gehabt die Experimentalbedingungen (also die Hate Intensität, das Geschlecht der Kommentierenden und das Reaktionsvolumen), Modell 2 integrierte die Elaborationsvariablen, Modell 3 berücksichtigte Ambivalenten Sexismus und Modell 4 nahm soziodemografische Variablen auf. In allen Modellen wurden Effekte sozialer Erwünschtheit kontrolliert. Zunächst werden die Ergebnisse der Prüfung der Voraussetzungen dargestellt: Die Inspektion der partiellen Regressionsdiagramme ließ auf lineare Zusammenhänge zwischen unabhängigen und abhängiger Variable schließen. Übliche Verfahren zur Identifikation von Ausreißern

(unstandardisierte, standardisierte oder studentisierte Residuen) führten zu keinem Ausschluss von Fällen. Die Kollinearitätsstatistik bescheinigte angemessene Toleranz- (> .7) und VIF-Werte (< 1.4) der unabhängigen Variablen, was somit keinen Anlass zu Bedenken bezüglich Multikollinearität der Daten gab. Aus den Residuenplots konnten annähernde Normverteilung und Homoskedastizität angenommen werden, welche zusätzlich über einen nicht-signifikanten Breusch-Pagan-Test bestätigt wurde.

Modell 1 konnte einen signifikanten Beitrag ( $F(6, 260) = 9.76, p < .001$ ) zur Aufklärung der Varianz leisten ( $R^2 = .18, R_{korr}^2 = .17$ ). Die kategorialen Variablen der Experimentalbedingung Hate Intensität wurden dummycodiert in die Regressionsgleichung aufgenommen mit sachlichen Kommentaren als Referenzkategorie. Dies hatte zum einen den Vorteil, dass die Regressionskoeffizienten der Vergleichskategorien in Beziehung zur Referenzkategorie gesetzt werden konnten, zum anderen konnte ein zu starker Einfluss der Referenzkategorie auf die übrigen Variablen in der Regressionsgleichung kontrolliert werden. Es zeigte sich nämlich, dass neben den Dummycodes für die Hate Intensität auch die Anzahl der Likes statistisch signifikant war ( $B = -0.13, p < 01$ ), was in der Varianzanalyse verdeckt wurde, da dort alle vier Stufen des Faktors Hate Intensität in die Berechnungen flossen. Da sachliche Kommentare am stärksten zur Varianzaufklärung beitrugen, entfiel auf die Likeanzahl nur ein geringer Anteil der Varianz und verdeckte den Haupteffekt von Likeanzahl auf den Anteil an negativen Gedanken. Der nun signifikante Regressionskoeffizient deutete darauf hin, dass in Anbetracht der nicht bewussten Wahrnehmung der Likeanzahl als experimentelle Manipulation (vgl. Kapitel 6.1.1.4) ein Mere Exposure-Effekt aufgetreten war. Der negative Regressionskoeffizient zeigte an, dass Stimuli mit vielen Likes einen geringeren Anteil an negativen Rezeptionsgedanken bewirkten als Stimuli mit wenigen Likes. In der Experimentalbedingung Geschlecht der Kommentierenden konnte kein Einfluss auf den Anteil an negativen Gedanken nachgewiesen werden, was mit den oben berichteten Befunden der Varianzanalyse übereinstimmt und die Verwandtschaft der Verfahren illustriert. Für die Hate Intensität wurde sichtbar, dass der Anteil an negativen Gedanken in den Vergleichskategorien signifikant von der Referenzkategorie verschieden und umso höher war, je höher die Intensitätsstufe war: Offensive Speech und Hate Speech lösten im Vergleich zu sachlichen Kommentaren 0.30 bzw. 0.31 und Dangerous Speech sogar 0.47 Einheiten mehr an negativen Gedanken aus. Soziale Erwünschtheit erwies sich als nicht signifikant, was die Interpretation der Befunde erleichtert, da an dieser Stelle nicht mit wesentlichen Verzerrungen zu rechnen ist. Die Aufnahme der Elaborationsvariablen in Modell 2, der Indexvariable Ambivalenter Sexismus in Modell 3 und der soziodemografischen Variablen in Modell 4 trug nur geringfügig

**Tabelle 11:** Hierarchische Regressionsanalyse zur Erklärung des negativen Gedankenanteils

	Variable	B	95% CI für B		SE B	$\beta$	$R^2$	$\Delta R^2$
			LL	UL				
Modell 1	Konstante	0.41***	0.28	0.55	0.07		.18	.18***
	Dummy OS	0.30***	0.19	0.40	0.05	0.35		
	Dummy HS	0.31***	0.20	0.41	0.05	0.36		
	Dummy DS	0.47***	0.32	0.62	0.08	0.40		
	Dummy Likes (1=viele)	-0.13**	-0.22	-0.04	0.05	-0.17		
	Dummy Geschlecht (1=m)	0.03	-0.05	0.12	0.04	0.04		
	SE	3.8e-4	-0.02	0.02	0.01	0.002		
Modell 2	Konstante	0.40**	0.13	0.67	0.14		.19	.001
	Dummy OS	0.30***	0.19	0.40	0.05	0.35		
	Dummy HS	0.31***	0.20	0.42	0.06	0.36		
	Dummy DS	0.47***	0.32	0.62	0.08	0.40		
	Dummy Likes (1=viele)	-0.13**	-0.22	-0.04	0.05	-0.17		
	Dummy Geschlecht (1=m)	0.03	-0.05	0.12	0.04	0.04		
	SE	2.1e-4	-0.02	0.02	0.01	0.002		
	NFC	-0.01	-0.06	0.04	0.03	-0.02		
	PI	0.01	-0.03	0.04	0.02	0.02		
	Vorwissen	0.01	-0.03	0.04	0.02	0.02		
	Modell 3	Konstante	0.56***	0.24	0.87	0.16		.20
Dummy OS		0.30***	0.19	0.40	0.05	0.35		
Dummy HS		0.31***	0.21	0.42	0.06	0.37		
Dummy DS		0.48***	0.32	0.63	0.08	0.41		
Dummy Likes (1=viele)		-0.14**	-0.23	-0.05	0.05	-0.18		
Dummy Geschlecht (1=m)		0.04	-0.04	0.13	0.04	0.06		
SE		0.002	-0.02	0.02	0.01	0.01		
NFC		-0.02	-0.07	0.04	0.03	-0.04		
PI		-0.005	-0.05	0.04	0.02	-0.02		
Vorwissen		0.01	-0.03	0.04	0.02	0.03		
Ambivalenter Sexismus		-0.04	-0.08	0.002	0.02	-0.11		
Modell 4		Konstante	0.45*	0.09	0.82	0.18		.21
	Dummy OS	0.28***	0.17	0.39	0.06	0.33		
	Dummy HS	0.30***	0.19	0.41	0.06	0.35		
	Dummy DS	0.47***	0.32	0.62	0.08	0.40		
	Dummy Likes (1=viele)	-0.12**	-0.22	-0.03	0.05	-0.17		
	Dummy Geschlecht (1=m)	0.04	-0.05	0.12	0.04	0.05		
	SE	-0.001	-0.02	0.02	0.01	-0.01		
	NFC	-0.02	-0.07	0.03	0.03	-0.04		
	PI	-0.005	-0.05	0.04	0.02	-0.02		
	Vorwissen	0.01	-0.03	0.04	0.02	0.03		
	Ambivalenter Sexismus	-0.04	-0.08	0.01	0.02	-0.11		
	Alter	0.003	1.4e-4	0.01	0.002	0.11		
	Geschlecht (1=m, 0=w)	-0.03	-0.12	0.07	0.05	-0.03		
	Bildung (1=hoch, 0=Rest)	0.02	-0.11	0.15	0.07	0.02		

Anmerkungen: CI = Konfidenzintervall, LL = untere Schranke (lower limit), UL = obere Schranke (upper limit), OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech, SE = Soziale Erwünschtheit, NFC = Kognitionsbedürfnis, PI = Persönliches Involvement. \* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

zur zusätzlichen Aufklärung der Streuung bei ( $\Delta R^2 = 2.4$  zwischen Modell 1 und Modell 4), was sich auch in den allesamt nicht-signifikanten Regressionskoeffizienten niederschlug und den Wert des korrigierten  $R^2$  nicht veränderte, z. T. sogar geringfügig senkte (s. Tabelle 11). Damit lassen sich H1b und H3b, nicht jedoch H2b vorläufig bestätigen.

Im nächste Schritt wurden Moderationsanalysen für den Anteil der negativen Gedanken an allen Rezeptionsgedanken berechnet (H1b, H2b und H3b). Hierzu wurde eine Moderationsanalyse mittels PROCESS-Software (Hayes, 2018) durchgeführt. Dazu wurden die drei Experimentalbedingungen jeweils einzeln als unabhängige Variable einer Moderationsanalyse durchgeführt und Effekte der beiden verbleibenden Experimentalbedingungen als Kovariate kontrolliert (Hayes, 2018, S. 141–144). Als Moderationsvariablen wurden das Kognitionsbedürfnis (das persönliche Involvement hatte sich vorab als nicht signifikant erwiesen) und das Vorwissen gewählt. Um die Interpretierbarkeit der Moderatorvariablen zu gewährleisten, wurden diese mittelwertzentriert in das Modell aufgenommen. Für die *Intensitätsstufe* bildeten sachliche Kommentare die Referenzkategorie, für *Geschlecht* wurden Frauen und für das *Reaktionsvolumen* wurden wenig Likes als Referenzkategorie gewählt.

#### *Moderation zwischen Kommentarmerkmalen und Anteil an negativen Rezeptionsgedanken*

Das Modell konnte den Zusammenhang zwischen *Hate Intensität* als Kommentarmerkmal, Kognitionsbedürfnis, Vorwissen und dem Anteil negativer Gedanken an allen Rezeptionsgedanken signifikant abbilden  $F(14, 311) = 11.07, p < .001$  und trug in hohem Maße zur Aufklärung der Varianz bei ( $R^2 = .33$ ). Der Anteil von negativen Gedanken an allen Rezeptionsgedanken unterschied sich statistisch signifikant in den drei Faktorstufen der unabhängigen Variablen im Vergleich zu sachlichen Kommentaren, OS ( $b = 0.47, p < .001$ ), HS ( $b = 0.46, p < .001$ ) und DS ( $b = 0.53, p < .001$ ). Die Regressionskoeffizienten der Prädiktorvariablen drücken die Differenz zwischen Referenzgruppe (SK) und Vergleichsgruppe (z. B. OS) bezüglich des Anteils an negativen Gedanken der Versuchspersonen mit durchschnittlichem Kognitionsbedürfnis (also durchschnittlicher Motivation) und durchschnittlichem Vorwissen (d. h. durchschnittlichen Fähigkeiten) aus, da die Moderatorvariable mittelwertzentriert wurde und bezüglich der Stimuli von Frauen (Kontrolle für Experimentalbedingung Geschlecht der Kommentierenden) und einer niedrigen Likeanzahl stammten. Der Regressionskoeffizient für die Moderatorvariable Kognitionsbedürfnis (NFC) spiegelt den konditionalen Effekt auf die abhängige Variable in der Referenzgruppe (SK) bei durchschnittlichem Vorwissen wider. Dieser war signifikant negativ ( $b = -0.11, p < .05$ ) und bedeutet, dass Versuchspersonen mit höherem Kognitionsbedürfnis in der SK-Bedingung einen signifikant geringeren Anteil an negativen Rezeptionsgedanken aufwiesen als Versuchspersonen mit einem

niedrigeren Kognitionsbedürfnis bei durchschnittlichem Vorwissen. Die Interaktion zwischen NFC und Offensive Speech war signifikant ( $b = 0.14, p < .05, 95\% \text{ CI } [0.01, 0.26]$ ), ebenso zwischen NFC und Hate Speech ( $b = 0.14, p < .05, 95\% \text{ CI } [0.01, 0.27]$ ), nicht jedoch für Dangerous Speech. Der Interaktionskoeffizient  $b_{\text{OS} \times \text{NFC}} = 0.14$  schätzt somit, wie stark sich zwei Versuchspersonen in der Experimentalbedingung OS im Vergleich zur Referenzkategorie SK bezüglich des Anteils an negativen Gedanken bei einer Differenz von einer Einheit auf der NFC-Skala und durchschnittlichem Vorwissen unterscheiden. Das positive Vorzeichen sagt aus, dass Personen in der Experimentalbedingung OS einen höheren Anteil an negativen Rezeptionsgedanken vortrugen als Personen in der Referenzkategorie SK. Die zweite Moderatorvariable Vorwissen war nicht signifikant, d. h. es konnte keine Aussage über den Einfluss von Vorwissen auf die abhängige Variable getroffen werden. Dafür war die Interaktion zwischen Vorwissen und Offensive Speech ( $b = -0.07, p < .05, 95\% \text{ CI } [-0.14, -0.004]$ ), nicht jedoch für Hate oder Dangerous Speech signifikant. Der Koeffizient zeigt an, dass im Vergleich zur Referenzkategorie weniger negative Gedanken in der Experimentalbedingung OS bei einem Unterschied von einer Einheit für Vorwissen und durchschnittlichem Kognitionsbedürfnis geäußert wurden. Damit lässt sich Hypothese H1b nur teilweise bestätigen. Die Hate Intensität beeinflusst zwar den Anteil an negativen Gedanken, aber nicht für alle Intensitätsstufen. Der Zusammenhang zwischen der Hate Intensität und dem Anteil an negativen Rezeptionsgedanken wurde von den Elaborationsvariablen, die die Fähigkeit und Motivation zur Elaboration messen moderiert, was allerdings nur für OS und teilweise HS, nicht jedoch DS galt.

Weitere Analysen zeigten, dass weder die un konditionale Interaktion aus Kognitionsbedürfnis und Hate Intensität ( $\Delta R^2 = .01, F(3, 311) = 2.13, p = .096$ ) noch aus Vorwissen und Hate Intensität ( $\Delta R^2 = .01, F(3, 311) = 1.80, p = .148$ ) signifikant waren. Zur Betrachtung des konditionalen Effekts der beiden Moderatorvariablen wurde die Signifikanz der Interaktionen für einzelne Perzentile<sup>51</sup> bestimmt (Hayes, 2018, 249ff.). Tabelle 12 gibt einen Überblick über die erzielten Befunde und zeigt, dass in den Daten für nahezu alle Perzentil-Kombinationen signifikante konditionale Effekte nachgewiesen werden konnten. Das angeforderte Johnson-Neyman-Verfahren<sup>52</sup> liefert keine Lösung, was in Kombination mit den Befunden aus Tabelle

<sup>51</sup> Hayes (2018) empfiehlt das 16., 50. und 84. Perzentil, weil damit bei einer Normalverteilung jeweils Werte in Höhe von einer Standardabweichung unterhalb bzw. oberhalb des Mittelwerts und des Mittelwerts selbst abgedeckt werden. Selbst wenn die Daten nicht normalverteilt sind, befinden sich die gewonnenen Werte innerhalb dieser Perzentile. Damit umgeht man problematische Ansätze, wie Mediansplits oder künstliche Gruppen (ebd.).

<sup>52</sup> Das Johnson-Neyman-Verfahren dient dazu, Moderationseffekte zu analysieren, indem nicht einzelne Werte des Moderators (z. B. Perzentile) auf Signifikanz zwischen X (unabhängige Variable) und Y (abhängige Variable) getestet werden, sondern Werte bzw. Wertebereiche des Moderators

12 dafür spricht, dass der Wertebereich für sämtliche Moderatorenwerte gilt (Hayes, 2018). Abbildung 11 visualisiert die Moderationseffekte und zeigt die Änderung im Anteil an negativen Gedanken in Abhängigkeit der o. g. Perzentilwerte für die beiden Moderatorvariablen.

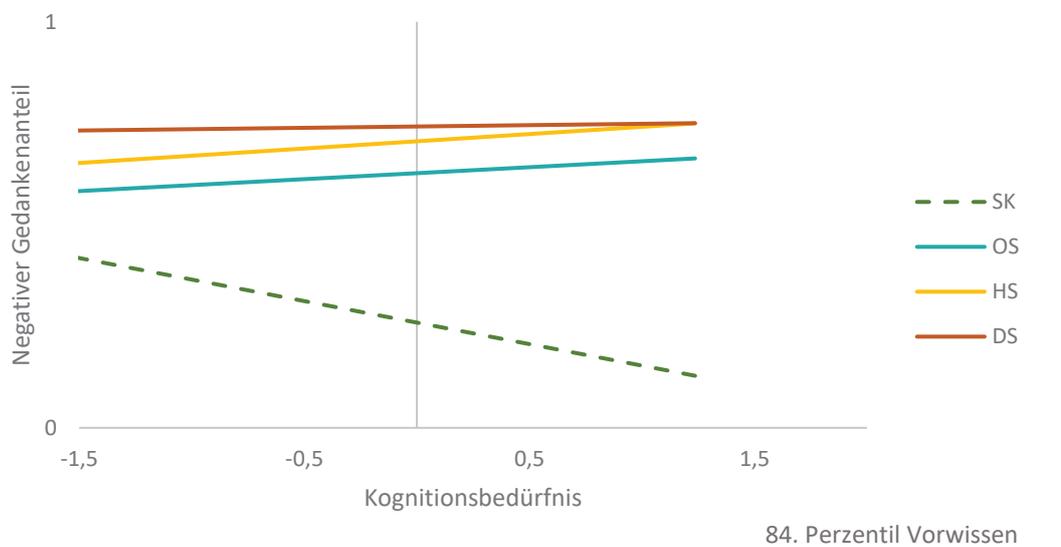
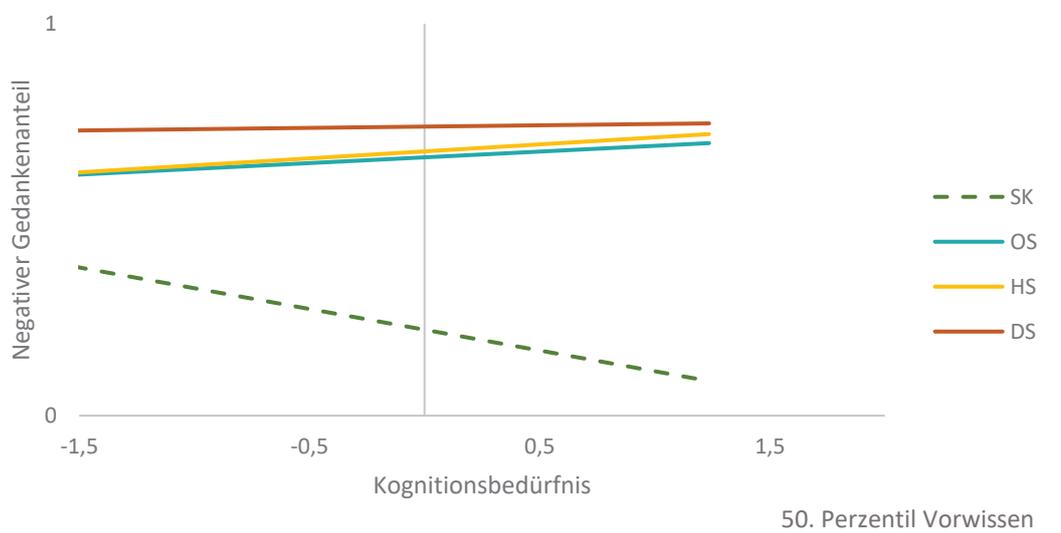
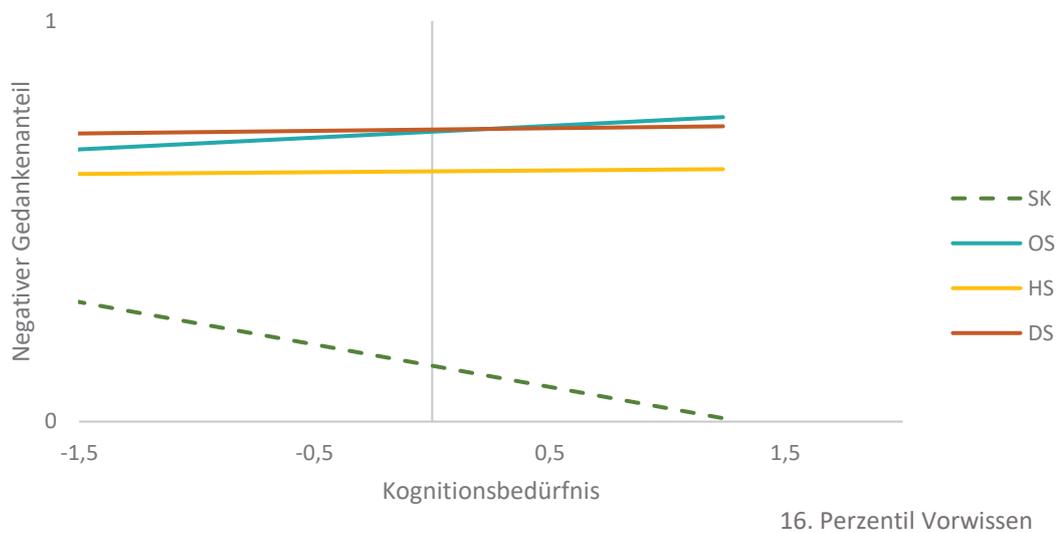
**Tabelle 12:** *Konditionaler Effekt der Hate Intensität auf den Anteil an negativen Gedanken in Abhängigkeit des Kognitionsbedürfnisses und Vorwissens in einem additiven multiplen Moderationsmodell*

Moderator	Perzentile	UV	Effekt	SE	t	p
NFC	16. Perzentil = -1.26	Dummy OS	.41	0.10	4.29	< .001
Vorwissen	16. Perzentil = -1.65	Dummy HS	.29	0.11	2.79	< .01
		Dummy DS	.45	0.10	4.61	< .001
NFC	16. Perzentil = -1.26	Dummy OS	.27	0.09	2.92	< .01
Vorwissen	50. Perzentil = 0.35	Dummy HS	.28	0.10	2.87	< .01
		Dummy DS	.38	0.09	4.00	< .001
NFC	16. Perzentil = -1.26	Dummy OS	.20	0.11	1.83	.07
Vorwissen	84. Perzentil = 1.35	Dummy HS	.27	0.11	2.41	< .05
		Dummy DS	.34	0.11	3.08	< .01
NFC	50. Perzentil = -0.26	Dummy OS	.55	0.07	7.47	< .001
Vorwissen	16. Perzentil = -1.65	Dummy HS	.44	0.08	5.46	< .001
		Dummy DS	.56	0.07	7.69	< .001
NFC	50. Perzentil = -0.26	Dummy OS	.40	0.05	7.56	< .001
Vorwissen	50. Perzentil = 0.35	Dummy HS	.42	0.06	7.57	< .001
		Dummy DS	.49	0.05	8.93	< .001
NFC	50. Perzentil = -0.26	Dummy OS	.33	0.07	4.65	< .001
Vorwissen	84. Perzentil = 1.35	Dummy HS	.41	0.07	5.59	< .001
		Dummy DS	.45	0.07	6.12	< .001
NFC	84. Perzentil = 0.74	Dummy OS	.68	0.10	7.17	< .001
Vorwissen	16. Perzentil = -1.65	Dummy HS	.58	0.10	5.94	< .001
		Dummy DS	.67	0.09	7.19	< .001
NFC	84. Perzentil = 0.74	Dummy OS	.54	0.07	7.75	< .001
Vorwissen	50. Perzentil = 0.35	Dummy HS	.56	0.07	8.26	< .001
		Dummy DS	.60	0.07	8.76	< .001
NFC	84. Perzentil = 0.74	Dummy OS	.46	0.08	5.92	< .001
Vorwissen	84. Perzentil = 1.35	Dummy HS	.55	0.08	7.03	< .001
		Dummy DS	.57	0.08	7.10	< .001

Anmerkungen: UV = Unabhängige Variable, NFC = Kognitionsbedürfnis, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

geprüft werden, bei denen der bedingte Effekt von X auf Y statistisch signifikant bzw. nicht signifikant wird (Hayes, 2018, S. 254f.).

**Abbildung 11:** Moderationseffekt von Hate Intensität in Abhängigkeit des Kognitionsbedürfnisses und Vorwissens



Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

Die grafische Darstellung in Abbildung 11 zeigt, dass Versuchspersonen in den drei Experimentalgruppen einen gleich bleibend hohen (DS) bzw. leicht ansteigenden negativen Gedankenanteil (OS und HS) äußerten, je höher das Kognitionsbedürfnis war. In der Gruppe SK sank mit höherem Kognitionsbedürfnis der Anteil an negativen Rezeptionsgedanken. Auffällig war, dass Personen mit steigendem Vorwissen mehr negative Gedanken äußerten, was auf einen Abgleich des Wissens mit den gelesenen Kommentaren deutet, was anhand der inhaltlichen Untersuchung der Rezeptionsgedanken im nächsten Abschnitt geprüft wird. Der Anteil an negativen Gedanken bewegte sich in den anderen drei Experimentalbedingungen auf einem sehr hohen Niveau und stieg nur geringfügig an.

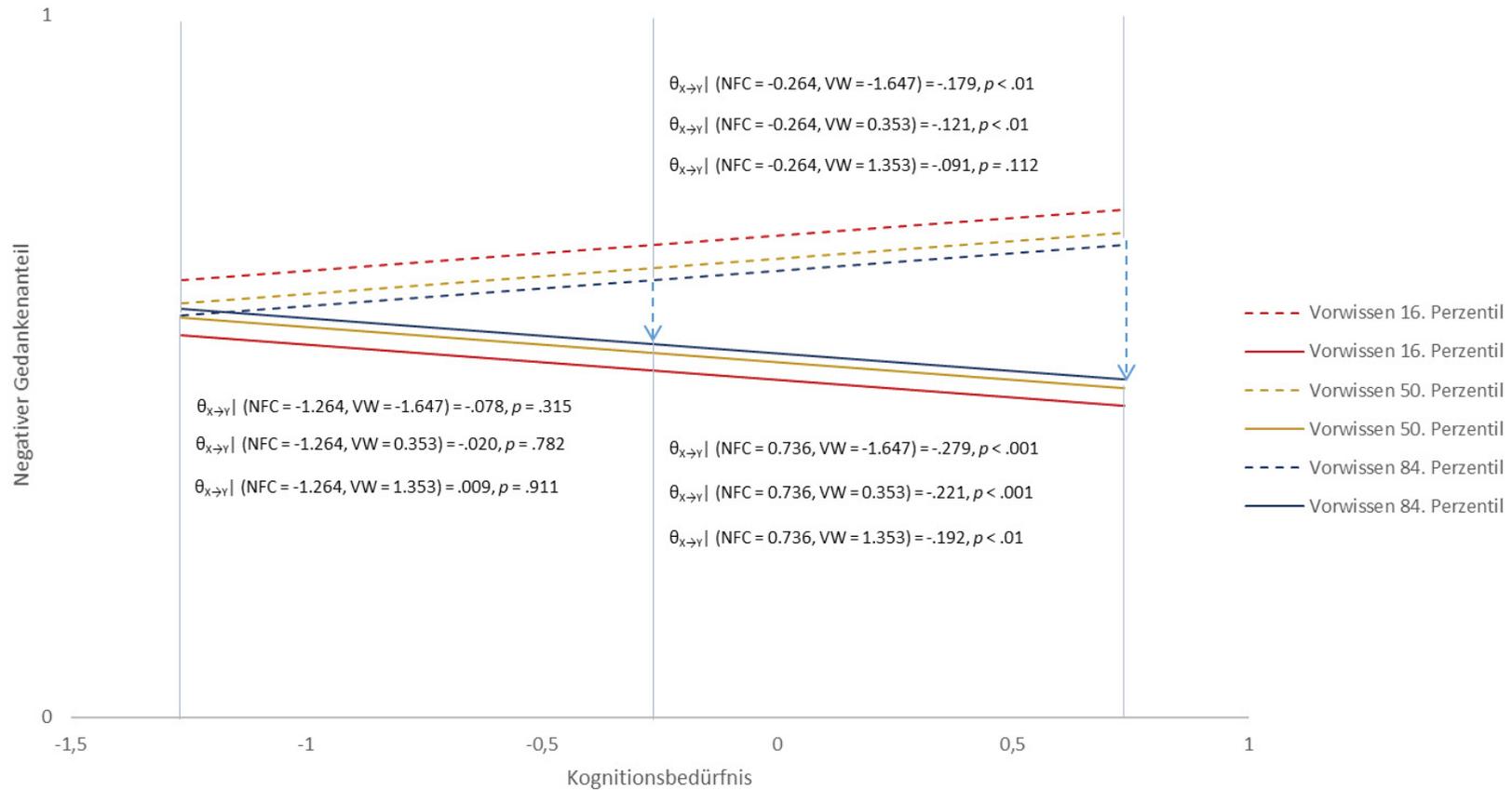
Als nächstes wurden analog zur Moderationsanalyse für Hate Intensität Berechnungen für das *Geschlecht* der Kommentierenden als UV durchgeführt. Obwohl das Modell signifikant war ( $F(10, 315) = 7.52, p < .001, R^2 = .19$ ), waren weder UV noch Moderatorvariablen, noch Interaktionen oder soziale Erwünschtheit als Kovariate signifikant. Damit muss die Hypothese H2b abgelehnt werden. Wie oben ausführlich dargestellt, wurden die beiden restlichen Bedingungen Hate Intensität (jeweils getrennt für die einzelnen Ausprägungen mit SK als Referenzkategorie) und Likeanzahl als dummycodierte Kovariate hinzugefügt, welche sich als signifikant erwiesen, d. h., dass diese Bedingungen den Zusammenhang zwischen dem Anteil an negativen Gedanken erklären konnten, die Geschlechterkategorie jedoch keine Rolle spielte.

#### *Moderation zwischen situativen Merkmalen und Anteil an negativen Gedanken*

Die dritte Experimentalbedingung *Reaktionsvolumen* wurde dem oben beschriebenen Moderationsverfahren unterworfen und ergab ein signifikantes Modell ( $F(10, 315) = 8.00, p < .001, R^2 = .20$ ). Bei Rezeption von vielen Likes wurden trotz unbewusster Wahrnehmung ein signifikant geringerer Anteil an negativen Gedanken im Vergleich zur Rezeption weniger Likes erzielt ( $b = -0.16, p < .001$ ). Das Kognitionsbedürfnis beeinflusste signifikant die Beziehung zwischen der Likeanzahl und dem Anteil an negativen Gedanken ( $b = -0.10, p < .05$ ). Unter der Bedingung viele Likes gingen die negativen Gedanken mit jeder Einheit auf der NFC-Skala zurück, während sie unter der Bedingung wenig Likes stiegen. Vorwissen hatte keinen Einfluss auf die Beziehung zwischen UV und Moderatorvariable. Damit lässt sich Hypothese H3b teilweise bestätigen. Auch die Kontrolle für soziale Erwünschtheit war nicht signifikant, ebenso wie die Geschlechterkategorie als Experimentalbedingung. Wie erwartet waren die dummycodierten einzelnen Ausprägungen der Hate Intensität hoch signifikant ( $b_{OS} = 0.31, p < .001, b_{HS} = 0.30, p < .001, b_{DS} = 0.46, p < .001$ ). Der unktionale Effekt des Kognitionsbedürfnisses auf den Anteil an negativen Rezeptionsgedanken war zwar gering, aber signifikant ( $\Delta R^2 = .01, F(1, 315) = 4.91, p < .05$ ), nicht jedoch zwischen Likeanzahl und Vorwissen

( $\Delta R^2 = .003$ ,  $F(1, 315) = 1.15$ ,  $p = .284$ ). Abbildung 12 zeigt eine grafische Darstellung der Prüfung des konditionalen Effekts auf der Basis des 16., 50. und 84. Perzentils der beiden Moderatoren. Bei unterdurchschnittlichem Kognitionsbedürfnis (16. Perzentil) waren die Unterschiede im negativen Gedankenanteil unabhängig vom Vorwissen zu gering, um Signifikanz zu erlangen. In der Bedingung *hohe Likeanzahl* nahmen die negativen Rezeptionsgedanken mit höherem Kognitionsbedürfnis ab, und nahmen zu, wenn Stimuli *wenige Likes* darboten. Dieser konditionale Effekt  $\theta$  wurde mit steigendem Kognitionsbedürfnis und gleichzeitig niedrigerem Vorwissen immer stärker.

**Abbildung 12:** Konditionaler Effekt des Reaktionsvolumens auf den negativen Gedankenanteil in Abhängigkeit vom Kognitionsbedürfnis und Vorwissen



Anmerkungen: NFC = Kognitionsbedürfnis, VW = Vorwissen; Gestrichelte Linie steht für X = 0 (Experimentalbedingung „wenige Likes“), durchgezogene Linie steht für X = 1 (Experimentalbedingung „viele Likes“). N = 326.

*Auswirkungen auf die Anzahl der Gedanken*

Rund 96 Prozent der Versuchspersonen ( $N = 326$ ) äußerten mindestens einen Rezeptionsgedanken, im Mittel 4.75 ( $SD = 2.40$ , Schiefe = 0.860, Wölbung = -0.020), was auf eine mittelstarke Informationsverarbeitung während des Experiments deutet. Die Anzahl der Likes und das Geschlecht der Kommentierenden im Stimulusmaterial wirkten sich nicht auf die Gedankenzahl aus, wohl aber die Hate Intensität (Welch-Test  $F(3, 178.14) = 10.05$ ,  $p < .001$ ). Aufgrund fehlender Varianzhomogenität wurde für die Gruppenvergleiche der Games-Howell-Test berechnet, der zeigte, dass Hate Speech (1.23, 95%-CI[0.32, 2.13]) und Dangerous Speech (1.90, 95%-CI[0.94, 2.86]) im Vergleich zu sachlichen Kommentaren signifikant die Quantität der Gedanken beeinflussten. Kein Unterschied hinsichtlich der Gedankenzahl konnte zwischen Offensive Speech und sachlichen Kommentaren gefunden werden. Die Hypothesen H2a und H3a müssen daher abgelehnt werden.

Zählvariablen wie die hier betrachtete Anzahl der Gedanken erfassen die Häufigkeit des Auftretens eines Ereignisses (Tutz, 2010). Da solche Variablen diskrete Werte  $> 0$  annehmen und anderen Verteilungen als der Normalverteilung folgen, sind Alternativen für regressionsanalytische Untersuchungen auf der Basis des Kleinstquadrateschätzers notwendig (Tutz, 2010). Die Poisson-Regression und die negative Binomialregression stellen solche Verfahren dar (Agresti, 2007). Die Poisson-Regression geht von Equidispersion aus, also (annähernd) gleichen Werten für Varianz und Mittelwert, während die negative Binomialregression eine flexiblere Modellierung mit einem zweiten Parameter erlaubt (A. C. Cameron & Trivedi, 2013). Vorliegend wurde zunächst ein linearer Term mit den drei Experimentalbedingungen, und mittelwertzentrierten metrischen Variablen, nämlich den drei Elaborationsvariablen, ambivalentem Sexismus sowie sozialer Erwünschtheit und als soziodemografische Variablen Alter, Geschlecht sowie Bildung der Versuchspersonen spezifiziert. Der Dispersionswert lag bei 1.033, was einer vernachlässigbar geringen Überdispersion entspricht, und wurde auch vom nicht signifikanten Likelihood Ratio Test bestätigt ( $\chi^2 = 0.304$ ,  $p = .291$ ). Mit der Hate Intensität stieg die Anzahl der geäußerten Rezeptionsgedanken signifikant um rund 19 Prozent in der OS- (IRR = 1.191, 95% CI [1.00, 1.41]), 37 Prozent in der HS- (IRR = 1.372, 95% CI [1.17, 1.64]) bzw. um 60 Prozent in der DS-Bedingung (IRR = 1.604, 95% CI [1.37, 1.90]) im Vergleich zu sachlichen Kommentaren. Der in Hypothese H1a postulierte Einfluss der Hate Intensität auf die Gedankenzahl kann bestätigt werden. Die übrigen Koeffizienten übten keinen Einfluss auf die Anzahl der Rezeptionsgedanken aus. Auch der erwartete Interaktionseffekt zwischen Experimentalbedingungen und Elaborationsvariablen lieferte keine signifikanten Ergebnisse. Was die Anzahl der Rezeptionsgedanken, also die rein quantitative Informationsverarbeitung

sexistischer Hasskommentare gegen Frauenkarrieren in der IT-Branche betrifft, muss festgehalten werden, dass diese nicht von der Motivation und Fähigkeit zur Verarbeitung beeinflusst wird. Es trat kein Moderationseffekt der Elaborationsvariablen auf, wie in Hypothese H1a genannt, weswegen diese daher teilweise abgelehnt werden muss.

**Tabelle 13:** *Poissonregression mit mittelwertzentrierten Variablen zur Erklärung der Anzahl der Rezeptionsgedanken*

Variable	B	SE	95 CI für B		IRR
			LL	UL	
Konstante	1.23***	0.12	0.98	1.46	3.408
Dummy OS	0.18*	0.09	0.004	0.35	1.191
Dummy HS	0.32***	0.09	0.15	0.49	1.381
Dummy DS	0.48***	0.08	0.31	0.64	1.611
Dummy Likes (1 = viele)	-0.02	0.06	-0.14	0.09	0.977
Dummy Geschlecht (1 = m, 0 = w)	0.07	0.06	-0.04	0.18	1.073
NFC	0.05	0.03	-0.01	0.12	1.055
VW	-0.02	0.02	-0.07	0.02	0.977
PI	0.05	0.03	-0.01	0.10	1.049
ASI	0.02	0.03	-0.03	0.08	1.024
SE	0.003	0.01	-0.02	0.03	1.003
Alter	-0.002	0.002	-0.01	0.002	0.998
Geschlecht (1 = m, 0 = w)	-0.01	0,06	-0.14	0.11	0.987
Bildung (1=hoch, 0 = Rest)	0.07	0.09	-0.10	0.25	1.075

*Anmerkungen:* CI = Konfidenzintervall, LL = untere Schranke (lower limit), UL = obere Schranke (upper limit), IRR = Incidence Rate Ratio, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 267$

AIC = 1182.85, BIC = 1233.07,  $-2\log L = 1154.85$

### *Auswirkungen auf die Ausgangsbasis der Gedanken*

Die quantitativ-inhaltliche Prüfung der Rezeptionsgedanken erfolgte mittels Auswertung des Gedankenursprungs. Dazu wurde die gedankliche Ausgangsbasis einer der vier Kategorien zugeordnet (vgl. Kapitel 6.1.1.5), um die kognitive und affektive Verarbeitung des Stimulusmaterials zu erfassen. Die so erzielten Zählwerte über die vier Kategorien hinweg konnten einen Wert zwischen 0 und 10 annehmen, je nachdem wie viele Rezeptionsgedanken eine Versuchsperson geäußert hatte (z. B. bei insgesamt fünf Rezeptionsgedanken konnten drei Gedanken auf die erste Kategorie, zwei auf die zweite und jeweils 0 auf die letzten beiden Kategorien entfallen). Gedanken, die die Experimentalsituation oder die Stimuluswahrnehmung thematisierten, also eine rein kognitive Perspektive darauf einnahmen, dominierten unter den Versuchspersonen, was sich daran zeigte, dass 89,6 Prozent mindestens einen

solchen Gedanken äußerten ( $M = 3.02$ ,  $SD = 2.10$ , Schiefe = 0.884, Wölbung = 1.050) und wurden als erstes analysiert. Da es sich hierbei um Zählvariablen handelt, empfahl sich eines der oben beschriebenen Verfahren. Allerdings bestand hier der Unterschied, dass der Wert 0 Bestandteil des Wertespektrums war, was die Anwendung von Zero-Inflated-Modellen bzw. Hurdle-Modellen erforderlich machte, da diese neben dem Zählen der Werte auch das Nicht-Eintreten von Ereignissen berücksichtigen. Im Zählmodell wird (1) die Häufigkeit des Auftretens des betrachteten Outcomes modelliert und das Hurdle-Modell berechnet (2) die Odds Ratios dessen Nicht-Eintretens, d. h. wie häufig Rezeptionsgedanken, die sich kognitiv mit dem Stimulus auseinandersetzen in Abhängigkeit der gewählten Prädiktoren auftreten *und* wie wahrscheinlich es ist, dass *kein* Gedanke dieser Kategorie geäußert wird.

Zero-Inflated Modelle unterscheiden *strukturelle* (z. B. Versuchsperson A nannte drei Rezeptionsgedanken, aber keiner davon ließ Vorwissen erkennen) und *Sampling-bezogene* Ursachen (z. B. Missings im Datengewinnungsprozess) für überschüssige 0-Werte (Hu et al., 2011). Hurdle-Modelle waren im vorliegenden Fall besser geeignet, da sie nur strukturelle Ursachen berücksichtigen und da die Antwortstruktur derjenigen, die sich äußern, relevant war. Methodisch betrachtet handelt es sich dabei um Verfahren, die parallel zur Poisson- bzw. negativen Binomialregression eine logistische Regression (für den Wert 0 bzw. die restlichen Zählwerte) anwenden (Huang & Cornell, 2012).

Da keine Überdispersion vorlag ( $\chi^2(1) = 0.60$ ,  $p = .439$ ), wurde mit dem R-Paket *pscl* ein Hurdle-Modell mit Poissonverteilung für die Zählwerte durchgeführt (Zeileis et al., 2008). Die verwendeten Modellparameter entsprachen denen in dem oben beschriebenen Modell zur Anzahl der Rezeptionsgedanken. In Tabelle 14 zeigen die Befunde des Zählmodells einen Anstieg der Gedanken, die sich kognitiv mit dem Stimulusmaterial auseinandersetzen, und zwar in der OS-Bedingung ein Plus von 45 Prozent (IRR = 1.452, 95% CI [1.13, 1.87]), rund 61 Prozent in der HS-Bedingung (IRR = 1.60, 95% CI [1.25, 2.06]) und sogar annähernd 98 Prozent in der DS-Bedingung, (IRR = 1.975, 95% CI [1.55, 2.51]) also fast doppelt so viele im Vergleich zu sachlichen Kommentaren. Damit kann der in Hypothese H1c postulierte Einfluss der Hate Intensität auf die gedankliche Ausgangsbasis vorerst als bestätigt angenommen werden. Des Weiteren lösten das Persönliche Involvement und Ambivalenter Sexismus eine gerinfüßige aber erkennbare *kognitive Auseinandersetzung der Versuchspersonen mit dem Stimulusmaterial* aus ( $p < .05$ ). Die beiden anderen Experimentalbedingungen beeinflussten die kognitive Verarbeitung ebenso wenig, wie die beiden anderen Elaborationsvariablen, soziale Erwünschtheit und die soziodemografischen Variablen. Im Hurdle-Modell waren die Modellparameter insgesamt nicht signifikant.

**Tabelle 14:** Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung der kognitiven Auseinandersetzung mit dem Stimulusmaterial als Ausgangsbasis für Rezeptionsgedanken

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	0.72***	0.17	0.39	1.05	2.056
	Dummy OS	0.37**	0.13	0.12	0.63	1.452
	Dummy HS	0.47***	0.13	0.23	0.72	1.608
	Dummy DS	0.68***	0.12	0.44	0.92	1.975
	Dummy Likes (1 = viele)	-0.07	0.08	-0.22	0.08	0.937
	Dummy Geschlecht (1 = Männer)	0.06	0.08	-0.09	0.22	1.065
	NFC	-0.02	0.05	-0.11	0.07	0.978
	VW	-0.04	0.03	-0.11	0.02	0.959
	PI	0.07*	0.04	0.003	0.15	1.078
	ASI	0.09*	0.04	0.02	0.17	1.098
	SE	0.002	0.02	-0.03	0.04	1.002
	Alter	-0.002	0.003	-0.01	0.00	0.998
	Geschlecht (1 = m, 0 = w)	0.03	0.08	-0.13	0.19	1.030
	Bildung (1 = hoch, 0 = Rest)	0.05	0.12	-0.18	0.28	1.029
Hurdle-Modell	Konstante	1.86*	0.81	0.26	3.45	6.392
	Dummy OS	0.25	0.52	-0.77	1.23	1.279
	Dummy HS	1.17	0.66	-0.12	2.47	3.233
	Dummy DS	1.06	0.64	-0.19	2.32	2.892
	Dummy Likes (1 = viele)	-0.12	0.44	-0.98	0.73	0.883
	Dummy Geschlecht (1 = Männer)	-0.37	0.45	-1.25	0.52	0.694
	NFC	0.29	0.26	-0.22	0.80	1.334
	VW	0.03	0.18	-0.31	0.37	1.032
	PI	0.07	0.21	-0.34	0.49	1.077
	ASI	-0.27	0.23	-0.71	0.18	0.764
	SE	-0.03	0.10	-0.22	0.16	0.968
	Alter	-0.01	0.02	-0.04	0.02	0.989
	Geschlecht (1 = m, 0 = w)	-0.29	0.46	-1.20	0.61	0.745
	Bildung (1 = hoch, 0 = Rest)	0.30	0.58	-0.84	1.45	1.350

Anmerkungen: OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 267$

AIC = 1113.01, BIC = 1213.45,  $-2\log L = 1057.006$

Im nächsten Schritt wurde das Modell zur Erklärung der Anzahl an Gedanken, die die kognitive Verarbeitung des Stimulusmaterials als Ausgangsbasis hatte, unter der Bedingung hinreichender Motivation und/oder<sup>53</sup> Fähigkeit vorbereitet. Dazu wurden zunächst die

<sup>53</sup> Petty und Cacioppo (1986b) sind uneindeutig, ob die Motivation und die Fähigkeit zur Informationsverarbeitung getrennt oder nur beide Elaborationsdimensionen gemeinsam berücksichtigt werden können. Im Folgenden wurden die Motivationsvariablen (Kognitionsbedürfnis und

Kovariate entfernt, da sie keinen Einfluss auf die Anzahl der Rezeptionsgedanken hatten. Enthalten waren die Experimentalbedingungen sowie die drei Elaborationsvariablen, dazu Interaktionsterme aus Experimentalbedingungen und Vorwissen, denn das Kognitionsbedürfnis und das Persönliche Involvement hatten sich als nicht signifikante Moderatoren herausgestellt. Im Zählmodell sorgten erwartungsgemäß die Bedingungen, die die Hate Intensität variierten, signifikant für einen Anstieg der Rezeptionsgedanken (OS: IRR = 1.340, 95% CI [1.04, 1.73], HS: IRR = 1.530, 95% CI [1.19, 1.97], DS: IRR = 1.914, 95% CI [1.50, 2.44]). Im Hurdle-Modell war die Wahrscheinlichkeit, sich *nicht* kognitiv zum Stimulusmaterial zu äußern, für Vorwissen als Prädiktor signifikant (IRR = 2.883, 95% CI [1.14, 7.28]). Signifikante Interaktionen gab es für OS und Vorwissen (IRR = 0.225, 95% CI [0.09, 1.97]) sowie für DS und Vorwissen (IRR = 0.168, 95% CI [0.05, 0.55]). Vorwissen erhöhte zwar generell die Wahrscheinlichkeit, sich *nicht* kognitiv zum Stimulusmaterial zu äußern, senkte aber in Verbindung mit OS und DS die Wahrscheinlichkeit, keine Gedanken dieser Kategorie zu äußern, m. a. W. Befragte mit Vorwissen äußerten sich verstärkt zum Stimulusmaterial. Der Moderationseffekt aus Hypothese H1c konnte teilweise bestätigt werden. Der Zusammenhang zwischen Reaktionsvolumen (H3c) bzw. Geschlecht der Kommentierenden (H2c) und der kognitiven Auseinandersetzung mit dem Stimulus wurde nicht von der Fähigkeit zur Elaboration beeinflusst.

Welche Faktoren beeinflussen die *Anzahl affektiv gefärbter Rezeptionsgedanken*, aber auch die Wahrscheinlichkeit keine derartigen Gedanken zu äußern? Um diese Frage zu beantworten, wurde ein Hurdle-Modell mit Poissonverteilung mit den oben beschriebenen Modellparametern durchgeführt, da der Dispersionswert nahe 1 lag (0.999 entspricht minimaler Underdispersion) und ein Hurdle-Modell mit negativer Binomialverteilung keinen signifikant besseren Modellfit vorweisen konnte ( $\chi^2(1) = 3.35, p = .067$ ). Tabelle 15 enthält die Befunde des Modells. Im Zählmodell konnte keiner der Prädiktoren die Anzahl affektiver Reaktionen auf das Stimulusmaterial erklären. Im logistischen Hurdle Modell waren Kognitionsbedürfnis, ambivalent sexistische Einstellungen und das Geschlecht der Versuchspersonen signifikante Prädiktoren für die Wahrscheinlichkeit *keinen* affektiv gefärbten Rezeptionsgedanken zu äußern. Während das Kognitionsbedürfnis die Odds Ratio um 44 Prozent erhöhte (IRR = 1.442, 95% CI [1.04, 2.00]), senkten ambivalent sexistische Einstellungen (IRR = 0.677, 95% CI [0.51, 0.90]) und männliche Versuchspersonen (IRR = 0.544, 95% CI [0.30, 0.99]) diese. Für die affektiven Reaktionen als gedankliche Ausgangsbasis kann kein Einfluss wie in den Hypothesen H1c, H2c und H3c postuliert, nachgewiesen werden.

---

Persönliches Involvement) wegen fehlender Signifikanz aus dem Modell entfernt und nur ein Moderationseffekt von Fähigkeit zur Elaboration (operationalisiert als Vorwissen) geprüft.

**Tabelle 15:** Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung affektiver Reaktionen auf das Stimulusmaterial

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	-0.24	0.55	-1.32	0.84	0.787
	Dummy OS	-0.09	0.36	-0.79	0.61	0.910
	Dummy HS	0.25	0.33	-0.39	0.89	1.286
	Dummy DS	0.21	0.32	-0.42	0.85	1.238
	Dummy Likes (1 = viele)	-0.06	0.22	-0.50	0.37	0.938
	Dummy Geschlecht (1 = Männer)	0.39	0.23	-0.05	0.83	1.474
	NFC	-4,5e-4	0.15	-0.29	0.29	1.000
	VW	-0.06	0.09	-0.23	0.10	0.938
	PI	0.10	0.10	-0.10	0.30	1.106
	ASI	-0.06	0.13	-0.30	0.19	0.942
	SE	0.04	0.05	-0.06	0.14	1.041
	Alter	-0.01	0.01	-0.02	0.01	0.994
	Geschlecht (1 = m, 0 = w)	-0.48	0.32	-1.11	0.15	0.620
	Bildung (1 = hoch, 0 = Rest)	0.22	0.40	-0.55	1.00	1.250
Hurdle-Modell	Konstante	-0.49	0.56	-1.59	0.60	0.611
	Dummy OS	0.16	0.39	-0.59	0.92	1.178
	Dummy HS	0.62	0.39	-0.15	1.38	1.851
	Dummy DS	0.65	0.39	-0.12	1.42	1.911
	Dummy Likes (1 = viele)	0.16	0.27	-0.36	0.69	1.177
	Dummy Geschlecht (1 = Männer)	-0.17	0.27	-0.70	0.36	0.845
	NFC	0.37*	0.17	0.04	0.69	1.441
	VW	-0.02	0.11	-0.24	0.19	0.975
	PI	-0.09	0.13	-0.34	0.16	0.917
	ASI	-0.39**	0.15	-0.68	-0.10	0.677
	SE	0.002	0.06	-0.12	0.12	1.002
	Alter	0.003	0.01	-0.02	0.02	1.003
	Geschlecht (1 = m, 0 = w)	-0.61*	0.30	-1.20	-0.01	0.544
	Bildung (1 = hoch, 0 = Rest)	-0.10	0.42	-0.93	0.73	0.902

Anmerkungen: OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 267$

AIC = 647.66, BIC = 748.10,  $-2\log L = 591.656$

In der Erweiterung des Modells mit Interaktionstermen wurde die beschriebene Modellierung eingesetzt mit den Experimentalbedingungen und den drei Elaborationsvariablen, dazu Interaktionsterme aus Experimentalbedingungen und Persönliches Involvement, da dies den besten Modellfit mit einem AIC = 643.18 (NFC-Interaktionen: AIC = 656.31, VW-Interaktionen: AIC = 651.52) vorweisen konnte. Persönliches Involvement beeinflusste signifikant die Beziehung zwischen der DS-Versuchsbedingung und der auftretenden Häufigkeit affektiver Reaktionen auf die Stimuli und erhöhte deren Anzahl um rund 76 Prozent (IRR = 1.757, 95%

CI [1.07, 2.89]). Der Moderationseffekt aus Hypothese H1c muss also teilweise bestätigt werden. Im Hurdle Modell erhöhte sich mit steigendem Kognitionsbedürfnis die Odds Ratio *keine* affektiv gefärbten Rezeptionsgedanken zu äußern, um 45 Prozent (IRR = 1.454, 95% CI [1.06, 2.00]).

**Tabelle 16:** Hurdle-Modell mit Poissonverteilung und mittelwertzentrierten Variablen zur Erklärung von Vorwissen und Erfahrungen in den Rezeptionsgedanken

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	0.62	0.46	-0.28	1.52	1.857
	Dummy OS	-0.41	0.25	-0.90	0.08	0.663
	Dummy HS	-0.61*	0.29	-1.18	-0.03	0.546
	Dummy DS	-0.42	0.27	-0.96	0.11	0.655
	Dummy Likes (1 = viele)	0.05	0.19	-0.32	0.42	1.051
	Dummy Geschlecht (1 = Männer)	0.42*	0.21	0.01	0.82	1.517
	NFC	0.39**	0.12	0.15	0.63	1.480
	VW	0.12	0.08	-0.03	0.27	1.132
	PI	-0.10	0.10	-0.29	0.09	0.908
	ASI	0.14	0.10	-0.05	0.33	1.150
	SE	-0.04	0.04	-0.12	0.03	0.957
	Alter	0.001	0.01	-0.01	0.01	1.001
	Geschlecht (1 = m, 0 = w)	0.16	0.20	-0.23	0.55	1.176
	Bildung (1 = hoch, 0 = Rest)	-0.25	0.35	-0.93	0.43	0.777
Hurdle-Modell	Konstante	-1.49*	0.59	-2.64	-0.34	0.225
	Dummy OS	0.43	0.36	-0.28	1.15	1.542
	Dummy HS	-0.24	0.38	-0.99	0.51	0.788
	Dummy DS	0.001	0.38	-0.74	0.75	1.001
	Dummy Likes (1 = viele)	0.14	0.26	-0.38	0.66	1.147
	Dummy Geschlecht (1 = Männer)	-0.18	0.27	-0.70	0.35	0.839
	NFC	0.18	0.16	-0.14	0.49	1.192
	VW	-0.01	0.11	-0.22	0.21	0.995
	PI	0.12	0.13	-0.13	0.37	1.129
	ASI	0.13	0.14	-0.14	0.41	1.140
	SE	0.01	0.06	-0.10	0.13	1.014
	Alter	0.01	0.01	-0.01	0.03	1.008
	Geschlecht (1 = m, 0 = w)	0.17	0.29	-0.40	0.73	1.181
	Bildung (1 = hoch, 0 = Rest)	1.07*	0.47	0.15	1.99	2.916

Anmerkungen: OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 267$

AIC = 679.50, BIC = 779.94,  $-2\log L = 623.498$

Die Modellierung des dritten Modells zur Erklärung des gedanklichen Ursprungs der Rezeptionsgedanken beinhaltete den *Vorwissens- und Erfahrungsabgleich*, den die Versuchspersonen in ihren schriftlichen Ausführungen erkennen ließen. Tabelle 16 enthält die Befunde. Es wurde wiederum ein Hurdle-Modell mit Poissonverteilung mit den oben beschriebenen Modellparametern gerechnet, da geringe Unterdispersion vorlag mit einem Wert von 0.939 und ein Hurdle-Modell mit negativer Binomialverteilung keinen signifikant besseren Modellfit versprach ( $\chi^2(1) = 2.87, p = .090$ ). In der Experimentalbedingung HS (IRR = 0.546, 95% CI [0.31, 0.97]) wurden signifikant weniger Rezeptionsgedanken kognitiver Art, die an bisherigem Wissen ansetzten, genannt (teilweise Bestätigung von H1c), während die Experimentalbedingung *Geschlecht* diese um 52 Prozent (IRR = 1.517, 95% CI [1.01, 2.28]) erhöhte (teilweise Bestätigung von H2c). Während ersteres für Reaktanz spricht und mit einer peripheren Verarbeitung zusammenhängt, lösen Kommentare von Männern im Vergleich zu Kommentaren von Frauen eine zentrale Informationsverarbeitung aus. Dies wird noch von dem hochsignifikanten Kognitionsbedürfnis unterstrichen, das ebenfalls die Anzahl der Gedanken dieses Ursprungs um 48 Prozent erhöhte (IRR = 1.480, 95% CI [1.16, 1.88]). Im Hurdle-Modell war die Odds Ratio für höher Gebildete stark erhöht, kein Vorwissen in die Rezeptionsgedanken einfließen zu lassen. Hier erhärtet sich der Verdacht auf aufgetretene Reaktanzeffekte der Versuchspersonen. Die restlichen Modellparameter übten keinen Einfluss auf die Gedankenanzahl dieser Art aus.

Das erweiterte Modell mit Interaktionstermen enthielt wie bereits oben dargestellt die Experimentalbedingungen, die drei Elaborationsvariablen und Interaktionsterme bestehend aus den Experimentalbedingungen und Kognitionsbedürfnis, da der Modellfit am besten war, AIC = 674.84 (Persönliches Involvement-Interaktionen: AIC = 677.61, Vorwissen-Interaktionen = 684.15). Im Zählmodell senkten Kommentare in der DS-Bedingung die Anzahl an Gedanken, die einen Wissensabgleich enthielten, signifikant (IRR = 0.450, 95% CI [0.23, 0.87]). Kommentare von Männern verursachten hingegen eine Steigerung an Rezeptionsgedanken verglichen mit Kommentaren von Frauen (IRR = 1.747, 95% CI [1.14, 2.69]). Das Kognitionsbedürfnis beeinflusste im Hurdle-Modell signifikant die Beziehungen zwischen den verschiedenen Intensitätsstufen von Hate Speech und der Anzahl an Rezeptionsgedanken aus dieser Kategorie (OS IRR = 2.976, 95% CI [1.22, 7.27]; HS IRR = 2.969, 95% CI [1.13, 7.79] und DS IRR = 3.279, 95% CI [1.22, 8.78]). Damit kann Hypothese H1c als teilweise bestätigt betrachtet werden. Da das Hurdle-Modell die Wahrscheinlichkeit von überschüssigen 0-Werten modelliert, können die Befunde dahingehend interpretiert werden, dass die *Hate-Intensität* zu signifikant weniger Rezeptionsgedanken führte, je höher das Kognitionsbedürfnis der

Versuchspersonen war. Es konnten keine Moderationseffekte für das Geschlecht der Kommentierenden (H2c) sowie das Reaktionsvolumen (H3c) nachgewiesen werden.

Von den insgesamt 326 Versuchspersonen, die mindestens einen Rezeptionsgedanken äußerten, ließen sich die Äußerungen von nur 21 Teilnehmer:innen der Sonstige-Kategorie zuordnen. Dabei handelte es sich um mindestens einen und bis zu sieben Gedanken ( $M = 3.14$ ,  $SD = 2.18$ ), die als irrelevante Gedanken bezeichnet werden, da sie weder die Experimentalsituation noch das Stimulusmaterial thematisierten, weder affektive Reaktionen hervorriefen noch Vorwissen vorbrachten. Folglich befand sich der überwiegende Teil der 326 Versuchspersonen mit keinem Gedanken in dieser Gedankenkategorie, was wiederum die Modellierung eines Hurdle-Modells notwendig machte. Die Modellierung mit den bekannten Modellparametern führte zu hohen Standardfehlern, was als Misspezifikation gewertet werden muss. Um jedoch zu verstehen, wie sich die Experimentalbedingungen auf die vierte Kategorie von Rezeptionsgedanken auswirken, wurde ein vereinfachtes Modell gerechnet. Tabelle 17 enthält eine Übersicht der eingesetzten Parameter und Koeffizienten. Einzig die Bedingung *Likeanzahl* erwies sich als signifikanter Prädiktor für sonstige Rezeptionsgedanken (IRR = 0.376, 95% CI[0.18, 0.79]), d. h., dass signifikant weniger irrelevante Gedanken in der Bedingung *viele Likes* geäußert wurden.

**Tabelle 17:** *Hurdle-Modell mit Poissonverteilung zur Erklärung sonstiger Gedanken*

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	1.41**	0.49	0.45	2.38	4.116
	Dummy OS	-0.39	1.04	-2.44	1.65	0.676
	Dummy HS	-0.19	0.53	-1.23	0.85	0.827
	Dummy DS	0.38	0.40	-0.40	1.17	1.464
	Dummy Likes (1 = viele)	-0.98*	0.38	-1.72	-0.23	0.376
	Dummy Geschlecht (1 = Männer)	-0.18	0.40	-0.97	0.61	0.836
Hurdle-Modell	Konstante	-2.37***	0.58	-3.51	-1.23	0.094
	Dummy OS	-1.05	0.85	-2.72	0.63	0.351
	Dummy HS	0.27	0.61	-0.93	1.48	1.316
	Dummy DS	0.27	0.61	-0.92	1.47	1.314
	Dummy Likes (1 = viele)	-0.41	0.46	-1.31	0.49	0.666
	Dummy Geschlecht (1 = Männer)	-0.17	0.46	-1.07	0.72	0.841

Anmerkungen: OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\*  $p < .01$ . \*\*\* $p < .001$ .

$N = 326$

AIC = 249.91, BIC = 295.35,  $-2\log L = 225.911$

Auf die Modellierung eines Hurdle-Modells mit Interaktionstermen zwischen den Elaborationsvariablen und den Versuchsbedingungen musste verzichtet werden, da das Modell nicht spezifiziert werden konnte. Die einzelne Aufnahme der Elaborationsvariablen und deren Interaktion mit den Versuchsbedingungen führte zu abnorm hohen Standardfehlern, worunter das Kognitionsbedürfnis die niedrigsten Standardfehler enthielt. Da die Schätzung der Koeffizienten folglich nicht verlässlich ist, musste auf deren Darstellung verzichtet werden.

### 6.1.2.3 *Auswirkung von Hate Speech auf die wahrgenommene persönliche Relevanz*

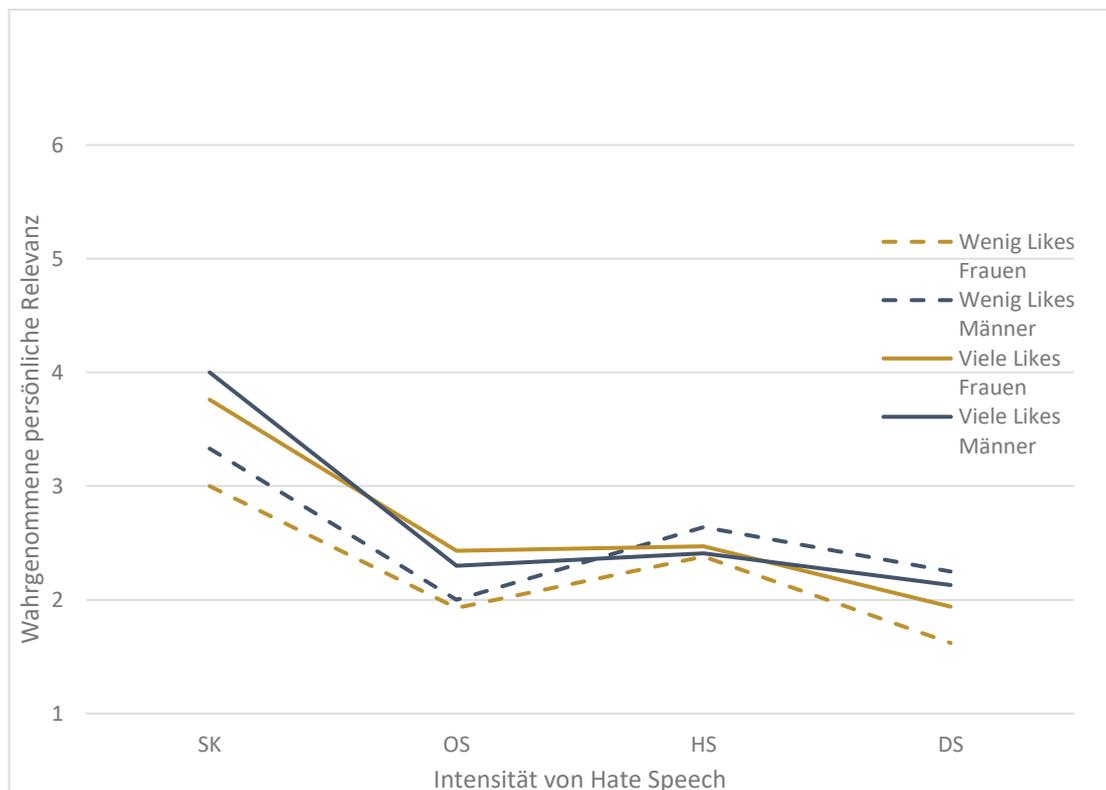
Um Einflüsse kategorialer Variablen auf metrisch skalierte Variablen nachzuweisen, sind v. a. in der experimentellen Forschung Varianzanalysen das Auswertungsverfahren der Wahl (Field, 2018). Zu den wichtigsten Voraussetzungen zählen neben der Homogenität der Varianz, die Unabhängigkeit der Messungen, was durch die Zufallsaufteilung der Versuchspersonen auf die Experimentalgruppen erfüllt wurde sowie die Normalverteilung der Residuen der abhängigen Variablen (Holling & Gediga, 2016). Die deskriptiven Ergebnisse im Abschnitt 6.1.1.5 zeigten, dass die wahrgenommene persönliche Relevanz rechtsschief verteilt ist und stark von der Normalverteilung abweicht. Bei ausreichend großen Stichproben gilt die Varianzanalyse aber als robust gegen die Verletzung der Normalverteilungsannahme. Des Weiteren konnte in einer Simulation von 5000 ANOVAs der Einfluss der Effektstärke auf die Robustheit der Analysemethode nachgewiesen werden (Schmider et al., 2010). Mittels G\*Power wurde eine Effektstärke  $f = 0.54$  in den Experimentalgruppen berechnet, was laut Schmider und Kollegen (2010) dahingehend interpretiert werden kann, dass die Durchführung einer varianzanalytischen Untersuchung robust gegen  $\alpha$ - und  $\beta$ -Fehler ist. Daraus folgend, werden nun die Ergebnisse varianzanalytischer Prüfungen dargestellt.

Die Prüfung des Einflusses von Kommentarmedien auf die wahrgenommene persönliche Relevanz wurde mittels mehrfaktorieller Varianzanalyse durchgeführt. Die Voraussetzung von Varianzhomogenität konnte auf Basis des nicht-signifikanten Levene-Tests ( $p = .117$ ) als erfüllt betrachtet werden. Es konnte ein signifikanter Haupteffekt der *Hate Intensität* auf die wahrgenommene persönliche Relevanz, wie in Hypothese H4a postuliert, nachgewiesen werden ( $F(3, 322) = 25.57, p < .001, \eta_p^2 = .19, \omega_p^2 = .18$ ). Dabei zeigte sich in Post-hoc-Tests mit Bonferroni-Korrektur, dass SK als persönlich am relevantesten wahrgenommen wurde ( $M = 3.59, SD = 1.18$ ) und damit statistisch signifikant ( $p < .001$ ) im Mittel mehr als ein Skalenpunkt höher als OS ( $M = 2.18, SD = 1.16$ ), HS ( $M = 2.48, SD = 1.32$ ) und DS ( $M = 2.02, SD = 1.27$ ) lag. Es lag kein signifikanter Unterschied zwischen OS, HS und DS vor. Für das Geschlecht der Kommentierenden als zweiter Faktor konnte kein signifikanter Haupteffekt gefunden werden ( $F(1, 322) = 1.92, p = 0.166$ ) und somit wird Hypothese H4b vorläufig

abgelehnt. Zwar wurden Kommentare von Männern insgesamt als persönlich relevanter wahrgenommen ( $M = 2.64, SD = 1.38$ ) als Kommentare von Frauen ( $M = 2.48, SD = 1.36$ ), aber der Unterschied war zu gering, um statistische Signifikanz zu erreichen.

Der Haupteffekt für die Likeanzahl als dritter Faktor erwies sich als signifikant und deutet in Anbetracht der nicht bewussten Wahrnehmung (vgl. Kapitel 6.1.1.4) auf das Eintreten eines Mere Exposure-Effekts hin ( $F(1, 322) = 6.75, p < .05$ ). Hypothese H4c kann folglich vorläufig bestätigt werden. Im Vergleich zu Kommentaren mit wenigen Likes ( $M = 2.40, SD = 1.26$ ) wurden Kommentare mit vielen Likes als persönlich relevanter wahrgenommen ( $M = 2.72, SD = 1.45$ ). Trotz der Signifikanz des Haupteffekts war die Effekstärke jedoch gering ( $\eta_p^2 = .02, \omega_p^2 = .02$ ).

**Abbildung 13:** Einfluss der Hate Intensität, Geschlecht der Kommentierenden und des Reaktionsvolumens auf die wahrgenommene persönliche Relevanz



Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.  $N = 338$

Abbildung 13 veranschaulicht die Auswirkungen der drei Experimentalbedingungen auf die wahrgenommene persönliche Relevanz. Der deutlich sichtbare Unterschied zwischen den Intensitätsstufen folgte in den beiden anderen Bedingungen (Likeanzahl und Geschlecht der Kommentierenden) einem konsistenten Muster und war am höchsten in der Bedingung SK und in der Bedingung DS am geringsten. Kommentare von Männern wurden tendenziell als relevanter wahrgenommen, auch wenn die Unterschiede zu gering waren, um statistische

Signifikanz zu erreichen (s. o.). Die Unterschiede in der Relevanzwahrnehmung der Likeanzahl wurden in der Intensitätsstufe SK am deutlichsten sichtbar, wobei sich der Effekt über die übrigen Stufen hinweg nivellierte. Dieser Effekt ist zwar im Schaubild visualisiert, ließ sich jedoch statistisch nach Bonferroni-Post-hoc-Korrektur nicht aufrechterhalten. Sämtliche Interaktionseffekte der Faktorstufen untereinander waren nämlich nicht signifikant ( $F(3, 322) = 0.08, p = .972$ ).

In den folgenden Ausführungen soll geprüft werden, inwiefern ambivalent sexistische Einstellungen den Einfluss zwischen den Experimentalbedingungen und der persönlichen Relevanzwahrnehmung moderieren. Geprüft wurden diese Annahmen mittels Moderationsanalyse, welche ihrerseits zunächst einige Voraussetzungen erfüllen muss: Hayes (2018, S. 69–73) nennt die lineare Beziehung zwischen abhängiger und unabhängiger Variable, normalverteilte Residuen, Homoskedastizität und die Unabhängigkeit der Messungen. Letztere Voraussetzung ist bereits im Studiendesign implementiert; Linearität konnte nach Inspektion der Streudiagramme der Residuen angenommen werden und da die Relevanz als abhängige Variable nicht normalverteilt war, wurden 10000 Bootstrapping-Samples berechnet, um Konfidenzintervalle ohne voraussetzungsvolle Verteilungsannahmen zu schätzen (Field, 2018). Um darüber hinaus verzerrten Schätzern durch Heteroskedastizität entgegenzuwirken, wurden robuste HC3-Standardfehler verwendet (R. Davidson & MacKinnon, 1993; Hayes & Cai, 2007).

Durchgeführt wurde die Moderationsanalyse mittels PROCESS-Makro für SPSS (Hayes, 2018). Dabei wurde jede der drei Experimentalbedingungen jeweils einzeln als unabhängige Variable einer Moderationsanalyse zugeführt und Effekte der beiden verbleibenden Experimentalbedingungen als Kovariate kontrolliert (Hayes, 2018, S. 141–144). Für die *Intensitätsstufe* bildeten sachliche Kommentare die Referenzkategorie, für *Geschlecht* wurden Frauen und für das *Reaktionsvolumen* wenig Likes als Referenzkategorie gewählt. Um die Interpretierbarkeit der Indexvariablen für Ambivalenten Sexismus (ASI) zu gewährleisten, wurde diese mittelwertzentriert in das Modell aufgenommen.

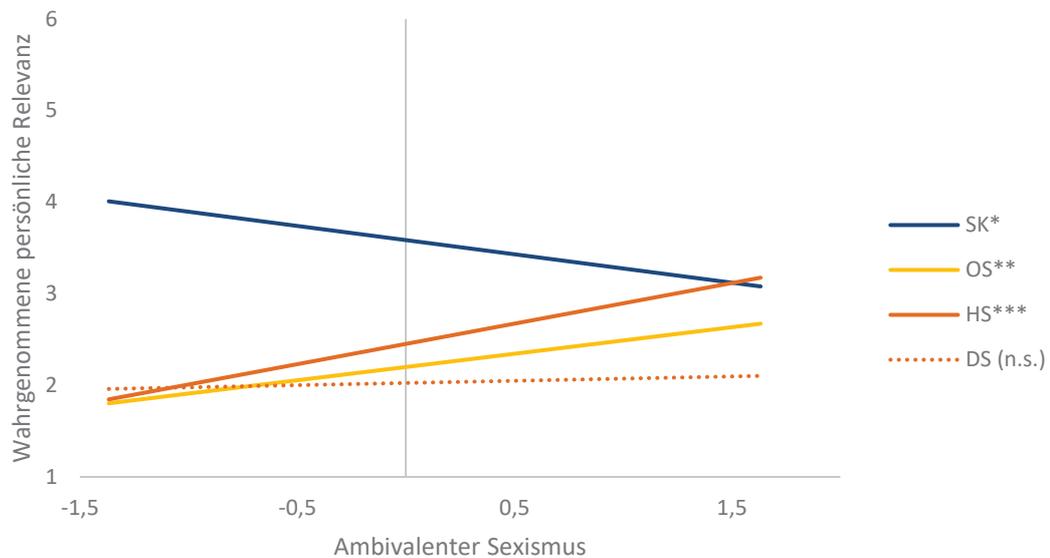
#### *Hate Intensität*

Das Modell konnte den Zusammenhang zwischen Intensität als Kommentarmerkmal, ambivalent sexistischen Einstellungen und der wahrgenommenen persönlichen Relevanz von Kommentaren signifikant abbilden  $F(9, 328) = 14.78, p < .001$  und trug in hohem Maße zur Aufklärung der Varianz bei ( $R^2 = .27$ ). Die drei Faktorstufen der unabhängigen Variablen wurden statistisch signifikant als weniger persönlich relevant wahrgenommen im Vergleich zu

sachlichen Kommentaren, OS ( $b = -1.39, p < .001$ ), HS ( $b = -1.13, p < .001$ ) und DS ( $b = -1.56, p < .001$ ). Die Regressionskoeffizienten der Prädiktorvariablen drücken die Differenz zwischen Referenzgruppe (SK) und Vergleichsgruppe (z. B. OS) in der Relevanzwahrnehmung der Versuchspersonen mit durchschnittlich ambivalenten Einstellungen aus, da die Moderatorvariable mittelwertzentriert wurde. Der Regressionskoeffizient für die Moderatorvariable Ambivalenter Sexismus (ASI) spiegelt den konditionalen Effekt für die Referenzgruppe (SK) wider, d. h. die Änderung in der Relevanzwahrnehmung für die Versuchspersonen in der Experimentalbedingung SK, die sich um eine Einheit auf der ASI-Skala unterscheiden. Dieser war signifikant negativ ( $b = -0.31, p < .05$ ). Der Moderationseffekt von Ambivalentem Sexismus und *Hate Intensität* war signifikant  $F(3, 328) = 5.95, p < .001$  und konnte das Modell geringfügig verbessern,  $\Delta R^2 = .04$ . Ein signifikanter Moderationseffekt von Ambivalentem Sexismus konnte zwischen OS und ASI ( $b = 0.60, p < .01$ ) sowie zwischen HS und ASI ( $b = 0.75, p < .001$ ), nicht jedoch zwischen DS und ASI ( $b = 0.36, p = .095$ ) festgestellt werden. So schätzte etwa der signifikante Interaktionskoeffizient  $b_{HS \times ASI} = 0.75$ , wie stark sich zwei Versuchspersonen in der Experimentalbedingung HS im Vergleich zur Referenzkategorie SK bezüglich der wahrgenommenen persönlichen Relevanz bei einer Abweichung von einer Einheit auf der Ambivalenter Sexismus Skala unterscheiden. Hypothese H4a kann also teilweise bestätigt werden und bedarf einer differenzierten Formulierung in folgenden Forschungsarbeiten, denn es konnte gezeigt werden, dass nicht alle Faktorstufen Signifikanz erlangten.

Abbildung 14 veranschaulicht den Zusammenhang und zeigt auch, dass sachliche Kommentare mit steigenden ambivalent sexistischen Einstellungen als persönlich weniger relevant wahrgenommen wurden. Dagegen wurden die drei Intensitätsstufen mit steigendem ASI-Score als immer relevanter wahrgenommen. Den höchsten Wert erzielte die HS-Bedingung. Die Steigung in der Experimentalbedingung DS war am geringsten und ähnelte einer horizontal flachen Linie. Ob Kommentare, die sich zu weiblichen Karrierechancen in der IT äußern, als persönlich relevant wahrgenommen wurden, hängt also zum Teil von ambivalent sexistischen Einstellungen der Versuchspersonen ab. Der Erklärungsbeitrag der Interaktion zwischen *Hate Intensität* und ASI erfuhr jedoch seine Grenzen, wenn die Aggressivität der Kommentare eine Schwelle überschritten, wie die nicht-signifikanten Befunde der DS-Bedingung signalisieren. Und schließlich zeigten die beiden als Kovariate eingefügten Experimentalbedingungen *Geschlecht* ( $b = 0.15, p = .269$ ) sowie *Anzahl der Likes* ( $b = 0.25, p = .062$ ) keine Signifikanz. Robuste Schätzungen mittels Bootstrapping bestätigten diese Befunde.

**Abbildung 14:** Einfluss von Hate Intensität auf die wahrgenommene persönliche Relevanz unter der Bedingung ambivalent sexistischer Einstellungen



\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ , n.s. = nicht signifikant

SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.  $N = 338$

Das Moderationsmodell wurde im nächsten Schritt um sieben Kovariate ergänzt, um Einflüsse auf den Zusammenhang aus der vorhergehenden Analyse auf die abhängige Variable zu kontrollieren. Berücksichtigt wurden soziodemografische Merkmale (Alter, Geschlecht und Bildung der Versuchspersonen als dummycodierte Variable, in der die mittlere und niedrige Bildung als Referenzkategorie dienten), die drei Elaborationsvariablen Kognitionsbedürfnis, persönliches Involvement, Vorwissen und schließlich soziale Erwünschtheit. Das Gesamtmodell war statistisch signifikant  $F(1, 258) = 10.70$ ,  $p < .001$  und konnte die Varianzaufklärung gegenüber dem Modell ohne Kovariate verbessern ( $R^2 = .34$ ). Die drei Vergleichsgruppen der Experimentalbedingung *Hate Intensität* unterschieden sich wieder signifikant von der Referenzkategorie SK, und zwar OS ( $b = -1.44$ ,  $p < .001$ ), HS ( $b = -1.01$ ,  $p < .001$ ) und DS ( $b = -1.49$ ,  $p < .001$ ). Ambivalent sexistische Einstellungen senkten die wahrgenommene persönliche Relevanz ( $b = -0.28$ ,  $p < .05$ ), hatten jedoch teilweise einen positiven Effekt auf die Beziehung zwischen der Rezeption der Hate Speech Intensität und der wahrgenommenen persönlichen Relevanz. OS wurde als relevanter wahrgenommen, je stärker ambivalent sexistische Einstellungen ausgeprägt waren ( $b = 0.63$ ,  $p < .01$ ). Ähnlich war der Befund für HS mit ( $b = 0.77$ ,  $p < .01$ ). Für DS konnte kein Interaktionseffekt nachgewiesen werden ( $b = 0.36$ ,  $p = .132$ ). Die Interaktion zwischen Prädiktor und Moderatorvariable trug signifikant zur Varianzaufklärung bei  $F(3, 258) = 6.04$ ,  $p < .001$ ,  $\Delta R^2 = .05$ . Die Kontrolle für Effekte

sozialer Erwünschtheit zeigten keinen Einfluss auf die Relevanzwahrnehmung ( $b = 0.01$ ,  $p = .713$ ), und führte daher nicht zur Einschränkung bei der Interpretation der Befunde. Von den zusätzlich zu den beiden Experimentalbedingungen (Geschlecht und Likeanzahl) eingefügten Kovariaten erhöhten das selbst eingeschätzte Vorwissen ( $b = 0.14$ ,  $p < .05$ ) und der Bildungsgrad ( $b = 0.44$ ,  $p < .05$ ) die wahrgenommene persönliche Relevanz. Das Alter der Versuchspersonen hatte einen negativen Einfluss auf die Kriteriumsvariable ( $b = -0.02$ ,  $p < .01$ ) und deutete darauf hin, dass Jüngere die Intensität von Hate Speech als signifikant relevanter wahrgenommen haben. Auch hier deckten sich wieder die Befunde der Konfidenzintervalle der Bootstrapping-Samples (d. h. der Wert 0 war nicht enthalten) mit den berechneten Signifikanzen auf dem 95%-Niveau.

#### *Geschlecht der Kommentierenden*

Analog zum ersten Modell enthielt das zweite Modell das *Geschlecht der Kommentierenden* als UV und die anderen beiden Experimentalbedingungen als Kovariate. Das Modell war signifikant,  $F(7, 330) = 9.00$ ,  $p < .001$  undklärte 16 Prozent der Varianz auf. Das Geschlecht der Kommentierenden als Experimentalbedingung war nicht signifikant ( $b = 0.14$ ,  $p = .316$ ), d. h., dass die Relevanzwahrnehmung nicht von der Rezeption dieses Kommentarmediums beeinflusst wurde. Ambivalent sexistische Einstellungen übten einen signifikanten Einfluss auf die Outcomevariable aus ( $b = 0.16$ ,  $p < .05$ ), d. h., dass Kommentare von Frauen bei einem Unterschied um eine Einheit auf der ASI-Skala als relevanter als Kommentare von Männern wahrgenommen wurden. Darüber hinaus war der Moderationseffekt zwischen dem Geschlecht der Kommentierenden und Ambivalentem Sexismus nicht signifikant  $\Delta R^2 = .003$ ,  $F(1, 330) = 0.85$ ,  $p = .356$ , 95% CI[-0.46, 0.15]. Hypothese H4b wurde daher abgelehnt. Für die einzelnen Kategorien der Experimentalbedingung *Intensität von Hate Speech* wurden Dummyvariablen mit SK als Referenzkategorie eingefügt. Es konnten wie erwartet signifikante Unterschiede nachgewiesen werden, OS ( $b = -0.98$ ,  $p < .001$ ), HS ( $b = -0.68$ ,  $p < .001$ ) und DS ( $b = -1.43$ ,  $p < .001$ ). Und schließlich erwies sich die dritte Experimentalbedingung *Likeanzahl*, für die statistisch kontrolliert wurde, als signifikant ( $b = 0.64$ ,  $p < .001$ ), was sich darin ausdrückt, dass viele SK-Likes von Frauen als persönlich relevanter wahrgenommen wurden als wenig Likes.

Die oben beschriebene Vorgehensweise bezüglich der Integration der Kovariate zur Kontrolle möglicher Störeffekte ergab ein signifikantes Modell,  $F(14, 260) = 6.44$ ,  $p < .001$  mit einer Varianzaufklärung von  $R^2 = .23$ . Das Geschlecht der Kommentierenden übte keinen Einfluss auf die Relevanzwahrnehmung der Versuchspersonen aus ( $b = 0.23$ ,  $p = .138$ ). Der Moderationseffekt zwischen Prädiktor und Moderator war wiederum nicht signifikant  $F(1, 260) = 1.38$ ,

$p = .242$ , 95% CI[-0.53, 0.11], obgleich Ambivalenter Sexismus als Moderatorvariable durchaus einen Einfluss auf die Relevanzwahrnehmung hatte ( $b = 0.24$ ,  $p < .05$ ) und zwar dergestalt, dass Kommentare von Frauen umso relevanter wahrgenommen wurden, je höher der Score auf der ASI-Skala war. Die dummycodierten Kategorien (SK als Referenzkategorie) der Experimentalbedingung *Hate Intensität* verringerten in diesem Modell die Relevanzwahrnehmung signifikant, OS ( $b = -1.07$ ,  $p < .001$ ), HS ( $b = -0.59$ ,  $p < .01$ ) und DS ( $b = -1.30$ ,  $p < .001$ ). Die als Kovariate eingefügte Experimentalbedingung *Reaktionsvolumen* erhöhte die Relevanzwahrnehmung signifikant ( $b = 0.58$ ,  $p < .001$ ). Auch Vorwissen ( $b = 0.16$ ,  $p < .05$ ) und Bildung ( $b = 0.46$ ,  $p < .05$ ) übten einen positiven Effekt aus. Jüngere Versuchspersonen nahmen die präsentierten Kommentare als relevanter wahr ( $b = -0.02$ ,  $p < .05$ ). Die übrigen Kovariaten (Geschlecht der Versuchspersonen, soziale Erwünschtheit, Kognitionsbedürfnis und Persönliches Involvement) waren nicht signifikant.

#### *Reaktionsvolumen*

Im dritten Modell wurde das Reaktionsvolumen – operationalisiert als Likeanzahl – als Prädiktor modelliert mit den beiden anderen Experimentalbedingungen als Kovariate. Das Modell war signifikant,  $F(7, 330) = 9.08$ ,  $p < .001$  mit einer Varianzaufklärung von  $R^2 = .16$ . Die Likeanzahl als Prädiktor hatte einen signifikanten Einfluss auf die Relevanzwahrnehmung ( $b = 0.65$ ,  $p < .001$ ), auch ambivalent sexistische Einstellungen nach Sichtung der robusten Bootstrapping-Intervalle ( $b = 0.16$ ,  $p = .051$ , 95% CI [0.003, 0.304]), wenn auch das Ergebnis streng genommen den Wert Null bis auf die zweite Kommastelle einschließt. Die Interaktion zwischen Moderator und Prädiktor ( $\Delta R^2 = .0003$ ,  $F(1, 330) = 0.08$ ,  $p = .775$ , 95% CI[-0.254, 0.354]) war nicht signifikant. Die Hypothese 4c wird daher teilweise abgelehnt. Die als Dummyvariablen eingefügten Kategorien für die Experimentalbedingung *Hate Intensität* senkten die Relevanzwahrnehmung signifikant: OS ( $b = -0.98$ ,  $p < .001$ ), HS ( $b = -0.68$ ,  $p < .001$ ) und DS ( $b = -1.44$ ,  $p < .001$ ). Die Experimentalbedingung *Geschlecht* übte in diesem Modell keinen Einfluss auf die wahrgenommene persönliche Relevanz der Versuchspersonen aus ( $b = 0.15$ ,  $p = .301$ ).

Die Kontrolle für störende Einflüsse durch die zusätzlichen Kovariate ergab ein signifikantes Modell,  $F(14, 260) = 6.20$ ,  $p < .001$  mit einer Varianzaufklärung von  $R^2 = .23$ . Signifikante positive Effekte auf die Relevanzwahrnehmung gingen von der Likeanzahl als Prädiktor ( $b = 0.58$ ,  $p < .001$ ), ambivalent sexistischen Einstellungen ( $b = 0.23$ ,  $p < .05$ ), Vorwissen ( $b = 0.16$ ,  $p < .05$ ) und Bildung ( $b = 0.48$ ,  $p < .05$ ) aus. Gesenkt wurde die Wahrnehmung der persönlichen Relevanz von den dummycodierten Kategorien der Hate Intensität, und zwar für OS ( $b = -1.07$ ,  $p < .001$ ), HS ( $b = -0.58$ ,  $p < .01$ ) und DS ( $b = -1.31$ ,  $p < .001$ ) sowie dem Alter

der Versuchspersonen ( $b = -0.02$ ,  $p < .05$ ). Die Moderation zwischen Prädiktor und Moderatorvariable war nicht signifikant ( $\Delta R^2 = .002$ ,  $F(1, 260) = 0.56$ ,  $p = .455$ , 95% CI[-0.187, 0.450], d. h. ambivalent sexistische Einstellungen konnten die Beziehung zwischen der Likeanzahl und der Relevanzwahrnehmung nach Kontrolle für Kovariate nicht beeinflussen.

#### 6.1.2.4 Einflüsse auf die Verhaltensintention

Im folgenden Abschnitt werden die Befunde zur Rezeptionswirkung auf die Verhaltensintention präsentiert. Sicherlich greifen mechanistische Rezeption-Wirkungsvorstellungen zu kurz, weswegen neben den Experimentalbedingungen, auch das habitualisierte Verhalten Berücksichtigung fand, das sich in der *Theory of Planned Behavior* als verlässlicher Prädiktor für Verhaltensintentionen erwiesen hat (Ajzen, 1991). Soziodemografische Merkmale wurden als Kovariate berücksichtigt, um für störende Einflüsse zu kontrollieren. Die habitualisierten Nutzungsweisen (HN) des sozialen Netzwerks Facebook wurden passend zur jeweiligen Verhaltensintention gewählt, beispielsweise wurde die Verhaltensintention VI4 „Ich unterstütze meine Vorredner mit einem sachlichen Kommentar“ sowohl darauf getestet, ob sie (1) gewohnheitsmäßig Beiträge auf öffentlichen Seiten kommentierten oder (2) gewohnheitsmäßig Beiträge von Freunden kommentierten. Einige habitualisierte Nutzungsweisen standen nicht im Zusammenhang mit den abgefragten Verhaltensintentionen (z. B. Tagging von Fotos und Beiträgen) und wurden daher nicht getestet.

Aufgrund von schwerwiegenden Verletzungen der Voraussetzungen für lineare Modellierungen, die auf OLS-Schätzungen basieren (nicht-lineare Beziehungen zwischen den Variablen und keine multivariaten Normalverteilungen) wurden nicht-parametrische Verfahren angewandt, und zwar Kruskal-Wallis-Tests mit Bonferroni-korrigierten Post-hoc-Tests für die *Hate Intensität* und Mann-Whitney-U-Tests für die dummycodierten Experimentalbedingungen *Geschlecht der Kommentierenden* und *Likeanzahl*, sowie die *habitualisierten Nutzungsweisen* von Facebook. Somit können die Hypothesen H7a bis H7c nicht mehr getestet werden, da sie eines Vergleichs der Akzeptanzrichtung und des Aktivitätsgrads bedürfen. Die Verhaltensintentionen wurden einzeln je Gruppe getestet, was sowohl Vor- als auch Nachteile hat. Dafür spricht, dass Effekte auf die einzelnen Faktoren zurückführbar sind, was die Interpretation der Befunde erleichtert. Ein wesentlicher Nachteil ist selbstredend die Kumulierung des  $\alpha$ -Fehlers aufgrund der Mehrfachtestung. Aus diesem Grund wurde das Signifikanzniveau auf  $\alpha = 0.01$  gesenkt, um für die Mehrfach-Tests für die acht Variablen zur Messung der Verhaltensintention zu korrigieren. Um auch Einflüsse von soziodemografischen Merkmalen (Alter, Geschlecht, Bildung) zu kontrollieren, wurde eine rangbasierte Kovarianzanalyse durchgeführt, welche von Dana Quade als Alternative zu voraussetzungsvollen linearen

Kovarianzanalysen vorgestellt wurde (Quade, 1967). Die Quade-ANCOVA erstellt Ränge zwischen mindestens ordinal-skalierten Zufallsvariablen (dichotome Variablen sind ebenfalls erlaubt) und ihre Teststatistik folgt asymptotisch einer F-Verteilung (ebd.). Dazu wird die Nullhypothese getestet, dass die konditionalen Verteilungen von Y auf X identisch sind (Quade, 1967, S. 1187). Fällt der F-Test signifikant aus, wird geprüft, welche der Kovariaten für die Unterschiede verantwortlich war.

#### *Liken als Zustimmung (VI1)*

Das *Liken als Zustimmung* der experimentell manipulierten Stimuli unterschied sich in den Gruppen *Hate Intensität* signifikant ( $H(3) = 86.18, p < .001$ ). In der Gruppe SK wurde das Liken als Verhaltensabsicht als wahrscheinlicher ( $M = 3.08, SD = 1.80$ ) als in den übrigen Experimentalgruppen angenommen. Die Unterschiede waren jeweils zwischen SK und OS ( $p < .001, r = .39$ ), HS ( $p < .001, r = .52$ ) und DS ( $p < .001, r = .68$ ) signifikant, nicht jedoch zwischen den Hate Gruppen untereinander. Der Einfluss sozial erwünschten Antwortverhaltens war nicht signifikant ( $H(3) = 6.25, p = .100$ ). Die Quade-ANCOVA ergab eine signifikante Teststatistik ( $F(3, 336) = 36.61, p < .001$ ). Nachfolgende Kruskal-Wallis-Tests zeigten, dass männliche und weibliche Versuchspersonen die Wahrscheinlichkeit mit einem Like zu reagieren, in den Gruppen unterschiedlich bewerteten (vgl. Tabelle 18). Weibliche Versuchspersonen unterschieden sich deutlich in den Gruppen SK und HS ( $p < .001, r = .39$ ) sowie SK und DS ( $p < .001, r = .51$ ) und auch in den Gruppen OS und DS ( $p < .001, r = .30$ ) hinsichtlich ihrer Absicht auf die Kommentare mit einem Like als Zustimmung zu reagieren. Diesbezügliche Unterschiede gab es bei männlichen Versuchspersonen jeweils in den Gruppe SK und OS ( $p < .001, r = .37$ ), SK und HS ( $p < .001, r = .35$ ) sowie SK und DS ( $p < .001, r = .47$ ). Das Alter der Versuchspersonen wirkte sich nicht signifikant auf die Verhaltensintention aus (Pearsons  $r = .02, p = .764$ ). Signifikante Gruppenunterschiede gab es bei Versuchspersonen mit hoher Bildung ( $H(3) = 80.47, p < .001$ ). Den größten Effekt gab es diesbezüglich zwischen den Gruppen SK und DS ( $p < .001, r = .65$ ), gefolgt von SK und HS ( $p < .001, r = .52$ ), SK und OS ( $p < .001, r = .37$ ). Der Unterschied zwischen Gruppen OS und DS war mittelstark mit ( $p < .01, r = .28$ ).

In den Experimentalgruppen, in denen das *Geschlecht der Kommentierenden* variiert wurde, konnten zwischen Kommentaren von Frauen ( $M = 2.01, SD = 1.49, Mdn. = 1$ ) und Kommentaren von Männern ( $M = 1.83, SD = 1.50, Mdn. = 1$ ) keine überzufälligen Unterschiede für das Liken als Zustimmung nachgewiesen werden ( $U = 12983.500, z = -1.68, p = .094, r = -.09$ ). Soziale Erwünschtheit beeinflusste das Antwortverhalten nicht ( $U = 45.000, z = 1.07, p = .391$ ). Die Quade-ANCOVA zeigte an, dass keine Gruppenunterschiede hinsichtlich einer der

Kovariaten Alter, Geschlecht und Bildung der Versuchspersonen vorlagen  $F(1, 338) = 2.83$ ,  $p = .094$ ).

**Tabelle 18:** Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Zustimmung für Hate Intensität mittels Like-Button auszudrücken

Gruppen	Gesamt		Weibliche Versuchspersonen		Männliche Versuchspersonen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SK ( <i>n</i> = 86)	3.08	1.80	2.93	1.80	3.42	1.79
OS ( <i>n</i> = 87)	1.83	1.31	2.02	1.43	1.38	0.80
HS ( <i>n</i> = 84)	1.56	1.24	1.53	1.20	1.63	1.35
DS ( <i>n</i> = 83)	1.13	0.60	1.10	0.59	1.18	0.63
<i>H</i> (3)	86.18***		50.07***		42.07***	

Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

Da das gewählte Signifikanzniveau bei  $\alpha = 0.01$  lag, musste der Mann-Whitney-U-Test für das Reaktionsvolumen auf dem Niveau  $p < .05$  abgelehnt werden ( $U = 16047.500$ ,  $z = 2.15$ ,  $p = .032$ ,  $r = .12$ ). Die Versuchspersonen gaben es als etwas wahrscheinlicher an, Kommentare mit vielen Likes selbst ebenfalls zu liken ( $M = 2.11$ ,  $SD = 1.65$ ,  $Mdn. = 1$ ) als Versuchspersonen, die Stimuli mit wenigen Likes sahen ( $M = 1.69$ ,  $SD = 1.27$ ,  $Mdn. = 1$ ). Es konnte kein signifikanter Einfluss sozial erwünschten Antwortverhaltens nachgewiesen werden ( $U = 51.500$ ,  $z = 0.93$ ,  $p = .432$ ). Auch in diesem Fall waren die Kovariaten auf dem Niveau  $\alpha = 0.01$  nicht für signifikante Gruppenunterschiede verantwortlich ( $F(1, 338) = 4.56$ ,  $p = .033$ ).

War die HN der Befragten davon geprägt, Likes für Beiträge und Kommentare zu vergeben, sahen sie es als wahrscheinlicher an, die experimentellen Stimuli zu liken ( $M = 2.09$ ,  $SD = 1.62$ ,  $Mdn. = 1$ ) im Vergleich zu Befragten, die eher seltener mit Inhalten per Like interagierten ( $M = 1.63$ ,  $SD = 1.24$ ,  $Mdn. = 1$ ). Dieser Befund war aber nicht signifikant auf dem Niveau  $\alpha = 0.01$  ( $U = 15633.500$ ,  $z = 2.46$ ,  $p = .014$ ,  $r = .13$ ).

Hypothese 8 konnte nicht überprüft werden, da die Voraussetzung der Linearität verletzt war. Der Zusammenhang zwischen ambivalent sexistischen Einstellungen und der artikulierten Wahrscheinlichkeit zustimmend mittels Like zu reagieren, wies nach Berechnung des Pearson Korrelationskoeffizienten eine mittlere Stärke auf ( $r = .20$ ,  $p < .05$ ).

*Emojis als Ausdruck der Ablehnung (VI2)*

Die Interaktion mit den Reaktionsbuttons bei Facebook, die als Ausdruck der Verärgerung, Missbilligung oder Ablehnung genutzt werden, wurde in den vier Gruppen der Bedingung *Hate Intensität* unterschiedlich wahrscheinlich eingestuft ( $H(3) = 17.33, p < .001$ ). In Verbindung mit bisherigen Ergebnissen, deutet die als gering eingeschätzte Wahrscheinlichkeit für DS ( $M = 1.72, SD = 1.40$ ) auf Reaktanz hin, bei der ein Schwellenwert an Aggression überschritten wird, den die Befragten noch als tolerabel wahrnehmen. Signifikante Unterschiede bestanden zwischen SK und HS ( $p < .01, r = -.28$ ) sowie zwischen DS und HS ( $p < .01, r = .27$ ). Soziale Erwünschtheit übte keinen signifikanten Einfluss aus ( $H(3) = 2.66, p = .448$ ). Die Prüfung der Einflüsse der Kovariate zeigte eine signifikante Statistik der Quade-ANCOVA  $F(3, 336) = 5.19, p < .01$  an und ergab einen schwach positiven Zusammenhang zwischen dem Alter und der artikulierten Wahrscheinlichkeit auf die rezipierten Kommentare mit einem negativen Reaktionsbutton zu antworten ( $r = .15, p < .01$ ). Die Verhaltensintentionen in den Gruppen SK und HS unterschieden sich mittelstark bei weiblichen ( $p < .001, r = .32$ ) und bei hoch gebildeten Versuchspersonen ( $p < .01, r = .25$ ).

**Tabelle 19:** Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Ablehnung für Hate Intensität mittels Reaktionsbutton auszudrücken

Gruppen	Gesamt		Weibliche Versuchspersonen		Männliche Versuchspersonen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SK ( $n = 86$ )	1.64	1.19	1.55	1.56	1.85	1.26
OS ( $n = 87$ )	1.70	1.10	1.79	1.14	1.50	0.99
HS ( $n = 84$ )	2.44	1.66	2.65	1.73	1.92	1.38
DS ( $n = 83$ )	1.72	1.40	2.08	1.67	1.21	0.59
$H(3)$	17.33***		17.70***		6.68, <i>n.s.</i>	

Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

In der Experimentalbedingung *Geschlecht* wurde die Wahrscheinlichkeit die Stimuluskommentare mit einem Missbilligungsemoji zu markieren als eher gering angegeben, denn sie lag bei  $M = 1.90$  ( $SD = 1.36, Mdn. = 1$ ) für weibliche und bei  $M = 1.85$  ( $SD = 1.41, Mdn. = 1$ ) für männliche Kommentierende. Die geringen Unterschiede waren nicht signifikant ( $U = 13655.500, z = -0.76, p = .447, r = .04$ ). Auch soziale Erwünschtheit verzerrte das Antwortverhalten nicht ( $U = 47.500, z = 1.50, p = .257$ ). Die Kontrolle für Einflüsse von Kovariaten ergab keine Signifikanz  $F(1, 338) = 0.72, p = .397$ ).

Auch die Experimentalbedingung *Reaktionsvolumen* lieferte per Mann-Whitney-U-Test ( $U = 14874.00$ ,  $z = 0.58$ ,  $p = .562$ ,  $r = .03$ ) und für die Kovariatenkontrolle per Quade-ANCOVA ( $F(1, 338) = 0.37$ ,  $p = .545$ ) keine signifikanten Befunde. Effekte sozial erwünschten Antwortverhaltens konnten nicht nachgewiesen werden ( $U = 52.000$ ,  $z = 1.09$ ,  $p = .432$ ).

Waren die Befragten darin routiniert, den *Reaktionsbutton* zu nutzen, gaben sie es als wahrscheinlicher an ( $U = 16722.00$ ,  $z = 3.80$ ,  $p < .001$ ,  $r = .21$ ), ein Missbilligungsemoji zu klicken ( $M = 2.11$ ,  $SD = 1.51$ ,  $Mdn. = 1$ ) als jene, die diese Interaktion nicht gewohnheitsmäßig nutzten ( $M = 1.51$ ,  $SD = 1.07$ ,  $Mdn. = 1$ ).

### *Sachliche Unterstützung mittels Kommentares (VI3)*

*Sachliche Kommentare zur Unterstützung* wurden als unterschiedlich wahrscheinliche Verhaltensintention in der Versuchsbedingung *Hate Intensität* angegeben ( $H(3) = 41.68$ ,  $p < .001$ ) und es ergaben sich im paarweisen Vergleich zwischen SK und DS ( $p < .001$ ,  $r = .47$ ), zwischen SK und HS ( $p < .01$ ,  $r = .29$ ) und OS und DS ( $p < .01$ ,  $r = .34$ ) signifikante mittelstarke bis starke Effekte. Die Antworten waren nicht durch soziale Erwünschtheit beeinflusst ( $H(3) = 1.46$ ,  $p = .692$ ). Die Berücksichtigung von Kovariaten in der Quade-ANCOVA ( $F(3, 336) = 15.93$ ,  $p < .001$ ) ergab einen signifikanten Befund. Weibliche Versuchspersonen unterschieden sich mittelstark im paarweisen Vergleich in den Gruppen SK und DS ( $p < .001$ ,  $r = .36$ ) sowie OS und DS ( $p < .001$ ,  $r = .35$ ) bezüglich der artikulierten Wahrscheinlichkeit mit einem sachlichen Kommentar zu antworten. Unterschiede bei männlichen Versuchspersonen waren nur in den Gruppen SK und DS ( $p < .001$ ,  $r = .32$ ) signifikant. Hoch gebildete Versuchspersonen unterschieden sich in den Gruppen SK und HS ( $p < .01$ ,  $r = .26$ ), SK und DS ( $p < .001$ ,  $r = .44$ ) sowie OS und DS ( $p < .001$ ,  $r = .32$ ) bezüglich der artikulierten Wahrscheinlichkeit sachlich und unterstützend zu kommentieren. Je älter die Befragten waren, desto eher gaben sie in der Experimentalbedingung *Hate Intensität* diese Verhaltensintention als wahrscheinlich an (Pearsons  $r = .16$ ,  $p < .01$ ), wobei die Effektstärke als gering einzustufen ist.

Die subjektive Wahrscheinlichkeit mit einem sachlichen Kommentar die Kommentierenden zu unterstützen, war in der Gruppe *Geschlecht* mit Kommentaren von Frauen ( $M = 1.79$ ,  $SD = 1.27$ ,  $Mdn. = 1$ ) verglichen mit Kommentaren von Männern ( $M = 1.71$ ,  $SD = 1.30$ ,  $Mdn. = 1$ ) minimal und daher statistisch nicht verschieden ( $U = 13423.500$ ,  $z = -1.11$ ,  $p = .268$ ,  $r = .06$ ). Es konnte kein Einfluss sozial erwünschten Antwortverhaltens festgestellt werden ( $U = 45.000$ ,  $z = 1.30$ ,  $p = .391$ ). Auch die Teststatistik der Quade-ANCOVA konnte in dieser Versuchsbedingung keine Einflüsse der Kovariaten nachweisen ( $F(1, 338) = 1.29$ ,  $p = .257$ ).

Das *Reaktionsvolumen*, also die Unterscheidung zwischen Kommentaren mit wenigen Likes ( $M = 1.67$ ,  $SD = 1.18$ ,  $Mdn. = 1$ ) im Vergleich zu Kommentaren mit vielen Likes ( $M = 1.81$ ,  $SD = 1.38$ ,  $Mdn. = 1$ ), beeinflusste nicht die Absicht mit einem sachlichen Kommentar auf die Stimuli zu antworten ( $U = 14968.00$ ,  $z = 0.73$ ,  $p = .464$ ,  $r = .04$ ) und es gab auch keine störenden Einflüsse von Alter, Geschlecht und Bildung ( $F(1, 338) = 1.18$ ,  $p = .278$ ) oder sozialer Erwünschtheit ( $U = 48.000$ ,  $z = 0.71$ ,  $p = .650$ ).

**Tabelle 20:** Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Zustimmung für Hate Intensität mittels sachlichen Kommentars auszudrücken

Gruppen	Gesamt		Weibliche Versuchspersonen		Männliche Versuchspersonen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SK ( $n = 86$ )	2.28	1.53	2.08	1.45	2.73	1.64
OS ( $n = 87$ )	1.90	1.30	2.07	1.39	1.50	0.99
HS ( $n = 84$ )	1.64	1.31	1.70	1.38	1.50	1.14
DS ( $n = 83$ )	1.13	0.44	1.08	0.34	1.21	0.54
$H(3)$	41.68***		28.77***		18.45***	

Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

Gehörte das *Kommentieren von Beiträgen auf öffentlichen Seiten* zu den HN, gaben die Befragten es als wahrscheinlicher ( $M = 2.36$ ,  $SD = 1.70$ ,  $Mdn. = 1$ ) an, auf die Kommentare mit einem sachlichen Kommentar zu antworten im Vergleich zu ungeübteren Befragten ( $M = 1.62$ ,  $SD = 1.16$ ,  $Mdn. = 1$ ). Der Unterschied war überzufällig mit ( $U = 9787.00$ ,  $z = 3.29$ ,  $p < .001$ ,  $r = .18$ ). Wer hingegen im privaten Rahmen *Kommentare von Freunden* beantwortete, unterschied sich nicht von jenen, die dies nicht gewohnheitsmäßig taten ( $U = 15657.00$ ,  $z = 2.01$ ,  $p = .045$ ,  $r = .11$ ).

#### *Sachlicher Widerspruch mittels Kommentares (VI4)*

*Sachlicher Widerspruch* ist im Zusammenhang mit Hate Speech als Gegenrede (*counter speech*) zu verstehen, welche mit sachlichen Gegenargumenten und faktischen Aussagen, Hate Speech entkräften und primäre Adressaten unterstützen kann und sekundären Adressaten Gegenpositionen aufzeigt. Vorliegend unterschied sich die Absicht, sachlichen Widerspruch zu äußern in den vier Gruppen, in denen die *Hate Intensität* variiert wurde ( $H(3) = 11.58$ ,  $p < .01$ ). Im paarweisen Vergleich unterschieden sich nur die Gruppen SK und HS ( $p < .01$ ,  $r = .25$ ), was einem kleineren bis mittelstarken Effekt entspricht. Die Antworten waren nicht durch soziale Erwünschtheit verzerrt ( $H(3) = 7.11$ ,  $p = .069$ ). Die Berücksichtigung der soziodemografischen Merkmale war auf dem Niveau von  $\alpha = 0.01$  in der robusten ANCOVA nicht signifikant ( $F(3, 336) = 3.67$ ,  $p = .013$ ).

**Tabelle 21:** Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention Widerspruch für Hate Intensität mittels sachlichen Kommentars auszudrücken

Gruppen	Gesamt		Weibliche Versuchspersonen		Männliche Versuchspersonen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SK ( <i>n</i> = 86)	1.84	1.39	1.70	1.29	2.15	1.57
OS ( <i>n</i> = 87)	2.38	1.56	2.41	1.56	2.31	1.57
HS ( <i>n</i> = 84)	2.22	1.57	2.60	1.70	2.63	1.86
DS ( <i>n</i> = 83)	2.61	1.74	2.31	1.66	2.09	1.44
<i>H</i> (3)	11.58**		11.82**		1.11, <i>n.s.</i>	

Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

In der Experimentalbedingung *Geschlecht* unterschieden sich die Versuchspersonen, die Stimuli mit Kommentaren von Frauen ( $M = 2.19$ ,  $SD = 1.43$ ,  $Mdn. = 2$ ) rezipiert hatten nicht von jenen, die Kommentare von Männern ( $M = 2.32$ ,  $SD = 1.70$ ,  $Mdn. = 1$ ) rezipiert hatten ( $U = 14151.00$ ,  $z = -0.12$ ,  $p = .906$ ,  $r = .007$ ). Soziale Erwünschtheit übte keinen signifikanten Einfluss auf das Antwortverhalten aus ( $U = 39.500$ ,  $z = 0.47$ ,  $p = .687$ ). Die Quade-ANCOVA konnte keine signifikanten Einflüsse der Kovariaten in dieser Experimentalbedingung nachweisen ( $F(1, 336) = 0.03$ ,  $p = .873$ ).

Auch hinsichtlich der Variation der *Likeanzahl* konnten keine Unterschiede in der Gruppe mit wenigen Likes ( $M = 2.27$ ,  $SD = 1.52$ ,  $Mdn. = 2$ ) verglichen mit vielen Likes ( $M = 2.25$ ,  $SD = 1.64$ ,  $Mdn. = 1$ ) nachgewiesen werden ( $U = 14001.00$ ,  $z = -0.49$ ,  $p = .622$ ,  $r = .03$ ). Soziale Erwünschtheit verzerrte die Antworten nicht ( $U = 37.500$ ,  $z = -0.43$ ,  $p = .711$ ). Die Kontrolle für Kovariate ergab keine störenden Einflüsse der getesteten soziodemografischen Variablen ( $F(1, 336) = 0.21$ ,  $p = .648$ ).

Das *Kommentieren auf öffentlichen Seiten* unterschied sich in einem mittelstarken Maße zwischen jenen, die dies gewohnheitsmäßig taten ( $M = 3.54$ ,  $SD = 2.05$ ,  $Mdn. = 4$ ) und jenen, die diesbezüglich nicht routiniert ( $M = 2.01$ ,  $SD = 1.35$ ,  $Mdn. = 1$ ) waren ( $U = 11299.500$ ,  $z = 5.33$ ,  $p < .001$ ,  $r = .29$ ). Das *Kommentieren bei Freunden* unterschied sich ebenfalls zwischen jenen, die das gewohnheitsmäßig taten ( $M = 2.66$ ,  $SD = 1.81$ ,  $Mdn. = 2$ ) und jenen, die diese Nutzungsweise nur gelegentlich oder seltener ( $M = 1.96$ ,  $SD = 1.32$ ,  $Mdn. = 1$ ) ausübten ( $U = 16860.500$ ,  $z = 3.22$ ,  $p < .01$ ,  $r = .19$ ).

*Beleidigender Widerspruch mittels Kommentares (VI5)*

*Beleidigender Widerspruch* als wahrscheinliche Verhaltensintention kann als Indikator für stimulusinduzierte Aggressivität verstanden werden. In der Experimentalbedingung *Hate Intensität* konnten keine Unterschiede zwischen den Gruppen nachgewiesen werden ( $H(3) = 5.81$ ,  $p = .121$ ). Die eingestufte Wahrscheinlichkeit, den Kommentaren mittels Beleidigungen zu widersprechen, bewegte sich über alle Gruppen am unteren Skalenende, d. h. die Befragten gaben diese Verhaltensintention als wenig wahrscheinlich an, SK ( $M = 1.16$ ,  $SD = 0.63$ ), OS ( $M = 1.22$ ,  $SD = 0.71$ ), HS ( $M = 1.40$ ,  $SD = 0.91$ ), DS ( $M = 1.30$ ,  $SD = 0.92$ ). Soziale Erwünschtheit beeinflusste die Antworten nicht ( $H(3) = 1.71$ ,  $p = .635$ ). Die Kontrolle für Kovariate ergab keine signifikanten Einflüsse soziodemografischer Merkmale ( $F(3, 336) = 1.89$ ,  $p = .132$ ).

Für die Experimentalbedingung *Geschlecht* waren die Befunde ebenfalls nicht signifikant ( $U = 14862.00$ ,  $z = 1.14$ ,  $p = .252$ ,  $r = .06$ ), was sich im geringen Unterschied zwischen Kommentaren von Frauen ( $M = 1.19$ ,  $SD = 0.62$ ,  $Mdn. = 1$ ) im Vergleich zu Kommentaren von Männern ( $M = 1.33$ ,  $SD = 0.92$ ,  $Mdn. = 1$ ) widerspiegelte. Sozial erwünschtes Antwortverhalten übte keinen Einfluss aus ( $U = 45.000$ ,  $z = 1.30$ ,  $p = .391$ ). Die Kontrolle für Kovariate ergab keine signifikanten Störeinflüsse ( $F(1, 338) = 1.27$ ,  $p = .261$ ).

Auch in den Gruppen die unterschiedlich viele Likes rezipiert hatten (Experimentalbedingung *Reaktionsvolumen*), wiederholte sich der Befund ( $U = 30176.500$ ,  $z = -0.32$ ,  $p = .748$ ,  $r = .02$ ). Die Befragten antworteten in beiden Gruppen nahezu identisch (viele Likes:  $M = 1.26$ ,  $SD = 0.77$ ,  $Mdn. = 1$ , wenig Likes:  $M = 1.28$ ,  $SD = 0.84$ ,  $Mdn. = 1$ ). Die Antworten wurden nicht von sozial erwünschten Antworttendenzen verzerrt ( $U = 36.000$ ,  $z = -0.71$ ,  $p = .650$ ). In dieser Experimentalbedingung machte sich kein störender Einfluss von Kovariaten bemerkbar ( $F(1, 338) = 0.03$ ,  $p = .872$ ).

Wer *gewohnheitsmäßig auf öffentlichen Seiten kommentierte*, gab als wahrscheinlicher an ( $M = 1.61$ ,  $SD = 1.22$ ,  $Mdn. = 1$ ), beleidigend zu widersprechen als Befragte, die seltener oder nie kommentierten ( $M = 1.20$ ,  $SD = 0.67$ ,  $Mdn. = 1$ ). Die Unterschiede waren was die Effektstärke betrifft, verhältnismäßig klein, jedoch groß genug, um Signifikanz zu erreichen ( $U = 9115.00$ ,  $z = 2.91$ ,  $p < .01$ ,  $r = .16$ ). Es zeigte sich, dass männliche Versuchspersonen, die routiniert darin waren, auf öffentlichen Seiten zu kommentieren ( $M = 2.06$ ,  $SD = 1.66$ ), es als deutlich wahrscheinlicher angaben, den rezipierten Kommentaren in beleidigender Weise zu widersprechen, im Vergleich zu männlichen Versuchspersonen, die diesbezüglich nicht routiniert waren ( $M = 1.21$ ,  $SD = 0.78$ ). Der Unterschied hatte eine mittlere Effektstärke und war signifikant ( $U = 1077.00$ ,  $z = 3.28$ ,  $p < .001$ ,  $r = .31$ ). Keine Unterschiede gab es in den Gruppen

die *gewohnheitsmäßig im eigenen Netzwerk kommentierten* ( $U = 14961.00$ ,  $z = 1.50$ ,  $p = .134$ ,  $r = .08$ ), mit ( $M = 1.37$ ,  $SD = 0.95$ ,  $Mdn. = 1$ ) für routinierte Kommentierende unter den Befragten und ( $M = 1.20$ ,  $SD = 0.65$ ,  $Mdn. = 1$ ) für jene, die angaben seltener mit ihrem Netzwerk per Kommentar zu interagieren.

#### *Beleidigende Unterstützung mittels Kommentares (VI6)*

Angenommen wird, dass eine *beleidigende Unterstützung der Kommentierenden* als latenter Indikator für stattgefundene Persuasion auf der Basis von rezipiertem Hate Speech dient. In den vier Gruppen, in denen die *Hate Intensität* variiert wurde, befanden sich die Angaben der Befragten am unteren Skalenende, d. h. also, dass sie es als unwahrscheinlich ansahen, selbst verbal aggressiv aufzutreten (SK:  $M = 1.10$ ,  $SD = 0.44$ , OS:  $M = 1.09$ ,  $SD = 0.33$ , HS:  $M = 1.12$ ,  $SD = 0.36$  und DS:  $M = 1.02$ ,  $SD = 0.34$ ). Die Unterschiede waren so gering, dass sie folglich keine Signifikanz erreichten ( $H(3) = 4.55$ ,  $p = .208$ ). Soziale Erwünschtheit beeinflusste die Antworten nicht ( $H(3) = 5.33$ ,  $p = .149$ ). Die Kontrolle für die soziodemografischen Kovariaten ergab eine nicht-signifikante Teststatistik der robusten Quade-ANCOVA ( $F(3, 336) = 1.65$ ,  $p = .177$ ).

Effekte in der Experimentalbedingung *Geschlecht der Kommentierenden* konnten nicht nachgewiesen werden. Kommentare von Männern ( $M = 1.05$ ,  $SD = 0.25$ ,  $Mdn. = 1$ ) riefen im Vergleich zu Kommentaren von Frauen ( $M = 1.13$ ,  $SD = 0.42$ ,  $Mdn. = 1$ ) keine signifikanten Unterschiede bezüglich der Verhaltensabsicht, die Kommentierenden mittels Beleidigungen zu unterstützen hervor ( $U = 13493.500$ ,  $z = -1.89$ ,  $p = .058$ ,  $r = -.10$ ). Soziale Erwünschtheit beeinflusste das Antwortverhalten der Befragten nicht signifikant ( $U = 37.500$ ,  $z = 0.60$ ,  $p = .823$ ).

Auch das *Reaktionsvolumen* bewirkte keine Unterschiede (viele Likes:  $M = 1.08$ ,  $SD = 0.37$ ,  $Mdn. = 1$ ; wenige Likes:  $M = 1.09$ ,  $SD = 0.30$ ,  $Mdn. = 1$ ) bezüglich der geäußerten Wahrscheinlichkeit als beleidigende Unterstützende der Kommentierenden in Erscheinung zu treten ( $U = 14166.500$ ,  $z = -0.63$ ,  $p = .531$ ,  $r = -.03$ ). Verzerrte Antworten aufgrund sozialer Erwünschtheit konnten nicht belegt werden ( $U = 36.000$ ,  $z = -1.31$ ,  $p = .650$ ). Die Quade-ANCOVA zur Kontrolle soziodemografischer Merkmale ergab keine Signifikanz ( $F(1, 338) = 0.28$ ,  $p = .597$ ).

Weder Befragte, die *gewohnheitsmäßig Kommentare auf öffentlichen Facebook-Seiten* verfassten, ( $M = 1.11$ ,  $SD = 0.37$ ,  $Mdn. = 1$ ) noch jene, die dies seltener oder nie ( $M = 1.08$ ,  $SD = 0.33$ ,  $Mdn. = 1$ ) taten ( $U = 8131.500$ ,  $z = -0.60$ ,  $p = .547$ ,  $r = -.03$ ), beabsichtigten die

Kommentierenden mittels beleidigenden Kommentares zu unterstützen. Die Kontrolle für Kovariate war ebenfalls nicht-signifikant ( $F(1, 338) = 0.43, p = .512$ ). Hinsichtlich des habitualisierten Kommentierens bei Freunden konnten keine Unterschiede festgestellt werden ( $U = 14269.500, z = -0.27, p = .787, r = -.01$ ). Die Quade-ANCOVA zur Kontrolle von Einflüssen der soziodemografischen Variablen war nicht signifikant ( $F(1, 338) = 0.17, p = .677$ ).

#### *Melden der Kommentare (VI7)*

Nehmen Nutzer:innen in sozialen Netzwerken Inhalte wahr, die sie als unangemessen, respektlos oder gesetzeswidrig empfinden, haben sie die Möglichkeit diese zu melden, was dann netzwerksintern geprüft wird. In der Folge können solche Inhalte gelöscht oder die Nutzenden bei Zuwiderhandeln von der Nutzung gesperrt werden. Diese Handhabe wurde mit dem NetzDG in eine juristische Handhabe zur Unterbindung von Hate Speech in sozialen Netzwerken umgesetzt. Nehmen die Versuchspersonen die Möglichkeit zur Beschwerde wahr? Dafür liefern die Befunde einige Anhaltspunkte.

So hielten es die Befragten mit steigender *Hate Intensität* für wahrscheinlicher, die rezipierten Kommentare zu melden (vgl. Tabelle 22) und dieser Befund erreichte denn auch hohe Signifikanz ( $H(3) = 39.66, p < .001$ ). Paarweise Vergleiche zeigten, dass schwache bis mittelstarke Effekte zwischen SK und HS ( $p < .01, r = -.26$ ), zwischen SK und DS ( $p < .001, r = -.48$ ), sowie zwischen OS und DS ( $p < .01, r = -.27$ ) auftraten. Die Antworten waren auf dem Signifikanzniveau  $\alpha = .01$  nicht durch sozial erwünschtes Antwortverhalten verzerrt ( $H(3) = 8.01, p = .046$ ). Die Quade-ANCOVA war signifikant ( $F(3, 336) = 14.75, p < .001$ ) und konnte signifikante paarweise Unterschiede bei weiblichen Versuchspersonen in den Gruppen SK und DS nachweisen ( $p < .001, r = -.43$ ). Auch der Bildungsgrad erwies sich als bedeutsam: Hoch gebildete Versuchspersonen unterschieden sich in den Gruppen SK und HS ( $p < .01, r = -.25$ ), SK und DS ( $p < .001, r = -.45$ ) sowie OS und DS ( $p < .01, r = -.26$ ) bezüglich der Absicht, die rezipierten Kommentare zu melden. Die Kontrolle für Einflüsse des Alters konnte keine signifikanten Unterschiede in den *Hate Intensitäts*-Gruppen nachweisen.

Die Befragten beabsichtigten mit größerer Wahrscheinlichkeit Kommentare von Männern zu melden ( $M = 2.08, SD = 1.63, Mdn. = 1$ ) als Kommentare von Frauen ( $M = 1.68, SD = 1.18, Mdn. = 1$ ), allerdings war der Unterschied in der Experimentalbedingung *Geschlecht* nicht signifikant ( $U = 15615.00, z = 1.75, p = .080, r = .09$ ). Soziale Erwünschtheit verzerrte das Antwortverhalten nicht ( $U = 33.000, z = -0.22, p = .893$ ). Keinen signifikanten Befund lieferte die Quade-ANCOVA ( $F(1, 338) = 3.10, p = .079$ ).

**Tabelle 22:** Globale Gruppenunterschiede und Unterschiede zwischen männlichen und weiblichen Versuchspersonen hinsichtlich der Verhaltensintention des Meldens in der Bedingung Hate Intensität

Gruppen	Gesamt		Weibliche Versuchspersonen		Männliche Versuchspersonen	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
SK ( <i>n</i> = 86)	1.31	0.87	1.28	0.83	1.38	0.98
OS ( <i>n</i> = 87)	1.80	1.36	1.80	1.36	1.81	1.39
HS ( <i>n</i> = 84)	1.83	1.26	1.87	1.38	1.75	0.90
DS ( <i>n</i> = 83)	2.70	1.87	3.02	2.06	2.24	1.46
<i>H</i> (3)	39.66***		30.90***		9.41, <i>n.s.</i>	

Anmerkungen: SK = Sachliche Kommentare, OS = Offensive Speech, HS = Hate Speech, DS = Dangerous Speech.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

Auch das *Reaktionsvolumen* beeinflusste nicht die selbsteingeschätzte Wahrscheinlichkeit zur Beschwerde ( $U = 14153.00$ ,  $z = -0.34$ ,  $p = .735$ ,  $r = .02$ ). Die Absicht die Kommentare zu melden, wurde in der Gruppe, in der die Versuchspersonen eine höhere Likeanzahl rezipierten, nur als geringfügig wahrscheinlicher angesehen ( $M = 1.94$ ,  $SD = 1.56$ ,  $Mdn. = 1$ ), verglichen mit der Gruppe, in der die Likeanzahl niedriger war ( $M = 1.87$ ,  $SD = 1.36$ ,  $Mdn. = 1$ ). Verzerrungen aufgrund von sozialer Erwünschtheit im Antwortverhalten konnten nicht nachgewiesen werden ( $U = 39.500$ ,  $z = -0.25$ ,  $p = .837$ ). Die Kovariatenkontrolle ergab keine signifikanten Unterschiede zwischen den Gruppen ( $F(1, 338) = 0.41$ ,  $p = .525$ ).

Befragte, die angaben regelmäßig unangemessene Inhalte bei Facebook zu melden, beabsichtigten dies auch in der Experimentalsituation zu tun ( $M = 2.94$ ,  $SD = 1.94$ ,  $Mdn. = 2.50$ ). Wer jedoch nicht regelmäßig von der Möglichkeit zur Beschwerde Gebrauch machte, gedachte dies weniger wahrscheinlich zu tun ( $M = 1.68$ ,  $SD = 1.23$ ,  $Mdn. = 1$ ). Der Unterschied hatte zwar eine schwache bis mittelstarke Effektgröße ( $r = .27$ ), war aber hochsignifikant ( $U = 11447.00$ ,  $z = 5.07$ ,  $p < .001$ ).

### *Ignorieren der Kommentare (VI8)*

Ein Großteil der Nutzer:innen in SNS nutzt diese als Lurker: Sie lesen Beiträge und Kommentare von anderen, tragen aber kaum zum Generieren neuer Inhalte bei (Han et al., 2014; Sun et al., 2014). Ob die Befragten in der Stichprobe auch nach Rezeption von Hate Speech dazu neigen, zu schweigen, wurde anhand des Items „Ich lese die Kommentare, ohne etwas zu tun“ abgefragt. Die Befragten gaben in allen vier Gruppen der Experimentalbedingung *Hate Intensität* an, die Kommentare in überwiegendem Maße zu ignorieren, und zwar SK ( $M = 5.03$ ,  $SD = 1.34$ ), OS ( $M = 4.78$ ,  $SD = 1.47$ ), HS ( $M = 4.76$ ,  $SD = 1.54$ ) und DS ( $M = 4.58$ ,

$SD = 1.67$ ). Die geringen Unterschiede waren folglich auch nicht signifikant ( $H(3) = 3.59$ ,  $p = .310$ ). Soziale Erwünschtheit verzerrte die Antworten nicht ( $H(3) = 6.31$ ,  $p = .098$ ). Die Kontrolle für soziodemografische Merkmale lieferte ein nicht-signifikantes Ergebnis per Quade-ANCOVA ( $F(3, 336) = 0.86$ ,  $p = .463$ ).

Die Absicht Kommentare von Männern ( $M = 4.66$ ,  $SD = 1.56$ ,  $Mdn. = 5$ ) verglichen mit Kommentaren von Frauen ( $M = 4.96$ ,  $SD = 1.44$ ,  $Mdn. = 6$ ) zu ignorieren, war auf dem Niveau  $\alpha = 0.01$  für die Experimentalbedingung *Geschlecht* nicht signifikant ( $U = 12426.00$ ,  $z = -2.16$ ,  $p = .031$ ,  $r = -.12$ ). Soziale Erwünschtheit beeinflusste das Antwortverhalten nicht ( $U = 22.000$ ,  $z = -1.25$ ,  $p = .257$ ). Für die Kontrolle der Kovariate war der Unterschied auf dem Niveau  $\alpha = .01$  ebenfalls nicht signifikant ( $F(1, 338) = 4.420$ ,  $p = .036$ ).

Geringe Unterschiede zwischen den Gruppen und damit ein nicht-signifikanter Gruppenvergleich wurde in der Experimentalbedingung *Reaktionsvolumen* nachgewiesen ( $U = 15419.00$ ,  $z = 1.18$ ,  $p = .239$ ,  $r = .06$ ). In der Gruppe, in der die Teilnehmer:innen viele Likes rezipierten ( $M = 4.90$ ,  $SD = 1.43$ ,  $Mdn. = 5.50$ ), wurde das Ignorieren der Kommentare als etwas wahrscheinlicher bewertet als in der Gruppe mit wenigen Likes ( $M = 4.67$ ,  $SD = 1.60$ ,  $Mdn. = 5$ ). Soziale Erwünschtheit konnte nicht belegt werden ( $U = 32.000$ ,  $z = -0.88$ ,  $p = .432$ ). Die Kontrolle für Kovariate war nicht signifikant ( $F(1, 338) = 0.68$ ,  $p = .409$ ).

Der Vergleich zwischen jenen, die zum einen das *gewohnheitsmäßige Lurken* betrieben, im Sinne eines passiven Rezipierens „Ich lese Beiträge und Kommentare“ ( $M = 4.74$ ,  $SD = 1.53$ ,  $Mdn. = 5$ ) und zum anderen jenen, die angaben, dies nicht regelmäßig zu tun ( $M = 4.94$ ,  $SD = 1.45$ ,  $Mdn. = 6$ ), ergab keine signifikanten Unterschiede ( $U = 10332.00$ ,  $z = -1.20$ ,  $p = .263$ ,  $r = .06$ ). Befragte, die regelmäßig ihren Newsfeed durchscrollten und die *meisten Beiträge ignorierten* ( $M = 4.94$ ,  $SD = 1.40$ ,  $Mdn. = 6$ ) unterschieden sich ebenfalls nicht signifikant ( $U = 15020.00$ ,  $z = 1.42$ ,  $p = .155$ ,  $r = .08$ ) in ihrem Antwortverhalten verglichen mit Befragten, die seltener so agierten ( $M = 4.67$ ,  $SD = 1.59$ ,  $Mdn. = 5$ ). Bei Befragten, die *gewohnheitsmäßig Kommentare ignorierten* ( $M = 4.98$ ,  $SD = 1.39$ ,  $Mdn. = 6$ ), war der Unterschied zu jenen, die dies seltener taten ( $M = 4.55$ ,  $SD = 1.63$ ,  $Mdn. = 5$ ) ebenfalls nicht signifikant ( $U = 15781.00$ ,  $z = 2.30$ ,  $p = .021$ ,  $r = .12$ ). Die Quade-ANCOVA, in der der Einfluss der soziodemografischen Kovariate kontrolliert wurde, ergab auf dem Niveau  $\alpha = .01$  keine Signifikanz ( $F(1, 333) = 5.31$ ,  $p = .022$ ).

Als zustimmende Verhaltensintentionen konnten aus der MDS die Variablen VI1 („Like“), VI3 („sachliche Unterstützung“) und VI6 („beleidigende Unterstützung“) abgeleitet werden. Die Beziehung zwischen ambivalent sexistischen Einstellungen und der geäußerten

Wahrscheinlichkeit zustimmender Verhaltensintentionen in Abhängigkeit der Versuchsbedingungen, wie in Hypothese H8 postuliert, konnte nicht mittels Moderationsanalyse überprüft werden, da die notwendige Voraussetzung der Linearität nach visueller Inspektion als grob verletzt betrachtet wurde. Ausgewichen wurde auf nicht-parametrische Korrelationen mit dem Korrelationskoeffizienten Kendall-Tau-b. Aufgrund der problematischen Fehlerkumulierung durch Mehrfachtestung wurde der Signifikanzwert  $\alpha$  auf das Niveau .01 gesenkt. Die Korrelationen zwischen Ambivalentem Sexismus und den drei VI-Items ergaben einen kleinen positiven Zusammenhang für VI1 („Like“) ( $\text{tau-b} = .13, p < .01$ ) und für VI3 („sachliche Unterstützung“) ( $\text{tau-b} = .18, p < .001$ ), jedoch keinen auf dem konservativen Signifikanzniveau für VI6 („beleidigende Unterstützung“) ( $\text{tau-b} = .10, p = .042$ ). Ob die Zusammenhänge in Abhängigkeit der drei experimentellen Treatments (Hate Intensität, Reaktionsvolumen, Geschlecht der Kommentierenden) stärker werden, wird im Folgenden dargestellt.

#### *Zustimmende Verhaltensintentionen unter der Bedingung der Hate Intensität*

Es bestand ein mittlerer positiver Zusammenhang zwischen ambivalent sexistischen Einstellungen und der Wahrscheinlichkeit mit einem Like auf die rezipierten Stimuli zu reagieren in den Gruppen OS ( $\text{tau-b} = .24, p < .01$ ), HS ( $\text{tau-b} = .45, p < .001$ ) und DS ( $\text{tau-b} = .28, p < .01$ ), d. h. mit jeder weiteren Einheit auf der Skala Ambivalenter Sexismus stieg die artikulierte Wahrscheinlichkeit Hate Kommentare mit einem Like zu versehen, um den Wert der Koeffizienten. Für sachliche Kommentare war der Befund negativ, jedoch nicht auf dem geforderten Niveau signifikant ( $\text{tau-b} = -.18, p = .045$ ).

Signifikante Befunde für den Zusammenhang zwischen Ambivalentem Sexismus und der Intention, sachlich kommentierend zu reagieren, konnten belegt werden, nämlich in den Gruppen HS ( $\text{tau-b} = .37, p < .001$ ) und DS ( $\text{tau-b} = .30, p < .01$ ), nicht jedoch für SK und OS. Für die Verhaltensabsicht mit einem beleidigenden Kommentar zu antworten, konnte die Beziehung zu vorhandenen ambivalent sexistischen Einstellungen nicht belegt werden.

#### *Zustimmende Verhaltensintentionen unter der Bedingung des Geschlechts der Kommentierenden*

In den Gruppen, in denen Kommentare von Frauen rezipiert wurden, konnte ein mittlerer positiver Zusammenhang zwischen ambivalent sexistischen Einstellungen und der artikulierten Wahrscheinlichkeit einer Like-Reaktion aufgezeigt werden ( $\text{tau-b} = .23, p < .01$ ). Für Kommentare von männlichen Versuchspersonen konnte dieser Zusammenhang nicht nachgewiesen werden ( $\text{tau-b} = .07, p = .305$ ).

Analog war der Befund für VI3 („sachliche Unterstützung“): Nach Rezeption von Kommentaren von Frauen wurde die Wahrscheinlichkeit für dieses Item als höher angegeben, je höher der Score für ambivalent sexistische Einstellungen war ( $\tau\text{-}b = .28, p < .001$ ). In den Gruppen, in denen Kommentare von Männern rezipiert wurden, konnte der Zusammenhang nicht nachgewiesen werden ( $\tau\text{-}b = .10, p = .108$ ).

Der Zusammenhang zwischen Ambivalentem Sexismus und beleidigender Unterstützung war weder nach Rezeption von Frauen-Kommentaren auf dem Niveau  $\alpha = .01$  ( $\tau\text{-}b = .18, p = .016$ ), noch nach Rezeption von Männer-Kommentaren signifikant ( $\tau\text{-}b = .03, p = .650$ ).

#### *Zustimmende Verhaltensintentionen unter der Bedingung des Reaktionsvolumens*

Auf dem Niveau  $\alpha = .01$  konnte weder nach Rezeption vieler Likes ( $\tau\text{-}b = .13, p = .037$ ) noch nach Rezeption weniger Likes ( $\tau\text{-}b = .14, p = .047$ ) der Zusammenhang zwischen Ambivalentem Sexismus und der artikulierten Wahrscheinlichkeit selbst mit einem Like zu reagieren, belegt werden.

In den Gruppen, in denen viele Likes rezipiert wurden, war der postulierte Zusammenhang aus Hypothese H8 zwischen Ambivalentem Sexismus und sachlicher Unterstützung mittelstark und positiv ( $\tau\text{-}b = .19, p < .01$ ). Die Beziehung war in den Gruppen mit wenigen Likes auf dem Niveau  $\alpha = .01$  nicht signifikant ( $\tau\text{-}b = .17, p = .015$ ).

Die Beziehung zwischen der Wahrscheinlichkeit beleidigender Unterstützung und Ambivalentem Sexismus war weder nach Rezeption vieler Likes ( $\tau\text{-}b = .13, p = .068$ ), noch nach Rezeption weniger Likes ( $\tau\text{-}b = .08, p = .298$ ) signifikant.

### **6.1.3 Einordnende Zusammenfassung der Befunde**

Das Elaboration Likelihood Model (ELM) postuliert umfassende Rezeptionswirkungen von persuasiven Botschaften, welche sich als positive Einstellungen gegenüber dem Einstellungsobjekt, als Übernahme der Argumentation, erhöhte wahrgenommene Relevanz darstellt und leitet daraus Verhaltensprognosen ab (Petty & Cacioppo, 1986). Als besonders wirkungsreich wird die vertiefte kognitive Verarbeitung angesehen, während die periphere Verarbeitung nur kurzfristiger Natur ist. Unter Elaboration wird im konzeptuellen Rahmen des ELM die Verarbeitung der Persuasionsbotschaft verstanden, und sie wird mittels inhaltsanalytischer Untersuchung der Rezeptionsgedanken eingehend geprüft. Abgeleitete Elaborations-

kategorien sind die absolute Gedankenzahl, deren Valenz und die inhaltliche Ausgangsbasis der Rezeptionsgedanken.

Hinsichtlich der Gedankenvalenz ergänzen sich die drei Forschungsstränge (Hate Speech-, Sexismus- und mit einigen Abstrichen die ELM-Forschung) gut, um die empirischen Ergebnisse einzuordnen. Die Befunde stehen insofern im Einklang zur Hate Speech-Forschung, als dass sowohl männliche Versuchspersonen als sekundäre Adressaten, als auch weibliche Befragte als prinzipiell primäre Adressatinnen mehrheitlich ablehnende Affekte und Kognitionen zum Ausdruck brachten. Als signifikanter Prädiktor erwies sich die *Hate Intensität*. Während sachliche Kommentare den geringsten Anteil an negativen Gedanken verursachten, nahm deren Anteil in den Gruppen Offensive Speech (OS), Hate Speech (HS) und Dangerous Speech (DS) zu. Die Einordnung des Befundes aus Perspektive der ELM-Forschung ist schwieriger. Im Hate Speech-Kontext ist die Begründung aus dem ELM heraus normativ sicherlich strittig, denn hier wird Persuasionserfolg mit positiven Gedanken zur Botschaft in Verbindung gebracht. Bei stattgefundenen Persuasionseffekten von Hate Speech würden die Rezeptionsgedanken eine objektive Verarbeitung erkennen lassen, da sie positiv auf die Botschaft, nicht jedoch auf das Einstellungsobjekt (hier: Frauen als primäre Adressatinnen) ausgerichtet wären. Entsprechend muss der Befund zur negativen Gedankenvalenz daher als verzerrte Verarbeitung von Hate Speech interpretiert werden, da die Reaktion auf die Botschaft und die Kommentierenden negativ ausfiel.

Das *Reaktionsvolumen* als Experimentalbedingung lieferte Anhaltspunkte für das Auftreten eines Mere-Exposure-Effekts: Zwar wurde die Likeanzahl als experimentelles Treatment von den Befragten nicht wahrgenommen, wurde jedoch aus methodischen Gründen in den Analysen berücksichtigt. Da der Anteil an negativen Gedanken in der Bedingung „viele Likes“ höher war als in der Bedingung „wenige Likes“ kann hier vorsichtig von sozialen Einflüssen gesprochen werden, also sogenannten „Bandwaggoning“-Effekten, die eine Orientierung an wahrgenommenen Mehrheiten bedingen und selbst bei unerschwelliger Wahrnehmung auftreten (H.-S. Kim & Sundar, 2014).

Keine Auswirkung auf den Anteil an negativen Gedanken hatte hingegen das *Geschlecht* als Experimentalbedingung, das Eigenschaften der Kommentierenden nicht salient hervorzuheben vermochte, trotz Wahrnehmung des Treatments durch die Versuchspersonen. Aus Perspektive der Sexismusforschung kann abgeleitet werden, dass das Stimulusmaterial als Verstoß gegen sowohl männliche als auch weibliche Globalstereotype hinsichtlich der Wärme-Kompetenz-Dimensionen wahrgenommen wurde, weswegen keine Unterschiede zu Tage traten. Diese Schlussfolgerung bedarf einer tragfähigen empirischen Grundlage, da

diesbezüglich keine Daten erhoben wurden. Für eine Interpretation aus ELM-Perspektive spricht, dass die Geschlechterkategorie zwar wahrgenommen, aber nur peripher verarbeitet wurde und daher die Gedankenvalenz nicht beeinflusst hat. Keine Erklärungskraft hatten soziodemografische Merkmale der Versuchspersonen und ambivalent sexistische Einstellungen. Des Weiteren konnten keine Einflüsse sozial erwünschten Antwortverhaltens auf die relativen Negativgedanken nachgewiesen werden.

Die Motivation und Fähigkeit zur Elaboration, operationalisiert als Kognitionsbedürfnis und Selbsteinschätzung des Vorwissens, beeinflussten zwar nicht die Beziehung zwischen *Geschlecht der Kommentierenden* und der Gedankenvalenz, wohl aber die Beziehung zwischen *Hate Intensität* und Gedankenvalenz. Mit höherem Kognitionsbedürfnis stieg der Anteil an negativen Gedanken in den OS- und HS-Gruppen an. Für DS gab es keine signifikanten Valenzänderungen, da diese über alle Kognitionsniveaus bereits einen hohen Anteil an negativen Rezeptionsgedanken ohne nennenswerte Steigerungen verursachten. Interessant war, dass sachliche Kommentare bei geringem Kognitionsbedürfnis deutlich mehr negative Gedanken verursachten als bei einem überdurchschnittlichen Kognitionsbedürfnis. Dies spricht für unterschiedliche Verarbeitungstiefen: Bei geringer Motivation wurden die Kommentare in der Gruppe SK oberflächlich und auf schematische Weise verarbeitet, während die Elaborationswahrscheinlichkeit deutlich zunahm, je ausgeprägter das Kognitionsbedürfnis angegeben wurde. Einen umgekehrten Effekt hatte das Vorwissen, denn bei konstantem Kognitionsbedürfnis, nahm der Anteil an negativen Gedanken in allen Gruppen zu, je stärker sich die Befragten als „kompetent, an einer Diskussion zum Thema ‚Karrierechancen von Frauen‘ teilzunehmen“ einschätzten. Vorwissen wird laut ELM als Fähigkeit zur Elaboration verstanden, also würde höheres Vorwissen eine vertiefte Verarbeitung bedingen. Erwartbar wäre, dass die Rezeptionsgedanken mit steigendem Vorwissen in der SK-Gruppe deutlich weniger negativ sind, was jedoch nicht bestätigt werden konnte. Als Erklärung kommt Reaktanz der Befragten in Betracht: Die Argumentation in den präsentierten Kommentaren und/oder die Studienthematik selbst könnten die Verarbeitung verzerrt haben, und zwar dahingehend, dass die sachlichen Argumentationen in den Kommentaren negativ wahrgenommen wurden. Alternativ könnten die Kommentare in der SK-Gruppe von den Befragten als strittig wahrgenommen worden sein und aus diesem Grund negative Rezeptionsgedanken ausgelöst haben. Das ist insofern weniger plausibel, als dass sonst deutliche Unterschiede hinsichtlich der Rezeptionswirkungen zwischen SK und OS, HS oder DS bestanden. Ob und wie das selbsteingeschätzte Vorwissen die Gedanken inhaltlich beeinflusste, wird in dem Abschnitt zu den gedanklichen Ursprüngen näher beleuchtet.

Der Zusammenhang zwischen Likeanzahl und dem Anteil an negativen Rezeptiongedanken wurde von dem Kognitionsbedürfnis (also der Motivation) der Befragten, nicht jedoch von deren Vorwissen (also der Fähigkeit zur Elaboration) beeinflusst. Während eine hohe Likezahl mit höherem Kognitionsbedürfnis eine Abnahme der negativen Rezeptionsgedanken bewirkte, steigerten geringe Likezahlen die negative Valenz. Der Befund deutet auf eine periphere Verarbeitung der Likezahl hin, was konsistent zum Ergebnis des Treatmentchecks und des mutmaßlichen Mere Exposure-Effekts ist.

Die Gedankenzahl wurde als korrelativer Indikator für eine vertiefte Informationsverarbeitung gewertet und gemessen daran, deuteten die Befunde diesbezüglich auf eine periphere bis mittelstarke Verarbeitung hin, da durchschnittlich etwas mehr als vier von bis zu zehn Gedanken niedergeschrieben wurden. Zwar beeinflussten die experimentellen Variationen von *Geschlecht* und *Likeanzahl* die Anzahl der Rezeptionsgedanken nicht, wohl aber die *Hate Intensität*. Je aggressiver die Kommentare waren, umso höher stieg die Gedankenzahl verglichen mit sachlichen Kommentaren, die als Referenzgruppe dienten. Andere Erklärfaktoren erwiesen sich als bedeutungslos, darunter soziodemografische Faktoren (Alter, Geschlecht, Bildung), ambivalent sexistische Einstellungen, die Motivation und Fähigkeit zur Elaboration (Kognitionsbedürfnis, Vorwissen, Persönliches Involvement) sowie sozial erwünschte Antworttendenzen. Darüber hinaus wurde der Zusammenhang zwischen den Experimentalbedingungen und der Gedankenzahl nicht von den Elaborationsvariablen beeinflusst.

Vertiefende Erkenntnisse zur Verarbeitungsweise der experimentellen Stimuli wie auch zu Einflussfaktoren auf die Verarbeitung konnten anhand der inhaltlichen Analyse der Rezeptionsgedanken gewonnen werden. Die wenigsten Gedanken – gerade einmal knapp jeder zehnte – wurden der Restkategorie „Sonstige Gedanken“ zugeordnet. Hier konnte nur das *Reaktionsvolumen* die Nennung dieser Gedanken erklären. Die Versuchsbedingung „viele Likes“ bewirkte eine geringere Gedankenzahl der Sonstige-Kategorie, verglichen mit der Versuchsbedingung „wenige Likes“. Ein Großteil der Gedanken – mehr als die Hälfte – hatte als Ausgangsbasis die kognitive Verarbeitung des Stimulusmaterials und wurde mehrheitlich von der Variation der *Hate Intensität* erklärt. Mit der Hate Intensität stiegen die Gedankenzahlen dieser Kategorie bis zur doppelten Anzahl verglichen mit der SK-Gruppe an. Darüber hinaus steigerten auch das Persönliche Involvement und ambivalent sexistische Einstellungen die kognitive Auseinandersetzung mit den Stimuli, wobei sowohl Signifikanz als auch Effekt gering waren und es sich hierbei auch um Zufallsbefunde handeln könnte. Vorwissen als Moderator beeinflusste die Beziehung zwischen OS bzw. DS und der Anzahl an Gedanken aus der Kategorie „kognitive Auseinandersetzung mit dem Stimulus“ signifikant. Damit kann der

Befund zum Interaktionseffekt von Vorwissen auf die Valenz der Gedanken erklärt werden: Folglich thematisierten Befragte mit höherem Vorwissen die Stimuli und kommentierten sie negativ.

Gedanken aus der Kategorie „affektive Verarbeitung des Stimulusmaterials“ ließen sich nicht durch Merkmale der Versuchsbedingungen, sondern durch Merkmale der Versuchspersonen erklären. Dieser Befund ist insofern überraschend, als dass höhere Empörung- und Entrüstungswerte vor allem in den OS-, HS und vor allem DS-Gruppen verglichen mit der SK-Gruppe erwartet wurden. Das Kognitionsbedürfnis senkte die Wahrscheinlichkeit Rezeptionsgedanken der affektiven Kategorie zu äußern, d. h. nicht Empörung, sondern der Wissens- und Erfahrungsabgleich waren für Befragte mit höherem Kognitionsbedürfnis Ausgangsbasis der Rezeptionsgedanken (s. u.). Für ambivalent sexistische oder männliche Versuchspersonen war dagegen die Wahrscheinlichkeit erhöht, affektiv auf das Stimulusmaterial zu reagieren. Dieser Befund bedurfte weiterer Analysen zur Einordnung. Waren die Befragten motiviert, die Stimuli zu verarbeiten, sprich mit höherem Persönlichen Involvement, stieg die Anzahl der affektiv gefärbten Rezeptionsgedanken in der DS-Bedingung. Generell gelten Emotionen als mentale Abkürzungen, die nicht durchdacht erfolgen, weswegen dieser Befund trotz vorhandener Motivation aus ELM-Perspektive im Elaborationskontinuum als periphere Verarbeitung zu bewerten ist.

Der *Vorwissens- und Erfahrungsabgleich* als dritte inhaltliche Kategorie für die Ausgangsbasis der Gedanken diente als weiterer Beleg für eine stattgefundene zentrale Verarbeitung des Stimulusmaterials. Sichtbar wurde dies bei Versuchspersonen mit höherem Kognitionsbedürfnis, denn dort flossen verstärkt eigene Überlegungen und neue Aspekte in die Rezeptionsgedanken ein. Auch das *Geschlecht der Kommentierenden* als Experimentalbedingung erhöhte leicht signifikant die Anzahl von Gedanken dieser Kategorie. Im Vergleich zu Kommentaren von Frauen bewirkten Kommentare von Männern ein Plus an Gedanken. Da es sich um kognitiv anspruchsvolle(re) Äußerungen handelt, kann für diese Kategorie durchaus von einer zentralen Verarbeitung der Geschlechterkategorie ausgegangen werden. Die *Hate Intensität*-Gruppen unterschieden sich nicht hinsichtlich der Gedankenzahl dieser Kategorie, abgesehen von der HS-Gruppe, welche leicht signifikant weniger Gedanken hervorrief als die Referenzgruppe SK. Ob der Befund mit Reaktanz der Versuchspersonen zusammenhängt oder ob es sich dabei um einen Zufallsbefund handelt (zumal OS und DS als Vergleichsgruppe nicht signifikant waren), konnte abschließend nicht mit Sicherheit geklärt werden. Anhaltspunkte für Reaktanzeffekte liefert die Bildungsvariable: Die Wahrscheinlichkeit Vorwissen und eigene Überlegungen in die Rezeptionsgedanken einfließen zu lassen, war für

Befragte mit höherer Bildung signifikant geringer. Darüber hinaus sank die Wahrscheinlichkeit in den Hate Intensität-Gruppen kognitiv elaboriert zu antworten, je höher das Kognitionsbedürfnis der Befragten war, was ebenfalls als Reaktanzeffekt gewertet werden muss.

Die Relevanzwahrnehmung als Rezeptionswirkung kann ebenfalls als Gradmesser für stattgefundenen Persuasion gewertet werden. Innerhalb des ELM-Theorierahmens wird davon ausgegangen, dass einstellungskongruente Kommentare die Relevanzwahrnehmung erhöhen. Allerdings ließen sich diese Annahmen nicht uneingeschränkt auf sexistischen Hate Speech zu Frauenkarrieren in der IT-Branche übertragen, da diesbezüglich eine defensive Verarbeitung erwartet wurde, bei der die persönliche Relevanzwahrnehmung abgestritten wird. Geprüft wurde, ob ambivalent sexistische Einstellungen die Relevanzwahrnehmung der rezipierten Kommentare beeinflussen.

Der bedeutendste Einfluss auf die wahrgenommene persönliche Relevanz ging von der *Hate Intensität* als Botschaftsmerkmal aus. Zunächst ging eine deutlich höhere Relevanz von sachlichen Kommentaren im Vergleich zu OS und HS-Komentaren aus. Unter der Bedingung ambivalent sexistischer Einstellungen stieg die Relevanz in den OS- und HS-Bedingungen signifikant, was darauf hindeutete, dass die rezipierten Stimuli als einstellungskongruente Botschaften wahrgenommen wurden. Das Stimulusmaterial in der SK-Gruppe wurde entsprechend mit höherem ambivalent sexistischen Einstellungsscore als weniger relevant registriert, da es als widersprüchlich zu bestehenden Einstellungen wahrgenommen wurde. In der DS-Bedingung war der Einfluss ambivalent sexistischer Einstellungen auf die Relevanzwahrnehmung nicht signifikant, was dahingehend als Abwehrmechanismus aufgrund der Aggressivität der Kommentare verstanden werden konnte.

Das *Geschlecht* der Kommentierenden als Botschaftsmerkmal beeinflusste die Relevanzwahrnehmung nicht, aber das *Reaktionsvolumen* konnte als hochsignifikanter Prädiktor identifiziert werden, was ein weiteres Mal als Beleg für den stattgefundenen Mere Exposure-Effekt gewertet werden konnte. Ambivalenter Sexismus moderierte weder die Beziehung zwischen Geschlecht noch zwischen Likeanzahl und Relevanzwahrnehmung. Positive Einflüsse gingen von jüngeren Befragten, Vorwissen und Bildung aus. Das Geschlecht der Teilnehmenden wirkte sich ebensowenig auf die Wahrnehmung persönlicher Relevanz aus, wie das Kognitionsbedürfnis oder das Persönliche Involvement. Auch soziale Erwünschtheit übte keinen Einfluss auf das Antwortverhalten aus, was darauf hindeutet, dass die Befragten die Privatheit der Befragungssituation als hinreichend gewährleistet ansahen (Stocké, 2004).

Mittels Multidimensionaler Skalierung (MDS) der abgefragten *Verhaltensintentionen* wurden zwei Dimensionen identifiziert, die den Befragten als Einordnungsmaßstab dienten, nämlich der *Aktivitätsgrad* und die *Akzeptanzrichtung*. Der *Aktivitätsgrad* beinhaltete unterschiedlich aufwendige Interaktionsniveaus, bei denen sich das Ignorieren (VI8 „*Ich lese die Kommentare, ohne etwas zu tun*“) am passiven Ende der Skala befand, während die übrigen sieben Variablen ausgeprägtere Aktivitätsniveaus erforderten (vom niedrighschwelligem Klicken von Like- und Reaktionsbuttons über das Melden bis hin zum Kommentieren). Die *Akzeptanzrichtung* bezeichnete die Positionierung der Items zum Stimulusmaterial, ob diese Ausdruck der Ablehnung oder eher Zustimmung waren. Die Akzeptanzrichtung lässt sich daher leicht aus der Itemformulierung ablesen (z. B. VI3 „*Ich unterstütze meine Vorredner mit einem sachlichen Kommentar*“ vs. VI5 „*Ich widerspreche meinen Vorrednern entschieden mit einem beleidigenden Kommentar*“). Die Anordnung im zweidimensionalen Raum positionierte VI8 als Verhaltensintention, die als Ausdruck von Ablehnung registriert wurde.

Als *zustimmende und aktive Verhaltensintentionen* konnten daher VI1 („*Ich reagiere mit einem Like*“), VI3 („*Ich unterstütze mit einem sachlichen Kommentar*.“) und VI6 („*Ich unterstütze mit einem beleidigenden Kommentar*.“) aus der MDS abgeleitet werden. Die Bereitschaft, den „Gefällt mir“-Button bei Facebook zu klicken, unterschied sich in den Experimentalbedingungen, in denen die *Hate Intensität* variiert worden war. SK erzielte hier eindeutig die höchste Zustimmung. Weder die übrigen Versuchsbedingungen, noch die *habitualisierte Nutzung des Like-Buttons*, noch soziodemografische Eigenschaften beeinflussten VI1. Ganz ähnlich war der Befund für die ausgedrückte Wahrscheinlichkeit, mit einem *sachlichen Kommentar* zu reagieren. Signifikante Unterschiede ließen sich hinsichtlich der *Hate Intensität*, nicht jedoch für die restlichen Versuchsbedingungen nachweisen. Das *gewohnheitsmäßige Kommentieren auf öffentlichen Seiten* (nicht jedoch das Kommentieren innerhalb des persönlichen Netzwerks) trug dazu bei, dass Befragte es als wahrscheinlicher ansahen, sachlich zu kommentieren. Die *beleidigende Unterstützung der Kommentierenden* wurde über sämtliche untersuchte Gruppen als nicht bedeutsam eingestuft. Die Streuung der Antworten war zu gering, um Überzufälligkeiten nachweisen zu können und insofern muss die mechanistische Stimulus-Antwort-Beziehung für die Produktion von Hate Speech abgelehnt werden. Dies steht auch im Einklang mit den Befunden aus der Literatur, wonach soziale und politische Ereignislagen die Produktion von Hate Speech befeuern oder wenn Konkurrenzbeziehungen zwischen sozialen Gruppen salient werden (vgl. Kapitel 4.5). In diesem Sinne hat das Stimulusmaterial kompetitive Wahrnehmungen nicht stimuliert.

*Ablehnende aktive Verhaltensintentionen* wurden mit den Items VI2 („*Ich reagiere mit einem missbilligenden Emoji.*“), VI4 („*Ich widerspreche mit einem sachlichen Kommentar.*“), VI5 („*Ich widerspreche mit einem beleidigenden Kommentar.*“) und VI7 („*Ich melde die Kommentare.*“) operationalisiert. Vergleichbar mit dem oben dargestellten Befund für VI6, war auch die geäußerte Verhaltensabsicht *beleidigenden Widerspruch* in Kommentare einfließen zu lassen, in keiner der Versuchsbedingungen signifikant. Das überraschte insofern, als dass die Rezeptionsgedanken überwiegend negative Gedanken enthielten und ein Teil davon durchaus ein Aggressivitätspotential erkennen ließen, das dieses Item abdeckt. Diese Inkonsistenz kann daher als Gedanken-Verhaltensschere verstanden werden. Die Bereitschaft zu kommentieren, wurde mit dem habitualisierten Kommentieren auf öffentlichen Seiten erklärt. Keinen Einfluss übte das habituelle Kommentieren im Facebook-Freundeskreis aus. Der Einfluss habitualisierten Verhaltens konnte auch die Bereitschaft erklären, Missbilligungsemojis zu betätigen, als unangemessen wahrgenommene Kommentare zu melden und sachlichen Widerspruch per Kommentar zu äußern. Diese drei Verhaltensintentionen unterschieden sich signifikant in den Gruppen, die die *Hate Intensität*, nicht jedoch die *Likeanzahl* oder das *Geschlecht der Kommentierenden* variiert hatten.

Das *Ignorieren* des rezipierten Stimulusmaterials (VI8) wurde als Verhaltensabsicht interpretiert, die *passive Ablehnung* artikuliert (s. MDS in Kapitel 6.1.1.5). Die Antworten lagen im mittleren bis höheren Skalenbereich, womit die Befragten die wahrscheinlichste Verhaltensweise zum Ausdruck brachten. Weder die Versuchsbedingungen noch die gewohnheitsmäßigen Verhaltensweisen konnten den Befund erklären. Bezüglich der soziodemografischen Variablen erwiesen sich höher Gebildete und Jüngere eher bereit, die Kommentare zu ignorieren. Vermutet werden konnten Gewöhnungseffekte bezüglich Hate Speech in sozialen Netzwerken. Unklar blieb, wie sich die Variablen gegenseitig beeinflussen, was angesichts der nicht-linearen Beziehungen nicht abschließend geklärt werden konnte.

Überraschend war der Befund, wonach sozial erwünschte Antworttendenzen die artikulierten Verhaltensintentionen in keiner Weise verzerrten. Dies ließe sich zum einen mit dem Erhebungsinstrument begründen. Da soziale Erwünschtheit mittels validierter Kurzskala erhoben wurde, kann diese Erklärung jedoch verworfen werden. Eine weitere Begründung für fehlende Belege sozialer Erwünschtheit besteht darin, Stichprobenmerkmale zu berücksichtigen. Bei der Stichprobe handelte es sich nämlich um mehrheitlich hochgebildete, studentische Versuchspersonen, die nicht nur routiniert in der Beantwortung von Fragebögen sind, sondern zudem wenig Hemmungen zeigen, entsprechend ihrer Disposition zu antworten.

Zustimmende Verhaltensintentionen hängen von bestehenden Einstellungen ab, wie gezeigt werden konnte. Ein niedrigschwelliges Like oder einen sachlichen Kommentar betrachteten Versuchspersonen als wahrscheinlichere Reaktion an, je stärker sie ambivalent sexistischen Einstellungsitems zustimmen. Insbesondere die Rezeption von Hate Speech und Dangerous Speech, aber auch von Kommentaren von Frauen erhöhten die Bereitschaft dieser Zustimmungformen. Viele Likes als peripherer Hinweisreiz verstärkten immerhin die artikulierte Wahrscheinlichkeit des sachlichen Kommentars. Der Zusammenhang zwischen Ambivalentem Sexismus und zustimmenden Verhaltensintentionen wurde nicht für die beleidigende Unterstützung nachgewiesen, denn nach Rezeption des Stimulusmaterials gaben die Versuchspersonen auf der Basis ihrer ambivalent sexistischen Einstellungen nicht an, selbst aggressiv auftreten zu wollen.

Zusammenfassend wurden in Studie I Rezeptionswirkungen von sexistischem Hate Speech aus der Perspektive des ELM, der Hate Speech- und Sexismusforschung untersucht. Diese wurden hinsichtlich der stattgefundenen Elaborationswahrscheinlichkeit, der Relevanzwahrnehmung und bekundeter Verhaltensintentionen analysiert. Noch ungeklärt sind Erkenntnisse bezüglich des Einflusses auf neu entstandene oder aktivierte Einstellungen zur Botschaft sowie zum Thema „Frauenkarrieren in der IT-Branche“, was Gegenstand von Studie II war.

## 6.2 Studie II – Elaboration, Relevanzwahrnehmungen und Einflüsse auf Einstellungen

### 6.2.1 Forschungsdesign und Methodik

#### 6.2.1.1 *Anpassung des Studiendesigns und Stimulusmaterial*

Im Mittelpunkt des Erkenntnisinteresses in Studie II standen Rezeptionswirkungen des Stimulusmaterials hinsichtlich seiner Relevanzwahrnehmung und die Untersuchung des Einflusses auf Einstellungen. Studie I lieferte die Erkenntnis, dass die Intensität von Hate Speech wesentlicher Erklärfaktor für die Elaboration und Relevanzwahrnehmung von Hate Speech, aber auch die Positionierung dazu in Form von artikulierten Verhaltensabsichten ist. Die Rolle von Geschlecht als Erklärfaktor für Rezeptionswirkungen von *sexistischem* Hate Speech blieb ungeklärt, was im Rahmen der zweiten Studie eingehend zu prüfen war. Im Wesentlichen wurde das Schema der Versuchsanordnung aus Studie I übernommen mit

1. randomisierter Zuordnung zu einer Versuchsbedingung,
2. Vorher-Messung von Ambivalentem Sexismus, soziodemografischen Angaben und habituellen Nutzungsweisen von Facebook,
3. Präsentation des Stimulusmaterials und
4. Nachher-Messung bezüglich der Elaboration, Relevanz, Bewertung, Einstellungen und Verhaltensintentionen.

Bei dem Studiendesign handelte es sich um ein 1 x 2-faktorielles Design mit Kontrollgruppe, bei dem das Geschlecht der Kommentierenden (weiblich, männlich) variiert wurde. Anders als in Studie I erhielt die Kontrollgruppe kein Treatment, das die Hate Intensität mittels sachlicher Kommentare von den Experimentalbedingungen abgrenzte, sondern es handelte sich insofern um eine Kontrollgruppe im engeren Sinne, als dass diese einen thematisch nicht verwandten Stimulus erhielt. Erwartet wurden stärkere Effekte zwischen Kontroll- und Experimentalgruppen.

In den beiden Experimentalbedingungen der zweiten Studie wurde Stimulusmaterial aus Studie I eingesetzt. Die Beschreibung der Stimuli ist in Kapitel 6.1.1.3 für Studie I dargestellt worden, weswegen sie aus Redundanzgründen hier nur cursorisch erfolgt. Konstant gehalten wurden das Reaktionsvolumen (operationalisiert als Likeanzahl der Kommentare) und die Hate Intensität, variiert wurde das Geschlecht als zweistufiger Faktor. In Studie I wurde deutlich, dass das Reaktionsvolumen ein bedeutsamer Prädiktor ist, selbst bei nicht bewusster Wahrnehmung. Die Befunde aus Studie I waren diesbezüglich erkenntnisreich, weswegen auf die Variation der Likeanzahl in Studie II verzichtet wurde. Da die Stärke der Effekte von einer

höheren Likeanzahl ausging, wurde die Variation „viele Likes“ gewählt. Hate Speech steht im Mittelpunkt des Erkenntnisinteresses dieser Arbeit, weswegen auf die Variation der Hate Intensität mit der Unterscheidung zwischen Offensive, Hate und Dangerous Speech verzichtet wurde. Dazu konnten starke Effekte der Hate Intensität nachgewiesen werden. Um gezielt Rezeptionswirkungen von Hate Speech zu analysieren, wurde dieses Botschaftsmerkmal konstant gehalten. Noch unzureichend geklärt ist die Rolle der Geschlechterkategorie von Hate Speech-Produzent:innen hinsichtlich ihrer Rezeptions- und Persuasionswirkung, weswegen das Geschlecht der Kommentierenden experimentell variiert wurde. Versuchspersonen erhielten folglich das selbe Stimulusmaterial aus Studie I, welches einem Facebook-Posting von *heise.de* nachempfunden war mit jeweils fünf Hate Speech-Kommentaren von weiblichen (Experimentalgruppe HS Frauen) und männlichen (Experimentalgruppe HS Männer) Kommentierenden (vgl. Abbildung 8). Für die Kontrollgruppe (KG) wurde ein Screenshot einer Facebook-Diskussion zur Marslandung des Rover „Perseverance“ im Februar 2021 gewählt. Es wurde darauf geachtet, dass die Länge der Kommentare vergleichbar mit denen in den Experimentalgruppen war und dass das Geschlecht der Diskussionsteilnehmer (männlich) konstant gehalten wurde.

Die forschungsethische Problematik wurde wiederum dahingehend adressiert, dass die Befragten eingangs auf die Freiwilligkeit der Teilnahme, die kurzfristige Beeinträchtigung der Stimmung und die vorhandene Möglichkeit des Abbruchs hingewiesen wurden und der Teilnahme aktiv zustimmen mussten. Ein ausführliches Debriefing inklusive Nennung von Anlaufstellen zur Beratung bei Betroffenheit wurde ebenfalls eingesetzt. Da in Studie I aufgrund des Exit-Buttons nach Stimuluspräsentation mehr als ein Viertel der Befragten die Befragung abgebrochen hatten, wurde in Studie II darauf verzichtet. Die mit dem Exit-Button verbundenen methodischen Schwierigkeiten wurden im Vergleich zu forschungsethischen Aspekten als schwerwiegender bewertet, zumal die getroffenen Vorkehrungen als umfassend zu betrachten sind (Rat für Sozial- und Wirtschaftsdaten, 2017).

#### 6.2.1.2 Aufbau des Fragebogens

Nach Begrüßung und der informierten Einwilligung der Befragten (s. o.) diente eine Screening-Frage dazu, regelmäßige Nutzer:innen von Facebook (Nutzung mindestens einmal im Monat) als relevante Befragte für die Studie zu identifizieren. Nicht-Nutzer:innen wurden zum Ende der Befragung geleitet („Vielen Dank für Ihre Teilnahme! Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen“). Nach diesen einleitenden Fragebogenseiten folgten die inhaltlichen Fragen. Zunächst wurde die habitualisierte Nutzungsweise abgefragt „Im Folgenden interessiert uns Ihr normales Nutzungsverhalten – nicht

*etwaige Ausnahmesituationen*“. Danach wurde um soziodemografische Angaben gebeten. Im Vergleich zu Studie I wurde die Abfrage soziodemografischer Angaben vorgezogen, um diese Informationen bei etwaigem Studienabbruch zu sichern.

Die Befragten gelangten sodann zu den Fragen zur Erhebung des Kognitionsbedürfnisses, welche gemeinsam mit den Fragen zu sozial erwünschtem Antwortverhalten gestellt wurden. Als nächstes wurde der Ambivalenter-Sexismus-Index abgefragt. Nach diesen extensiven Skalen wurden kürzere Fragen gestellt, um die Befragten kognitiv zu entlasten. Dazu gehörte die Aussage „Ich bin in der IT-Branche tätig“, die Selbsteinschätzung von Vorwissen und das persönliche Involvement. Es folgte die Einleitungsseite vor der Stimuluspräsentation, welche aus dem Studiendesign in Studie I beibehalten wurde. Die Zuweisung zu einer der drei Gruppen (HS von Frauen, HS von Männern, Kontrollgruppe) erfolgte randomisiert. Auch in diesem Experiment oblag die Entscheidung zur Rezeptionsdauer und -intensität den Befragten. Es folgte der Treatmentcheck, um dieses Mal direkte Eindrücke aus der Rezeptionssituation zu gewinnen.

Als nächstes folgten Messungen zur Rezeptionswirkung. Dabei handelte es sich zunächst um Relevanzeinschätzungen, gefolgt von der Nennung von Rezeptionsgedanken, die teilstandardisiert erhoben wurde, d. h. die Befragten wurden gebeten, bis zu zehn Gedanken einzutragen, die im Zusammenhang mit der Rezeptionssituation aufgetreten waren. Des Weiteren wurden die Befragten gebeten, das Stimulusmaterial zu bewerten. Fragen zu Einstellungen gegenüber modernen Frauenkarrieren wurden nur Befragten aus den beiden Experimentalbedingungen (HS von Frauen, HS von Männern) gestellt, da die Frage für die Befragten in der Kontrollgruppe keine Relevanz hatte. Als letzte inhaltliche Frage wurden Verhaltensintentionen erhoben. Die Qualität der Antworten wurde als Selbsteinschätzungssitem formuliert „*Ich habe alle Aufgaben, wie in den Instruktionen verlangt, bewältigt*“. Zuletzt wurde das Debriefing im gleichen Wortlaut wie in Studie I präsentiert.

### 6.2.1.3 Stichprobenbeschreibung

Die Stichprobe wurde aus dem kommerziellen Online-Access-Panel der *respondi* AG rekrutiert, was gewisse Vorteile gegenüber einem studentisch geprägten Panel wie dem SoSci-Panel hat. So können Verzerrungen hinsichtlich der Soziodemografie (Alter, Geschlecht, Bildung) abgefangen werden, indem Quotenvorgaben bei der Rekrutierung umgesetzt werden. Des Weiteren spricht für die Wahl eines nicht-studentischen Panels, dass Einstellungs- und Persönlichkeitsmerkmale eine höhere Variabilität aufweisen können und eher die relevante Population deutschsprachiger Facebook-Nutzer:innen abdecken, welche zum Zeitpunkt der

Erhebung aus breiten Bevölkerungsschichten besteht (Statista, 2021). Rekrutierungsrelevant war die Nutzung des sozialen Netzwerks, darüber hinaus wurde bezüglich des Geschlechts quotiert (50:50-Verteilung).

Vorab wurde die Befragung auf technische Mängel von 14 Personen aus dem persönlichen Umfeld der Autorin geprüft. Eine inhaltliche Prüfung war bereits in Studie I erfolgt und wurde hier als redundant bewertet. Nachdem technische Probleme ausgeschlossen werden konnten, wurde die Feldphase zwischen dem 18.03. und dem 24.03.2021 realisiert. In dieser Zeit wurde der Fragebogen insgesamt 393-mal gestartet, wozu auch versehentliche doppelte Klicks der Versuchsteilnehmenden oder Testaufrufe durch die Autorin zu zählen sind. Ein knappes Drittel, nämlich 122 Personen, wurden von der Befragung ausgeschlossen, da sie nicht als Facebook-Nutzende zur Verfügung standen. Weitere 23 Personen wurden von der Auswertung ausgeschlossen, da sie die Teilnahme an der Befragung frühzeitig abgebrochen hatten, 11 davon unmittelbar nach der Stimuluspräsentation. Wie schon in Studie I dargelegt, wurden 38 Versuchspersonen aus dem Datensatz entfernt, da sie eine Bearbeitungszeit unter 6 Minuten vorwiesen. Durchschnittlich dauerte die Beantwortung des Fragebogens rund 10 Minuten ( $M = 10.53$ ,  $SD = 4.57$ ). Als weiteres Maß zur Beurteilung der Antwortqualität wurde eine Selbsteinschätzungsfrage zur Gewissenhaftigkeit bei der Beantwortung des Fragebogens gestellt und führte zum weiteren Ausschluss von zwei Teilnehmenden, da sie „häufig irgendwas angeklickt“ hatten<sup>54</sup>. Damit befinden sich in der vorläufigen Stichprobe insgesamt 208 Versuchspersonen. Nach der abschließenden Prüfung, ob die experimentelle Manipulation erfolgreich war, blieben in der finalen Nettostichprobe 133 Teilnehmende, deren Antworten ausgewertet wurden.

Die Befragten waren zwischen 20 und 69 Jahre alt, und mit durchschnittlich knapp 48 Jahren ( $M = 47.76$ ,  $SD = 13.91$ ) fast 14 Jahre älter als die Befragten in Studie I. Die Quotenvorgabe bezüglich des Geschlechts wurde nach Datenbereinigung verpasst, aber es waren immerhin 56.4 Prozent Frauen und 43.6 Prozent Männer in der Stichprobe verblieben. Nur sechs Befragte waren in der IT-Branche tätig, was gerade einmal fünf Prozent der Teilnehmenden entspricht. Diese Unterscheidung fand daher in den weiteren Berechnungen keine Berücksichtigung. 26.4 Prozent der Teilnehmenden verfügen über einen Hochschulabschluss und jede:r Vierte über ein Abitur oder Fachabitur (27.8 Prozent). Eine abgeschlossene Lehre gab ein weiteres Viertel an (24.8 Prozent) und rund 15 Prozent verfügen über die Mittlere Reife. Sechs Prozent der Befragten haben einen niedrigeren oder keinen Schulabschluss. Rund zwei

---

<sup>54</sup> Solche Verfahren dienen der Reduzierung von Messfehlern und werden in der Literatur extensiv diskutiert (Shamon & Berning, 2020).

Drittel der Befragten (63.1 Prozent) sind berufstätig, die meisten in Vollzeit (38.3 Prozent), in Teilzeit (15.8 Prozent), als Beamte (2.3 Prozent) oder Selbständige (6.8 Prozent). Der Anteil an Studierenden ist deutlich geringer (5.3 Prozent) als in Studie I und der Anteil an Rentner:innen und Müttern und Vätern in Elternzeit (Kategorie „Sonstiges“) mit 25.6 Prozent wiederum deutlich höher. Im Vergleich zur Stichprobe aus dem SoSci-Panel sind in dieser Stichprobe doppelt so viele Personen, nämlich sechs Prozent arbeitssuchend. Damit bildet die Stichprobe aus dem *respondi*-Access-Panel im Vergleich zu der Stichprobe aus dem SoSci-Panel für Studie I ein breiteres Bevölkerungsspektrum ab.

#### 6.2.1.4 Operationalisierung zentraler Konstrukte

Die Operationalisierung der Konstrukte wurde im Wesentlichen aus Studie I übertragen, wobei an einigen Stellen Anpassungen notwendig geworden waren, entweder durch Rückgriff auf andere Skalen oder durch Neuerhebung von Konstrukten, wie im Folgenden ausführlich dargestellt wird.

##### *Messungen zur Erfassung kognitiver und affektiver Effekte*

Die *wahrgenommene persönliche Relevanz* der Stimuli wurde mit zwei Items der in Studie I verwendeten vier-Item-Message-Relevance-Skala (Laczniak & Muehling, 1993) auf einer sechsstufigen endbepunkteten Skala von „*stimme überhaupt nicht zu*“ bis „*stimme voll und ganz zu*“ gemessen. Die verwendeten Items lauteten „*Solche Kommentare sind wichtig für mich.*“ ( $M = 1.50$ ,  $SD = 0.97$ ) und „*Solche Kommentare sind es wert, dass ich ihnen Aufmerksamkeit schenke.*“ ( $M = 1.61$ ,  $SD = 1.11$ ). Als neues Item wurde die Erwartung auf die Meinungsbildung formuliert, „*Solche Kommentare werden meine Meinung beeinflussen.*“ ( $M = 1.50$ ,  $SD = 1.03$ ). In Studie I wurde die wahrgenommene persönliche Relevanz von den Befragten auffällig stark abgelehnt. Um ergänzend zu den Hypothesen auch mögliche Third-Person-Effekte zu berücksichtigen, wurde der Wortlaut der Items auf Dritte ausgeweitet, und zwar 1) Nutzer:innen des Netzwerks und 2) die „gesamte Gesellschaft“. Die Verteilung der Daten über sämtliche Relevanz-Items war rechtsschief (Tabelle 23), so dass Friedman-Tests durchgeführt wurden, um einen Vergleich für die Paarungen (Relevanz auf individueller Ebene, auf Dritte und auf die Gesellschaft) über die drei Items (1. „wichtig“, 2. „Aufmerksamkeit schenken“, 3. Meinungseffekte) anzustellen.

**Tabelle 23:** Deskriptive Kennwerte für die Relevanz-Items ("Solche Kommentare...")

Item	<i>M</i>	<i>SD</i>	Schiefe	Wölbung
... sind wichtig für mich.	1.50	0.97	2.36	5.62
... sind wichtig für andere Nutzerinnen und Nutzer.	2.06	1.32	1.06	0.15
... sind für die gesamte Gesellschaft wichtig.	1.66	1.13	2.02	4.05
... sind es wert, dass ich ihnen Aufmerksamkeit schenke.	1.61	1.11	2.13	4.32
... sind es wert, dass ihnen andere Nutzerinnen und Nutzer Aufmerksamkeit schenken.	1.82	1.31	1.73	2.18
... sind es wert, dass Ihnen die gesamte Gesellschaft Aufmerksamkeit schenkt.	1.69	1.15	1.85	2.96
... werden meine Meinung beeinflussen.	1.50	1.03	2.60	6.87
... werden die Meinung anderer Nutzerinnen und Nutzer beeinflussen.	2.69	1.46	0.39	-0.88
... werden die gesamtgesellschaftliche Meinung beeinflussen	2.36	1.38	0.84	-0.11
Index wahrgenommene persönliche Relevanz	1.62	1.01	1.97	4.03
Index wahrgenommene Relevanz für andere Nutzer:innen	2.06	1.28	1.13	0.59
Index wahrgenommene gesellschaftliche Relevanz	1.80	1.13	1.52	1.75

*N* = 133

Der Unterschied zwischen den drei Bezugsgruppen (individuell, Dritte, Gesellschaft) war für die erste Itemgruppe „Solche Kommentare sind wichtig“ signifikant  $\chi^2(2) = 40.72, p < .001$ . Dabei waren die Unterschiede zwischen der persönlich-individuellen und der wahrgenommenen Relevanz für andere Nutzer:innen hoch signifikant  $T = -0.47, p < .001$ , jedoch mit einer minimalen Effektstärke von  $r = .04$ . Wahrgenommene persönliche und gesellschaftliche Relevanz unterschieden sich zwar nicht, dafür war der Unterschied zwischen gesellschaftlicher und der Relevanz für andere Nutzer:innen signifikant ( $T = 0.31, p < .05, r = .03$ ). Die zweite Itemgruppe („Solche Kommentare sind es wert, ihnen Aufmerksamkeit zu schenken“) konnte zwar globale Unterschiede feststellen  $\chi^2(2) = 10.92, p < .01$ , die sich jedoch nicht paarweise nachweisen ließen. Der Vergleich der Items in der dritten Gruppe („Solche Kommentare werden Meinungen beeinflussen.“) förderte signifikante Unterschiede zutage ( $\chi^2(2) = 95.53, p < .001$ ). Paarweise Vergleiche mit Bonferroni-Korrektur zeigten, dass der Unterschied zwischen persönlicher und der wahrgenommenen Relevanz für die Meinungsbildung Dritter hochsignifikant, jedoch minimal war ( $T = -0.87, p < .001, r = .08$ ). Ähnlich ist der Befund für den Vergleich zwischen persönlicher und gesamtgesellschaftlicher Relevanz für die Meinungsbildung einzuordnen ( $T = -0.60, p < .001, r = .05$ ). Es konnte kein signifikanter Unterschied zwischen der wahrgenommenen Relevanz für andere Nutzer:innen und der wahrgenommenen gesamtgesellschaftlichen Relevanz gefunden werden.

Ein Durchschnittsindex für die wahrgenommene persönliche Relevanz ( $\alpha = .79$ ,  $M = 1.62$ ,  $SD = 1.01$ ), die wahrgenommene Relevanz für andere Nutzer:innen ( $\alpha = .80$ ,  $M = 2.06$ ,  $SD = 1.28$ ) sowie die Gesellschaft insgesamt ( $\alpha = .84$ ,  $M = 1.80$ ,  $SD = 1.13$ ) wurde jeweils aus den beiden Message-Relevance-Items gebildet. Da sich der wahrgenommene Einfluss auf die Meinungsbildung ungünstig auf die Reliabilität auswirkte, wurde dieses Item separat zur Einordnung der Befunde herangezogen.

Die *kognitive Bewertung* des Stimulusmaterials wurde anhand von sieben Adjektiven auf einer sechsstufigen endbepunkteten Skala von „*stimme überhaupt nicht zu*“ bis „*stimme voll und ganz zu*“ gemessen und diente dazu, den kognitiven Anteil der Einstellung gegenüber Hate Speech zu messen. Erhoben wurden jeweils drei Items mit einer positiven (*konstruktiv, sachlich, rational*) und drei mit einer negativen Valenz (*beleidigend, aggressiv, diskriminierend*). Tabelle 24 gibt einen Überblick über die deskriptiven Kennwerte. Des Weiteren wurde das Item *realistisch* erhoben, was als Indikator für die externe Validität der Stimuli bewertet werden kann. Die Stimuli wurden in der Kontrollgruppe ( $M = 3.33$ ,  $SD = 1.71$ ) im Vergleich zu Hate-Kommentaren von Frauen ( $M = 2.05$ ,  $SD = 1.51$ ) oder von Männern ( $M = 2.21$ ,  $SD = 1.46$ ) als realistischer wahrgenommen; ein Unterschied, der sich als signifikant erwies ( $F(2, 132) = 22.05$ ,  $p < .01$ ,  $\eta_p^2 = .25$ ,  $\omega_p^2 = .24$ ). Post-hoc-Tests mit Bonferroni-Korrektur zeigten, dass der Unterschied zwischen der Kontrollgruppe und den beiden Experimentalgruppen signifikant war (KG und HS<sub>Frauen</sub> 1.28, 95%-CI [0.45, 2.11] sowie KG und HS<sub>Männer</sub> 1.13, 95%-CI [0.35, 1.90]), nicht jedoch zwischen den beiden Experimentalgruppen. Insgesamt muss jedoch festgehalten werden, dass die Kommentare als eher wenig bis durchschnittlich realistisch wahrgenommen wurden.

**Tabelle 24:** Deskriptive Kennwerte für die kognitiven Bewertungs-Items ("Die Kommentare waren...")

Item	<i>M</i>	<i>SD</i>	Schiefe	Wölbung
... beleidigend	3.97	2.08	-0.47	-1.49
... aggressiv	3.70	2.04	-0.25	-1.59
... diskriminierend	4.02	2.10	-0.47	-1.51
... konstruktiv*	4.97	1.42	-1.11	-0.12
... sachlich*	4.71	1.71	-1.08	-0.24
... rational*	4.51	1.70	-0.67	-0.96
... realistisch	2.57	1.66	0.69	-0.84
Index Bewertung	4.46	1.58	-0.72	-0.76

Items mit \* sind umgekehrt gepolt.

$N = 133$

Die drei positiven Items wurden umcodiert, so dass niedrige Werte als Zustimmung und hohe Werte als Ablehnung zu verstehen sind. Je höher der erzielte Wert, desto negativer fällt die kognitive Bewertung aus. In die Berechnung des Durchschnittsindexes flossen die umcodierten (positiven) sowie die negativen Items ein und erzielten eine ausgezeichnete Reliabilität von  $\alpha = .92$  und eine deutlich negative Bewertung ( $M = 4.46, SD = 1.58$ ).

Die *Elaboration* des Stimulusmaterials wurde dem in Studie I beschriebenen inhaltsanalytischen Verfahren zugeführt. In der Stichprobe von 133 Befragten wurden 410 Rezeptionsgedanken genannt, d. h. also, dass im Durchschnitt drei Gedanken ( $M = 3.13, SD = 2.02$ ) geäußert wurden. Um die Intrarater-Reliabilität zu bestimmen, wurden 93 Rezeptionsgedanken von insgesamt 33 Befragten im Abstand von acht Wochen zur Erstcodierung von der Autorin der Arbeit codiert. Die absolute Gedankenanzahl wurde mittels Auszählung im Statistikprogramm SPSS ermittelt und bedurfte damit nicht der Reliabilitätsprüfung. Als Reliabilitätsmaß wurde das Holsti-CR gewählt. Die Reliabilitätswerte bewegten sich in einem ausgezeichneten Übereinstimmungsbereich: Für die Valenz wurde ein Holsti-CR von .90 erreicht und für die vier Kategorien für die Gedankenursprung betrug das Holsti-CR .85.

In dieser Stichprobe handelte es sich mehrheitlich um affektive Reaktionen auf das Stimulusmaterial, die 39 Prozent der Rezeptionsgedanken umfassten. Kognitive Reaktionen, die von der Stimuluswahrnehmung ausgelöst wurden, wurden in 28 Prozent der Gedanken genannt, in ähnlichem Ausmaß kam das Vorwissen der Versuchspersonen zur Geltung (26 Prozent) und nur 7 Prozent der Rezeptionsgedanken behandelten irrelevante Gedanken der „Sons-tige“-Kategorie. Was die Valenz der Gedanken betrifft, waren nur knapp ein Prozent der Gedanken positiv, dafür enthielten mehr als ein Drittel (38 Prozent) neutrale Aussagen und ein Großteil der Rezeptionsgedanken wies eine negative Valenz auf, nämlich 61 Prozent. Insgesamt war der Anteil an negativen Gedanken deutlich geringer im Vergleich zu den Befunden in Studie I ( $M = -0.24, SD = 0.81$ ).

Sechs Items dienten dazu, *Einstellungen zu modernen Frauenkarrieren* auf einer sechsstufigen endbepunkteten Skala von „*stimme überhaupt nicht zu*“ bis „*stimme voll und ganz zu*“ zu erfassen. Hierzu wurde aus der GRO-Skala (Skala zur Messung der normativen Geschlechtsrollenorientierung) von Krampen (1979) das zweite Item in abgewandelter Form verwendet („*Die Idee, dass Frauen in der IT-Branche Karriere machen, ist einfach lächerlich*“). Der Anomie-Skala von Adnan Gümüs und Kolleg:innen (2014) wurden zwei Items entnommen, um Überforderung mit Prozessen sozialen Wandels zu messen: „*In der heutigen Zeit durchschaue ich nicht mehr, was eigentlich passiert*“ und „*Heute ändert sich alles so schnell, dass ich oft nicht weiß, woran ich mich halten soll*“. Drei Items aus der Modern-Sexism-Scale

von Janet Swim und Kolleg:innen (1995) wurden in der deutschen Übersetzung von Eckes und Six-Materna (1998) in leicht abgewandelter Form verwendet, um die Leugnung von Diskriminierung zu messen, und zwar „Die Forderungen von Frauen nach Gleichberechtigung sind nicht nachzuvollziehen“, „Diskriminierung von Frauen ist heute kein Problem mehr“ und schließlich „Frauen haben auf allen gesellschaftlichen Ebenen schon so viel erreicht, dass Sonderprogramme zu ihrer Förderung im Berufsleben überflüssig sind“. Tabelle 25 enthält eine Übersicht der deskriptiven Befunde. Die geringste Zustimmung erhielten die beiden Items EF1 und EF6, die traditionell sexistische Aussagen enthielten (EF1) bzw. Gleichstellungsmaßnahmen ablehnten (EF6). Die höchste Zustimmung entfiel auf die beiden Anomie-Items EF2 und EF3 und bewegte sich in einem mittleren Bereich. Das Item EF1 wurde aus der Berechnung des Durchschnittsindex‘ entfernt, da es sich negativ auf die Reliabilität auswirkte. Die Reliabilität für den Einstellungsindex aus den verbliebenen fünf Items lag bei Cronbachs  $\alpha = .82$  ( $M = 2.11$ ,  $SD = 1.04$ ).

**Tabelle 25:** Deskriptive Kennwerte für Einstellungen zu modernen Frauenkarrieren

Item	<i>M</i>	<i>SD</i>	Schiefe	Wölbung
Die Idee, dass Frauen in der IT-Branche Karriere machen, ist einfach lächerlich. (EF1)	1.18	0.56	4.62	26.64
In der heutigen Zeit durchschaue ich nicht mehr, was eigentlich passiert. (EF2)	2.32	1.27	0.83	0.11
Heute ändert sich alles so schnell, dass ich oft nicht weiß, woran ich mich halten soll. (EF3)	2.49	1.44	0.63	-0.75
Die Forderungen von Frauen nach Gleichberechtigung sind nicht nachzuvollziehen. (EF4)	2.09	1.44	1.17	0.24
Diskriminierung von Frauen ist heute kein Problem mehr. (EF5)	2.02	1.15	1.05	0.80
Frauen haben auf allen gesellschaftlichen Ebenen schon so viel erreicht, dass Sonderprogramme zu ihrer Förderung im Berufsleben überflüssig sind. (EF6)	1.81	1.20	1.43	1.32
Index Einstellungen zu modernen Frauenkarrieren	2.11	1.04	0.90	1.05

*N* = 85

### Messungen zur Erfassung der Verhaltensintention

Studie I zeigte, dass die Items zur Erfassung der Verhaltensintention nicht anfällig für Effekte sozial erwünschten Verhaltens waren, weswegen auf die Unterscheidung zwischen sachlichen und beleidigenden Kommentaren in Studie II verzichtet wurde. Stattdessen wurden vier Items erhoben, drei davon als eindeutig zustimmende (VI1 „Ich reagiere zustimmend mit einem Kommentar oder einem Like“) bzw. ablehnende Positionierung (VI2 „Ich reagiere ablehnend mit einem Kommentar“ und VI3 („Ich melde Facebook die Kommentare“) deklariert und das vierte Item dazwischen (VI4 „Ich lese die Kommentare, ohne etwas zu tun“). Erneut wurde

die Einschätzung der Verhaltensintention auf einer sechsstufigen endbepunkteten Skala von „sehr unwahrscheinlich“ bis „sehr wahrscheinlich“ gemessen.

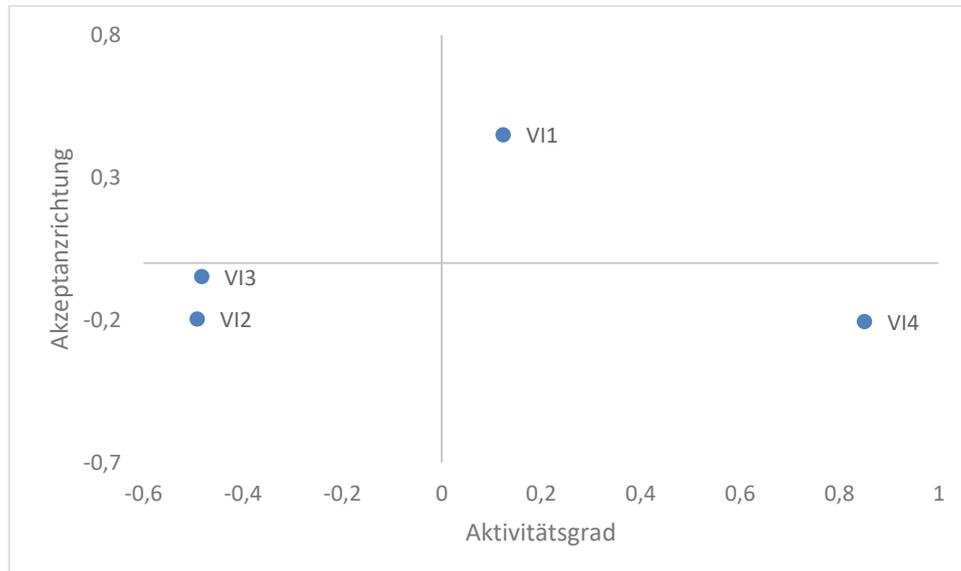
**Tabelle 26:** Deskriptive Kennwerte für die Items Verhaltensintention (VI)

Item	M	SD	Schiefe	Wölbung
Ich reagiere zustimmend mit einem Kommentar oder einem Like. (VI1)	1.67	1.35	2.12	3.50
Ich reagiere ablehnend mit einem Kommentar. (VI2)	2.13	1.66	1.27	0.22
Ich melde Facebook die Kommentare. (VI3)	2.14	1.69	1.36	0.43
Ich lese die Kommentare, ohne etwas zu tun. (VI4)	4.21	1.98	-0.68	-1.17

N = 133

Die beiden ablehnenden Items VI2 und VI3 waren deutlich miteinander korreliert (Pearsons Korrelationskoeffizient  $r = .41$ ,  $p < .01$ ). Die Wahrscheinlichkeit die Kommentare zu ignorieren, war negativ mit den ablehnenden Items korreliert, für die Intention einen ablehnenden Kommentar zu schreiben mit  $r = -.40$ ,  $p < .01$  und die Kommentare zu melden mit  $r = -.38$ ,  $p < .01$ . Es gab keinen statistisch signifikanten Zusammenhang zwischen den drei Items und der Verhaltensintention, zustimmend zu reagieren.

**Abbildung 15:** Zweidimensionale Anordnung der Verhaltensintention im Wahrnehmungsraum (Studie II)



Anmerkungen: VI1 = Zustimmung per Like oder Kommentar, VI2 = Ablehnung per Kommentar, VI3 = Melden, VI4 = Ignorieren

N = 133

Die Ergebnisse der Korrelation zwischen den Items liefert bereits erste Hinweise auf den Wahrnehmungsraum der Versuchspersonen, und zwar hinsichtlich der Akzeptanzrichtung. Um den gewonnenen Eindruck zu validieren, wurde analog zum angewandten Verfahren in

Studie I hierzu eine Multidimensionalen Skalierung mittels PROXSCAL-Verfahren mit der Euklidischen Distanz mit standardisierten Z-Scores als Distanzmaß durchgeführt. Der Stress-I-Wert war ausgezeichnet mit .02 und einer erklärten Streuung von .99. Es ließen sich erneut zwei Dimensionen identifizieren, die auf der Ordinate als *Akzeptanzrichtung* (VI1 eindeutig als Zustimmung und VI2-4 als Ablehnung) und auf der Abszisse als *Aktivitätsgrad* identifiziert werden konnten (Abbildung 15). Die Koordination auf der Abszisse lässt sich leicht identifizieren, wenn die Achse um 45 Grad gegen den Uhrzeigersinn gekippt wird. Erkennbar sind dann die aktiven Items (VI1-3) als deutlich abgesetzt von VI4. Die Erklärung für das Zustandekommen für die Wertebereiche auf den Achsen wurde in den entsprechenden Abschnitten von Studie I ausführlich erläutert.

#### *Messungen von Kovariaten und Moderatorvariablen*

Die Messung *ambivalent sexistischer Einstellungen* erfolgte wiederum mittels ASI von Glick und Fiske (1996) in der deutschen Übersetzung von Eckes und Six-Materna (1999). Verwendet wurde eine endbepunktete sechsstufige Skala von „stimme überhaupt nicht zu“ bis „stimme voll und ganz zu“. Die Reliabilität der Skala war mit  $\alpha = .80$  akzeptabel ( $M = 3.11$ ,  $SD = 0.94$ ), wobei die Unterskala *Benevolenter Sexismus* eine gerade noch zufriedenstellende ( $\alpha = .73$ ,  $M = 3.53$ ,  $SD = 1.08$ ) und die Unterskala *Hostiler Sexismus* eine gute Reliabilität erreichte ( $\alpha = .85$ ,  $M = 2.77$ ,  $SD = 1.21$ ). Die Korrelation der beiden Subskalen fiel wider Erwarten niedrig aus, war jedoch signifikant ( $r = .25$ ,  $p < .01$ ).

Da für Studie II im Vergleich zu Studie I eine deutlich heterogenere Stichprobe hinsichtlich der Bildung erwartet (und realisiert) wurde, konnte das Kognitionsbedürfnis der Versuchspersonen nicht anhand der Kurzskala für *Need for Cognition* (NFC) abgefragt werden. Wie in Kapitel 6.1.1.5 begründet wurde, bestehen Zweifel ob der Eignung des Reliabilitätsmaßes Cronbachs Alpha für Kurzskalen bei heterogenen Stichproben. Aus diesem Grund wurde die 16 Items umfassende Kurzversion der NFC-Skala von Cacioppo und Petty (1982) in der deutschen Übersetzung von Herbert Bless und Kolleg:innen (1994) eingesetzt. Die Versuchspersonen wurden gebeten, ihre Antworten auf einer sechsstufigen endbepunkteten Skala von „stimme gar nicht zu“ bis „stimme voll und ganz zu“ abzugeben. Die Reliabilität des Durchschnittsindex erwies sich als sehr gut mit  $\alpha = .84$  ( $M = 4.31$ ,  $SD = 0.77$ ).

Die Messung des subjektiv wahrgenommenen *Vorwissens* (VW) wurde wie in Studie I mit einem Item „*Ich fühle mich kompetent an einer Diskussion zum Thema ‚Karrierechancen von Frauen‘ teilzunehmen.*“ realisiert ( $M = 3.37$ ,  $SD = 1.61$ ). Auch das *persönliche Involvement* (PI) wurde analog zu Studie I abgefragt. Die Befragten wurden gebeten, mittels Mehrfachauswahl

(Auswahl aller zutreffenden Antworten) alle genutzten Informationsquellen zum Thema ‚*Karriereschancen von Frauen*‘ zu nennen. Im Durchschnitt wurden zwei Informationsquellen genutzt ( $M = 2.36$ ,  $SD = 1.12$ ), und zwar hauptsächlich *mediale Quellen* (Fernsehsendungen wurden von knapp 40 Prozent, Bücher/Zeitungsartikel von 35 Prozent und Internetquellen von einem Viertel der Befragten genannt) oder *Gespräche* (38 Prozent der Befragten). Vorträge wurden seltener genannt, nämlich von sechs Prozent der Versuchspersonen. In der offenen Antwortkategorie wurden Lebenserfahrungen, wissenschaftliche Publikationen und das berufliche Umfeld als weitere Informationsquellen von sieben Prozent der Befragten genannt. 48 Personen, also rund ein Drittel der Befragten gaben an, sich nicht zum Thema zu informieren. Die Stärke des PI floss als Summenscore in die Berechnungen ein.

Die *habitualisierte Nutzung von Facebook* (HN) wurde mit einer verkürzten selbst erstellten Skala von sechs Items (elf Items in Studie I) auf der endbepunkteten Skala von 1 („nie“) bis 6 („immer“) gemessen. Dabei standen Nutzungsweisen, die eine Interaktion mit anderen Nutzenden beinhalteten, im Vordergrund (Nutzung des Like- und Reaktionsbuttons, Modus = 5), Teilen von Inhalten (Modus = 1), Verfassen von Kommentaren (Modus = 4), Melden von Kommentaren (Modus = 1). Zusätzlich wurde auch das Lurken (Lesen von Beiträgen und Kommentaren (Modus = 5) und das Ignorieren (Durchscrollen des Newsfeeds (Modus = 3) erfasst. Die Kürzung der Skala diente dazu, einer Ermüdung und Reaktanz der Befragten durch ähnlich klingende Items vorzubeugen. Darüber hinaus hatte sich in Studie I die Unterscheidung zwischen eigenen Inhalten und Inhalten auf öffentlichen Seiten als nicht bedeutsam erwiesen. Um die Ordinalskala parametrischen Analyseverfahren zugänglich zu machen, wurde die sechsstufige Skala wie in Studie I dichotomisiert. Skalenwerte von 1 bis 3 wurden zum Wert 0 („keine HN“) und Skalenwerte von 4 bis 6 zum Wert 1 („HN“) umcodiert.

*Soziale Erwünschtheit* (SE) wurde anders als in Studie I anhand einer Skala gemessen, welche die Gemein Sinn-orientierten Aspekte sozial erwünschten (Antwort-)verhaltens (KSE-G) fokussiert und damit die moralischen Maßstäbe der Befragten aufzeigt (Kemper et al., 2014). Für die Untersuchung der Rezeptionswirkung von Hate Speech-Kommentaren wurde die sechs Items umfassende Skala mit sechs endbepunkteten Stufen von „stimme gar nicht zu“ bis „stimme voll und ganz zu“ als geeignet betrachtet. Die Überprüfung der Reliabilität ergab jedoch ein unzureichendes Cronbachs  $\alpha$  von .59, weswegen nur das Item „*Im Streit bleibe ich stets sachlich und objektiv.*“ als Indikator für soziale Erwünschtheit behalten wurde, da es annähernd Normalverteilung erreichte ( $M = 4.14$ ,  $SD = 1.29$ ) und somit die Anwendung parametrischer Verfahren erlaubt. Unterschiede zwischen männlichen und weiblichen Versuchspersonen bestanden nicht ( $F(1, 132) = 0.94$ ,  $p = .334$ ).

## 6.2.2 Ergebnisse

Im Folgenden werden die Ergebnisse der zweiten Studie präsentiert. Wie auch für Studie I orientierte sich die Reihenfolge der Bearbeitung im Wesentlichen an der Reihenfolge der jeweiligen Hypothesen.

### 6.2.2.1 Prüfung der Zufallsaufteilung

Die Prüfung auf Zufallsaufteilung erfolgte für die metrisch skalierten unabhängigen Variablen mittels einfaktorieller Varianzanalyse und für die kategoriellen Variablen mittels  $\chi^2$ -Test. Die Befunde für die metrisch skalierten Variablen sind in Tabelle 27 aufgeführt. Die Randomisierung des *Geschlechts* der Befragten war erfolgreich mit ( $\chi^2(2) = 2.27, p = .321$ ). Auch was die Bildung betrifft, konnte eine erfolgreiche Zufallsaufteilung erreicht werden, und zwar hinsichtlich eines *geringen* ( $\chi^2(2) = 2.06, p = .358$ ), *mittleren* ( $\chi^2(2) = 0.73, p = .696$ ), wie auch *hohen Bildungsgrads* ( $\chi^2(2) = 0.09, p = .958$ ).

**Tabelle 27: Prüfung auf Randomisierung zentraler metrischer Variablen**

		<i>M</i>	<i>SD</i>	<i>F</i>
NFC	Kontrollgruppe	4.31	.83	$F(2, 131) = 0.40, ns$
	Hate Speech Frauen	4.22	.80	
	Hate Speech Männer	4.38	.70	
VW	Kontrollgruppe	3.63	1.55	$F(2, 132) = 2.40, ns$
	Hate Speech Frauen	2.89	1.49	
	Hate Speech Männer	3.48	1.70	
PI	Kontrollgruppe	2.61	1.26	$F(2, 84) = 1.21, ns$
	Hate Speech Frauen	2.25	1.03	
	Hate Speech Männer	2.20	1.03	
Alter	Kontrollgruppe	51.27	13.39	$F(2, 132) = 2.64, ns$
	Hate Speech Frauen	46.81	14.77	
	Hate Speech Männer	44.98	13.25	
SE	Kontrollgruppe	4.17	1.40	$F(2, 132) = 0.60, ns$
	Hate Speech Frauen	4.16	1.30	
	Hate Speech Männer	4.08	1.18	
ASI	Kontrollgruppe	3.17	1.02	$F(2, 132) = 0.17, ns$
	Hate Speech Frauen	3.05	0.85	
	Hate Speech Männer	3.08	0.94	

NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, SE = Soziale Erwünschtheit, ASI = Ambivalenter Sexismus.  
 $p > .05$

### 6.2.2.2 Auswirkungen von Hate Speech auf Elaboration

Auch in Studie II wurden die mittels Thought-Listing-Technik erhobenen Rezeptionsgedanken nach inhaltsanalytischer Auswertung tiefergehenden Datenanalysen unterworfen, deren Befunde im Folgenden dargestellt werden.

#### *Auswirkungen auf die Gedankenvalenz*

Von durchschnittlich drei Gedanken wiesen zwei Gedanken eine negative Valenz auf ( $M = 1.92$ ,  $SD = 1.73$ ) und ein Gedanke eine neutrale Valenz ( $M = 1.18$ ,  $SD = 1.53$ ). Vernachlässigbar war der Anteil an positiven Gedanken, da diese gerade einmal ein Prozent der Rezeptionsgedanken ausmachten. Der Vergleich zwischen den beiden Experimentalgruppen und der Kontrollgruppe ergab signifikante Unterschiede hinsichtlich des Anteils an negativen Gedanken (Welch-Test,  $F(2, 77.14) = 19.45$ ,  $p < .001$ ). Games-Howell-Post-hoc Tests konnten zeigen, dass der Anteil an negativen Rezeptionsgedanken im Vergleich zur Kontrollgruppe in der Gruppe *Hate Speech von Frauen* deutlich höher lag (0.52, 95%-CI[0.11, 0.94]) und in der Gruppe *Hate Speech von Männern* noch übertroffen wurde (0.90, 95%-CI[0.56, 1.25]). Dabei rezipierten die Versuchspersonen Hate Speech von männlichen Kommentierenden negativer als Hate Speech von weiblichen Kommentierenden (0.38, 95%-CI[0.03, 0.73]).

Studie I konnte zeigen, dass der Einfluss der experimentellen Stimuli ursächlich für den Anteil an negativen Gedanken verantwortlich war. Da die Versuchspersonen in Studie I einem homogenen Milieu mit hohem Bildungsniveau und ähnlichen Einstellungen entstammen, wurden die gleichen hierarchischen Regressionsmodelle zur Auswertung der Daten in Studie II herangezogen. In *Modell 1* wurden die beiden Experimentalbedingungen (HS Frauen und HS Männer) dummycodiert aufgenommen (KG war Referenzgruppe) und für soziale Erwünschtheit kontrolliert. *Modell 2* enthielt die drei Elaborationsvariablen Kognitionsbedürfnis, Vorwissen und persönliches Involvement. Ambivalenter Sexismus wurde in *Modell 3* und soziodemografische Variablen (Alter, Geschlecht, Bildung) in *Modell 4* aufgenommen. Partielle Regressionsdiagramme zeigten an, dass Linearität zwischen abhängiger und unabhängigen Variablen angenommen werden konnte. Die Analyse der Residuen (Backhaus et al., 2018) ergab für standardisierte bzw. studentisierte ausgeschlossene Residuen keine Ausreißer von  $\pm 3$  Standardabweichungen. Hebelwerte als Maß für Ausreißer bewegten sich in einem akzeptablen Toleranzbereich (Field, 2018), was noch von den Cook-Distanzwerten ( $< .07$ ) unterstrichen wurde. Die Kollinearitätsstatistik ergab keinen Hinweis auf problematische Zusammenhänge zwischen den unabhängigen Variablen, denn der Toleranzwert nahm Werte  $> .65$  und VIF-Werte lagen bei  $< 1.53$ . Normalverteilung konnte nach Sichtung des P-P-Plots des standardisierten Residuendiagramms angenommen werden. Da Heteroskedastizität aufgrund

des Residuendiagramms nicht ausgeschlossen werden konnte, wurde Bootstrapping mit 5000 Iterationen zur Schätzung robuster Standardfehler durchgeführt.

**Tabelle 28:** Einfluss auf den Anteil an negativen Gedanken mittels hierarchischer Regressionsanalyse

	Variable	B	SE B	95% CI für B		$\beta$	R <sup>2</sup>	$\Delta R^2$
				LL	UL			
Modell 1	Konstante	-0.52	0.28	-1.10	0.12		.27	.27***
	Dummy HS Frauen	-0.41*	0.20	-0.85	0.03	-0.23		
	Dummy HS Männer	-0.85***	0.19	-1.17	-0.51	-0.50		
	SE	0.18**	0.06	0.06	0.30	0.29		
Modell 2	Konstante	-0.06	0.53	-1.10	1.06		.30	.03
	Dummy HS Frauen	-0.46*	0.20	-0.89	-0.02	-0.26		
	Dummy HS Männer	-0.84***	0.19	-1.19	-0.47	-0.50		
	SE	0.20**	0.06	0.07	0.31	0.31		
	NFC	-0.02	0.11	-0.23	0.20	-0.02		
	PI	-0.04	0.08	-0.18	0.11	-0.05		
	Vorwissen	-0.09	0.07	-0.21	0.04	-0.15		
Modell 3	Konstante	-0.35	0.63	-1.47	0.90		.31	.01
	Dummy HS Frauen	-0.47*	0.20	-0.90	-0.01	-0.26		
	Dummy HS Männer	-0.83***	0.20	-1.19	-0.44	-0.50		
	SE	0.19**	0.06	0.05	0.30	0.29		
	NFC	0.001	0.12	-0.22	0.22	0.001		
	PI	-0.02	0.08	-0.17	0.14	-0.03		
	Vorwissen	-0.10	0.07	-0.22	0.03	-0.16		
	ASI	0.08	0.09	-0.08	0.27	0.09		
Modell 4	Konstante	-0.30	0.69	-1.58	0.96		.33	.02
	Dummy HS Frauen	-0.48*	0.21	-0.91	-0.03	-0.26		
	Dummy HS Männer	-0.87***	0.20	-1.23	-0.46	-0.51		
	SE	0.19**	0.06	0.05	0.31	0.30		
	NFC	0.02	0.12	-0.21	0.26	0.02		
	PI	-0.03	0.08	-0.18	0.13	-0.04		
	Vorwissen	-0.09	0.07	-0.21	0.04	-0.15		
	ASI	0.04	0.10	-0.14	0.25	0.05		
	Alter	6.63e-5	0.01	-0.01	0.01	0.001		
	Geschlecht (1 = m, 0 = w)	0.13	0.17	-0.20	0.47	0.08		
	Bildung (1 = hoch, 0 = Rest)	-0.17	0.19	-0.53	0.20	-0.10		

Anmerkungen: CI = Konfidenzintervall, LL = untere Schranke (lower limit), UL = obere Schranke (upper limit), HS = Hate Speech, SE = Soziale Erwünschtheit, NFC = Need for Cognition, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

N = 133

Hate Speech von Männern erwies sich als beständiger hochsignifikanter Prädiktor für negative Rezeptionsgedanken, und zwar deutlich stärker als Hate Speech von Frauen im Vergleich zur Kontrollgruppe, die als Referenzgruppe diente. Abgesehen von sozialer Erwünschtheit,

das einen hochsignifikanten Einfluss in allen vier Modellen ausübte, waren alle übrigen Prädiktoren nicht signifikant. Damit decken sich die Befunde mehrheitlich mit denen in Studie I. Zwar traten dort keine Effekte sozialer Erwünschtheit auf, was mit der Stichprobenszusammensetzung und der Messung sozialer Erwünschtheit zusammenhängen könnte. Die Modelle trugen insgesamt zu einer guten Varianzaufklärung bei; wobei die Aufnahme weiterer Variablen die Modelle nur geringfügig verbesserte ( $\Delta R^2$  zwischen Modell 1 und Modell 4 = .06), allein Modell 1 konnte mit  $R^2 = .27$  ( $R^2_{\text{korrr}} = .25$ ) einen mittleren Anteil an Varianz aufklären. Damit konnte die Hypothese H2b vorläufig bestätigt werden.<sup>55</sup>

Um der Frage nachzugehen, ob der Anteil an negativen Rezeptionsgedanken in den Experimentalgruppen unter der Bedingung unterschiedlicher Fähigkeits- und Motivationsniveaus variiert, wurde eine Moderationsanalyse mit PROCESS (Hayes, 2018) gerechnet. Für die Versuchsbedingungen als Prädiktor wurden orthogonale Kontraste mit Helmert-Codierung aufgesetzt, was sowohl den Vergleich zwischen Kontroll- und Hate Speech-Gruppen (Vergleich 1) als auch den Vergleich der Hate Speech-Gruppen untereinander (Vergleich 2) erlaubt. Als Moderatoren wurden das Kognitionsbedürfnis und Vorwissen gewählt und mittelwertzentriert, weil beide laut ELM jeweils die Fähigkeit und Motivation zur Elaboration abbilden. Um robuste Schätzer zu erzielen, wurden 10000 Bootstrapping-Iterationen und HC3-Schätzer berechnet.

Das Modell war signifikant ( $F(8, 121) = 5.20, p < .001$ ) und trug in hohem Maß zur Varianzaufklärung bei ( $R^2 = .26$ ), was aber auf den Einfluss der Experimentalgruppen ( $b_{\text{Vergleich 1}} = -0.75, p < .001$  und  $b_{\text{Vergleich 2}} = -0.34, p < .05$ ) auf den Anteil an negativen Rezeptionsgedanken zurückzuführen war, denn es konnte nicht nachgewiesen werden, dass die Beziehung zwischen Prädiktor und Kriterium von unterschiedlichen Elaborationsniveaus abhängig war (Moderator Kognitionsbedürfnis:  $\Delta R^2 = .02, F(2, 121) = 1.53, p = .222, 95\% \text{ CI}$  [Vergleich 1: -0.5, 0.23 und Vergleich 2: -0.04, 0.68], Moderator Vorwissen: ( $\Delta R^2 = .002, F(2, 121) = 0.12, p = .884, 95\% \text{ CI}$  [Vergleich 1: -0.17, 0.22 und Vergleich 2: -0.16, 0.24])). Die Befunde sprechen dafür, die postulierte Interaktion aus Hypothese H2b damit abzulehnen.

#### *Auswirkungen von Hate Speech auf die Anzahl an Gedanken*

Die Stimulusdarbietungen wurden von nahezu allen Befragten (98 Prozent) rege diskutiert mit durchschnittlich drei Rezeptionsgedanken ( $M = 3.13, SD = 2.02, \text{ Schiefe} = 1.51, \text{ Wölbung} = 2.66$ ). Entgegen den Erwartungen zeigte sich zunächst, dass sich die Anzahl der Gedanken in den Experimentalgruppen ( $M_{\text{HS Frauen}} = 3.32, SD_{\text{HS Frauen}} = 1.99$  bzw.  $M_{\text{HS Männer}} = 3.44,$

<sup>55</sup> Die Hypothesen H1a-c und H3a-c werden in Studie II nicht geprüft, da die Hate Intensität und das Reaktionsvolumen nicht variiert wurden.

$SD_{HS \text{ Männer}} = 2.26$ ) und in der Kontrollgruppe ( $M = 2.65, SD = 1.70$ ) nicht voneinander unterschieden ( $F(2, 128) = 2.05, ns$ ). Im nächsten Schritt wurde eine Poissonregression gerechnet, was als Mittel der Wahl bei Zählvariablen gilt (s. Kapitel 6.1.2.2 für die Begründung und Befunde aus Studie I). Das Modell beinhaltete die dummycodierten Experimentalbedingungen (Kontrollgruppe als Referenzkategorie), dazu die Variablen zur Messung von Fähigkeit und Motivation zur Elaboration (Kognitionsbedürfnis, Vorwissen und persönliches Involvement), den Ambivalenter-Sexismus-Index, soziale Erwünschtheit und die soziodemografischen Angaben Alter, Geschlecht und Bildung (s. Tabelle 29). Alle metrischen Variablen flossen mittelwertzentriert in das Modell ein, um die Interpretation der Befunde zu erleichtern. Die Pearson-Dispersion lag bei einem Wert von 0.998, was erwartungsgemäß im Dispersionstest als nicht verschieden von 1 bestätigt wurde  $\chi^2(1) = 0.01, p = .454$ . Hypothese H2a wurde wegen der fehlenden Rezeptionswirkungen in den Experimentalgruppen vorläufig abgelehnt. Die Anzahl der Rezeptionsgedanken konnte erfolgreich erklärt werden durch Vorwissen und der Neigung sozial erwünscht zu antworten, welche jeweils knapp 12 Prozent mehr Rezeptionsgedanken bewirkten. Männliche Versuchspersonen waren im Vergleich zu weiblichen Versuchspersonen weniger aktiv, und zwar um den Faktor 0.71, was bedeutet, dass sie 1.4 weniger Gedanken als Frauen äußerten.

**Tabelle 29:** *Poissonregression mit mittelwertzentrierten Kovariaten zur Erklärung der Anzahl von Rezeptionsgedanken (Studie II)*

Variable	B	SE	95%-CI für B		IRR
			LL	UL	
Konstante	1.04***	0.15	0.74	1.32	2.825
Dummy HS Frauen	0.26	0.16	-0.06	0.58	1.294
Dummy HS Männer	0.19	0.17	-0.13	0.52	1.213
NFC	-0.03	0.09	-0.21	0.15	0.970
VW	0.11*	0.05	0.003	0.22	1.115
PI	0.08	0.07	-0.06	0.21	1.079
ASI	-0.07	0.07	-0.22	0.08	0.933
SE	0.11*	0.05	0.01	0.21	1.115
Dummy Bildung (1 = hoch, 0 = Rest)	-0.04	0.15	-0.33	0.25	0.961
Alter	-0.001	0.01	-0.01	0.01	0.999
Geschlecht (1 = m, 0 = w)	-0.34*	0.14	-0.61	-0.07	0.714

*Anmerkungen:* CI = Konfidenzintervall, LL = untere Schranke (lower limit), UL = obere Schranke (upper limit), HS = Hate Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 82$

AIC = 325.51, BIC = 351.98,  $-2\log L(11) = 303.509$

Das gerechnete Poissonregressionsmodell aus Tabelle 29 wurde um Interaktionen der Elaborationsvariablen (Kognitionsbedürfnis, Vorwissen und Persönliches Involvement) mit den dummycodierten Experimentalgruppen (Kontrollgruppe als Referenzgruppe) ergänzt. Männliche Versuchspersonen äußerten im Vergleich zu weiblichen Versuchspersonen weniger Rezeptionsgedanken, nämlich um den Faktor 0.71. In der Experimentalgruppe Hate Speech von Männern war die Anzahl der Gedanken im Vergleich zur Kontrollgruppe um 49 Prozent erhöht. Auch das Vorwissen erhöhte die Gedankenzahl signifikant, und zwar um 26 Prozent. Die Gedankenzahl wurde nach Rezeption von Hate Speech jedoch nicht von der Fähigkeit und Motivation zur Elaboration beeinflusst, da die Interaktionen keine Signifikanz erreichen konnten. Damit wird Hypothese H2a abgelehnt.

#### *Auswirkungen von Hate Speech auf den Gedankenursprung*

Der Gedankenursprung als Ausgangsbasis für die einzelnen Rezeptionsgedanken wurde jeweils einer der vier Kategorien zugeordnet, wie bereits ausführlich in Studie I dargelegt wurde. Diese sind 1) kognitive Auseinandersetzungen mit dem Stimulusmaterial, 2) affektive Reaktionen darauf, 3) kognitiver Erfahrungsabgleich und Vorwissen und 4) sonstige Gedanken ohne eindeutigen kognitiv-affektiven Hintergrund. Zwischen null und fünf von zehn möglichen Rezeptionsgedanken je Versuchsperson entfielen auf die erste Kategorie. Auch in diesem Fall wurden aufgrund des gehäuftten Vorkommens von Nullen Hurdle-Modelle notwendig, die sowohl das Zählen der Gedanken als auch die Wahrscheinlichkeit des Nicht-eintretens (=kein Rezeptionsgedanke dieser Kategorie) berücksichtigen. In die Modelle flossen dieselben Variablen ein, wie in Kapitel 6.1.2.2 (s. a. Tabelle 30-32 für einen Überblick der Modelle) dargestellt. Um den Zusammenhang zwischen der Rezeption von Hate Speech und der Verarbeitung des Stimulusmaterials in Abhängigkeit von der Motivation und Fähigkeit zur Elaboration zu untersuchen, wurden in die einfachen Hurdle-Modelle Interaktionsterme zwischen den Experimentalgruppen (HS von Frauen und HS von Männern, Kontrollgruppe als Referenzkategorie) und den Elaborationsvariablen aufgenommen.

#### *Kognitive Auseinandersetzung mit dem Stimulusmaterial*

Die Prüfung auf Überdispersion mittels Likelihood-Ratio-Test (Perumean-Chaney et al., 2013) ergab, dass eine Hurdle-Modellierung mit Poisson-Verteilung die Daten angemessen abbildete ( $\chi^2(1) = 0.005, p = .946$ ). Die Sichtung des Zählmodells zeigte, dass keines der Modellparameter die kognitive Verarbeitung des Stimulus als Ursprung für die notierten Rezeptionsgedanken erklären konnte. Im Hurdle-Modell war einzig das Persönliche Involvement signifikant: Die Wahrscheinlichkeit *keine* Gedanken aus dieser Kategorie zu äußern, war mit

jeder zusätzlichen Informationsquelle zu weiblichen Karrieren (also höherem PI) um 76 Prozent erhöht. Die Hypothese H2c wird vorläufig abgelehnt.

Die Interaktionen zwischen den Elaborationsvariablen und den Versuchsbedingungen waren nicht signifikant. Da der Zusammenhang zwischen dem Geschlecht der Kommentierenden und der kognitiven Auseinandersetzung mit dem Stimulusmaterial nicht von der Fähigkeit und Motivation zur Elaboration beeinflusst wurde, muss Hypothese H2c vorläufig abgelehnt werden. Im Hurdle-Modell war die Wahrscheinlichkeit für höher Gebildete signifikant erhöht, *keinen* Rezeptionsgedanken zu äußern, der eine kognitive Verarbeitung der Stimuli zum Ausgangspunkt hatte (IRR = 3.856,  $p < .05$ ) und für männliche Befragte sank diese signifikant (IRR = 0.215,  $p < .05$ ). Die übrigen Modellparameter erreichten keine Signifikanz.

**Tabelle 30:** Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Kognitive Auseinandersetzung mit dem Stimulusmaterial)

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	0,68	0,35	-0,01	1,37	1,968
	Dummy HS Frauen	-0,41	0,46	-1,32	0,50	0,665
	Dummy HS Männer	0,03	0,40	-0,76	0,82	1,028
	NFC	-0,08	0,21	-0,49	0,34	0,927
	VW	0,12	0,15	-0,17	0,42	1,131
	PI	-0,41	0,23	-0,87	0,05	0,662
	ASI	-0,28	0,23	-0,74	0,18	0,756
	SE	0,18	0,16	-0,13	0,50	1,202
	Dummy Bildung (1 = hoch, 0 = Rest)	-0,58	0,47	-1,51	0,35	0,558
	Alter	-0,01	0,02	-0,04	0,03	0,994
	Geschlecht (1 = m, 0 = w)	-0,22	0,50	-1,20	0,76	0,803
Hurdle-Modell	Konstante	-0,36	0,53	-1,41	0,68	0,695
	Dummy HS Frauen	0,47	0,64	-0,78	1,72	1,607
	Dummy HS Männer	-0,09	0,62	-1,30	1,12	0,914
	NFC	0,001	0,35	-0,68	0,69	1,001
	VW	0,12	0,21	-0,29	0,53	1,129
	PI	0,57*	0,28	0,02	1,11	1,763
	ASI	0,02	0,30	-0,58	0,61	1,015
	SE	-0,15	0,21	-0,55	0,25	0,861
	Dummy Bildung (1 = hoch, 0 = Rest)	1,01	0,58	-0,11	2,14	2,759
	Alter	0,005	0,02	-0,04	0,05	1,005
	Geschlecht (1 = m, 0 = w)	-0,97	0,52	-2,00	0,05	0,379

Anmerkungen: HS = Hate Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalentler Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 82$

AIC = 231.03, BIC = 283.97,  $-2\log L = 187.027$

*Affektive Verarbeitung des Stimulusmaterials*

Überdispersion lag in den Modellen zur Untersuchung affektiv gefärbter Rezeptionsgedanken nicht vor ( $\chi^2(1) = 0.61, p = .435$ ). Das Zählmodell konnte keine signifikanten Aussagen treffen, um die Äußerung von Gedanken dieser Kategorie zu prognostizieren. Im Hurdle-Modell senkten ambivalent sexistische Einstellungen ( $IRR = 0.470, p < .05$ ) und soziale Erwünschtheit ( $IRR = 0.572, p < .05$ ) die Wahrscheinlichkeit, nicht affektiv zu reagieren. Anders ausgedrückt: Die Wahrscheinlichkeit sich nicht empört (negative Gedankenvalenz erhöht, s. o.) zu äußern, sank mit jeder Einheit auf der Skala zu Ambivalentem Sexismus bzw. sozialer Erwünschtheit um jeweils rund zwei Gedanken.

**Tabelle 31:** Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Affektive Auseinandersetzung mit dem Stimulusmaterial)

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	-0.25	0.44	-1.11	0.61	0.780
	Dummy HS Frauen	0.57	0.51	-0.43	1.57	1.770
	Dummy HS Männer	0.59	0.48	-0.35	1.53	1.808
	NFC	-0.02	0.25	-0.50	0.47	0.983
	VW	0.15	0.14	-0.13	0.42	1.156
	PI	-0.18	0.16	-0.49	0.13	0.836
	ASI	0.13	0.17	-0.21	0.47	1.135
	SE	0.28	0.15	-0.01	0.57	1.322
	Dummy Bildung (1 = hoch, 0 = Rest)	0.22	0.37	-0.51	0.94	1.243
	Alter	0.003	0.01	-0.02	0.03	1.003
	Geschlecht (1 = m, 0 = w)	-0.45	0.38	-1.18	0.29	0.641
Hurdle-Modell	Konstante	-0.35	0.56	-1.45	0.76	0.706
	Dummy HS Frauen	0.54	0.66	-0.76	1.84	1.708
	Dummy HS Männer	1.00	0.68	-0.33	2.32	2.706
	NFC	0.10	0.38	-0.64	0.84	1.104
	VW	0.06	0.23	-0.39	0.51	1.060
	PI	0.20	0.27	-0.32	0.72	1.216
	ASI	-0.75*	0.35	-1.43	-0.08	0.470
	SE	-0.56*	0.23	-1.00	-0.12	0.572
	Dummy Bildung (1 = hoch, 0 = Rest)	0.39	0.59	-0.76	1.54	1.477
	Alter	-0.01	0.02	-0.05	0.03	0.991
	Geschlecht (1 = m, 0 = w)	0.10	0.54	-0.95	1.15	1.107

Anmerkungen: HS = Hate Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenter Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

$N = 82$

AIC = 241.08, BIC = 294.03,  $-2\log L = 197.080$

Das Interaktionsmodell konnte keinen Moderationseffekt der Elaborationsvariablen auf die Beziehung zwischen der Rezeption des experimentellen Stimulus und deren affektive Verarbeitung identifizieren. Hypothese H2c wird vorläufig als nicht bestätigt betrachtet. Im Zählmodell war der Einfluss von sozialer Erwünschtheit auf die mengenmäßige Anzahl der Gedanken dieser Kategorie erhöht (IRR = 1.379,  $p < .05$ ). Mit jeder Einheit für das soziale Erwünschtheitsitem wurden 38 Prozent mehr Gedanken artikuliert, die eine affektive Verarbeitung des Stimulusmaterials erkennen ließen. Auch im Hurdle-Modell machte sich soziale Erwünschtheit bemerkbar: Die Wahrscheinlichkeit, sich *nicht* affektiv zu äußern, sank mit jeder zusätzlichen Einheit im Item zur Messung von sozialer Erwünschtheit (IRR = 0.556,  $p < .05$ ). Der nachgewiesene Einfluss von Ambivalentem Sexismus aus dem einfachen Modell blieb im Interaktionsmodell unter gleichen Vorzeichen bestehen (IRR = 0.493,  $p < .05$ ), d. h. mit jeder Einheit auf der Skala für ambivalent sexistische Einstellungen wurden rund zwei Gedanken dieser Kategorie weniger geäußert.

#### *Anwendung von Vorwissen als Verarbeitungsbasis*

Floss vorhandenes Vorwissen in die Rezeptionsgedanken ein, wurden sie dieser Kategorie zugeordnet. Dabei konnte es sich um einen Erfahrungsabgleich, Vorwissen oder eigene Überlegungen handeln. Eine typische Aussage lautete „Über Varianten positiver Diskriminierung wie eine Frauenquote zu diskutieren, ist wichtig - schade, dass es in der Form aber zu keinem echten Diskurs kommt“. Auch für diese Testung konnte keine Überdispersion nachgewiesen werden ( $\chi^2(1) = 0.002$ ,  $p = .998$ ). Versuchspersonen mit sozial erwünschtem Antwortverhalten schrieben signifikant mehr Rezeptionsgedanken nieder, die Vorwissen beinhalteten (IRR = 1.432,  $p < .05$ ). Im Hurdle-Modell trat ein signifikanter Effekt von *HS von Frauen* auf, hinsichtlich der Wahrscheinlichkeit keinen Rezeptionsgedanken zu äußern, dem Vorwissen der Proband:innen als Verarbeitungsbasis zugrundelag. Die Wahrscheinlichkeit dafür war stark erhöht (IRR = 4.005,  $p < .05$ ). Der Befund spricht für eine teilweise Bestätigung von H2c.

Auch in diesem Modell konnten keine signifikanten Moderationseffekte der Elaborationsvariablen auf die Beziehung zwischen der Rezeption der Stimuli und der Anzahl an Gedanken, die ein Vorwissen der Befragten erkennen ließen, identifiziert werden, was wiederum zur teilweisen Ablehnung von H2c führt. Einflussreich war im Vergleich zur Kontrollgruppe nur Hate Speech von Frauen im Hurdle-Modell, welches die Wahrscheinlichkeit *keinen* Gedanken dieser Kategorie niederzuschreiben deutlich erhöhte (IRR = 6.799,  $p < .05$ ).

**Tabelle 32:** Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Vorwissen)

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	0.12	0.57	-1.00	1.25	1.132
	Dummy HS Frauen	1.03	0.53	-0.01	2.08	2.807
	Dummy HS Männer	0.88	0.59	-0.28	2.04	2.405
	NFC	0.38	0.35	-0.30	1.06	1.463
	VW	-0.14	0.17	-0.46	0.19	0.872
	PI	0.36	0.25	-0.14	0.85	1.430
	ASI	-0.05	0.24	-0.52	0.41	0.948
	SE	0.36*	0.17	0.03	0.68	1.432
	Dummy Bildung (1 = hoch, 0 = Rest)	-0.98	0.52	-1.99	0.03	0.374
	Alter	0.02	0.02	-0.02	0.05	1.016
	Geschlecht (1 = m, 0 = w)	-0.54	0.35	-1.23	0.15	0.583
Hurdle-Modell	Konstante	-1.30*	0.58	-2.44	-0.16	0.272
	Dummy HS Frauen	1.39*	0.65	0.11	2.67	4.005
	Dummy HS Männer	0.07	0.64	-1.18	1.32	1.073
	NFC	-0.34	0.37	-1.06	0.38	0.714
	VW	0.11	0.22	-0.30	0.53	1.121
	PI	0.45	0.26	-0.05	0.96	1.576
	ASI	-0.34	0.31	-0.94	0.26	0.712
	SE	0.07	0.20	-0.33	0.46	1.069
	Dummy Bildung (1 = hoch, 0 = Rest)	0.55	0.57	-0.56	1.66	1.733
	Alter	0.001	0.02	-0.04	0.04	1.001
	Geschlecht (1 = m, 0 = w)	0.59	0.52	-0.42	1.60	1.806

Anmerkungen: HS = Hate Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenten Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\*  $p < .01$ . \*\*\* $p < .001$ .  $N = 82$

AIC = 220.29, BIC = 273.23,  $-2\log L = 176.287$

### Sonstige Gedankenursprünge

Die Modellierung des Einflusses auf die Nennung von Gedanken, die sich keiner der vorherigen Kategorien zuordnen ließ, war von abnorm hohen Standardfehlern im Zählmodell geprägt, was auf eine Missspezifikation des Modells hindeutete. Erklärt wurde dies mit der geringen Anzahl von insgesamt nur 28 Gedanken, die in dieser Kategorie zu finden waren, und mit der Menge an getesteten Variablen nicht vereinbar war. Aus diesem Grund wurde das Modell vereinfacht, indem im Zählmodell nur die dummycodierte Variable HS von Frauen (restliche Versuchsbedingungen als Referenzkategorie) aufgenommen und im Hurdle-Modell die bekannte Modellierung mit negativer Binomialverteilung durchgeführt wurde, da der Likelihood Ratio Test für diese Modellierung sprach,  $\chi^2(1) = 5.61$ ,  $p < .05$ . Im Zählmodell waren die Standardfehler für die Konstante weiterhin abnorm hoch, was die Aussagekraft der

Befunde einschränkt (vgl. Tabelle 33). Zwar zeigt sich ein Einfluss sozial erwünschten Antwortverhaltens im Hurdle-Modell, aber das Konfidenzintervall befindet sich bis auf die dritte Stelle vor dem Komma nahe am Wert Null, so dass dieser Befund nicht als gesichert gelten kann.

**Tabelle 33:** Vereinfachtes Hurdle-Modell mit mittelwertzentrierten Variablen zur Erklärung des Gedankenursprungs (hier: Sonstige Gedanken)

	Variable	B	SE	95%-CI für B		IRR
				LL	UL	
Zählmodell	Konstante	-9.22	132.91	-2.70	251.27	9.94e-5
	Dummy HS Frauen	-0.82	1.52	-3.80	2.16	0.440
Hurdle-Modell	Konstante	-0.65	0.74	-2.09	0.79	0.523
	Dummy HS Frauen	-1.44	0.99	-3.39	0.51	0.236
	Dummy HS Männer	-1.70	0.98	-3.62	0.22	0.183
	NFC	0.95	0.62	-0.26	2.16	2.593
	VW	0.05	0.31	-0.55	0.66	1.054
	PI	-0.47	0.44	-1.33	0.39	0.625
	ASI	0.87	0.51	-0.12	1.87	2.395
	SE	0.67*	0.34	0.003	1.34	1.955
	Dummy Bildung (1 = hoch, 0 = Rest)	-1.25	0.95	-3.11	0.61	0.286
	Alter	0.001	0.03	-0.06	0.06	1.001
	Geschlecht (1 = m, 0 = w)	-0.43	0.79	-1.99	1.12	0.649

Anmerkungen: HS = Hate Speech, NFC = Need for Cognition, VW = Vorwissen, PI = Persönliches Involvement, ASI = Ambivalenten Sexismus, SE = Soziale Erwünschtheit.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .  $N = 82$

AIC = 220.29, BIC = 273.23,  $-2\log L = 176.287$

Zusammenfassend lässt sich Hypothese H2c nur dahingehend bestätigen, dass die Rezeption von Hate Speech von Frauen, den gedanklichen Ursprung dahingehend beeinflusst, dass die Wahrscheinlichkeit steigt, keine Rezeptionsgedanken zu artikulieren, die Vorwissen beinhalten. Für die anderen Kategorien des gedanklichen Ursprungs konnten keine Rezeptionswirkungen nachgewiesen werden. Erwartet wurde zudem, dass der Gedankenursprung nach Rezeption von Hate Speech von Frauen bzw. Hate Speech von Männern von der Fähigkeit und Motivation zur Elaboration moderiert wird. Diese Annahme wurde wegen fehlender Evidenz verworfen.

### 6.2.2.3 Auswirkungen von Hate Speech auf die wahrgenommene Relevanz

In Studie I wurde gezeigt, dass die persönliche Relevanz gegenüber Hate Speech nur bedingt auf die Personengruppe der unbeteiligten Dritten anwendbar ist. In diesem Zusammenhang dient der Third-Person-Effekt von W. Phillips Davison (Davison, 1983; Gunther, 1991; Huck &

Brosius, 2007) als Erklärungsansatz für wahrgenommene Einflussrichtungen auf die wahrgenommene Relevanz von Hate Speech. Demnach schätzen Rezipierende den Einfluss von negativen (Medien-)Botschaften auf sich selbst als gering und auf Dritte als hoch ein. Diese Diskrepanz von Selbst- und Fremdwahrnehmung spricht für die Plausibilität der Annahme, wonach sekundäre Adressat:innen Hate Speech für sich persönlich als nicht relevant, wohl aber eine Relevanz für andere Nutzer:innen oder die Gesellschaft annehmen.

Da die Effektstärken (Persönliche Relevanz:  $f = .35$ , Gesellschaftliche Relevanz  $f = .24$ ) für die Berechnung von parametrischen Varianzanalysen zu gering waren (Schmider et al., 2010) und die Daten stark von der Normalverteilung abwichen bzw. zwar eine ausreichende Effektstärke erzielten (Relevanz für Andere:  $f = .51$ ), aber keine Varianzhomogenität vorlag, wurden nicht-parametrische Kruskal-Wallis-Tests zwischen den Versuchsbedingungen und den drei Relevanzvariablen durchgeführt, um robuste Ergebnisse zu erzielen. Die paarweisen Vergleiche erfolgten unter Anwendung einer Bonferroni-Korrektur. Zunächst konnten signifikante Unterschiede hinsichtlich der drei Relevanzvariablen festgestellt werden ( $\chi^2(2) = 35.95$ ,  $p < .001$ ), die sich auf den Vergleich zwischen persönlicher Relevanz und Relevanz für andere Nutzer:innen zurückführen ließen ( $T = -0.46$ ,  $p < .01$ ).

Das Stimulusmaterial in den Versuchsbedingungen beeinflusste signifikant die *wahrgenommene persönliche Relevanz (RZ)*  $H(2) = 20.42$ ,  $p < .001$ . Die RZ war am höchsten in der Kontrollgruppe ( $M = 2.06$ ,  $SD = 1.25$ ) und deutlich am unteren Skalenende in der Gruppe HS von Frauen ( $M = 1.54$ ,  $SD = 0.80$ ) und HS von Männern ( $M = 1.25$ ,  $SD = 0.70$ ). Der Unterschied zwischen der Kontrollgruppe und HS von Männern war signifikant ( $p < .001$ ,  $r = -0.48$ ), nicht jedoch zwischen HS von Frauen und der Kontrollgruppe ( $p = .080$ ,  $r = -0.24$ ) oder zwischen den beiden Hate Speech-Bedingungen ( $p = .137$ ,  $r = 0.21$ ). Hypothese H4b kann damit teilweise bestätigt werden.

Die *wahrgenommene Relevanz für andere Nutzer:innen (RZN)* war in den drei Experimentalbedingungen signifikant verschieden ( $H(2) = 22.10$ ,  $p < .001$ ). Die Kommentare in der Kontrollgruppe wurden als besonders relevant für andere Nutzer:innen bewertet ( $M = 2.65$ ,  $SD = 1.38$ ), gefolgt von HS von Frauen ( $M = 2.08$ ,  $SD = 1.28$ ) und deutlich weniger für HS von Männern ( $M = 1.20$ ,  $SD = 0.87$ ). Der Unterschied zwischen Kontrollgruppe und HS von Männern war groß und hoch signifikant ( $p < .001$ ,  $r = -0.50$ ) und nur von mittlerer Stärke zwischen den beiden Hate Speech-Bedingungen ( $p < .05$ ,  $r = 0.26$ ), nicht jedoch zwischen Hate Speech von Frauen und der Kontrollgruppe ( $p = .151$ ,  $r = -0.21$ ).

Keine signifikanten Unterschiede konnten bezüglich der *wahrgenommenen gesellschaftlichen Relevanz (RZG)* mittels Kruskal-Wallis-Test ermittelt werden ( $H(2) = 5.97$ ,  $p = .051$ ). Die

RZG bewegte sich in den Experimentalbedingungen auf einem ähnlichen Niveau; in der Kontrollgruppe war die Zustimmung im Mittel bei 2 Skalenpunkten ( $SD = 1.24$ ), nur knapp darunter in der Gruppe HS von Frauen ( $M = 1.95$ ,  $SD = 1.27$ ) und HS von Männern ( $M = 1.48$ ,  $SD = 0.83$ ).

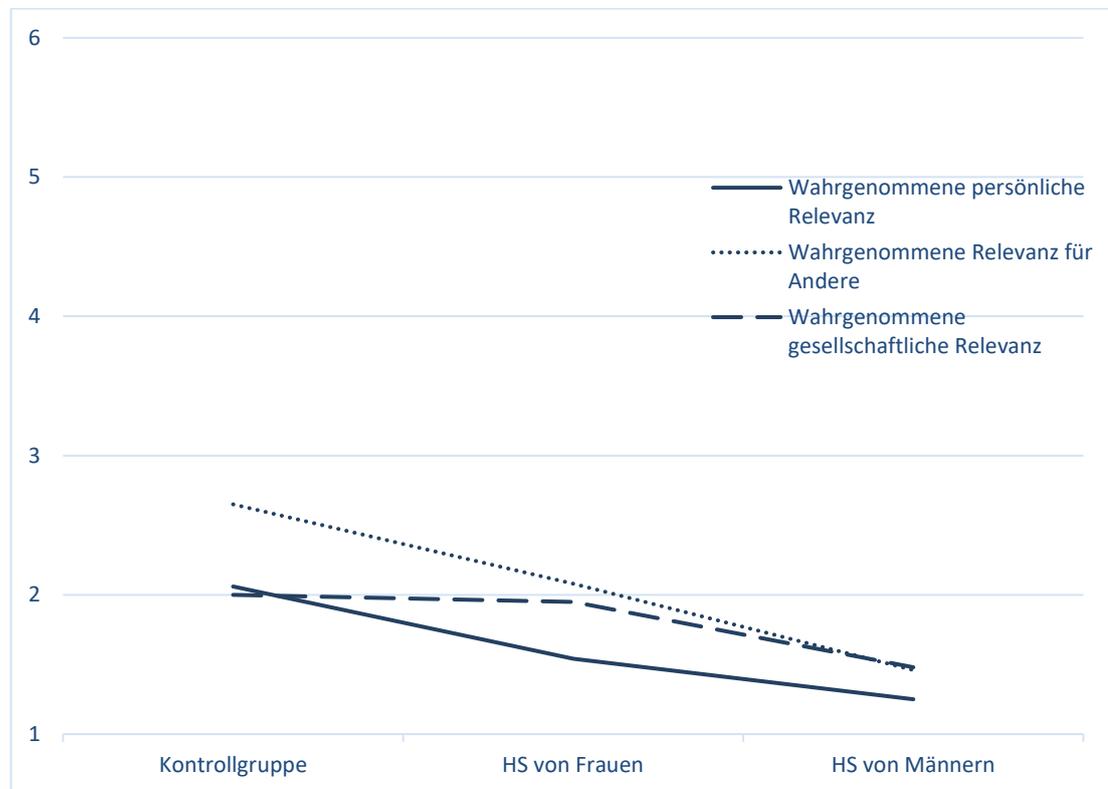
Im nächsten Schritt wurde geprüft, ob Ambivalenter Sexismus die Beziehung zwischen den Versuchsbedingungen und den drei Relevanzvariablen moderiert. Die Moderatorvariable wurde mittelwertzentriert und die Versuchsbedingungen als orthogonale Kontraste mit Helmert-Codierung in das Modell aufgenommen, d. h. die paarweisen Vergleiche beinhalteten im ersten Vergleich die Kontrollgruppe und die beiden Versuchsbedingungen gemeinsam und im zweiten Vergleich die beiden HS-Bedingungen untereinander. Um auch den Einfluss von Sozialer Erwünschtheit und soziodemografischen Merkmalen (Alter, Geschlecht und Bildung) zu kontrollieren, wurde ein Modell mit diesen Kovariaten gerechnet. Trotz der schiefen Verteilungen der abhängigen Variablen konnte Linearität auf Basis der Residuenplots und annähernde Normalverteilung der Residuen angenommen werden. Durchgeführt wurden die Analysen mit PROCESS von Hayes (2018) unter Anwendung von 10000 Bootstrapping-Iterationen, um robuste Konfidenzintervalle und Standardfehler für die Koeffizienten zu gewinnen. Zusätzlich wurden robuste HC3-Schätzer zur Modellberechnung eingesetzt.

#### *Wahrgenommene persönliche Relevanz*

Das Basismodell erwies sich als hoch signifikant mit  $F(5, 127) = 10.80$ ,  $p < .001$  und konnte einen mittelstarken Beitrag zur Varianzaufklärung leisten ( $R^2 = .23$ ). Ambivalenter Sexismus erhöhte die RZ signifikant ( $b = 0.34$ ,  $p < .001$ ). Ein Moderationseffekt von Ambivalentem Sexismus auf die Beziehung zwischen den Experimentalbedingungen und der wahrgenommenen persönlichen Relevanz konnte jedoch nicht nachgewiesen werden ( $\Delta R^2 = .01$ ,  $F(2, 127) = 0.56$ ,  $p = .575$ , 95% CI [Vergleich 1: -0.60, 0.22 und Vergleich 2: -0.42, 0.29]). Das Modell mit Kovariaten war ebenfalls signifikant mit  $F(9, 123) = 4.92$ ,  $p < .001$ ,  $R^2 = .29$ , es konnte auch hier kein Moderationseffekt nachgewiesen werden ( $\Delta R^2 = .05$ ,  $F(2, 123) = 0.33$ ,  $p = .723$ , 95% CI [Vergleich 1: -0.53, 0.24 und Vergleich 2: -0.42, 0.26]). Soziale Erwünschtheit erhöhte die persönliche Relevanz signifikant ( $b = 0.13$ ,  $p < .05$ ).

Angenommen wurde in Hypothese H4b, dass das Geschlecht der Kommentierenden die wahrgenommene persönliche Relevanz beeinflusst. Dies konnte für HS von Männern im Vergleich zur Kontrollgruppe nachgewiesen werden, nicht jedoch für HS von Frauen. Die postulierte Interaktion in Hypothese H4b zwischen dem Geschlecht von Hate Speech-Produzent:innen und ambivalent sexistischen Einstellungen konnte nicht belegt werden. Hypothese H4b muss daher teilweise abgelehnt werden.

**Abbildung 16:** Zusammenhang zwischen Experimentalbedingungen und wahrgenommener persönlicher Relevanz



Anmerkungen: HS = Hate Speech; Messung der wahrgenommenen Relevanz erfolgte als Zustimmung (1 = „stimme überhaupt nicht zu“ – 6 = „stimme voll und ganz zu“).  $N = 133$ .

#### Wahrgenommene Relevanz für andere Nutzer:innen

Das vereinfachte Moderationsmodell konnte einen signifikanten Beitrag zur Varianzaufklärung leisten mit  $F(5, 127) = 10.06$ ,  $p < .001$ ,  $R^2 = .23$ . Ambivalent sexistische Einstellungen beeinflussten auch die wahrgenommene RZN positiv ( $b = 0.37$ ,  $p < .01$ ), nicht jedoch den Zusammenhang zwischen Experimentalbedingungen und RZN ( $\Delta R^2 = .003$ ,  $F(2, 127) = 0.30$ ,  $p = .740$ , 95% CI [Vergleich 1: -0.58, 0.34 und Vergleich 2: -0.67, 0.42]). Im erweiterten Modell mit Kovariaten ( $F(9, 123) = 5.21$ ,  $p < .01$ ,  $R^2 = .29$ ) hatten Ambivalenter Sexismus ( $b = 0.30$ ,  $p < .05$ ) und Soziale Erwünschtheit ( $b = 0.17$ ,  $p < .05$ ) einen positiven Effekt auf RZN, während höher Gebildete die experimentellen Stimuli als weniger relevant für Andere wahrnahmen ( $b = -0.42$ ,  $p = .051$ , heteroskedastisch robustes Bootintervall 95% CI [-0.84, -0.02]). Der Moderationseffekt zwischen RZN und Ambivalentem Sexismus war nicht signifikant ( $\Delta R^2 = .002$ ,  $F(2, 123) = 0.16$ ,  $p = .856$ , 95% CI [Vergleich 1: -0.50, 0.43 und Vergleich 2: -0.71, 0.42]).

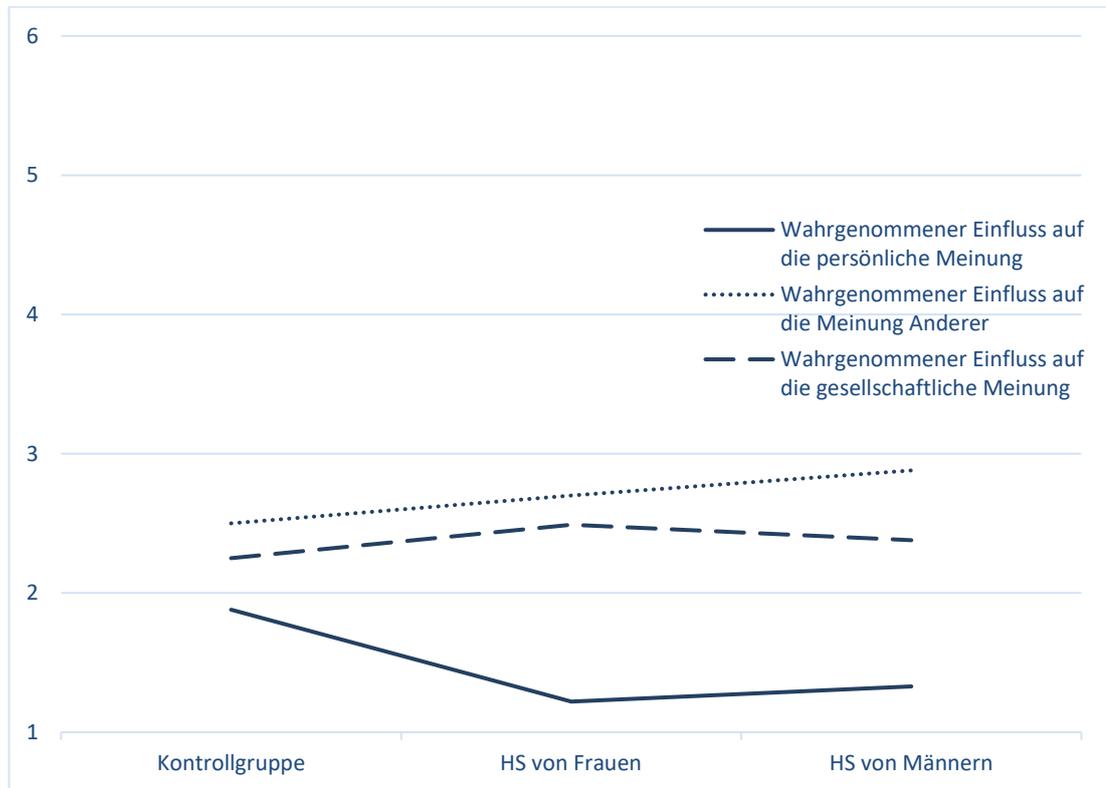
*Wahrgenommene gesellschaftliche Relevanz*

Das vereinfachte Moderationsmodell trug signifikant zur Varianzaufklärung bei ( $F(5, 127) = 8.00, p < .001, R^2 = .13$ ). Auch für diesen Fall distanzierender Relevanzwahrnehmung wirkte sich Ambivalenter Sexismus positiv aus ( $b = 0.34, p < .01$ ). Ein signifikanter Moderationseffekt von Ambivalentem Sexismus auf den Zusammenhang zwischen dem Geschlecht der Hate Speech-Produzent:innen und der wahrgenommenen gesellschaftlichen Relevanz konnte nicht nachgewiesen werden ( $\Delta R^2 = .003, F(2, 127) = 0.18, p = .837, 95\% \text{ CI}$  [Vergleich 1: -0.51, 0.32 und Vergleich 2: -0.40, 0.60]). Für das Modell mit Kovariaten konnte ebenfalls Signifikanz erreicht werden ( $F(9, 123) = 4.20, p < .001, R^2 = .18$ ). Der Moderationseffekt war für diese Beziehung ebenfalls nicht signifikant ( $\Delta R^2 = .002, F(2, 123) = 0.19, p = .831, 95\% \text{ CI}$  [Vergleich 1: -0.46, 0.39 und Vergleich 2: -0.38, 0.62]).

*6.2.2.4 Auswirkungen von Hate Speech auf den wahrgenommenen Meinungsfluss*

Da für die Effektstärke ein mittlerer Wert  $f = 0.29$  berechnet wurde, diese jedoch nicht ausreicht, um Varianzanalysen zu berechnen (Schmider et al., 2010), wurden nicht-parametrische Kruskal-Wallis-Tests mit Bonferroni-Korrektur für die paarweisen Vergleiche durchgeführt. Der Einfluss auf die persönliche Meinung unterschied sich signifikant in den Versuchsgruppen,  $H(2) = 11.20, p < .01$ . Geringsten Zuspruch erhielten die experimentellen Stimuli in der Gruppe HS von Frauen ( $M = 1.22, SD = 0.48$ ), wenn auch der Durchschnittswert in der Gruppe HS von Männern nur ein Zehntel einer Skaleneinheit höher lag ( $M = 1.33, SD = 0.88$ ) und der Experimentalstimulus der Kontrollgruppe nur knapp einen halben Skaleneinheit mehr Zustimmung erzielte ( $M = 1.88, SD = 1.35$ ). Paarweise Vergleiche zeigten, dass der Unterschied zwischen Kontrollgruppe und HS von Männern eine mittlere Stärke aufwies und signifikant war ( $p < .01, r = 0.32$ ), ähnlich zwischen der Kontrollgruppe und HS von Frauen ( $p < .05, r = 0.29$ ), nicht jedoch zwischen den beiden HS-Gruppen ( $p = .980, r = 0.003$ ). Unterschiede zwischen den Versuchsbedingungen waren weder für den wahrgenommenen Einfluss auf die Meinung anderer Nutzer:innen signifikant ( $H(2) = 1.75, p = .416$ ) noch für den wahrgenommenen Einfluss auf die gesellschaftliche Meinung ( $H(2) = 1.30, p = .523$ ). Die Inspektion der Residuendiagramme konnte nicht-lineare Zusammenhänge aller drei Meinungsvariablen nicht ausschließen, weswegen auf die Berechnung von Moderationsanalysen verzichtet wurde.

**Abbildung 17:** Zusammenhang zwischen den Experimentalbedingungen und dem wahrgenommenen Einfluss auf Meinungen

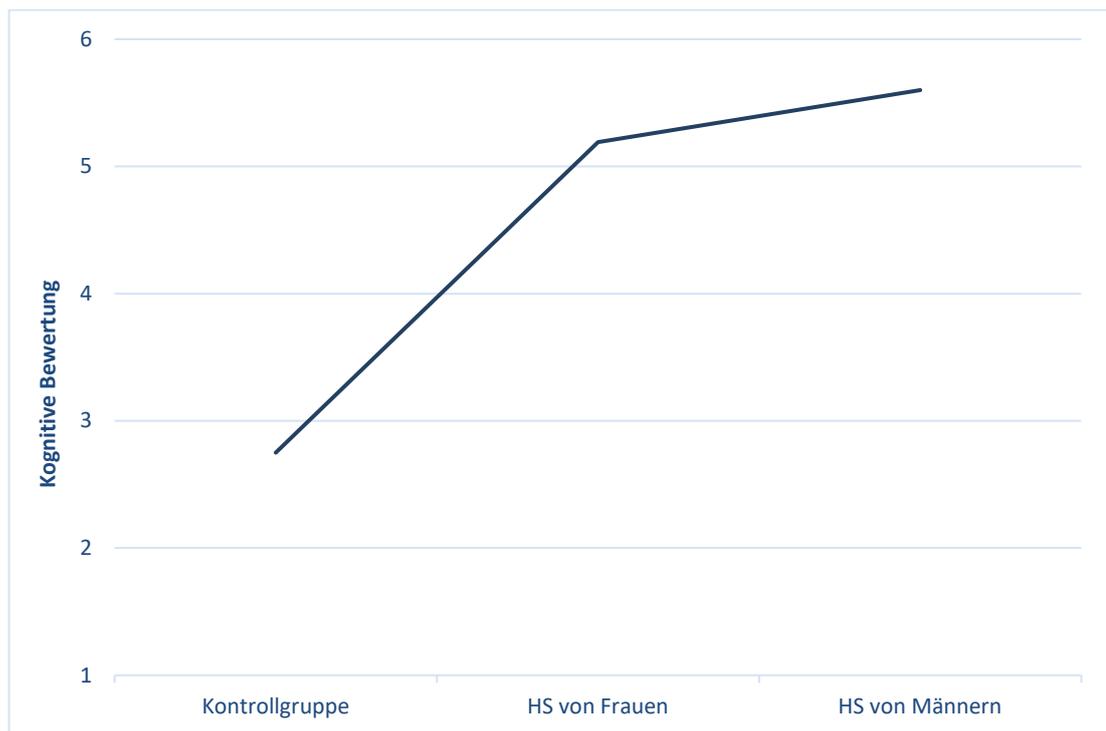


Anmerkungen: HS = Hate Speech; Messung des wahrgenommenen Einflusses auf Meinungen erfolgte als Zustimmung (1 = „stimme überhaupt nicht zu“ – 6 = „stimme voll und ganz zu“).  $N = 133$

#### 6.2.2.5 Zusammenhang zwischen wahrgenommener Relevanz und Bewertung

Die *kognitive Bewertung* der experimentellen Stimuli diente dazu, die Einstellung gegenüber Hate Speech zu erfassen. Diese war in den beiden Hate Speech-Bedingungen deutlich negativer ( $M_{HS\_Frauen} = 5.19$ ,  $SD_{HS\_Frauen} = 0.94$ ;  $M_{HS\_Männer} = 5.60$ ,  $SD_{HS\_Männer} = 0.61$ ) als in der Kontrollgruppe ( $M = 2.75$ ,  $SD = 1.10$ ). Folglich war der Unterschied zwischen den Gruppen signifikant (Welch-Test  $F(2) = 123.39$ ,  $p < .001$ ). Mittels G\*Power wurde eine Effektstärke von  $f = 1.44$  berechnet, was einer sehr starken Effektstärke nach Cohen entspricht. Paarweise Vergleiche mit Games-Howell-Korrektur zeigten, dass HS von Männern (2.44, 95%-CI[1.91, 2.97]) bzw. HS von Frauen im Vergleich zur Kontrollgruppe (2.85, 95%-CI[2.42, 3.29]) signifikant (jeweils  $p < .001$ ) negativer bewertet wurde. Die Unterschiede zwischen den beiden Hate Speech-Bedingungen waren zu gering, um Signifikanz zu erreichen.

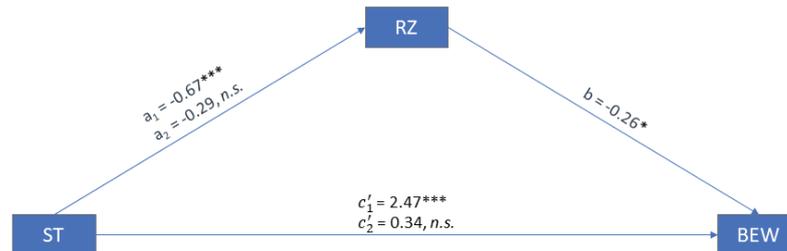
Abbildung 18: Kognitive Bewertung von Hate Speech



Anmerkungen: HS = Hate Speech; Bewertung erfolgte als Zustimmung (1 = „stimme überhaupt nicht zu“ – 6 = „stimme voll und ganz zu“) zu Adjektiven mit negativer Valenz und Umpolung positiver Adjektive.  $N = 133$ .

In Hypothese 5 wird ein medierter Einfluss der persönlichen Relevanzwahrnehmung auf die Bewertung der Stimuli nach deren Rezeption angenommen, was aus dem ELM-Modell abgeleitet wurde. Anhand des SPSS-Makros PROCESS von Hayes (2018) wurde ein einfaches Mediationsmodell aufgestellt, das die Beziehung zwischen den Versuchsbedingungen (UV) und deren kognitiver Bewertung (AV) durch die wahrgenommene persönliche Relevanz vermittelt (M). Die Vergleiche zwischen den Versuchsbedingungen wurden als orthogonale Kontraste mittels Helmert-Codierung implementiert. Verglichen wurden zunächst Kontrollgruppe und die HS-Bedingungen (Vergleich 1) und schließlich die beiden HS-Bedingungen untereinander (Vergleich 2). Dies hatte den Vorteil, dass nicht nur der Vergleich mit einer Referenzkategorie berücksichtigt, sondern zwischen allen Bedingungen ermöglicht wurde. Um Verletzungen der Normalverteilungsannahme und Heteroskedastizität entgegenzuwirken, wurden 10000 Bootstrapping-Iterationen durchgeführt und robuste HC3-Schätzer eingesetzt. Abbildung 19 visualisiert das konzeptuelle Diagramm des Modells und bildet in der Beschriftung die Regressionskoeffizienten der Pfade ab.

**Abbildung 19:** Einfaches Mediationsmodell zur Erklärung des Zusammenhangs von Geschlecht über die wahrgenommene Relevanz auf die kognitive Bewertung



*Anmerkungen:* ST = Stimulusmaterial, RZ = Wahrgenommene persönliche Relevanz, BEW = Bewertung des Stimulus. Angegeben sind unstandardisierte Regressionskoeffizienten.

$N = 133$

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$

Die Koeffizienten  $a$ ,  $b$  und  $c'$  geben die Pfade im Mediationsmodell wieder.  $c'_1$  und  $c'_2$  stellen dabei den relativen direkten Effekt für die beiden Vergleiche zwischen dem variierten Geschlecht im Stimulusmaterial für die Hate Speech-Experimentalgruppen und der Kontrollgruppe (1) sowie zwischen den beiden Hate Speech-Experimentalgruppen untereinander (2) dar. Die Indexierung der Pfade  $a$  und  $c'$  ist den beiden Vergleichen geschuldet. Beispielfhaft: Der Koeffizient für den Pfad  $a_1$  steht für den Vergleich zwischen den Hate Speech-Experimentalgruppen und der Kontrollgruppe, der Koeffizient für den Pfad  $a_2$  steht für den Vergleich zwischen den Hate Speech-Experimentalgruppen untereinander.

Indirekter Effekt: HS  $\rightarrow$  RZ  $\rightarrow$  BEW, ( $b = .18$ , 95% CI[.03, .36]) und ( $b = .08$ , 95% CI[-.01, .19]).

Hate Speech geht im Vergleich zur Kontrollgruppe (Vergleich 1) negativ mit der persönlichen Relevanzwahrnehmung einher ( $b = -.67$ ,  $p < .001$ ), während der Vergleich zwischen der Geschlechterkategorie für Hate Speech-Produzent:innen untereinander (Vergleich 2) keine Signifikanz aufweist ( $b = -.29$ ,  $p = .087$ ). Die Relevanzwahrnehmung ist negativ mit der Bewertung des Stimulusmaterials korreliert ( $b = -.26$ ,  $p < .05$ ). Da höhere Werte für eine negativere Bewertung stehen, heißt dies, dass mit jeder zusätzlichen Einheit auf der Relevanzskala die Bewertung um 0.26 Einheiten auf der Bewertungsskala positiver ist.

Im Mediationsmodell quantifiziert der *indirekte Effekt*  $ab$  den Einfluss der UV auf die AV über die Pfade UV  $\rightarrow$  Mediator(en) (Pfad  $a$ )  $\rightarrow$  AV (Pfad  $b$ ). Daraus ergibt sich ein signifikanter indirekter Effekt  $a_1b$  für den Vergleich 1 von HS von Männern bzw. HS von Frauen im Vergleich zur Kontrollgruppe über die persönliche Relevanzwahrnehmung auf die kognitive Bewertung des Stimulusmaterials ( $b = .18$ , 95% CI[.03, .36]). Der indirekte Effekt für den zweiten Vergleich  $a_2b$  war nicht signifikant ( $b = .08$ , 95% CI[-.01, .19]). Die Bewertung des Stimulusmaterials HS von Männern wurde im Vergleich zu HS von Frauen nicht über die persönliche Relevanzwahrnehmung vermittelt.

Der *direkte Effekt*  $c'$  quantifiziert den Unterschied bezüglich der AV zwischen zwei Fällen aus der UV, bei konstant gehaltenen Werten für die Mediatorvariablen. Der direkte Effekt  $c'_1$  des Stimulusmaterials war für den ersten Vergleich signifikant ( $t(128) = 12.865, p < .001$  und 95% CI[2.09, 2.85]). Bei konstant gehaltener persönlicher Relevanzwahrnehmung war die kognitive Bewertung der Stimuli in den beiden Hate Speech-Bedingungen um 2.47 Einheiten negativer als in der Kontrollgruppe. Der direkte Effekt  $c'_2$  der Geschlechterkategorie der Hate Speech-Kommentierenden auf die Bewertung des Stimulusmaterials war hingegen nicht signifikant ( $t(128) = 1.967, p = .051$  und 95% CI[-.002, .68]).

Der *totale Effekt*  $c$  im Mediationsmodell drückt die Summe aus indirekten und direkten Effekten aus, also die Auswirkung des Prädiktors auf die abhängige Variable. Der totale Effekt  $c$  wird für beide Vergleiche ebenfalls getrennt ausgegeben: Vergleich 1 weist einen deutlichen totalen Effekt auf mit  $c_1 = 2.65$ , der signifikant von Null verschieden ist ( $t(130) = 14.377, p < .001$  und 95% CI[2.28, 3.01]). Der totale Effekt für den Vergleich 2 ist mit  $c_2 = 0.42$  geringer, aber ebenfalls signifikant von Null verschieden ( $t(130) = 2.307, p < .05$  und 95% CI[.06, .77]).

Hypothese 5 wird abgelehnt, da der Einfluss der Geschlechterkategorie auf die Bewertung von Hate Speech nicht nachgewiesen werden konnte. Auch die erwartete Mediation der persönlichen Relevanzwahrnehmung auf das Verhältnis zwischen der Geschlechterkategorie der Hate Speech-Kommentierenden und der kognitiven Bewertung konnte nicht belegt werden. Signifikante Unterschiede hinsichtlich des Mediationseinflusses der Relevanzwahrnehmung auf die Bewertung des Stimulusmaterials konnten nur aus dem Vergleich der Hate Speech-Gruppen mit der Kontrollgruppe nachgewiesen werden, d. h. Hate Speech beeinflusst sowohl die Relevanzwahrnehmung als auch die Bewertung signifikant von der Kontrollbedingung verschieden.

#### *Erweitertes Mediationsmodell*

Der Ursprung des ELM in der Konsumentenforschung wird daran deutlich, dass in den klassischen Studien nur der Einfluss der persönlichen Relevanz auf die kognitive Bewertung geprüft wurde. Da in Studie II neben der wahrgenommenen persönlichen Relevanz auch die wahrgenommene Relevanz für andere Nutzer:innen sowie die wahrgenommene gesellschaftliche Relevanz erhoben wurden, können die Beziehungen zwischen diesen Variablen in einem erweiterten Mediationsmodell analysiert werden. Die folgenden Analysen werden allerdings nicht mittels Hypothesentest geprüft, sondern sind ergänzend zu verstehen.

Mittels SPSS-Makro *PROCESS* von Hayes (2018) wurde ein Modell aufgesetzt, welches die drei Relevanz-Variablen als Mediatoren aufnahm, um die Beziehung zwischen den Versuchsbedingungen (UV) und der kognitiven Bewertung (AV) als Einstellungsvariable zu untersuchen. Ein zweites Modell inkludierte drei soziodemografische Kovariate, um für deren Einfluss zu kontrollieren. Diese waren Alter und Geschlecht sowie die dummycodierte Variable für hohe Bildung. Die Vergleiche zwischen den Versuchsbedingungen wurden wiederum als orthogonale Kontraste mittels Helmert-Codierung implementiert (Vergleich 1: Kontrollgruppe vs. HS von Frauen und HS von Männern und Vergleich 2: die beiden HS-Bedingungen untereinander). Dies hatte den Vorteil, dass nicht nur der Vergleich mit einer Referenzkategorie berücksichtigt, sondern zwischen allen Bedingungen ermöglicht wurde. Um Verletzungen der Normalverteilungsannahme und Heteroskedastizität entgegenzuwirken, wurden 10000 Bootstrapping-Iterationen durchgeführt und robuste HC3-Schätzer eingesetzt. Tabelle 34 enthält die Ergebnisse der Modelle ohne und mit Kovariaten. Abbildung 20 visualisiert das konzeptuelle Diagramm des vereinfachten Modells ohne Kovariate und bildet in der Beschriftung die Regressionskoeffizienten der Pfade ab.

Hate Speech von Frauen bzw. von Männern unterschied sich nicht hinsichtlich der wahrgenommenen persönlichen oder gesellschaftlichen Relevanz, es konnte aber ein signifikanter Befund für die wahrgenommene Relevanz für andere Nutzende nachgewiesen werden: Demnach wurde Hate Speech von Männern (HS Frauen = 0) als weniger relevant für Andere wahrgenommen ( $b = -0.36, p < .05, 95\% \text{ CI } [-0.66, -0.07]$ ). Im Vergleich zur Kontrollgruppe wurde Hate Speech signifikant als persönlich weniger relevant ( $b = -0.67, p < .001$  und  $95\% \text{ CI } [-1.08, -0.30]$ ), dafür als gesellschaftlich relevanter wahrgenommen ( $b = 0.41, p < .001, 95\% \text{ CI } [-1.08, -0.30]$ ). Dieser Vergleich war für die wahrgenommene Relevanz für andere Nutzende nicht signifikant. Die drei Mediatoren übten keinen statistisch signifikanten Einfluss auf die abhängige Variable aus, beeinflussten aber einander stark positiv ( $p < .001$ ). Dies ist insofern überraschend, als dass zwischen den Relevanzvariablen und der Bewertung des Stimulusmaterials ein mittelstarker negativer Zusammenhang besteht (Persönliche Relevanz und Bewertung:  $r = -.43, p < .01$ ; Relevanz für Andere und Bewertung:  $r = -.49, p < .01$ ; Gesellschaftliche Relevanz und Bewertung:  $r = -.30, p < .01$ ). Darüber hinaus nahmen höher Gebildete einen geringeren Einfluss von Hate Speech auf die Relevanz anderer Nutzende wahr ( $b = -0.30, p < .05, 95\% \text{ CI } [-0.57, -0.02]$ ).

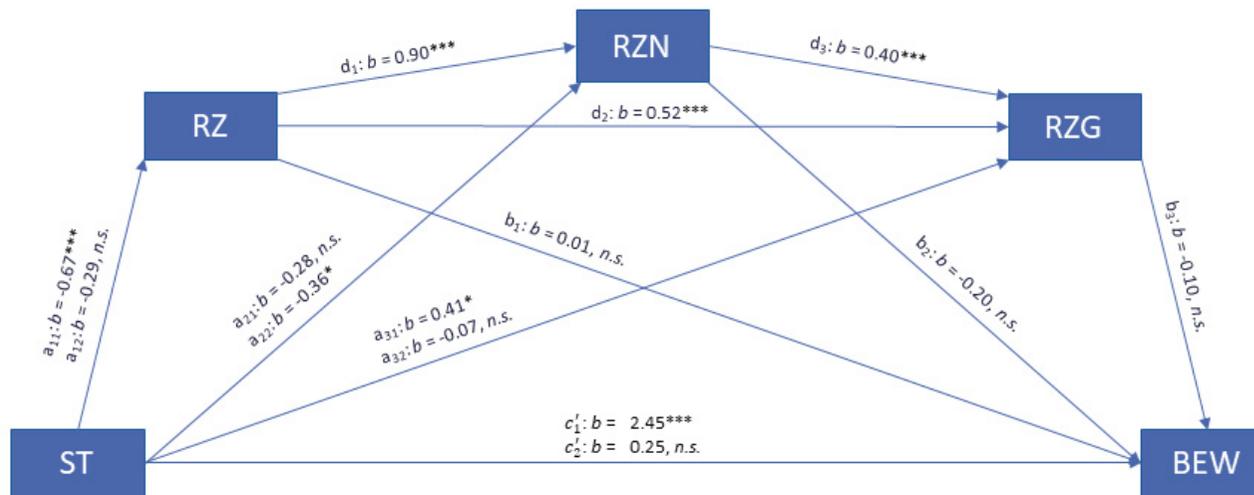
Vorliegend war der *direkte Effekt* der experimentellen Stimuli auf deren kognitive Bewertung stark und hoch signifikant  $\Delta R^2 = .46, F(2, 127) = 73.01, p < .001$ , was auf den ersten Vergleich zurückgeführt werden kann ( $b = 2.45, p < .001, 95\% \text{ CI } [2.04, 2.85]$ ), nicht jedoch auf den

Vergleich von HS von Frauen und HS von Männern untereinander. Bei konstant gehaltenen Relevanzwahrnehmungen war die kognitive Bewertung der Stimuli in den beiden Hate Speech-Bedingungen um 2.45 Einheiten negativer als in der Kontrollgruppe.

Der *totale Effekt* in der vorliegenden Untersuchung war signifikant mit  $\Delta R^2 = .68$ ,  $F(2, 130) = 121.87$ ,  $p < .001$ . Da die UV als kategorielle Variable in das Modell einfluss, wurden relative totale Effekte berechnet. Der relative totale Effekt war sowohl zwischen Kontrollgruppe und HS von Frauen sowie HS von Männern signifikant ( $b = 2.65$ ,  $p < .001$ , 95% CI [2.28, 3.01]) als auch zwischen den beiden Hate Speech-Bedingungen (Frauen vs. Männer) ( $b = 0.42$ ,  $p < .05$ , 95% CI [0.06, 0.77]).

Vorliegend waren die *indirekten Effekte* zwischen den experimentellen Stimuli, der Relevanzwahrnehmung und der resultierenden kognitiven Bewertung minimal, da sie sich zwischen 0.002 und 0.12 bewegten. Alle Bootstrapping-Konfidenzintervalle für die indirekten Effekte enthielten den Wert 0, weswegen von nicht signifikanten Effekten ausgegangen werden musste. Da die indirekten Effekte als eigentliche Mediationseffekte verstanden werden, spricht dies für die Nicht-Berücksichtigung von wesentlichen Mediatoren (Hayes, 2018; Zhao et al., 2010) und bedarf weiterer empirischer Untersuchungen. In diesem Zusammenhang sprechen Zhao und Kolleg:innen von einer „direct nonmediation“ (Zhao et al., 2010, S. 200), welche durch signifikante direkte Effekte gekennzeichnet sind, ohne dass indirekte Effekte auftreten.

**Abbildung 20:** Erweitertes Mediationsmodell zur Erklärung des Zusammenhangs von Geschlecht über die wahrgenommene Relevanz auf die kognitive Bewertung



**Anmerkungen:**  $N = 133$ . ST = Stimulusmaterial, RZ = Wahrgenommene persönliche Relevanz, RZN = Wahrgenommene Relevanz für andere Nutzer:innen, RZG = Wahrgenommene gesellschaftliche Relevanz, BEW = Bewertung von Hate Speech. Die Koeffizienten  $a$ ,  $b$ ,  $d$ , und  $c'$  geben die Pfade im Mediationsmodell wieder.  $c'_1$  und  $c'_2$  stellen dabei den relativen direkten Effekt für die beiden Vergleiche zwischen den beiden Hate Speech Experimentalgruppen und der Kontrollgruppe (1) sowie zwischen den beiden Hate Speech Experimentalgruppen (2) dar. Die Indexierung der Pfade ausgehend von HS als unabhängige Variable zu den drei Mediationsvariablen ist zweistellig. Beispielhaft steht  $a_{21}$  für den Pfad, der von HS zur zweiten Mediationsvariable RZN geht und den ersten Gruppenvergleich zwischen den beiden Hate Speech Experimentalgruppen und der Kontrollgruppe darstellt. Der Pfad  $a_{32}$  bildet die Beziehung von HS und der dritten Mediationsvariable RZG sowie den Gruppenvergleich zwischen den beiden Experimentalgruppen ab. Die indirekten Effekte sind gering ( $< .12$ ) und statistisch nicht signifikant. Angaben für  $b$  sind unstandardisierte Regressionskoeffizienten.  $*p < .05$ .  $**p < .01$ .  $***p < .001$ .

**Tabelle 34: Mediationsanalysen zur Erklärung des Einflusses von Relevanz auf die Bewertung von sexistischem Hate Speech**

Variable	Totaler Effekt von X auf Y				M <sub>1</sub> Wahrgenommene persönliche Relevanz				M <sub>2</sub> Wahrgenommene Relevanz für Andere				M <sub>3</sub> Wahrgenommene gesellschaftliche Relevanz				Y Kognitive Bewertung der experimentellen Stimuli			
	B	SE	95% CI		B	SE	95% CI		B	SE	95% CI		B	SE	95% CI		B	SE	95% CI	
			LL	UL			LL	UL			LL	UL			LL	UL			LL	UL
Konstante	4.51***	0.08	4.36	4.67	1.62***	0.08	1.46	1.78	0.61***	0.15	0.30	0.91	0.15	0.12	-0.06	0.39	5.10***	0.17	4.78	5.44
KG vs HS (X <sub>1</sub> )	2.65***	0.18	2.28	3.01	-0.67**	0.20	-1.08	-0.30	-0.28	0.19	-0.66	0.09	0.41*	0.13	0.16	0.64	2.45***	0.20	2.07	2.83
HS <sub>Frauen</sub> vs. HS <sub>Männer</sub> (X <sub>2</sub> )	0.42*	0.18	0.06	0.77	-0.29	0.17	-0.61	0.03	-0.36*	0.15	-0.66	-0.07	-0.07	0.14	-0.35	0.20	0.25	0.17	-0.07	0.56
RZ (M <sub>1</sub> )	-	-	-	-	-	-	-	-	0.90***	0.09	0.72	1.09	0.52***	0.11	0.29	0.72	0.01	0.18	-0.30	0.37
RZN (M <sub>2</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	0.40***	0.10	0.21	0.59	-0.20	0.14	-0.48	0.03
RZG (M <sub>3</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.10	0.19	-0.45	0.21
	$R^2 = .68$				$R^2 = .12$				$R^2 = .60$				$R^2 = .67$				$R^2 = .72$			
	$F(2, 130) = 121.87, p < .001$				$F(2, 130) = 7.79, p < .001$				$F(3, 129) = 52.93, p < .001$				$F(4, 128) = 38.15, p < .001$				$F(5, 127) = 59.85, p < .001$			
Konstante	4.43***	0.35	3.73	5.13	1.16***	0.29	0.58	1.71	1.00**	0.35	0.30	1.63	-0.04	0.29	-0.56	0.55	4.97***	0.43	4.23	5.81
KG vs HS (X <sub>1</sub> )	2.66***	0.20	2.26	3.06	-0.65**	0.20	-1.04	-0.27	-0.31	0.21	-0.71	0.09	0.40**	0.15	0.12	0.67	2.46***	0.23	2.03	2.89
HS <sub>Frauen</sub> vs. HS <sub>Männer</sub> (X <sub>2</sub> )	0.43*	0.19	0.06	0.81	-0.32	0.18	-0.66	0.01	-0.39*	0.15	-0.68	-0.10	-0.08	0.14	-0.34	0.19	0.25	0.18	-0.08	0.59
RZ (M <sub>1</sub> )	-	-	-	-	-	-	-	-	0.89***	0.09	0.72	1.07	0.50***	0.12	0.27	0.71	-0.001	0.19	-0.31	0.38
RZN (M <sub>2</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	0.40***	0.10	0.21	0.59	-0.19	0.14	-0.48	0.04
RZG (M <sub>3</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-0.11	0.19	-0.46	0.21
Alter (C <sub>1</sub> )	0.001	0.01	-0.01	0.01	0.01	0.01	-0.002	0.02	-0.01	0.01	-0.02	0.01	0.003	0.01	-0.01	0.01	0.002	0.01	-0.01	0.01
Geschlecht (C <sub>2</sub> )	-0.10	0.18	-0.46	0.26	0.34	0.19	-0.01	0.71	0.03	0.15	-0.26	0.33	0.11	0.13	-0.13	0.37	0.01	0.18	-0.32	0.37
Dummy Bildung (1=hoch) (C <sub>3</sub> )	0.19	0.19	-0.19	0.57	-0.25	0.18	-0.59	0.11	-0.30*	0.14	-0.57	-0.02	0.04	0.10	-0.15	0.24	0.06	0.19	-0.30	0.41
	$R^2 = .68$				$R^2 = .19$				$R^2 = .61$				$R^2 = .67$				$R^2 = .72$			
	$F(5, 127) = 52.31, p < .001$				$F(5, 127) = 3.69, p < .01$				$F(6, 126) = 27.23, p < .001$				$F(7, 125) = 27.22, p < .001$				$F(8, 124) = 38.81, p < .001$			

Anmerkungen: N = 133. HS = Hate Speech, KG = Kontrollgruppe, RZ = Wahrgenommene persönliche Relevanz, RZN = Wahrgenommene Relevanz für andere Nutzende, RZG = Wahrgenommene gesellschaftliche Relevanz.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

### 6.2.2.6 *Zusammenhang zwischen dem antizipierten Einfluss auf die Meinung und Bewertung*

In diesem Abschnitt werden analog zu dem oben präsentierten erweiteren Mediationsmodell weitere Ergebnisse präsentiert, welche den Einfluss auf die persönliche sowie die gesellschaftliche Meinung und schließlich den Einfluss auf die Meinung Dritter als Mediatoren aufzeigen sollen. Die Analysen im vorigen Abschnitt zeigten eine „direct nonmediation“ der Relevanzwahrnehmungen an. Diese sind gekennzeichnet durch signifikante direkte Effekte, ohne dass indirekte Effekte auftreten (s. o.). Die Befunde für die Mediation der Rezeption von Hate Speech und deren Bewertung durch den wahrgenommenen Einfluss auf Meinungen müssen ebenfalls als eine solche „direct nonmediation“ bewertet werden. Die indirekten Effekte waren gering ( $< .09$ ) und statistisch nicht signifikant. Da das Modell dennoch wichtige Erkenntnisse zu diesen Zusammenhängen liefert, werden die Befunde dargestellt. Gewählt wurde die gleiche Vorgehensweise wie im vorigen Abschnitt dargestellt, jedoch mit den drei Variablen als Mediatoren, die einen Meinungseinfluss auf die drei bekannten Perspektiven (individuell, andere Nutzer:innen betreffend sowie gesellschaftlich) annehmen.

Im Vergleich zur Kontrollgruppe antizipierten Versuchspersonen in den beiden HS-Bedingungen (HS von Frauen und HS von Männern) einen signifikant geringeren Einfluss auf die persönliche Meinung ( $b = -0.60, p < .01, 95\% \text{ CI } [-1.02, -0.22]$ ). Höher Gebildete lehnten einen Einfluss auf die eigene Meinung besonders intensiv ab ( $b = -0.52, p < .01, 95\% \text{ CI } [-0.90, -0.16]$ ). Antizipiert wurde hingegen ein signifikant höherer Einfluss auf die Meinung anderer Nutzer:innen ( $b = 0.75, p < .001, 95\% \text{ CI } [0.35, 1.16]$ ), jedoch nicht auf die gesellschaftliche Meinung ( $95\% \text{ CI } [-0.19, 0.65]$ ). Hinsichtlich eines Einflusses auf Meinungen (individuell, auf Dritte sowie gesellschaftlich) konnten nach Rezeption von HS von Männern keine Unterschiede im Vergleich zu HS von Frauen belegt werden. Der Einfluss der Mediatoren auf die kognitive Bewertung des Stimulusmaterials war nicht signifikant. Auch in dieser Analyse beeinflussten sich die Mediatoren stark positiv (vgl. Tabelle 35).

Der direkte Effekt der rezipierten Stimuli auf deren Bewertung war hoch signifikant  $\Delta R^2 = .55, F(2, 127) = 81.24, p < .001$ . Der Unterschied zwischen Kontrollgruppe und HS-Bedingungen ergab einen hoch signifikanten direkten Effekt ( $b = 2.56, p < .001, 95\% \text{ CI } [2.18, 2.97]$ ) für die Bewertung des Stimulusmaterials. Auch der Vergleich zwischen den beiden HS-Bedingungen erbrachte signifikante Unterschiede in der Bewertung ( $b = 0.46, p < .05, 95\% \text{ CI } [0.13, 0.81]$ ), d. h. unter Konstanthaltung der Mediatorvariablen bewerteten Versuchspersonen in der Bedingung HS von Männern das Stimulusmaterial um 0.46 Einheiten negativer als Versuchspersonen, die HS von Frauen rezipierten.

**Tabelle 35: Mediationsanalysen zur Erklärung des wahrgenommenen Einflusses von Meinungen auf die Bewertung von sexistischem Hate Speech**

Variable	Totaler Effekt von X auf Y				M <sub>1</sub> Wahrgenommener Einfluss auf die persönliche Meinung				M <sub>2</sub> Wahrgenommene Einfluss auf die Meinung Anderer				M <sub>3</sub> Wahrgenommener Einfluss auf die gesellschaftliche Meinung				Y Kognitive Bewertung der experimentellen Stimuli			
	95% CI <sup>†</sup>				95% CI				95% CI				95% CI				95% CI			
	B	SE	LL	UL	B	SE	LL	UL	B	SE	LL	UL	B	SE	LL	UL	B	SE	LL	UL
Konstante	4.51***	0.08	4.36	4.67	1.47***	0.08	1.32	1.64	1.56***	0.17	1.22	1.88	0.57**	0.21	0.21	1.00	4.73***	0.18	4.40	5.06
KG vs HS (X <sub>1</sub> )	2.65***	0.18	2.28	3.01	-0.60**	0.21	-1.02	-0.22	0.75***	0.21	0.35	1.16	0.23	0.22	-0.19	0.65	2.56***	0.21	2.18	2.97
HS <sub>Frauen</sub> vs. HS <sub>Männer</sub> (X <sub>2</sub> )	0.42*	0.18	0.06	0.77	0.12	0.15	-0.16	0.42	0.08	0.30	-0.50	0.67	-0.23	0.23	-0.68	0.21	0.46*	0.18	0.13	0.81
PM (M <sub>1</sub> )	-	-	-	-	-	-	-	-	0.77***	0.07	0.64	0.93	0.33**	0.10	0.11	0.52	-0.15	0.14	-0.41	0.09
MN (M <sub>2</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	0.49***	0.11	0.28	0.69	-0.08	0.09	-0.27	0.08
GM (M <sub>3</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.10	0.09	-0.06	0.28
	R <sup>2</sup> = .68 F(2, 130) = 121.87, p < .001				R <sup>2</sup> = .08 F(2, 130) = 4.84, p < .01				R <sup>2</sup> = .29 F(3, 129) = 42.11, p < .001				R <sup>2</sup> = .45 F(4, 128) = 27.46, p < .001				R <sup>2</sup> = .69 F(5, 127) = 48.48, p < .001			
Konstante	4.43***	0.35	3.73	5.13	1.83***	0.37	1.19	2.59	1.41**	0.52	0.39	2.38	0.50	0.43	-0.26	1.39	4.67***	0.48	3.80	5.56
KG vs HS (X <sub>1</sub> )	2.66***	0.20	2.26	3.06	-0.63**	0.21	-1.05	-0.25	0.76**	0.24	0.32	1.22	0.24	0.24	-0.21	0.70	2.58***	0.24	2.14	3.04
HS <sub>Frauen</sub> vs. HS <sub>Männer</sub> (X <sub>2</sub> )	0.43*	0.19	0.06	0.81	0.07	0.16	-0.23	0.38	0.08	0.31	-0.50	0.67	-0.23	0.24	-0.68	0.23	0.47*	0.19	0.13	0.85
PM (M <sub>1</sub> )	-	-	-	-	-	-	-	-	0.77***	0.08	0.64	0.95	0.35**	0.10	0.14	0.54	-0.13	0.15	-0.41	0.12
MN (M <sub>2</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	0.49***	0.11	0.28	0.69	-0.08	0.10	-0.27	0.09
GM (M <sub>3</sub> )	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.09	0.09	-0.06	0.28
Alter (C <sub>1</sub> )	0.001	0.01	-0.01	0.01	-0.004	0.01	-0.02	0.01	0.002	0.01	-0.02	0.02	-4e-04	0.01	-0.01	0.01	3e-04	0.01	-0.01	0.01
Geschlecht (C <sub>2</sub> )	-0.10	0.18	-0.46	0.26	0.21	0.20	-0.16	0.60	0.05	0.24	-0.43	0.49	-0.03	0.20	-0.40	0.34	-0.07	0.19	-0.43	0.30
Dummy Bildung (1=hoch) (C <sub>3</sub> )	0.19	0.19	-0.19	0.57	-0.52**	0.19	-0.90	-0.16	0.05	0.25	-0.43	0.54	0.16	0.20	-0.21	0.52	0.11	0.21	-0.27	0.51
	R <sup>2</sup> = .68 F(5, 127) = 52.31, p < .001				R <sup>2</sup> = .14 F(5, 127) = 2.64, p < .05				R <sup>2</sup> = .29 F(6, 126) = 21.13, p < .001				R <sup>2</sup> = .45 F(7, 125) = 17.91, p < .001				R <sup>2</sup> = .69 F(8, 124) = 30.97, p < .001			

Anmerkungen: N = 133. HS = Hate Speech, KG = Kontrollgruppe, PM = Wahrgenommener Einfluss auf die persönliche Meinung, MN = Wahrgenommener Einfluss auf die Meinung Anderer, GM = Wahrgenommener Einfluss auf die gesellschaftliche Meinung. \*p < .05. \*\* p < .01. \*\*\*p < .001.

Der totale Effekt bildet den Zusammenhang zwischen UV und AV ab und ist daher identisch mit den Befunden im vorigen Abschnitt, weswegen aus Redundanzgründen auf die Darstellung verzichtet wird.

#### 6.2.2.7 *Einflüsse auf Einstellungen zu Frauenkarrieren*

Innerhalb des ELM werden Effekte auf Einstellungen erwartet, die sich als Einstellungsbildung, -verstärkung oder -änderungen offenbaren (Petty & Haugvedt, 1992), wobei davon ausgegangen wird, dass starke Einstellungen widerstandsfähig gegenüber Beeinflussungen sind. Vorliegend wurden ambivalent sexistische Einstellungen als solche vorhandenen Voreinstellungen verstanden. Geprüft wurde zunächst, ob der starke Zusammenhang zwischen ambivalent sexistischen Einstellungen und Einstellungen zu Frauenkarrieren ( $r = .51, p < .01$ ) sich zwischen den beiden Experimentalgruppen unterscheidet (Hypothese H6a). Da auch Geschlechtereffekte denkbar sind, wurde als nächstes die Interaktion zwischen Ambivalentem Sexismus und dem Geschlecht der Versuchspersonen untersucht (Hypothese H6b). In einem dritten Schritt wurde geprüft, ob der Zusammenhang zwischen ambivalent sexistischen Einstellungen und Einstellungen zu Frauenkarrieren für männliche und weibliche Versuchspersonen anders ausfällt, je nachdem welchem Hate Speech Stimulus sie ausgesetzt wurden (Hypothese H6c). Aus methodischen Gründen wurden die Items zur Erhebung der Einstellungen zu Frauenkarrieren der Kontrollgruppe nicht vorgelegt, da deren Stimulus keinen thematischen Bezug hatte.

Nach visueller Prüfung der Residuendiagramme konnte approximative Linearität für sämtliche Beziehungen angenommen werden. Die Moderationsanalysen wurden mittels PROCESS-Makros für SPSS von Hayes (2018) durchgeführt. Dazu wurde Bootstrapping mit 10000 Stichproben angewandt, um Abweichungen von der Normalverteilung auszugleichen und HC3-Schätzverfahren für robuste Standardfehler. Die mehrfache Testung der Variablen erforderte ein Absenken des Signifikanzniveaus auf  $\alpha = 0.01$ , um eine  $\alpha$ -Fehlerkumulierung zu minimieren.

Im ersten Modell wurde die Beziehung zwischen ambivalent sexistischen Einstellungen (UV) und Einstellungen zu Frauenkarrieren (AV), Hate Speech als dummymodierter Moderator (HS von Frauen als Referenzkategorie) und soziale Erwünschtheit als Kovariate modelliert. Das Modell war signifikant und konnte zu einer mittelstarken Varianzaufklärung beitragen,  $F(4, 80) = 13.50, p < .001, R^2 = .26$ . Ambivalent sexistische Einstellungen der Versuchspersonen führten in der Experimentalgruppe HS von Frauen (Moderator = 0) zu einem hoch signifikanten Anstieg der Zustimmung zu ablehnenden Aussagen, welche dazu dienten, die

Einstellungen zu Frauenkarrieren zu messen ( $b = 0.59, p < .001, 99\% \text{ CI } [0.38, 0.83]$ ). Weder die Moderatorvariable noch die Interaktion zwischen UV und dieser waren signifikant ( $\Delta R^2 = .002, F(1, 80) = 0.28, p = .596$ ). Hypothese H6a wird daher abgelehnt. Soziale Erwünschtheit beeinflusste die Einstellungen zu Frauenkarrieren nicht ( $b = -0.07, p = .456, 99\% \text{ CI } [-0.33, 0.16]$ ).

Um Haupteffekte des Modells zu untersuchen, wurde eine lineare Regression ohne Interaktionsterm durchgeführt ( $F(3, 81) = 9.60, p < .001, R^2 = .26$ ), die zeigte, dass der größte Anteil der Varianzaufklärung vom Einfluss ambivalent sexistischer Einstellungen auf negative Einstellungen gegenüber Frauenkarrieren ausging, ( $b = 0.58, p < .001$ ). Weder die dummycodierte Versuchsbedingung Hate Speech noch die Kovariate soziale Erwünschtheit war signifikant.

Analog zum ersten Modell wurde die Beziehung zwischen ambivalent sexistischen Einstellungen (UV) und Einstellungen zu Frauenkarrieren (AV) und dem Geschlecht der Versuchspersonen als dummycodierter Moderator (weibliche Befragte als Referenzkategorie) modelliert. Das Modell war signifikant ( $F(4, 80) = 13.97, p < .001$ ) mit einem mittelstarken Beitrag zur Varianzaufklärung,  $R^2 = .30$ . Die Prädiktoren wurden zur besseren Interpretierbarkeit der Koeffizienten mittelwertzentriert. Ambivalent sexistische Einstellungen bei Frauen (Moderator = 0) führten zu einem signifikanten Anstieg der Zustimmung auf der Skala zur Messung von Einstellungen zu Frauenkarrieren ( $b = 0.60, p < .001, 99\% \text{ CI } [0.37, 0.82]$ ). Bei durchschnittlich ambivalenten Einstellungen gab es keinen Unterschied zwischen den Geschlechtern hinsichtlich ihrer Zustimmung zu den Items der abhängigen Variablen. Es konnte jedoch eine schwach signifikante Moderation nachgewiesen werden,  $\Delta R^2 = .04, F(1, 80) = 6.54, p < .05, 99\% \text{ CI } [-0.97, -0.02]$ . Hypothese H6b wird daher angenommen. Es konnte kein Einfluss sozialer Erwünschtheit nachgewiesen werden. Der konditionale Effekt von ambivalent sexistischen Einstellungen auf Einstellungen zu Frauenkarrieren war für Frauen deutlich größer als für Männer. Weibliche Versuchspersonen, die sich um einen Punkt auf der ASI-Skala unterschieden, erzielten auf der EF-Skala 0.81 Einheiten mehr ( $p < .001$ ), während für Männer dieser Wert mit 0.36 Einheiten deutlich geringer war ( $p < .01$ ). Weibliche Versuchspersonen, die sich um einen Punkt auf der ASI-Skala unterschieden, wiesen also verstärkt negativere Einstellungen gegenüber Frauenkarrieren auf.

Das dritte Modell wurde als moderierte Moderation zwischen ambivalent sexistischen Einstellungen (UV), Einstellungen zu Frauenkarrieren (AV), Experimentalgruppen (Moderator 1) und dem Geschlecht der Befragten (Moderator 2) modelliert, welches sich als signifikant erwies ( $F(8, 76) = 7.31, p < .001$ ) und einen mittleren Beitrag zur Varianzaufklärung hatte,

$R^2 = .32$ . Die Prädiktoren wurden zur besseren Interpretierbarkeit der Koeffizienten mittelwertzentriert. Ambivalent sexistische Einstellungen von weiblichen Versuchspersonen (Moderator 2 = 0) in der Experimentalbedingung HS von Frauen (Moderator 1 = 0) erhöhten die Zustimmung zur Skala, welche feindselige Einstellungen zu Frauenkarrieren maß, um 0.59 Einheiten ( $p < .001$ , 99% CI [0.34, 0.86]). Die Koeffizienten der beiden Moderatoren erreichten keine Signifikanz. Keine Signifikanz erreichten auch die Interaktionen zwischen ambivalent sexistischen Einstellungen und Hate Speech, zwischen Hate Speech und dem Geschlecht der Versuchspersonen, sowie die unktionale Interaktion zwischen der UV und den beiden Moderatoren,  $\Delta R^2 = .002$ ,  $F(1, 76) = 0.24$ ,  $p = .628$ . Schwache Signifikanz wurde für die Interaktion aus UV und dem Geschlecht der Versuchspersonen nachgewiesen: Männliche Befragte zeigten bei durchschnittlichen ambivalent sexistischen Einstellungen signifikant geringere Zustimmung zur Skala zur Messung von Einstellungen zu Frauenkarrieren ( $b = -0.48$ ,  $p < .05$ , 99% CI [-1.12, -0.04]), waren also weniger feindselig gegenüber modernen Frauenkarrieren. Hypothese H6c wird daher teilweise angenommen.

#### 6.2.2.8 Einflüsse auf Verhaltensintentionen

Der Zusammenhang zwischen Hate Speech und jeweils den vier Variablen für Verhaltensintentionen enthielt nicht-lineare Beziehungen und konnte auch keine multivariate Normalverteilung vorweisen, weswegen statt dem multivariaten MANOVA-Verfahren als nicht-parametrisches Verfahren in einem ersten Schritt Kruskal-Wallis-Tests mit Bonferroni-korrigierten Post-hoc-Tests durchgeführt wurden. Als zweiten Schritt wurden vier nicht-parametrische Quade-ANCOVAs (Quade, 1967) gerechnet, um für die Kovariaten Alter, Bildung und Geschlecht zu kontrollieren. Aufgrund der mehrfachen Testung der Variablen wurde das Signifikanzniveau auf  $\alpha = 0.01$  gesenkt, um einer Kumulierung des  $\alpha$ -Fehlers zuvorzukommen. Hypothesentests, um Hypothese 7b zu testen, konnten daher nicht erfolgen. Präsentiert werden die Ergebnisse der nicht-parametrischen Tests.

*Zustimmende Verhaltensintentionen in Form von Likes und Kommentaren* wurden in der Kontrollgruppe als geringfügig wahrscheinlicher bewertet ( $M = 1.83$ ,  $SD = 1.42$ ) als in den beiden Experimentalgruppen, denn HS-Kommentare von Männern ( $M = 1.46$ ,  $SD = 1.18$ ) und HS-Kommentare von Frauen ( $M = 1.73$ ,  $SD = 1.47$ ) lagen nur knapp darunter. Dieser knappe Unterschied war nicht signifikant ( $H(2) = 2.24$ ,  $p = .327$ ). Soziale Erwünschtheit beeinflusste die zustimmenden Verhaltensintentionen nicht ( $\text{tau-b} = .10$ ,  $p = .188$ ). Die Kontrolle für Kovariate ergab keine signifikante Test-Statistik der Quade-ANCOVA ( $F(2, 130) = 1.52$ ,  $p = .223$ ).

*Kommentare, die Ablehnung ausdrückten*, konnten zwischen den Gruppen hoch signifikante Unterschiede aufdecken ( $H(2) = 18.08$ ,  $p < .001$ ). In der Kontrollgruppe bewegte sich die

Ablehnung im unteren Skalenbereich ( $M = 1.42$ ,  $SD = 1.45$ ) und lag in den beiden Experimentalgruppen rund eine Skaleneinheit höher ( $M_{HS\_Frauen} = 2.49$ ,  $SD_{HS\_Frauen} = 1.81$  und  $M_{HS\_Männer} = 2.56$ ,  $SD_{HS\_Männer} = 1.66$ ). Quantifiziert wurde die Stärke des Unterschieds mittels Pearsons Korrelationskoeffizient  $r$ . Zwischengruppenvergleiche wiesen signifikante Unterschiede von mittlerer Stärke zwischen Kontrollgruppe und HS von Frauen nach ( $p < .01$ ,  $r = -0.36$ ). Der Unterschied zwischen Kontrollgruppe und HS von Männern war noch deutlicher ( $p < .001$ ,  $r = -0.41$ ). Der Effekt zwischen den beiden Experimentalgruppen war unbedeutend gering und folglich nicht signifikant ( $p = 1.000$ ,  $r = 0.02$ ). Soziale Erwünschtheit beeinflusste die Wahrscheinlichkeit dieser Verhaltensintention nicht ( $\text{tau-b} = .03$ ,  $p = .685$ ). Die Kontrolle für die soziodemografischen Kovariate war signifikant  $F(2, 130) = 9.86$ ,  $p < .001$ , weswegen weitere Prüfungen notwendig wurden. Zwar zeigten sich Signifikanzen auf dem Niveau  $\alpha = 0.05$  bei Versuchspersonen mit höherer Bildung, diese muss aber aus oben genannten Gründen abgelehnt werden. Signifikante Unterschiede wurden bei weiblichen Versuchspersonen festgestellt: Diese sahen es als wahrscheinlicher an, HS von Frauen im Vergleich zur Kontrollgruppe mit einem ablehnenden Kommentar zu begegnen ( $p < .01$ ,  $r = -0.41$ ). Der Bildungsgrad und das Alter beeinflussten in den Experimentalgruppen nicht die Bereitschaft, Ablehnung per Kommentar auszudrücken.

In dem *Melden* von Social-Media-Aktivitäten anderer Nutzer:innen drückt sich ebenfalls Ablehnung aus, welche sich als statistisch signifikant erwies  $H(2) = 22.71$ ,  $p < .001$ . Keinen Anlass zur Ablehnung gaben die Kommentare in der Kontrollgruppe ( $M = 1.33$ ,  $SD = 0.93$ ), wohl aber HS-Kommentare von Frauen ( $M = 2.43$ ,  $SD = 1.89$ ) und HS-Kommentare von Männern ( $M = 2.71$ ,  $SD = 1.82$ ). Im Vergleich der Gruppen untereinander erreichten die Effekte eine mittlere bis hohe Stärke, bei der die Absicht, Kommentare in der Kontrollgruppe zu melden, überzufällig weniger wahrscheinlich eingestuft wurde als in der Gruppe, in der HS von Frauen ( $p < .01$ ,  $r = -0.36$ ) bzw. HS von Männern ( $p < .001$ ,  $r = -0.48$ ) rezipiert wurde. Keine signifikanten Unterschiede konnten im Vergleich zwischen den beiden HS-Gruppen aufgedeckt werden ( $p = 1.000$ ,  $r = 0.09$ ). Soziale Erwünschtheit beeinflusste die Wahrscheinlichkeit, die Kommentare zu melden, nicht ( $\text{tau-b} = -.03$ ,  $p = .638$ ). Die nicht-parametrische Quade-ANCOVA ergab eine signifikante Teststatistik für Einflüsse soziodemografischer Variablen ( $F(2, 126) = 18.12$ ,  $p < .001$ ). Keine Signifikanz auf dem Niveau  $\alpha = 0.01$  konnte die Untersuchung nachweisen, ob weibliche Versuchspersonen, eher die Absicht verfolgen, die rezipierten Kommentare zu melden. Dagegen artikulierten männliche Probanden die Absicht, Hate von Männern im Vergleich zu dem Stimulusmaterial aus der Kontrollgruppe zu melden ( $p < .001$ ,  $r = -0.50$ ), als signifikant wahrscheinlicher. Versuchspersonen mit geringerer Bildung waren ebenfalls bereiter, HS von Männern im Vergleich zur Kontrollgruppe zu melden

( $p < .001$ ,  $r = -0.53$ ). Der Befund hatte eine hohe Effektstärke und war hochsignifikant. Die Korrelation zwischen der Rezeption der Stimuli und der Intention sich darüber zu beschweren, sank nach Kontrolle für das Alter von  $r = .35$  ( $p < .01$ ) auf  $r = .31$  ( $p < .01$ ).

Die größte Zustimmung erhielt die Aussage „*Ich lese die Kommentare, ohne etwas zu tun*“. In der Kontrollgruppe lag die Zustimmung im oberen Skalenbereich ( $M = 4.92$ ,  $SD = 1.83$ ) und mehr als eine Skaleneinheit niedriger in der Gruppe, in der HS von Frauen ( $M = 3.65$ ,  $SD = 2.00$ ) bzw. von Männern ( $M = 3.94$ ,  $SD = 1.94$ ) rezipiert wurden. Diese Unterschiede waren statistisch signifikant ( $H(2) = 14.97$ ,  $p < .01$ ). Die Unterschiede zwischen Kontrollgruppe und HS von Frauen erreichten eine mittlere Effektstärke und waren signifikant ( $p < .01$ ,  $r = 0.38$ ), ebenso die Befunde des Vergleichs zwischen Kontrollgruppe und HS von Männern ( $p < .01$ ,  $r = 0.31$ ). Die geringe Effektstärke zwischen HS von Frauen und HS von Männern war nicht signifikant  $p = 1.000$ ,  $r = -0.09$ . Soziale Erwünschtheit beeinflusste das Ignorieren als Verhaltensintention nicht ( $\text{tau-b} = -.09$ ,  $p = .222$ ). Die signifikante Test-Statistik der Quade-ANCOVA für soziodemografische Merkmale erforderte weitergehende Überprüfungen ( $F(2, 130) = 7.48$ ,  $p < .01$ ). Es gab keine Unterschiede hinsichtlich des Bildungsgrads und des Geschlechts der Versuchspersonen bezüglich der Absicht die Stimuli zu ignorieren. Allerdings erreichte der negative Korrelationskoeffizient nach Prüfung der Altersvariable Signifikanz  $r = -.21$  ( $p = .015$ ) auf  $r = -.25$  ( $p < .01$ ).

Hypothese H8 konnte nicht mittels Moderationsanalyse überprüft werden, da die Voraussetzung der Linearität verletzt war. Der Zusammenhang zwischen ambivalent sexistischen Einstellungen und der artikulierten Wahrscheinlichkeit Zustimmung auszudrücken wurde mittels nicht-parametrischer Kendall-tau-b-Korrelation untersucht ( $\text{tau-b} = .17$ ,  $p < .05$ ,  $N = 133$ ). Diese zeigte einen schwach positiven Zusammenhang an, d. h. je ausgeprägter ambivalent sexistische Einstellungen waren, desto höher wurde die Wahrscheinlichkeit zuzustimmen bewertet. Bei näherer Betrachtung des Zusammenhangs zwischen Ambivalentem Sexismus und der Wahrscheinlichkeit der Zustimmung in den einzelnen Versuchsbedingungen (KG:  $\text{tau-b} = .13$ ,  $p = .320$ ,  $N = 48$ ; HS Frauen:  $\text{tau-b} = .29$ ,  $p = .055$ ,  $N = 37$ ; HS Männer:  $\text{tau-b} = .16$ ,  $p = .217$ ,  $N = 48$ ) konnte keine statistische Signifikanz belegt werden.

Die Prüfung von Hypothese H9 war ebenfalls davon betroffen, dass keine Linearität zwischen den Variablen angenommen werden konnte. Erwartet wurde, dass der Zusammenhang zwischen ambivalent sexistischen Einstellungen und der artikulierten Wahrscheinlichkeit mit Zustimmung auf das Stimulusmaterial zu reagieren, von Einstellungen zu Frauenkarrieren moderiert wird. Die nicht-parametrische Kendall-tau-b-Korrelation zwischen ambivalent sexistischen Einstellungen und Einstellungen zu modernen Frauenkarrieren konnte einen

mittelstarken bis starken Zusammenhang nachweisen ( $\tau\text{-}b = .48, p < .001, N = 85$ ). Der Zusammenhang zwischen der Wahrscheinlichkeit zustimmend zu reagieren und Einstellungen zu Frauenkarrieren, war nicht signifikant ( $p = .122, N = 85$ ).

Auch für die Prüfung der Hypothese H10 wurden zunächst die Tests der Voraussetzungen durchgeführt. Es zeigte sich, dass von den drei VI-Variablen, die Ablehnung ausdrücken, nämlich VI2 *ablehnender Kommentar*, VI3 *Melden* und VI4 *Ignorieren* nur letzterer annähernde Linearität nach der visuellen Überprüfung der Residuendiagramme versprach. Daher werden die Befunde nur für den Zusammenhang zwischen persönlicher Relevanzwahrnehmung (UV) und dem Ignorieren als bekundeter Verhaltensabsicht (AV), moderiert von den Einstellungen zu Frauenkarrieren, dargestellt. UV und Moderator wurden mittelwertzentriert in das Modell aufgenommen, um die Interpretation der Befunde zu vereinfachen. Um auch den Einfluss der soziodemografischen Variablen Alter, Geschlecht, Bildung sowie sozialer Erwünschtheit zu berücksichtigen, wurde ein weiteres Moderationsmodell gerechnet. Durchgeführt wurden die Berechnungen mit dem SPSS-Makro PROCESS von Hayes (2018). Es wurden jeweils 10000 Bootstrapping-Iterationen gerechnet, um robuste Konfidenzintervalle zu erhalten, und für heteroskedastisch konsistente Schätzer der Standardfehler kamen HC3-Schätzer zum Einsatz.

Das Basismodell war signifikant ( $F(3, 81) = 4.53, p < .01$ ), konnte jedoch nur einen geringen Anteil zur Varianzaufklärung beitragen mit  $R^2 = .10$ . Die wahrgenommene persönliche Relevanz als Prädiktor übte keinen Einfluss auf die artikulierte Wahrscheinlichkeit aus, die rezipierten Kommentare zu ignorieren. Diese Verhaltensintention wurde von Einstellungen zu modernen Frauenkarrieren positiv beeinflusst ( $b = .58, p < .01$ ), d. h. bei durchschnittlicher Relevanzwahrnehmung gaben die Versuchspersonen mit jeder zusätzlichen Einheit auf der Skala zu modernen Frauenkarrieren es als wahrscheinlicher an, die Kommentare zu ignorieren. Erwartet wurde eine Interaktion zwischen Relevanzwahrnehmung und Einstellungen zu modernen Frauenkarrieren, was jedoch nicht belegt werden konnte ( $\Delta R^2 = .01, (1, 81) = 0.96, p = .330$ ). Das Modell mit Kovariaten war signifikant mit einer mittleren Varianzaufklärung,  $R^2 = .18, F(7, 77) = 3.13, p < .01$ . Weder Prädiktor noch Kovariaten übten auf die AV einen signifikanten Einfluss aus. Höhere Werte auf der Skala zu Einstellungen zu Frauenkarrieren erhöhten die artikulierte Wahrscheinlichkeit, die rezipierten Kommentare zu ignorieren, signifikant mit ( $b = .57, p < .01$ ). Ein Moderationseffekt konnte auch im erweiterten Modell mit Kovariaten nicht belegt werden ( $\Delta R^2 = .004, (1, 77) = 0.43, p = .514$ ).

### 6.2.3 Einordnende Zusammenfassung der Befunde

In der ersten Studie hatte sich gezeigt, dass die Geschlechterkategorie als Botschaftsmerkmal in soziodemografisch homogen(er)en Populationen nur minimale Effekte bewirkt. Aufbauend auf den Befunden aus Studie I wurden im Rahmen der zweiten Studie Rezeptionswirkungen von sexistischem Hate Speech von Männern und Frauen an einer soziodemografisch heterogenen Stichprobe untersucht.

Die Anzahl der Rezeptionsgedanken lag mit durchschnittlich gerade einmal drei von bis zu zehn möglichen Gedanken etwas unter dem Output in Studie I, was auf eine eher periphere Verarbeitung des Stimulusmaterials hindeutet. Gestützt wird diese Interpretation vom positiven Einfluss sozial erwünschten Antwortverhaltens auf die Gedankenanzahl. Die Erwartung der Forscherin möglichst viele Rezeptionsgedanken zu lesen, wurde von den Teilnehmenden antizipiert und führte zu einem erhöhten Aufkommen an Eintragungen in die offenen Antwortfelder. Des Weiteren schrieben männliche im Vergleich zu weiblichen Versuchspersonen weniger Rezeptionsgedanken nieder. Es gab aber auch Versuchspersonen, die bei durchschnittlicher Motivation zur Elaboration fähig waren und für einen Anstieg an Rezeptionsgedanken sorgten: Wer sich kompetent fühlte an einer Diskussion zum Thema Frauenkarrieren in der IT-Branche teilzunehmen, stellte dies auch im Rahmen des Experiments unter Beweis, zumindest was die reine Gedankenanzahl betraf.

Auch in dieser Stichprobe herrschten negative Rezeptionsgedanken vor, wenn sie auch von sozial erwünschten Antworttendenzen beeinflusst wurden. Die Valenz der Rezeptionsgedanken war in den beiden Experimentalbedingungen deutlich negativer als in der Kontrollgruppe und negativer gegenüber Hate Speech von Männern im Vergleich zu Hate Speech von Frauen. Argumentiert man auf Basis der Sexismusforschung, deutet der Befund darauf hin, dass Hate Speech von Männern als schwerwiegenderer Verstoß gegen das Globalstereotyp des kompetenten Mannes, denn gegen das Globalstereotyp der freundlichen und empathischen Frau wahrgenommen wurde. In Verbindung mit der eher geringen Zahl der Rezeptionsgedanken liegt aus ELM-Perspektive der Schluss nahe, dass die Stimuli schnell verfügbare Stereotype zu misogynen Männern aktivierten und der Widerspruch aus weiblichem Globalstereotyp und Hate Speech nur peripher verarbeitet wurde.

Die inhaltliche Betrachtung der Rezeptionsgedanken untermauert zusätzlich den Befund der peripheren Verarbeitung. Gerade einmal ein Viertel der Gedanken ließ Vorwissen erkennen oder stellte eigene Überlegungen auf. Der überwiegende Teil war durch den Stimulus induziert, indem der Stimulus beschrieben wurde, vor allem indem Empörung und Ablehnung zum Ausdruck gebracht oder indem irrelevante Gedanken geäußert wurden. Darüber hinaus

hatten die Motivation und Fähigkeit zur Informationsverarbeitung keinen Einfluss auf die Beziehung zwischen der Rezeption von Hate Speech und der inhaltlichen Ausgangsbasis der Rezeptionsgedanken.

Sozial erwünschtes Antwortverhalten spielte eine dominante Rolle bei der Erklärung der inhaltlichen Ausrichtung der Rezeptionsgedanken. Es senkte die Wahrscheinlichkeit irrelevante Gedanken zu äußern und erhöhte maßgeblich die Anzahl der Gedanken mit Vorwissen, und auch die Wahrscheinlichkeit die Stimuli affektiv zu verarbeiten. Die Befragten antizipierten den sozialen Konsens, Hate Speech mit Missbilligung zu begegnen und antworteten entsprechend. Der oben angesprochene unverarbeitete Widerspruch von weiblichem Globalstereotyp und Hate Speech ließ sich auch daran erkennen, dass HS von Frauen signifikant weniger Gedanken mit Vorwissen auslöste. Die kognitive Auseinandersetzung mit dem Stimulusmaterial als Ausgangsbasis sank umso deutlicher, je mehr Informationsquellen zum Thema Frauenkarrieren die Befragten angaben oder je höher der Bildungsgrad war. Der Befund kann als Reaktanz gegenüber dem Stimulusmaterial verstanden werden: Die Bereitschaft Hate Speech-Stimuli zu beschreiben, war umso geringer, je gebildeter die Befragten waren und je stärker sie sich mit Frauenkarrieren beschäftigten. Für männliche Versuchspersonen war sie erhöht. Die meisten Rezeptionsgedanken brachten Unmut und Empörung zum Ausdruck und neben dem oben erwähnten Einfluss sozial erwünschten Antwortverhaltens, erhöhten auch ambivalent sexistische Einstellungen die Wahrscheinlichkeit für eine affektive Verarbeitung der Stimuli.

Beim Third-Person-Effekt geht es darum, antizipierte Wirkungen von medialen Stimuli bei anderen zu überschätzen, v. a. wenn der Effekt als wenig wünschenswert gilt (Huck & Brosius, 2007). Wie erwartet konnte ein Third-Person-Effekt dahingehend festgestellt werden, dass sich Selbst- und Fremdwahrnehmung bezüglich der Relevanz deutlich unterschieden. Die persönliche Relevanz wurde über alle Gruppen hinweg als geringer eingeschätzt als die Relevanz für andere Nutzer:innen. Unterschiede zur wahrgenommenen gesellschaftlichen Relevanz konnten nicht nachgewiesen werden, was möglicherweise mit dem Abstraktionsgrad von Relevanzwirkungen auf die Gesellschaft zusammenhängt und methodisch zu diskutieren ist. Des Weiteren wurden Unterschiede in der Relevanzwahrnehmung im Intergruppenvergleich sichtbar: Kommentare in der Kontrollgruppe wurden insgesamt als relevanter wahrgenommen als Hate Speech von Frauen oder Hate Speech von Männern. Sozial erwünscht antwortende Versuchspersonen nahmen die Kommentare als relevanter für Dritte wahr. Abgesehen vom Bildungsgrad hatten soziodemografische Merkmale wie Alter und Geschlecht keinen Einfluss auf die Relevanzwahrnehmungen. Im Vergleich zu geringer Gebildeten nahmen höher Gebildete die Kommentare als weniger relevant für Dritte an. Ambivalent

sexistische Einstellungen erhöhten sowohl die wahrgenommene persönliche Relevanz als auch die wahrgenommene Relevanz für andere Nutzer:innen und die Gesellschaft. Der Zusammenhang zwischen Ambivalentem Sexismus und der Zunahme der wahrgenommenen persönlichen Relevanz wird als Hinweis gewertet, dass einstellungskongruente Beziehungen zwischen den beiden Variablen bestehen. Ambivalenter Sexismus beeinflusste jedoch nicht die Beziehung zwischen Rezeption der präsentierten Stimuli und der Relevanzwahrnehmung.

Auch hinsichtlich des wahrgenommenen Einflusses auf die persönliche Meinung im Vergleich zum Einfluss auf die Meinung Dritter bzw. der Gesellschaft konnte ein Third-Person-Effekt nachgewiesen werden. Eine Diskrepanz von Selbst- und Fremdwahrnehmung bestand diesbezüglich zwischen dem wahrgenommenen Einfluss auf die persönliche Meinung und auf die Meinung Anderer, sowie zwischen dem wahrgenommenen Einfluss auf die persönliche Meinung und auf die gesellschaftliche Meinung, wenn auch die Effekte nur minimal waren.

Der nächste Analyseteil setzte an dem Mechanismus zur Erklärung des Zusammenhangs zwischen Stimulusrezeption und Bewertung des Stimulus. Festgestellt wurde zunächst ein starker Unterschied der Bewertung des Stimulusmaterials. Die Kommentare in der Kontrollgruppe wurden deutlich positiver als die Kommentare in den Hate Speech-Bedingungen bewertet, was zugleich als Treatmentcheck für die Validität der Stimuli gewertet kann. Wie die Bewertung des Stimulusmaterials zustande kommt, konnte abschließend nur teilweise beantwortet werden. Abgeleitet aus der ELM-Forschung steht die Bewertung von Persuasionsbotschaften mit der wahrgenommenen persönlichen Relevanz in einem Zusammenhang. Diese Annahme fand im vorliegenden experimentellen Aufbau Berücksichtigung, indem die Relevanzvariablen vor der Bewertung des Stimulusmaterials abgefragt wurden. Vorliegend konnte die Annahme jedoch nicht für die experimentell variierte Geschlechterkategorie bestätigt werden. Gezeigt werden konnte, dass Hate Speech-Kommentare unabhängig vom Geschlecht der Hate Speech-Produzent:innen über die persönliche Relevanzwahrnehmung negativer als Kommentare in der Kontrollgruppe bewertet werden. Die weitergehende Berücksichtigung der Relevanzwahrnehmung auf Andere und auf die Gesellschaft als weitere Mediatoren konnte in dem Mediationsmodell von Stimulusrezeption über persönliche Relevanz auf die kognitive Bewertung keine Mediationsbeziehungen nachweisen. Die Relevanzvariablen hatten keinen Einfluss auf die Bewertung des Stimulusmaterials, trotz hochsignifikanter mittelstarker negativer Zusammenhänge zwischen den Variablen. Vermutet werden Verschiebungen im Kausalmodell, nämlich dass die Bewertung des Stimulusmaterials unmittelbar im Rezeptionsprozess erfolgt, noch bevor diese abgefragt wird und sich die wahrgenommene Relevanz erst im Anschluss festigt. Was die antizipierten Meinungsflüsse betrifft, erwiesen sie sich ebenfalls als ungeeignet, den Wirkungsprozess zwischen Rezeption

und Bewertung zu erklären. Die Bewertung der experimentellen Stimuli erfolgte unabhängig von soziodemografischen Einflussgrößen wie Alter und Geschlecht. Ein Effekt konnte für den Bildungsgrad nachgewiesen werden: Höher Gebildete lehnten besonders stark einen Einfluss der Stimuli auf die eigene Meinung ab.

Dass persuasive Botschaften Einstellungsänderungen bewirken, lässt sich nur unter bestimmten Voraussetzungen erreichen (für Hate Speech z. B. politische Ereignislagen oder situative Bedingungen) und ist ungleich schwerer als die Festigung bestehender Einstellungen. Festgestellt wurde ein starker positiver Zusammenhang zwischen ambivalent sexistischen Einstellungen und feindseligen Einstellungen zu modernen Frauenkarrieren. Ambivalent sexistische Einstellungen erhöhten die Zustimmung zu feindseligen Einstellungen zu Frauenkarrieren. Die Beziehung zwischen diesen beiden Variablen war jedoch nicht von der Rezeption von HS von Männern oder HS von Frauen abhängig. Da jedoch der Vergleich zur Kontrollgruppe fehlt, kann ein Einfluss von Hate Speech vorerst nicht gänzlich ausgeschlossen werden. Die postulierte Beziehung zwischen ambivalent sexistischen Einstellungen und Einstellungen zu modernen Frauenkarrieren unterschied sich in Abhängigkeit des Geschlechts der Versuchspersonen: Je ausgeprägter ambivalent sexistische Einstellungen bei weiblichen Befragten waren, desto stärker stimmten sie feindseligen Einstellungen zu Frauenkarrieren zu. Der Effekt trat auch bei männlichen Befragten auf, war aber dort weniger stark ausgeprägt. Dieser Befund lässt sich als Festigung bestehender Einstellungen verstehen. Aus Perspektive der Sexismus-Forschung ist hier festzuhalten, dass Frauenkarrieren umso negativer bewertet werden, je stärker der Ambivalente Sexismus ist und dass sich ambivalent sexistische Frauen deutlich negativer zu Frauenkarrieren positionieren, als ambivalent sexistische Männer. Der Nachweis, ob ambivalent sexistische Frauen oder Männer aufgrund der Rezeption von Hate Speech stärker feindseligen Einstellungen zu Frauenkarrieren zustimmen, konnte nicht erbracht werden.

Die Interaktionsmöglichkeiten mit Inhalten auf sozialen Netzwerkseiten sind vielseitig, wenn auch passive Verhaltensweisen (Lurken) dominieren. Dies konnte auch vorliegend bestätigt werden. Wenn auch die Befragten sich in den Rezeptionsgedanken empört äußerten, Vorwissen vorbrachten, die persönliche Relevanz abstritten oder die Stimuli negativ bewerteten, so dominierte das Ignorieren von Hate Speech von Männern wie von Frauen im Vergleich zur Kontrollgruppe. In der Kontrollgruppe wurde das Ignorieren als wahrscheinlicher als in den Experimentalbedingungen bewertet. Dafür war die selbsteingeschätzte Wahrscheinlichkeit die Kommentare zu Melden oder einen ablehnenden Kommentar zu schreiben in den Experimentalgruppen gegenüber der Kontrollgruppe erhöht. Über alle abgefragten Intentionen hinweg gab es keine signifikanten Unterschiede zwischen den beiden Hate Speech-Gruppen.

Sowohl die Wahrscheinlichkeit die Kommentare zu melden als auch einen ablehnenden Kommentar zu verfassen, war für HS von Männern geringfügig erhöht, aber nicht signifikant im Vergleich zu HS von Frauen. Die Wahrscheinlichkeit, die dargebotenen Kommentare mit einem Like positiv zu bejahen, war in allen drei Versuchsbedingungen ausgesprochen gering. Die Unterschiede zwischen den Gruppen waren minimal und konnten keine Signifikanz erreichen. Weder soziodemografische Merkmale, noch soziale Erwünschtheit, noch die habitualisierte Nutzungsweise beeinflussten die Verhaltensintentionen.

Vorhandene ambivalent sexistische Einstellungen wiesen einen mittleren Zusammenhang mit der artikulierten Verhaltsabsicht, dem rezipierten Stimulusmaterial zuzustimmen, auf. Dies unterschied sich jedoch nicht in Abhängigkeit des Stimulus, was wiederum eher für einstellungsverstärkende, denn -verändernde Effekte spricht. Wenn auch vorhandene ambivalent sexistische Einstellungen mit den in der Experimentalsituation aktivierten Einstellungen zu modernen Frauenkarrieren hoch korrelieren, konnte keine Beziehung zwischen den aktivierten Einstellungen und zustimmenden Verhaltensabsichten belegt werden. Vorhandene Einstellungen und aktivierte Einstellungen hängen also zusammen, während sich aber *vorhandene* Einstellungen auf die Verhaltensabsicht zuzustimmen auswirken, konnte kein Effekt für *aktivierte* Einstellungen und die Verhaltensabsicht zuzustimmen, belegt werden. Dass die aktivierte Einstellung zu modernen Frauenkarrieren auf Verhaltensabsichten durchaus einen Einfluss hat, konnte für das Ignorieren des rezipierten Stimulusmaterials gezeigt werden.

Zusammenfassend lässt sich konstatieren, dass Unterschiede in der Wahrnehmung und Bewertung von Hate Speech von Männern im Vergleich zu Hate Speech von Frauen minimal sind. Es lassen sich eher Unterschiede zur Kontrollgruppe nachweisen, was sich mit den Befunden aus Studie I zur Intensitätsstufe von Hate Speech deckt. Hinweise auf *Einstellungsänderungen* konnten in der vorliegenden Studie nicht erbracht werden, was sich mit einer Querschnittsstudie sowieso nicht bewerkstelligen lässt. Allerdings ließen vereinzelte Hinweise auf *Einstellungsfestigungen* durchaus den Schluss zu, dass persuasive Effekte von Hate Speech beobachtet werden konnten.

## 7 Fazit

Welche Verarbeitungsprozesse bei Hate Speech ablaufen, dazu müssen zunächst deren Rezeptions- und Persuasionseffekte auf ein unbeteiligtes Publikum systematisch erforscht werden, wie die bisherigen Ausführungen belegt haben. Diese Arbeit zielte auf eine theoriegeleitete Modellierung von Hate Speech und deren empirische Überprüfung ab, und der Ertrag der vorliegenden Forschungsarbeit wird im Folgenden zusammengefasst und bewertet.

### 7.1 Zusammenfassung der zentralen Ergebnisse

Die übergeordnete Fragestellung widmete sich der Klärung des persuasiven Potenzials von sexistischem Hate Speech in Abgrenzung zu Inzivilität und Dangerous Speech. Diese Begriffe erscheinen in dem noch jungen Forschungsfeld „Dark Participation“ austauschbar, da aufgrund der rasanten Entwicklung von neuen Internettechnologien bislang zwar permanent neue Phänomene als Neologismen Eingang in den Fachwortschatz finden, diese allerdings noch nicht in Arbeiten systematisiert wurden, um Ordnung in die terminologische Vielfalt zu bringen. Die vorliegende Arbeit stellte sich dieser Aufgabe, indem zunächst der Forschungsstand zu Hate Speech aus kommunikationswissenschaftlicher Sicht reflektiert wurde, um danach zu klären, inwiefern Hate Speech differenziert zu betrachten ist.

Da der Fokus der Arbeit auf Hate Speech im Internet lag, wurden die verwandten Phänomene Inzivilität und Dangerous Speech diskutiert, die in der theoretischen und empirischen Bearbeitung als Intensitätsabstufungen von Hate Speech verstanden wurden. Operationalisiert wurde Inzivilität über das messbare Konstrukt der Offensive Speech, wohingegen Dangerous Speech analog zur Definition als Hate Speech mit Gewaltandrohung messbar gemacht wurde. Sexismus als Einstellungskonstrukt diente als Bindeglied zwischen den variierten Intensitätsabstufungen von Hate Speech und den erwarteten Rezeptionswirkungen, abgeleitet aus dem Elaboration Likelihood Model (ELM). Dadurch ließen sich Anknüpfungspunkte an sozialpsychologische Fragestellungen aus der Sexismusforschung finden. So erzeugen Abweichungen von der männlichen oder weiblichen Geschlechterrolle häufig heftige Gegenreaktionen in Form von Ablehnung, Misstrauen oder Empörung. Solche Abweichungen von der Geschlechterrolle wurden im empirischen Teil der Arbeit auf zwei Arten akzentuiert: Zum einen wurde exemplarisch der Themenbereich „Frauen in der männerdominierten IT-Branche“ herangezogen und zum anderen wurde vermutet, dass Hate Speech von Männern und Frauen als abweichend von der jeweiligen Geschlechterrolle betrachtet wird.

Die Ergebnisse der ersten Untersuchung stützten die Annahme, dass bei Hate Speech sowohl eine starke Rezeptionswirkung als auch ein schwaches persuasives Potenzial besteht, wenn sich die Kommentare eines Abwertungsmechanismus' bedienen. Als wichtigster Prädiktor erwies sich dabei die *Hate Intensität*, die die meisten und stärksten Effekte hervorrief. Die Rezeptionswirkungen variierten je nach Intensitätsabstufung: Während Inzivilität schwächere Effekte als Hate Speech hervorrief, löste Hate Speech wiederum schwächere Effekte als Dangerous Speech aus. Nachweisen ließen sich kognitive und affektive Effekte der Botschaftsmerkmale auf die Elaborationstiefe, auf die Relevanzwahrnehmung und die bekundeten Verhaltensintentionen der Rezipierenden. Dass die Hate Intensität ausschlaggebend für die Rezeptionswirkung ist, hängt mit den einzelnen Abstufungen zusammen, die unterschiedliche Effekte je nach Motivation und Fähigkeit zur Elaboration bewirken: Während Dangerous Speech unabhängig vom Kognitionsbedürfnis stets einen hohen Anteil an negativen Rezeptionsgedanken hervorrief, gilt dies nicht für Inzivilität und Hate Speech. Ein höheres Kognitionsbedürfnis bewirkte einen höheren Anteil an Negativgedanken, die Hate Speech bzw. Inzivilität rezipiert hatten. Überraschenderweise war die Interaktion zwischen Vorwissen und Inzivilität umgekehrt: Je stärker die Versuchspersonen der Angabe zustimmten, sich kompetent zu fühlen an einer Diskussion zum Thema „Karrierechancen von Frauen“ zu partizipieren, desto geringer war der Anteil an Negativgedanken. Für Hate Speech und Dangerous Speech war die Interaktion nicht nachzuweisen. Hier kann also von einem normativen Schwellenwert gesprochen werden, der den Intensitätsabstufungen von Hate Speech anhaftet und den Nachweis eines davon abhängigen Einflusses auf die Tonalität der Rezeptionsgedanken je nach Motivation und Fähigkeit erbringt.

Allerdings liegt nicht immer eine lineare Steigerung der Rezeptionswirkungen bei Zunahme der Hate Intensität vor: So offenbarte etwa die Relevanzwahrnehmung, dass die Unterschiede zwischen Inzivilität und Hate Speech geringer waren als zwischen den Experimental- und den Kontrollgruppen. Persuasive Effekte sind zwar in einer Querschnittsstudie nur bedingt nachweisbar, denn bekundete Verhaltensintentionen ließen keine Unterstützung bzw. Befürwortung der rezipierten Kommentare nach der Rezeption erkennen. Indes konnten durchaus Anhaltspunkte für Persuasionseffekte identifiziert werden und zwar dann, wenn der Einfluss vorhandener Einstellungen als Bedingung berücksichtigt wurde, um den Zusammenhang zwischen Rezeption von Hate Speech und seinen Wirkungen zu erklären. So zeigte sich in der Kontrollgruppe, dass sachliche Kommentare dort als weniger relevant wahrgenommen wurden, je stärker ambivalent sexistische Einstellungen waren. Für Inzivilität und Hate Speech kehrte sich dieser Effekt um: Ambivalenter Sexismus erhöhte die wahrgenommene Relevanz, was in der Argumentation des ELM für persuasive Effekte von Hate Speech

spricht. Auch hier wird wieder ein Schwelleneffekt belegt: Ambivalent sexistische Einstellungen beeinflussen die wahrgenommene Relevanz von Dangerous Speech nicht. Vermutet werden kann, dass die Ablehnung nach Rezeption von Dangerous Speech die Einstellungsstärke übersteigt.

Die Frage danach, inwiefern sich das ELM bei der persuasiven Wirkungsforschung von sexistischem Hate Speech eignet, lässt sich in der Gesamtschau kaum zufriedenstellend und eindeutig beantworten. Für den Einsatz des Modells spricht nicht nur, dass es seit vierzig Jahren bei einer Vielzahl von Fragestellungen empirisch erprobt ist, sondern auch, dass es die Prognose stark ausdifferenzierter Rezeptions- und Persuasionswirkungen ermöglicht. Da das Modell diesbezüglich Wirkungspfade modelliert, erlaubte dies eine Anwendung für diese Rezeptionsstudie im Sinne des ELM. Anhand des ELM wurden deshalb folgende Effekte und ihre Beziehung zueinander detailliert untersucht: *kognitive* Effekte (absolute Gedankenanzahl, kognitive Ausrichtung der Rezeptionsgedanken, Relevanzwahrnehmung, Bewertung des Stimulusmaterials, Einstellungen zu Frauenkarrieren) und *affektive* (Gedankenvalenz, affektive Ausrichtung der Rezeptionsgedanken) sowie *verhaltensmäßige* Effekte (subjektiv wahrgenommene Wahrscheinlichkeiten von Verhaltensintentionen).

Bestätigt wurden im Sinne des ELM Unterschiede in der Elaborationstiefe von Hate Speech, je nachdem wie ausgeprägt die Fähigkeit und Motivation zur Informationsverarbeitung bei den Rezipierenden war, was sich bei der Analyse der Gedankenvalenz oder der kognitiven bzw. affektiven Auseinandersetzung mit dem Stimulusmaterial in den Rezeptionsgedanken abzeichnete. Erkennbar wurden diese Einflüsse bei der kognitiven Bewertung des Stimulusmaterials als auch bei den Einstellungseffekten zu modernen Frauenkarrieren. Unvereinbar mit Hate Speech ist hingegen die Annahme des ELM, dass die Argumentstärke mit einer höheren Persuasivität einhergeht. Bei Hate Speech lässt sich – analog zur Argumentation im theoretischen Teil der Arbeit – sinnvollerweise nicht von Argumentstärke sprechen als vielmehr von einer Abstufung der Hate Intensität, die sich hier als tragfähiges Konstrukt erwiesen hat.

Kritisch anzumerken ist jedoch, dass dabei der Ursprung des Modells in der Konsumentenforschung deutlich wird, denn die Argumentstärke prognostiziert im Vergleich zur Hate Intensität gegensätzliche Effekte, wenngleich beide ihre Ausgangsbasis in der Botschaft haben: Nach dem ELM sind bei der Verarbeitung auf der zentralen Route positive Rezeptionsgedanken bei starken Argumenten erwartbar und negative Rezeptionsgedanken bei schwachen Argumenten. Studien zur Wirkung von Hate Speech setzen ebenfalls bei der Botschaft an, wobei von unterschiedlichen Effekten je nach Variation der Hate Intensität ausgegangen wird.

Die Wirkungsrichtung berücksichtigt zusätzlich die sekundären Adressat:innen von Hate Speech, wobei hier angenommen wurde, dass sich Sympathisant:innen und unbeteiligte Dritte in ihren Einstellungsmerkmalen unterscheiden: Bei einstellungskongruenter Wahrnehmung der Hate Intensität sollten die Rezeptionsgedanken positiver sein. Vorliegend bestätigte sich der Einfluss von vorhandenen ambivalent sexistischen Einstellungen auf den Zusammenhang zwischen Hate Intensität und Gedankenvalenz. Inzivilität und Hate Speech riefen weniger negative Gedanken hervor, je stärker ambivalent sexistische Einstellungen bei den Befragten ausgeprägt waren. Für Dangerous Speech war dieser Effekt nicht nachweisbar, was vermuten lässt, dass hier wiederum normative Schwellenwerte überschritten wurden, die nicht in der Einstellungsstruktur der Teilnehmenden verankert waren.

Neben der Hate Intensität wurde hinsichtlich der Persuasionswirkungen noch das *Geschlecht* der Kommentierenden als Botschaftsmerkmal sowie das *Reaktionsvolumen* als situatives Merkmal analysiert. Wider Erwarten zeigten sich bei der experimentellen Variation des Geschlechts keine Unterschiede in den postulierten Rezeptions- oder Persuasionswirkungen auf die Befragten. Erwartet wurde ein Einfluss der Geschlechterkategorie auf die *Elaboration*, d.h. die Anzahl der Rezeptionsgedanken, deren Valenz und die gedankliche Ausgangsbasis, was jedoch nicht bestätigt werden konnte. Auch bei der *wahrgenommenen Relevanz* und den *Einstellungen zur Botschaft* sowie zu *Frauenkarrieren* ließ sich kein Unterschied hinsichtlich der rezipierten Geschlechterkategorie feststellen. Dies ist vermutlich darauf zurückzuführen, dass die Geschlechterkategorie als peripherer Hinweisreiz und nicht als Argument zentral verarbeitet wird, wie von der Sexismusforschung behauptet (Chang & Hitchon, 2004; Johnston & Coolen, 1995). Dieser Befund überrascht insofern, als die Befragten die Geschlechterkategorie als experimentell manipuliertes Treatment durchaus erkannt hatten. Die Verarbeitung des Geschlechts der Kommentierenden als Botschaftsmerkmal könnte durch verzerrte Verarbeitung – jedoch nicht nachweisbar in Form von sozial erwünschtem Antwortverhalten – erklärt werden. Zudem ist das einzige signifikante Ergebnis bezüglich der Geschlechterkategorie als Botschaftsmerkmal inkonsistent über die beiden Experimentalstudien: Die Analyse der Rezeptionsgedanken zeigte, dass Kommentare von Männern die Quantität der Rezeptionsgedanken erhöhten, die auf eigenen Überlegungen fußten (Studie I). In Studie II gab es diesen quantitativen Einfluss auf Rezeptionsgedanken dieser Art zwar nicht, jedoch erhöhte sich die Wahrscheinlichkeit, dass Kommentare von Frauen ebensolche Gedanken bewirkten. Die Widersprüchlichkeit der Befunde lässt sich mit den geringen Effektstärken in beiden Studien sowie Unterschieden in der Untersuchungsanlage erklären. Ist es plausibel anzunehmen, dass kein Sexismus vorherrscht, weil sich keine Geschlechtereffekte nachweisen ließen – d.h. dass weder Unterschiede zwischen männlichen und weiblichen

Versuchspersonen noch bei der experimentellen Manipulation des Geschlechts der Kommentierenden ausgemacht werden konnten? Erlaubt dies den Schluss, dass Sexismus in der heutigen Gesellschaft überwunden ist? Dagegen sprechen nicht nur realweltliche Erfahrungen und zahlreiche Arbeiten aus der Sexismusforschung, sondern auch vorliegende Ergebnisse, nach denen sich Ambivalenter Sexismus einstellungskongruent mit sexistischem Hate Speech und seinen Intensitätsabstufungen verhält.

Mit dem *Reaktionsvolumen* wurde die qualitative Variation der Anzahl von Likes für die Kommentare erfasst (viele vs. wenige Likes). Dies gilt als situativer Indikator für die wahrgenommene persönliche Verantwortung in der Rezeptionssituation. Die experimentelle Manipulation dieser Bedingung war nicht erfolgreich, allerdings rechtfertigten methodische Gründe dennoch die Berücksichtigung des Treatments bei den Analysen, da nicht ausgeschlossen werden konnte, dass sich auch hier Rezeptionseffekte einstellen. Und mehrfach beeinflusste das Reaktionsvolumen die Rezeption der Versuchspersonen, sodass dies den sozialen Einflüssen geschuldet ist. Diese sogenannten „Bandwaggoning-Effekte“ bestätigten sich jedoch nur für die Referenzgruppe, d.h. für die sachlichen Kommentare von Frauen, nicht jedoch für die Abstufungen von Hate Speech und auch nicht bei männlichen Kommentarschreibern. Diese Orientierung an wahrgenommenen Mehrheiten hatte zur Folge, dass weniger negative Rezeptionsgedanken in der genannten Referenzgruppe vorkamen, wenn die Likeanzahl hoch war, und umso mehr auftraten, je höher die Elaborationstiefe war. Auch die wahrgenommene persönliche Relevanz wurde positiv von der unbewussten Wahrnehmung vieler Likes beeinflusst, was laut ELM für persuasive Effekte in der Referenzgruppe spricht. Keine Effekte des Reaktionsvolumens zeigten sich bei der weiteren Analyse der Rezeptionsgedanken hinsichtlich der Anzahl und dem gedanklichen Ursprung. Auch unter Berücksichtigung ambivalent sexistischer Einstellungen bewirkte das Reaktionsvolumen keinen Einfluss auf die wahrgenommene Relevanz.

Aufgrund dieser verschiedenen Befunde erweist sich die Forderung nach einer differenzierten Prüfung von Rezeptions- und Persuasionswirkungen von Hate Speech und seinen Intensitätsabstufungen als berechtigt. Demnach muss eine theoriegeleitete Operationalisierung von Hate Speech, Inzivilität und Dangerous Speech erfolgen, da bei diesen Phänomenen unterschiedliche Rezeptionseffekte eintraten. Das Thema „Frauenkarrieren in der IT-Branche“ wurde aufgrund des reichhaltigen Forschungsvolumens in der Forschungsliteratur ausgewählt, schien jedoch für die Untersuchung von sexistischem Hate Speech weniger geeignet gewesen zu sein. Für Rezeptions- und Persuasionswirkungen der Geschlechterkategorie fehlten jedenfalls stichhaltige Belege, weshalb sich eine empirisch geleitete Auswahl an Themen

anbietet, die als Prime für Sexismus dient. Der Mere Exposure-Effekt bei der Rezeption des Reaktionsvolumens in der Referenzgruppe legt nahe, dass eine Verantwortungsübernahme bei der Rezeption von Hate Speech abgelehnt wird. Dieser Befund deckt sich mit den artikulierten Verhaltensintentionen der Befragten, die mehrheitlich entweder ein Ignorieren der Kommentare bekundeten oder Hate Speech ablehnten. Zusammenfassend lässt sich festhalten, dass die detaillierten Annahmen und postulierten Zusammenhänge zwischen den Variablen des ELM in dieser Querschnittsstudie differenzierte Aussagen über die Rezeptions- und Persuasionseffekte von Hate Speech erlauben.

## 7.2 Methodische und theoretische Limitationen

Limitationen bestehen für diese Arbeit dennoch auf theoretischer sowie auf empirischer Ebene, die im Folgenden diskutiert werden. Eines der Hauptanliegen dieser Arbeit war es, die Eignung des ELM zur Erforschung von Rezeptions- und Persuasionseffekten von Hate Speech zu prüfen. Als theoretisches Rahmenwerk ermöglichte das ELM eine differenzierte Analyse der kognitiven und affektiven Dimensionen von Rezeptionsgedanken, von Einstellungseffekten bis hin zu Verhaltensintentionen, sodass diese Arbeit einen Beitrag zur Grundlagenforschung geleistet hat. Da das Modell jedoch für die Konsumentenforschung entwickelt wurde, musste es hinsichtlich seiner Annahmen und Voraussagen für die Analyse von sexistischem Hate Speech modifiziert werden. Zentral für dieses Modell ist die Konzeption der Argumentstärke, die jedoch als Botschaftsmerkmal von Hate Speech nicht adäquat operationalisiert werden kann, wie dies im vorigen Unterkapitel sowie in Kapitel 4.3.2.1 ausführlich begründet wurde. Stattdessen wurde der Intensitätsgrad von negativen Aussagen herangezogen (Inzivilität, Hate Speech und Dangerous Speech), was aus Perspektive der Hate-Speech-Forschung durchaus angemessen sein kann. Allerdings ist dies im ELM-Kontext nicht unumstritten, weil dies die ursprüngliche Wirkungsrichtung während der objektiven zentralen Verarbeitung auflöst (stichhaltige Argumente → positive Rezeptionsgedanken bzw. schwache Argumente → negative Rezeptionsgedanken). Ausgangspunkt in der Arbeit war, dass die Rezeptionsgedanken mit steigender Intensität von Hate Speech immer negativer werden, was in der ELM-Logik einem misslungenen Persuasionsversuch gleichkommt, d. h., dass die Gedankenvolanz im Kontext von Hate Speech demnach nicht als Indikator für Persuasion fungiert, sondern deskriptiv das Rezeptionsergebnis erfasst.

Das ELM beschreibt und strukturiert sozialpsychologische Beziehungen des Individuums im Rahmen einer sozialen Situation (Aronson et al., 2014), was keine Aussagen über gesellschaftliche Auswirkungen erlaubt. Die gesellschaftliche Dimension wird in der Hate Speech-Forschung mehrheitlich aus theoretischen Überlegungen abgeleitet und ist nur selten

Gegenstand von empirischen Studien, wie z.B. im Kontext politischer Umbruchphasen (z. B. zwischen verfeindeten Volksgruppen während innerstaatlicher Konflikte). Insofern steht die Hate Speech-Forschung noch am Anfang der Erkenntnisse, welche Prozesse auf individueller, überindividueller oder gesellschaftlicher Ebene stattfinden, weil sich diese bislang hauptsächlich auf die primären Adressatengruppen konzentrierte.

Die massive Verbreitung von Hate Speech in sozialen Netzwerken wirft jedoch vermehrt Fragen nach dessen persuasiver Sogwirkung auf ein unbeteiligtes Publikum auf, das sich als sekundärer Adressatenkreis bezeichnen lässt. Dass Rezeptionsgedanken mit steigender Hate Speech-Intensität umso negativer sind, bestätigte diese Arbeit, nicht jedoch die eigene Bereitschaft, selbst Hate Speech einzusetzen, um den eigenen Standpunkt zu verteidigen. Mit den Erkenntnissen der vorliegenden Grundlagenarbeit zur individuellen Verarbeitung von Hate Speech ließen sich nun Analyseebenen verknüpfen (Quandt & Scheufele, 2011), um „Medienwirkungen jenseits des einzelnen Rezipienten“ (Scheufele, 2011) zu modellieren und zu erklären. Als Anknüpfungspunkt auf individueller Ebene kann die wahrgenommene Relevanz von Botschaften dienen, denn die subjektive Relevanzwahrnehmung von Hate Speech fällt im Vergleich zur Relevanzwahrnehmung für andere SNS-Nutzende sowie für die Gesellschaft geringer aus, wie Befunde dieser Arbeit darlegen. Wird die persönliche Relevanz von Hate Speech abgestritten und dessen Ignorieren als bevorzugte Verhaltensabsicht deklariert, sind potenziell unerwünschte dysfunktionale Effekte auf einer höheren Ebene (z. B. innerhalb des SNS, aber auch gesellschaftsweit) zu erwarten, etwa dass attackierte Gesellschaftsgruppen oder einzelne Nutzende keine Unterstützung erhalten oder dass auf Gegenrede verzichtet wird. Hieran können weitere Arbeiten anknüpfen.

In methodischer Hinsicht gilt das Querschnittsdesign der beiden experimentellen Untersuchungen als weitere Limitation dieser Arbeit, denn Petty und Kollegen zeigen den Mehrwert von ELM-Längsschnittuntersuchungen auf. Mit einer Erhebung zu mehreren Messzeitpunkten lässt sich der Einfluss der rezipierten Botschaften auf die Persistenz und Resistanz von Einstellungen nachweisen (Petty et al., 1995). Die Befunde dieser Arbeit gelten nur für die unmittelbare Rezeptionssituation. Ob sich die Relevanzwahrnehmung im zeitlichen Verlauf ändert oder verstärkt, die Gedankenvalenz oder die Ausgangsbasis der Rezeptionsgedanken konstant bleiben oder weitere Elaborationsprozesse durchlaufen und sich möglicherweise wandeln, muss in weiteren Longitudinalstudien geprüft werden.

Dass das Geschlecht keinen Effekt zeitigte, liegt möglicherweise an der Thematik der Studie: „Frauenkarrieren in der der IT-Branche“ empfanden die Versuchspersonen als weniger polarisierend, als dies nach Sichtung der Forschungsliteratur zu erwarten gewesen wäre. Dieses

Ergebnis weicht auch von den bisherigen Beobachtungen ab, denn unter realweltlichen Bedingungen findet sich durchaus sexistisches Hate Speech in den Kommentarspalten thematisch verwandter Postings, wobei auch die Gegenrede auf derartige Kommentare von großer Aggressivität zeugt. Die Befragten verneinten in dieser Arbeit jedoch, derart zu handeln. Es kann nicht ausgeschlossen werden, dass Effekte sozialer Erwünschtheit aufgetreten sind, die sich dennoch nicht nachweisen ließen. Dies kann mit dem Erhebungsinstrumentarium zusammenhängen oder mit der Zusammensetzung der Stichprobe (z. B. hinsichtlich soziodemografischer Merkmale). Dass Effekte von Sexismus bei Hate Speech vermehrt bei Frauen oder bei Männern auftreten, ließ sich ebenfalls nicht belegen. Im Geschlechtervergleich zeigten sich keine Unterschiede, wenn das Geschlecht der Kommentierenden variiert wurde oder beim Vergleich zwischen dem Geschlecht der Befragten. Vermutet wird, dass die Hate Intensität sämtliche andere Effekte oder Variablen wie das Geschlecht als „Killer“-Manipulation dominiert, sodass möglicherweise geringe Effekte der Geschlechterkategorie unentdeckt geblieben sind. Eine Vorabtestung von polarisierenden Themen zwischen den Geschlechtern (z. B. Abtreibungsrecht) bietet sich im Rahmen weiterer Untersuchungen an.

Die hohe Zahl der Ausschlüsse nach Datenbereinigung und die Abbruchzahlen bedeuten eine weitere Einschränkung dieser Studie. Ein großer Anteil der Daten musste ausgeschlossen werden, weil die Befragten unaufmerksam oder zu schnell durch die Fragen „durchklickten“. Das ist zum einen der mangelnden Qualität der Stichprobe geschuldet, kann aber zum anderen auch auf Reaktanzeffekte zurückgeführt werden, die durch die Befragung selbst ausgelöst wurden (Thematik, Länge des Fragebogens und Befragungssituation etc.). Die meisten Abbrüche wurden nach der Stimuluspräsentation in Studie I registriert, was mit der Option eines Exit-Buttons unmittelbar nach Rezeption zusammenhängt, um aufkommenden etwaigen emotionalen Stress aufzufangen. In Studie II wurde auf den Exit-Button verzichtet, was die Abbruchquote merklich verringerte.

„Yet another Facebook study.“ Die Befunde sind insofern nicht verallgemeinerbar, als sich die Ergebnisse zu den Rezeptions- und Persuasionswirkungen von sexistischem Hate Speech nicht auf das Internet oder die sozialen Netzwerke übertragen lassen. Die Befragten wurden nach ihrer aktiven Facebook-Nutzung ausgewählt. Die Stimuli zeigten einen Ausschnitt eines Facebook-Postings mitsamt den dazugehörigen Kommentaren. Damit sind die Ergebnisse nur bedingt auf andere Plattformen generalisierbar, zumal verschiedene Interaktionskulturen in den unterschiedlichen sozialen Netzwerken vorherrschen, was sich auf die Befunde auswirken kann (Chandrasekharan et al., 2017; König & König, 2016; Su et al., 2018). Die Wahl fiel auf Facebook, weil es bis dato das größte soziale Netzwerk weltweit ist und in Deutschland

das meistgenutzte Medium bei den bevölkerungsstärksten Altersgruppen darstellt (Statista, 2021a). Dadurch konnte eine möglichst große Stichprobe realisiert werden.

Nach methodischen Überlegungen wurde sich gegen eine Integration des Stimulusmaterials als interaktive Oberfläche in ein Facebook-Profil entschieden und für die Präsentation eines bildlichen Stimulus. Damit befanden sich die Versuchspersonen in keiner natürlichen Umgebung, weil mögliche Verhaltensweisen wie z. B. die Nutzung des Reaktionsbuttons oder das Verfassen eines Kommentars lediglich als hypothetische Verhaltensintention abgefragt werden konnten und keine Interaktion mit dem Stimulusmaterial unter Echtzeitbedingungen stattfand. Damit wurde zwar in Kauf genommen, dass die externe und die Inhaltsvalidität nur eingeschränkt erfüllt war. Allerdings blieb die interne Validität insofern aufrechterhalten, als alle Versuchspersonen demselben Stimulusmaterial ausgesetzt waren. Damit wurde verhindert, dass bereits kleine Unterschiede einen unvorhergesehenen Effekt auf das Antwortverhalten auslösen, wenn z.B. eine andere Uhrzeit angezeigt wird oder durch das Datum des Postings.

Die angeführten Limitationen schmälern jedoch nicht den Verdienst dieser Arbeit, denn diese lieferte als Grundlagenforschung von sexistischem Hate Speech wertvolle Erkenntnisse zu seiner Informationsverarbeitungsweise. Insbesondere die experimentelle Forschung erlaubt Kausalschlüsse, womit sich die Befunde zu den Relevanzwahrnehmungen, Einstellungseffekten und Rezeptionsgedanken gewinnbringend zur Prävention und Unterstützung primärer Adressatengruppen einsetzen lassen. So können Nutzende in SNS verstärkt sensibilisiert werden, primären Adressat:innen mittels Gegenrede (Schieb & Preuss, 2018) beizustehen, auch wenn sie die persönliche Relevanz von Hate Speech ablehnen. In theoretischer Hinsicht erfuhr das Konstrukt der Argumentstärke im ELM-Theoriegebäude (Erb et al., 2007; O'Keefe, 2003) eine Erweiterung bzw. Ergänzung um zusätzliche Aspekte, indem dessen Anwendbarkeit im Hate-Speech-Kontext intensiv diskutiert wurde. Es wurde gezeigt, dass die theoriegeleitete qualitative Abstufung der Hate Intensität in Verbindung mit rezeptions-situativen Merkmalen erkenntnisbringend sind. Daran können zukünftige Arbeiten anknüpfen.

### 7.3 Eine forschungsethische Schlussbetrachtung von Hate Speech

Experimentelle Forschung mit dem Ziel, Rezeptions- und Persuasionswirkungen von sexistischem Hate Speech in SNS zu beforschen, ist herausfordernd – nicht nur, was die Umsetzung des Untersuchungsdesigns, sondern auch die Forschungsethik betrifft. Abschließend wird daher die Verwobenheit von theoretischem Zugang, Methodik und forschungsethischen Überlegungen diskutiert, womit nicht nur die Forschungsleistung dieser Arbeit eingeordnet, sondern auch ein Anreiz für zukünftige Forschungsanstrengungen gegeben wird.

Angesichts der Verbreitung von toxischen Verhaltensweisen im Internet, zu denen Hate Speech zählt, ist der gesamtgesellschaftliche öffentliche wie auch wissenschaftliche Diskurs darauf ausgerichtet, deren Ursachen, Entstehungsbedingungen und ablaufende Rezeptionsprozesse zu verstehen. Diesen als „Dark Participation“ (Quandt, 2018) bezeichneten Verhaltensweisen wird eine gesellschaftlich dysfunktionale Wirkung attestiert, die Bestrebungen zur Herstellung gesellschaftlicher Integration torpedieren bzw. den Erhalt gefährden. Die Untersuchung von Hate Speech impliziert einerseits den normativen Anspruch, gesellschaftliche Probleme zu untersuchen, um mögliche Lösungsansätze aufzuzeigen. Andererseits befinden sich Forschende in dem Dilemma, dass sie mit ihrem Forschungsvorhaben selbst aufgrund der Thematisierung von Hate Speech – hier genauer sexistischem Hate Speech – an der Perpetuierung des Problems mitwirken. Insofern enthält die vorliegende Arbeit eine implizit hohe normative Aufladung, wobei diese mit der sozialpsychologischen Herangehensweise an das Thema zu umgehen versucht wird. Dennoch wird sie davon eingeholt, da forschungsethische Fragen bereits in einer frühen Konzeptionierungsphase aufkommen, die im Zusammenhang mit der methodischen Umsetzung zur Erforschung der Rezeptions- und Persuasionswirkung von sexistischem Hate Speech entstehen.

Gemäß der Forschungsliteratur handelt es sich bei Hate Speech um einen gruppenbezogenen Abwertungsmechanismus, weshalb Forschende zunächst über die Auswahl der Adressatengruppe entscheiden müssen. Gängige Ansätze im sozialwissenschaftlichen Forschungsprozess, nämlich forschungsrelevante sowie forschungsökonomische Abwägungen, spielen für Hate Speech dahingehend keine große Rolle, denn forschungsrelevant sind alle Adressatenkreise und forschungsökonomische Fragen sind hier nachrangig, weil Onlinebefragungen keinen erheblichen Mehraufwand bedeuten. Die vorliegende Arbeit versteht sich als Beitrag zur Grundlagenforschung, wobei auch hier forschungsethische Fragen von Bedeutung sind, weil der experimentelle Stimulus die Anfeindung einer sowieso schon marginalisierten sozialen Gruppe enthält. Dies birgt einerseits ethische Probleme, da den Versuchspersonen Hate Speech dadurch als legitim, verbreitet und akzeptiert aufgrund der Einbettung in einen

wissenschaftlichen Kontext erscheinen könnte. Andererseits legitimiert das Forschungs- und gesellschaftliche Interesse das Forschungsvorhaben, wenn auch die Wahl der Adressat:innen nach forschungsethischen Überlegungen erfolgen sollte.

Da Frauen allerdings keine gesellschaftliche Minorität darstellen, obwohl deren Statusdisparität unbestreitbar soziale Realität ist und sie aufgrund ihres Geschlechts von Anfeindungen im Internet überproportional betroffen sind, gilt es die Untersuchung von Rezeptions- und Persuasionswirkungen von sexistischem Hate Speech gegenüber Frauen kritisch zu reflektieren. Deshalb ist Hate Speech gegen Frauen keineswegs unproblematischer als Hate Speech gegen z. B. eine ethnische oder religiöse Minderheit, da Hate Speech auf der wahrgenommenen Statusdisparität fußt. Forschungsethisch heikler erscheint Hate Speech bei Minoritäten wie z.B. Ethnien und Religionszugehörigkeiten. Ist deshalb eine Untersuchung von Rezeptions- und Persuasionswirkungen bei sekundären Adressat:innen von Hate Speech gegen Ethnien und Religionszugehörigkeiten tabu, um dem Vorwurf der Produktion und Reproduktion von Abwertung und Diskriminierung gegenüber diesen zu entgehen? Wird der wissenschaftliche Ertrag für die Grundlagenforschung als hinreichend eingeschätzt und überwiegt er moralisch-ethische Bedenken (beispielsweise bei komparativen Ansätzen, in denen mehrere Adressatengruppen als Vergleichsgruppen herangezogen werden)? Muss bereits die Konzeptionierung einer solchen Rezeptionsstudie allerhöchste Sorgfalt genießen, um sie verantwortungsvoll umzusetzen?

Bei der Auswahl des kommunikativen Inhalts für das Stimulusmaterial müssen sowohl bisherige Forschungserkenntnisse zu Hate Speech bzw. dysfunktionaler/inzivilen Formen der Online-Kommunikation und realweltliche Erfahrungen mit sexistischem Hate Speech einfließen. Zur Sicherstellung der internen sowie externen Validität des Stimulusmaterials wurden eigene neue Kommentare für die experimentelle Variation der Hate Intensität konstruiert, welche jedoch echten Hasskommentaren in SNS nachempfunden waren. Zu diesem Zweck wurden einzelne Begriffe selektiert, um diese in einem neuen Kontext zu reformulieren. In wissenschaftlichen Studien zu Hate Speech im Internet ist es verbreitet, derartige Aussagen als Anschauungsmaterial (z. B. als Screenshot) zu verwenden, ohne die Urheber:innen nach Veröffentlichungserlaubnis zu fragen, da sich diese mit ihren Profilen auf den SNS in der Öffentlichkeitosphäre bewegen. Bei der Formulierung der Stimuli wurde größtmögliche Sorgfalt darauf gelegt, dass die Aussagen einerseits dem Anspruch an theoriegeleiteter Manipulation der Hate Intensität und externer Validität genügten, ohne Bedenken hinsichtlich persönlichkeitsrechtlicher Aspekte aufkommen zu lassen.

Wie die Ergebnisse der LfM-Studie (2021) belegen, bestehen bei der SNS-Nutzung im Alltag vielfältige Berührungspunkte mit Hate Speech. Deshalb war davon auszugehen, dass die Befragten bereits vor der Erhebung mit mehr oder minder starken Ausprägungen von Inzivilität, Hate Speech oder Dangerous Speech konfrontiert gewesen waren. Dennoch waren nach der Rezeption von Hate oder Dangerous Speech im Rahmen dieser Studie zumindest kurzfristige psychische Beeinträchtigungen erwartbar. Um den möglichen Schaden bei den Befragten möglichst gering zu halten, wurden mehrere Maßnahmen integriert, um die Selbstwirksamkeit der Befragten zu aktivieren. Dazu gehörte vorab die informierte Einwilligung, ein Exit-Button unmittelbar nach der Stimuluspräsentation und ein ausführliches Debriefing, welches abschließend über das Forschungsinteresse informierte. Während Letzteres methodisch keinerlei Probleme nach sich zieht, da es nach Abschluss der Erhebung stattfindet, kann bereits die informierte Einwilligung einen Prime für das Forschungsthema aktivieren. Wie von einschlägigen Kodizes in der Umfrageforschung gefordert, wurden die Befragten darin bezüglich der üblichen Rahmenbedingungen in Kenntnis gesetzt (Gewährleistung der Anonymität und Vertraulichkeit der gewonnenen Daten, Freiwilligkeit der Teilnahme, Kontaktadresse). Was die Ziele der Untersuchung betrifft, erfolgte keinerlei Täuschung, wenngleich das Studienthema mit „Diskussionen in sozialen Medien“ eher vage gehalten wurde. Der Hinweis auf mögliche problematische Facebook-Kommentare war für jene Interessierte gedacht, die bereits vor Beginn der Erhebung ihre Teilnahme zurückziehen wollten.

Methodisch am umstrittensten ist sicherlich der Exit-Button, den Befragte nach der Stimuluspräsentation betätigen konnten, um die Befragung abubrechen. Dieser wurde nur in Studie I eingesetzt, da er zu massiven Abbruchquoten führte. In Studie II wurde dieser entfernt, wodurch sich der Prozentsatz an Abbrüchen erwartungsgemäß erheblich verringerte. Der Verzicht auf den Exit-Button schien in Studie II gerechtfertigt, da es sich um mehr oder minder erfahrene Umfrageteilnehmende handelte, für die der Abbruch einer Befragung kein Novum<sup>56</sup> darstellte, wobei auch die Operationalisierung von Hate Speech in der Studie nicht das übliche Maß an toxischer Kommunikation in SNS überstieg. Trotzdem stellt sich die Frage, ob die Erfahrung der Befragten im Umgang mit als aversiv empfundenen Befragungsthemen ausreicht, um sich als selbstwirksam in der Befragungssituation zu erleben oder ob die Thematik so schwerwiegend ist, dass sie von einem Exit-Button optional Gebrauch machen müssen. Diesen methodischen Fragestellungen gilt es im Kontext von Hate Speech und ähnlichen

---

<sup>56</sup> Auch in Studie I handelt es sich um erfahrene Panelteilnehmende, aber der Panelbetreiber hatte den Exit-Button zur Bedingung des Zugangs zum Panelpool gemacht.

---

Online-Phänomenen zukünftig nachzugehen, um empirische Lösungen für das Forscherdilemma zu finden.

Welche Rezeptions- und Persuasionswirkungen von sexistischem Hate Speech auf sekundäre Adressat:innen ausgehen, war Untersuchungsziel dieser Arbeit, womit sich der Informationsverarbeitungsprozess nachvollziehen lässt. Dazu wurden die Versuchspersonen in eine als belastend empfundene Lage versetzt, denn sie rezipierten Hate Speech, was sie womöglich unter realweltlichen Umständen vermieden hätten, weil es etwa nicht ihren Nutzungsgewohnheiten entspricht, die Ansicht der Kommentarspalten aufzuklappen. Dennoch ist die Konfrontation mit Hate Speech bei der Nutzung von SNS eher die Regel als die Ausnahme, weil selbst in „harmlosen“ Netzwerken wie Nachbarschaftsplattformen Hate Speech vorkommen kann. Rechtfertigt das Forschungsinteresse die Vorgehensweise aus forschungsethischer Sicht? Die Antwort muss aus Sicht der Autorin bejaht werden, da zum einen zahlreiche Vorkehrungen getroffen wurden, um Auswege aus der psychischen Belastungssituation aufzuzeigen. Zudem ist aufgrund der Ubiquität dysfunktionaler Online-Kommunikation anzunehmen, dass bei allen Proband:innen hinreichend Erfahrungen in der Rezeption von Hate Speech oder zumindest von Inzivilität bestehen. Zum anderen sind die Befunde nützlich, um medienpädagogische Aufklärungsarbeit zu unterstützen. Hate Speech in SNS richtet für primäre Adressatengruppen immense Schäden an, die auch in Hassverbrechen münden können. Allerdings gedeiht Hate Speech in SNS in einem Nährboden, der als gesellschaftlicher Spiegel bewertet werden muss. Sexismus entsteht nicht primär im Netz, sondern ist als Einstellungsstruktur in Individuen ebenso wie Rassismus verankert. Die Befunde der Arbeit sind diesbezüglich eindeutig: Sexistische Einstellungen verstärken die Wirkung von Hate Speech. Dies mindert jedoch nicht die tragende Rolle von SNS bei der Verbreitung von Hate Speech durch deren Nutzungsalgorithmen – genauso wenig wie die dysfunktionalen Wirkungen, die von Hate Speech ausgehen. Plädiert wird deshalb für eine integrative Betrachtung von Lebenswelt und Abwertungsmechanismus im gesellschaftlichen Diskurs wie in der Wissenschaft, ohne Lebenswelt und Abwertungsmechanismus gegeneinander auszuspielen.

Forschungsethische Herausforderungen für die methodisch-empirische Herangehensweise ergeben sich zwangsläufig aus der Rezeptionsweise von Hate Speech als Untersuchungsgegenstand, was letztendlich auch die Ergebnisse beeinflussen kann. Der Forderung nach verantwortungsvoller Forschung und dem Bestreben, schwerwiegenden Schaden von Versuchspersonen fernzuhalten und ihr Wohlergehen im Blick zu haben, wurde in dieser Untersuchung Rechnung getragen. Wenn jedoch Einbußen in der Datenqualität und in den Resultaten als Konsequenz aus forschungsethischer Rücksichtnahme zu erwarten sind,

müssten dann eher nichtreaktive Erhebungsmethoden bevorzugt zum Einsatz kommen? Inhaltsanalysen eignen sich, das Aufkommen von Hate Speech zu dokumentieren, indem sowohl die Ausprägung, angegriffene Adressatengruppen oder kommunikative Inhalte erfasst werden. Allerdings bleibt mit dieser Methode die Rezeptionswirkung weiter Gegenstand von Spekulation und Plausibilitätsüberlegungen. Die vorliegende Arbeit hat trotz widriger Umstände diese Herausforderung gewagt, um primär den Kenntnisstand über die Rezeptions- und Persuasionswirkung von Hate Speech zu erhöhen.

## Literaturverzeichnis

- Abele, A. E. (2000). A dual-impact model of gender and career-related processes. In T. Eckes & H. M. Trautner (Hrsg.), *The developmental social psychology of gender* (S. 361–389). Lawrence Erlbaum Associates.
- Abele, A. E. (2003). The dynamics of masculine-agentive and feminine communal traits: Findings from a prospective study. *Journal of Personality and Social Psychology, 85*(4), 768–776. <https://doi.org/10.1037/0022-3514.85.4.768>
- Abele, A. E. & Spurk, D. (2011). The dual impact of gender and the influence of timing of parenthood on men's and women's career development: Longitudinal findings. *International Journal of Behavioral Development, 35*(3), 225–232. <https://doi.org/10.1177/0165025411398181>
- Adams, C. (2018). "They go for gender first": The nature and effect of sexist abuse of female technology journalists. *Journalism Practice, 12*(7), 850–869. <https://doi.org/10.1080/17512786.2017.1350115>
- Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. & Sanford, R. N. (1969). *The authoritarian personality*. Norton.
- Adya, M. P. (2008). Women at work: Differences in IT career experiences and perceptions between South Asian and American women. *Human Resource Management, 47*(3), 601–635. <https://doi.org/10.1002/hrm.20234>
- Afridi, F., Li, S. X. & Ren, Y. (2015). Social identity and inequality: The impact of China's hukou system. *Journal of Public Economics, 123*, 17–29. <https://doi.org/10.1016/j.jpubeco.2014.12.011>
- Agresti, A. (2007). *An introduction to categorical data analysis*. John Wiley & Sons.
- Ahuja, M., Ogan, C., Herring, S. & Robinson, J. C. (2006). Gender and career choice determinants in information systems professionals: A comparison with computer science. In F. Niederman & T. W. Ferratt (Hrsg.), *IT Workers: Human capital issues in a knowledge-based environment* (S. 277–302). Information Age Publishing.
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- Ajzen, I. & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Prentice-Hall.
- Ajzen, I. & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracin, B. T. Johnson & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 173–221). Psychology Press.
- Ajzen, I., Joyce, N., Sheikh, S. & Cote, N. G. (2011). Knowledge and the prediction of behavior: The role of information accuracy in the theory of planned behavior. *Basic and Applied Social Psychology, 33*(2), 101–117. <https://doi.org/10.1080/01973533.2011.568834>
- Al-Ansari, J., Honkala, E. & Honkala, S. (2003). Oral health knowledge and behavior among male health sciences college students in Kuwait. *BMC Oral Health, 3*(1). <https://doi.org/10.1186/1472-6831-3-2>
- Alhabash, S., Baek, J.-H., Cunningham, C. & Hagerstrom, A. (2015). To comment or not to comment? How virality, arousal level, and commenting behavior on YouTube

- videos affect civic behavioral intentions. *Computers in Human Behavior*, 51, 520–531. <https://doi.org/10.1016/j.chb.2015.05.036>
- Alhabash, S., McAlister, A. R., Hagerstrom, A., Quilliam, E. T., Rifon, N. J. & Richards, J. I. (2013). Between likes and shares: Effects of emotional appeal and virality on the persuasiveness of anticyberbullying messages on Facebook. *Cyberpsychology, Behavior, and Social Networking*, 16(3), 175–182. <https://doi.org/10.1089/cyber.2012.0265>
- Alhabash, S., McAlister, A. R., Lou, C. & Hagerstrom, A. (2015). From clicks to behaviors: The mediating effect of intentions to like, share, and comment on the relationship between message evaluations and offline behavioral intentions. *Journal of Interactive Advertising*, 15(2), 82–96. <https://doi.org/10.1080/15252019.2015.1071677>
- Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.
- Allport, G. W. (1985). The historical background of social psychology. In G. Lindzey & E. Aronson (Hrsg.), *Handbook of social psychology. Volume I. Theory and method* (3. Aufl., S. 1–46). Random House.
- Altemeyer, B. (1981). *Right-wing authoritarianism*. University of Manitoba Press.
- Altemeyer, B. (1988). *Enemies of freedom: Understanding right wing authoritarianism*. Jossey-Bass.
- Alvarez-Benjumea, A. & Winter, F. (2018). Normative change and culture of hate: An experiment on online environments. *SSRN Electronic Journal*. Vorab-Onlinepublikation. <https://doi.org/10.2139/ssrn.3126089>
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A. & Ladwig, P. (2014). The “nasty effect:” Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19, 373–387. <https://doi.org/10.1111/jcc4.12009>
- Angst, C. M. & Agarwald, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly*, 33(2), 339–370.
- Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N. & Evershed, J. K. (2020). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*. Vorab-Onlinepublikation. <https://doi.org/10.3758/s13428-020-01501-5>
- Archer, J. (1996). Sex differences in social behavior: Are the social role and evolutionary explanations compatible? *American Psychologist*, 51(9), 909–917.
- Arendt, F., Brosius, H.-B. & Hauck, P. (2017). Die Auswirkung des Schlüsselereignisses „Silvesternacht in Köln“ auf die Kriminalitätsberichterstattung. *Publizistik*, 62(2), 135–152. <https://doi.org/10.1007/s11616-017-0322-z>
- Aro, J. (2016). The cyberspace war: Propaganda and trolling as warfare tools. *European View*, 15(1), 121–132. <https://doi.org/10.1007/s12290-016-0395-5>
- Aronson, E., Wilson, T. D. & Akert, R. M. (2014). *Sozialpsychologie* (8., aktualisierte Aufl.). Pearson.
- Asbury, M. B. & Haas, J. (2008). *An exploratory investigation of whether individuals differentiate between hate speech and offensive language* [Paper submitted to the Applied Communication Division for presentation at the 2008 annual meeting of the National Communication Association, San Diego, CA.].

- Ashmore, R. D. & Del Boca, F. K. (1979). Sex stereotypes and implicit personality theory: Toward a cognitive—Social psychological conceptualization. *Sex Roles, 5*(2), 219–248. <https://doi.org/10.1007/BF00287932>
- Atkinson, D. C. (2018). Charlottesville and the alt-right: a turning point? *Politics, Groups, and Identities, 6*(2), 309–315. <https://doi.org/10.1080/21565503.2018.1454330>
- Austin, J. L. (1975 [1955]). *How to do things with words: The William James Lectures delivered at Harvard University in 1955* (2. ed. (reprint)). Oxford University Press.
- Autorengruppe Bildungsberichterstattung (Hrsg.). (2020). *Bildung in Deutschland 2020: Ein indikatorengestützter Bericht mit einer Analyse zu Bildung in einer digitalisierten Welt*. <https://doi.org/10.3278/6001820gw>
- Averbeck-Lietz, S. (2015). *Soziologie der Kommunikation: Die Mediatisierung der Gesellschaft und die Theoriebildung der Klassiker*. De Gruyter Oldenbourg.
- Awan, I. (2016). Islamophobia on social media: A qualitative analysis of the Facebook's walls of hate. *International Journal of Cyber Criminology, 10*(1), 1–20. <https://doi.org/10.5281/zenodo.58517>
- Axsom, D., Yates, S. & Chaiken, S. (1987). Audience response as a heuristic cue in persuasion. *Journal of Personality and Social Psychology, 53*(1), 30–40. <https://doi.org/10.1037/0022-3514.53.1.30>
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2018). *Multivariate Analysemethoden: Eine anwendungsorientierte Einführung* (15., vollständig überarbeitete Auflage). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-56655-8>
- Baek, J. & Bullock, L. M. (2014). Cyberbullying: a cross-cultural perspective. *Emotional and Behavioural Difficulties, 19*(2), 226–238. <https://doi.org/10.1080/13632752.2013.849028>
- Bak, P. (2017). Offene und versteckte Aggression im Gebrauch von Dysphemismen und Euphemismen. In S. Bonacchi (Hrsg.), *Verbale Aggression: Multidisziplinäre Zugänge zur verletzenden Macht der Sprache* (S. 145–168). De Gruyter Mouton. <https://doi.org/10.1515/9783110522976-007>
- Bandura, A., Underwood, B. & Fromson, M. E. (1975). Disinhibition of aggression through diffusion of responsibility and dehumanization of victims. *Journal of Research in Personality, 9*(4), 253–269. [https://doi.org/10.1016/0092-6566\(75\)90001-X](https://doi.org/10.1016/0092-6566(75)90001-X)
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effect. In S. Chaiken & Y. Trope (Hrsg.), *Dual-process theories in social psychology* (S. 361–382). Guilford Press.
- Barlett, C. (2015). Anonymously hurting others online: The effect of anonymity on cyberbullying frequency. *Psychology of Popular Media Culture, 4*(2), 70–79. <https://doi.org/10.1037/a0034335>
- Barlett, C. & Coyne, S. M. (2014). A meta-analysis of sex differences in cyber-bullying behavior: The moderating role of age. *Aggressive Behavior, 40*(5), 474–488. <https://doi.org/10.1002/ab.21555>
- Barlett, C., DeWitt, C. C., Maronna, B. & Johnson, K. (2018). Social media use as a tool to facilitate or reduce cyberbullying perpetration: A review focusing on anonymous and nonanonymous social media platforms. *Violence and Gender, 5*(3), 147–152. <https://doi.org/10.1089/vio.2017.0057>
- Barnes, A. & Ephross, P. H. (1994). The impact of hate violence on victims: Emotional and behavioral responses to attacks. *Social Work, 39*(3), 247–251.

- Barreto, M. & Ellemers, N. (2005). The burden of benevolent sexism: How it contributes to the maintenance of gender inequalities. *European Journal of Social Psychology*, 35(5), 633–642. <https://doi.org/10.1002/ejsp.270>
- Barrett, L. F. & Bliss-Moreau, E. (2009). She's emotional. He's having a bad day: Attributional explanations for emotion stereotypes. *Emotion*, 9(5), 649–658. <https://doi.org/10.1037/a0016821>
- Barthels, I. (2021). Nicht jede Kritik an Baerbock ist frauenfeindlich. *Tagesspiegel*. <https://www.tagesspiegel.de/politik/alles-taktik-bei-den-gruenen-nicht-jede-kritik-an-baerbock-ist-frauenfeindlich/27263962.html>
- Baruch, Y. (2005). Bullying on the net: Adverse behavior on e-mail and its impact. *Information & Management*, 42(2), 361–371. <https://doi.org/10.1016/j.im.2004.02.001>
- Beck, K. (2013a). Handlungstheorie. In G. Bentele, H.-B. Brosius & O. Jarren (Hrsg.), *Lexikon Kommunikations- und Medienwissenschaft* (2. Aufl., S. 112–113). Springer VS.
- Beck, K. (2013b). Kommunikation. In G. Bentele, H.-B. Brosius & O. Jarren (Hrsg.), *Lexikon Kommunikations- und Medienwissenschaft* (2. Aufl., S. 155–156). Springer VS.
- Becker, J. C. (2014). Subtile Erscheinungsformen von Sexismus. *Aus Politik und Zeitgeschichte*, 64(8), 29–34.
- Beisch, N., Koch, W. & Schäfer, C. (2019). ARD/ZDF-Onlinestudie 2019: Mediale Internetnutzung und Video-on-Demand gewinnen weiter an Bedeutung. *Media Perspektiven*(9), 374–388.
- Beißert, H., Köhler, M., Rempel, M. & Beierlein, C. (2014). *Eine deutschsprachige Kurzsкала zur Messung des Konstrukts Need for Cognition: Die Need for Cognition Kurzsкала (NFC-K)*. GESIS - Leibniz-Institut für Sozialwissenschaften.
- Beißwenger, M. (2007). *Sprachhandlungskoordination in der Chat-Kommunikation*. De Gruyter.
- Ben-David, A. & Matamoros-Fernández, A. (2016). Hate speech and covert discrimination on social media: Monitoring the Facebook pages of extreme-right political parties in Spain. *International Journal of Communication*, 10, 1167–1193.
- Benesch, S. (2013). *The Dangerous Speech Project: Dangerous Speech: A Proposal to Prevent Group Violence*. [www.worldpolicy.org/sites/default/files/Dangerous%20Speech%20Guidelines%20Benesch%20January%202012.pdf](http://www.worldpolicy.org/sites/default/files/Dangerous%20Speech%20Guidelines%20Benesch%20January%202012.pdf)
- Benesch, S. (2014). Flower Speech: New Responses to Hatred Online. *Internet Monitor 2014: Reflections on the Digital World*, 100–104. [https://dash.harvard.edu/bitstream/handle/1/13632937/IM2014\\_ReflectionsontheDigitalWorld%5B1%5D.pdf?sequence=1](https://dash.harvard.edu/bitstream/handle/1/13632937/IM2014_ReflectionsontheDigitalWorld%5B1%5D.pdf?sequence=1)
- Berlet, C. (2001). *When hate went online*. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.552.239&rep=rep1&type=pdf>
- Beyer, S. (1999). The accuracy of academic gender stereotypes. *Sex Roles*, 40(9/10), 787–813.
- Bielefeldt, H. (2010). Das Diskriminierungsverbot als Menschenrechtsprinzip. In U. Hormel & A. Scherr (Hrsg.), *Diskriminierung: Grundlagen und Forschungsergebnisse* (S. 21–34). VS Verlag für Sozialwissenschaften.
- Biernat, M. & Vescio, T. K. (2002). She swings, she hits, she's great, she's benched: Implications of gender-based shifting standards for judgment and behavior. *Personality and Social Psychology Bulletin*, 28(1), 66–77.

- Bilewicz, M., Soral, W., Marchlewska, M. & Winiewski, M. (2017). When authoritarians confront prejudice. Differential effects of SDO and RWA on support for hate-speech prohibition. *Political Psychology, 38*(1), 87–99. <https://doi.org/10.1111/pops.12313>
- Bishop, J. (2007). Increasing participation in online communities: A framework for human–computer interaction. *Computers in Human Behavior, 23*, 1881–1893. <https://doi.org/10.1016/j.chb.2005.11.004>
- Bitkom. (2018). *Frauenanteil in der ITK-Branche wächst langsam*. <https://www.bitkom.org/Presse/Presseinformation/Frauenanteil-in-der-ITK-Branche-waechst-langsam.html>
- Bitner, M. J. & Obermiller, C. (1985). The elaboration likelihood model: Limitations and extensions in marketing. *Advances in Consumer Research, 12*(1), 420–425.
- Blakemore, B. (2016). Online hate and political activist groups. In I. Awan (Hrsg.), *Islamophobia in Cyberspace: Hate Crimes Go Viral* (S. 63–83). Ashgate Publishing.
- Blaya, C. & Audrin, C. (2019). Toward an understanding of the characteristics of secondary school cyberhate perpetrators. *Frontiers in Education, 4*, Artikel 46, 179. <https://doi.org/10.3389/educ.2019.00046>
- Blazak, R. (2001). White boys to terrorist men. *American Behavioral Scientist, 44*(6), 982–1000. <https://doi.org/10.1177/00027640121956629>
- Blee, K. (2004, 20. März). *Positioning hate*. Paper presented at the conference to establish the field of Hate Studies. <http://hdl.handle.net/20.500.12389/20060>
- Bloch, R. H. (1987). Medieval misogyny. *Representations, 20*, 1–24.
- Bloch, R. H. (1991). *Medieval misogyny and the invention of Western romantic love*. University of Chicago Press; EBSCO Industries, Inc.
- Bodenhausen, G. V. (1993). Emotions, arousal, and stereotyping judgements: A heuristic model of affect and stereotyping. In D. M. Mackie & D. L. Hamilton (Hrsg.), *Affect, cognition and stereotyping: Interactive processes in group perception* (S. 13–37). Academic Press.
- Bodenhausen, G. V., Kramer, G. P. & Süsler, K. (1994). Happiness and stereotypic thinking in social judgment. *Journal of Personality and Social Psychology, 66*(4), 621–632. <https://doi.org/10.1037/0022-3514.66.4.621>
- Bovermann, P. (2021). Unerträglich nah. <https://www.sueddeutsche.de/kultur/querdenken-atilla-hildmann-alles-dicht-machen-hate-speech-verschwörungstheorien-corona-hass-im-netz-1.5289254>
- Boyd, D. M. & Ellison, N. B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication, 13*(1), 210–230. <https://doi.org/10.1111/j.1083-6101.2007.00393.x>
- Boyle, K. (2001). Hate Speech - the United States versus the rest of the world? *Maine Law Review, 53*(2), 488-502.
- Bradshaw, S. & Howard, P. N. (2017). *Troops, trolls and troublemakers: A global inventory of organized social media manipulation*. University of Oxford. <https://ora.ox.ac.uk/objects/uuid:cef7e8d9-27bf-4ea5-9fd6-855209b3e1f6>
- Brandt, J. (2014). “As seen on TV”: Gender, television, and popular culture. In A. M. Trier-Bieniek & P. Leavy (Hrsg.), *Gender & pop culture: A text-reader* (S. 103–120). SensePublishers.
- Brandt, M. J. (2011). Sexism and gender inequality across 57 societies. *Psychological Science, 22*(11), 1413–1418. <https://doi.org/10.1177/095679761142044>

- Brasse, C. (2003). *Junge Branche, alte Muster. Vom Arbeiten und Leben in den Neuen Medien: Daten und Analysen zur Arbeitssituation der Beschäftigten in der Multimediabranche. Ergebnisse der bundesweiten Umfrage von connexx.av.* [https://www.connexx-av.de/upload/m47287322eafab\\_verweis1.pdf](https://www.connexx-av.de/upload/m47287322eafab_verweis1.pdf)
- Breckler, S. J. & Wiggins, E. C. (1991). Cognitive responses in persuasion: Affective and evaluative determinants. *Journal of Experimental Social Psychology, 27*(2), 180–200. [https://doi.org/10.1016/0022-1031\(91\)90021-W](https://doi.org/10.1016/0022-1031(91)90021-W)
- Brehm, J. W. (1966). *A theory of psychological reactance*. Academic Press.
- Brescoll, V. L., Okimoto, T. G. & Vial, A. C. (2018). You've come a long way...maybe: How moral emotions trigger backlash against women leaders. *Journal of Social Issues, 74*(1), 144–164. <https://doi.org/10.1111/josi.12261>
- Brescoll, V. L. & Uhlmann, E. L. (2008). Can an angry woman get ahead? Status conferral, gender, and expression of emotion in the workplace. *Psychological Science, 19*(3), 268–275.
- Brescoll, V. L., Uhlmann, E. L., Moss-Racusin, C. & Sarnell, L. (2012). Masculinity, status, and subordination: Why working for a gender stereotype violator causes men to lose status. *Journal of Experimental Social Psychology, 48*, 354–357. <https://doi.org/10.1016/j.jesp.2011.06.005>
- Brett, J. M., Olekalns, M., Friedman, R., Goates, N., Anderson, C. & Lisco, C. C. (2007). Sticks and stones: Language, face, and online dispute resolution. *Academy of Management Journal, 50*(1), 85–99. <https://doi.org/10.5465/amj.2007.24161853>
- Brewer, G. & Kerslake, J. (2015). Cyberbullying, self-esteem, empathy and loneliness. *Computers in Human Behavior, 48*, 255–260. <https://doi.org/10.1016/j.chb.2015.01.073>
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love or outgroup hate? *Journal of Social Issues, 55*(3), 429–444.
- Brewer, M. B. (2001). Ingroup identification and intergroup conflict. When does ingroup love become outgroup hate? In R. D. Ashmore, L. J. Jussim & D. Wilder (Hrsg.), *Social identity, intergroup conflict, and conflict reduction* (S. 17–41). Oxford University Press.
- Briñol, P. & Petty, R. E. (2012). A history of attitudes and persuasion research. In A. Kruglanski & W. Stroebe (Hrsg.), *Handbook of the history of social psychology* (S. 285–320). Psychology Press.
- Brosius, H.-B., Haas, A. & Koschel, F. (2016). *Methoden der empirischen Kommunikationsforschung: Eine Einführung* (7., überarbeitete und aktualisierte Auflage). VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-531-19996-2>
- Brown, R. & Gaertner, S. L. (Hrsg.). (2001). *Blackwell handbook of social psychology: Intergroup processes*. Blackwell.
- Buckels, E. E., Jones, D. N. & Paulhus, D. L. (2013). Behavioral confirmation of everyday sadism. *Psychological Science, 24*(11), 2201–2209. <https://doi.org/10.1177/0956797613490749>
- Buckels, E. E., Trapnell, P. D., Andjelovic, T. & Paulhus, D. L. (2019). Internet trolling and everyday sadism: Parallel effects on pain perception and moral judgment. *Journal of Personality, 87*(2), 328–340. <https://doi.org/10.1111/jopy.12393>
- Buckels, E. E., Trapnell, P. D. & Paulhus, D. L. (2014). Trolls just want to have fun. *Personality and Individual Differences, 67*, 97–102. <https://doi.org/10.1016/j.paid.2014.01.016>

- Bugarski, R. (2001). Language, nationalism and war in Yugoslavia. *International Journal of the Sociology of Language*(151), 69–87. <https://doi.org/10.1515/ijsl.2001.048>
- Bundeskriminalamt. (2020). *Partnerschaftsgewalt. Kriminalstatistische Auswertung. Berichtsjahr 2019*. [https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Partnerschaftsgewalt/Partnerschaftsgewalt\\_2019.html;jsessionid=501E45DE2544069A9D59DA39DC8B5B83.live302?nn=63476](https://www.bka.de/SharedDocs/Downloads/DE/Publikationen/JahresberichteUndLagebilder/Partnerschaftsgewalt/Partnerschaftsgewalt_2019.html;jsessionid=501E45DE2544069A9D59DA39DC8B5B83.live302?nn=63476)
- Burgoon, J. K., Dunbar, N. E. & Segrin, C. (2002). Nonverbal influence. In J. P. Dillard & M. Pfau (Hrsg.), *The persuasion handbook: Developments in theory and practice* (S. 446–474). SAGE Publications. <https://doi.org/10.4135/9781412976046.n23>
- Burgoon, J. K., Floyd, K. & Guerrero, L. K. (2010). Nonverbal communication theories of interaction adaptation. In C. R. Berger, M. E. Roloff & D. R. Roskos-Ewoldsen (Hrsg.), *The handbook of communication science* (2. Aufl., S. 93–108). Sage. <https://doi.org/10.4135/9781412982818.n6>
- Burkart, R. (2002). *Kommunikationswissenschaft. Grundlagen und Problemfelder. Umrisse einer interdisziplinären Sozialwissenschaft* (4. überarbeitete und aktualisierte Auflage). Böhlau.
- Burkart, R. (2003). Kommunikationstheorien. In G. Bentele, H.-B. Brosius & O. Jarren (Hrsg.), *Öffentliche Kommunikation: Handbuch Kommunikations- und Medienwissenschaft* (S. 169–192). VS Verlag für Sozialwissenschaften.
- Burnap, P. & Williams, M. L. (2015). Cyber hate speech on Twitter: An application of machine classification and statistical modeling for policy and decision making. *Policy and Internet*, 7(2), 223–242. <https://doi.org/10.1002/poi3.85>
- Burris, V., Smith, E. & Strahm, A. (2000). White supremacist networks on the internet. *Sociological Focus*, 33(2), 215–235. <https://doi.org/10.1080/00380237.2000.10571166>
- Butler, J. (2018). *Haß spricht: Zur Politik des Performativen* (6. Auflage). Suhrkamp.
- Buyse, A. (2014). Words of violence: "Fear speech", or how violent conflict escalation relates to the freedom of expression. *Human Rights Quarterly*, 36(4), 779–797. <https://doi.org/10.1353/hrq.2014.0064>
- Cacioppo, J. T. & Petty, R. E. (1979). Effects of message repetition and position on cognitive response recall, and persuasion. *Journal of Personality and Social Psychology*, 37(1), 97–109.
- Cacioppo, J. T. & Petty, R. E. (1981). Social psychology procedures for cognitive response assessment: The thought-listing technique. In T. V. Merluzzi, C. R. Glass & M. Genest (Hrsg.), *Cognitive assessment* (S. 309–342). Guilford Press.
- Cacioppo, J. T. & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131.
- Cacioppo, J. T., Petty, R. E. & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment*, 48(3), 306–307.
- Cacioppo, J. T., Petty, R. E., Kao, C. F. & Rodriguez, R. (1986). Central and peripheral routes to persuasion: An individual difference perspective. *Journal of Personality and Social Psychology*, 51(5), 1032–1043.
- Cacioppo, J. T., Petty, R. E. & Morris, K. J. (1983). Effects of need for cognition on message evaluation, recall, and persuasion. *Journal of Personality and Social Psychology*, 45(4), 805–818.

- Cameron, A. C. & Trivedi, P. K. (2013). *Regression analysis of count data* (2. Aufl.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139013567>
- Cameron, J. E. (2001). Social identity, modern sexism, and perceptions of personal and group discrimination by women and men. *Sex Roles, 45*(11/12), 743–766. <https://doi.org/10.1023/A:1015636318953>
- Cammaerts, B. (2009). Radical pluralism and free speech in online public spaces: The case of North Belgian extreme right discourses. *International Journal of Cultural Studies, 12*(6), 555–575. <https://doi.org/10.1177/1367877909342479>
- Campbell, B., Schellenberg, E. G. & Senn, C. Y. (1997). Evaluation measures of contemporary sexism. *Psychology of Women Quarterly, 21*, 89–102.
- Caprariello, P. A., Cuddy, A. J. & Fiske, S. T. (2009). Social structure shapes cultural stereotypes and emotions: A causal test of the stereotype content model. *Group Processes & Intergroup Relations, 12*(2), 147–155. <https://doi.org/10.1177/1368430208101053>
- Carli, L. L., Alawa, L., Lee, Y., Zhao, B. & Kim, E [Elaine] (2016). Stereotypes about gender and science. *Psychology of Women Quarterly, 40*(2), 244–260. <https://doi.org/10.1177/0361684315622645>
- Cason, K. L. & Wenrich, T. R. (2002). Health and nutrition beliefs, attitudes, and practices of undergraduate college students: A needs assessment. *Topics in Clinical Nutrition, 17*(3), 52–70.
- Cassino, D. & Besen - Cassino, Y. (2019). Race, threat and workplace sexual harassment: The dynamics of harassment in the United States, 1997–2016. *Gender, Work & Organization, 26*(9), 1221–1240. <https://doi.org/10.1111/gwao.12394>
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*(5), 752–766. <https://doi.org/10.1037/0022-3514.39.5.752>
- Chaiken, S. & Eagly, A. H. (1983). Communication modality as a determinant of persuasion: The role of communicator salience. *Journal of Personality and Social Psychology, 45*(2), 241–256.
- Chaiken, S. & Stangor, C. (1987). Attitudes and attitude change. *Annual Review of Psychology, 38*(1), 575–630. <https://doi.org/10.1146/annurev.ps.38.020187.003043>
- Chandrasekharan, E., Pavalanathan, U., Srinivasan, A., Glynn, A., Eisenstein, J. & Gilbert, E. (2017). You can't stay here: The efficacy of Reddit's 2015 ban examined through hate speech. *Proceedings ACM Human-Computer Interaction, 1*(2), 1–22. <https://doi.org/10.1145/3134666>
- Chang, C. & Hitchon, J. C. B. (2004). When does gender count? Further insights into gender schematic processing of female candidates' political advertisements. *Sex Roles, 51*(3/4), 197–208. <https://doi.org/10.1023/B:SERS.0000037763.47986.c2>
- Cheryan, S., Plaut, V. C., Davies, P. G. & Steele, C. M. (2009). Ambient belonging: How stereotypical cues impact gender participation in computer science. *Journal of Personality and Social Psychology, 97*(6), 1045–1060. <https://doi.org/10.1037/a0016239>
- Cheryan, S., Plaut, V. C., Handron, C. & Hudson, L. (2013). The stereotypical computer scientist: Gendered media representations as a barrier to inclusion for women. *Sex Roles, 69*, 58–71. <https://doi.org/10.1007/s11199-013-0296-x>

- Chetty, N. & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior, 40*, 108–118. <https://doi.org/10.1016/j.avb.2018.05.003>
- Cho, C.-H. (1999). How advertising works on the WWW: Modified elaboration likelihood model. *Journal of Current Issues & Research in Advertising, 21*(1), 34–50. <https://doi.org/10.1080/10641734.1999.10505087>
- Cho, D. & Kwon, K. H. (2015). The impacts of identity verification and disclosure of social cues on flaming in online user comments. *Computers in Human Behavior, 51*, 363–372. <https://doi.org/10.1016/j.chb.2015.04.046>
- Christie, R. & Geis, F. L. (Hrsg.). (1970). *Studies in machiavellianism*. Academic Press.
- Christofides, E., Muise, A. & Desmarais, S. (2009). Information disclosure and control on Facebook: Are they two sides of the same coin or two different processes? *Cyberpsychology & Behavior, 12*(3), 341–345. <https://doi.org/10.1089/cpb.2008.0226>
- Cialdini, R. B. & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Hrsg.), *The handbook of social psychology* (Bd. 2, S. 151–192). McGraw-Hill.
- Citron, D. K. (2014). *Hate crimes in cyberspace*. Harvard University Press.
- Citron, D. K. & Norton, H. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review, 91*, 1435–1484.
- Claus, R. (2014). *Maskulismus. Antifeminismus zwischen vermeintlicher Salonfähigkeit und unverhohlenem Frauenhass*. Friedrich Ebert Stiftung. Forum Politik und Gesellschaft. <http://library.fes.de/pdf-files/dialog/10861.pdf>
- Cloetta, B. (1983). Der Fragebogen zur Erfassung von Machiavellismus und Konservatismus. *Schweizerische Zeitschrift für Psychologie und ihre Anwendungen*(2-3), 127–159. <https://doi.org/10.6102/zis82>
- Coe, K., Kenski, K. & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication, 64*(4), 658–679. <https://doi.org/10.1111/jcom.12104>
- Coester, M. (2007). *Hate crimes*. Peter Lang.
- Coles, B. A. & West, M. (2016). Trolling the trolls: Online forum users constructions of the nature and properties of trolling. *Computers in Human Behavior, 60*, 233–244. <https://doi.org/10.1016/j.chb.2016.02.070>
- Collani, G. & Werner, R. (2002). *Eine deutsche Skala zu ambivalent-sexistischen Einstellungen (Hostilität und Benevolenz) gegenüber Männern (ASEM)*. Technische Universität Braunschweig, Braunschweig.
- Connelly, K. & Heesacker, M. (2012). Why is benevolent sexism appealing? *Psychology of Women Quarterly, 36*(4), 432–443. <https://doi.org/10.1177/0361684312456369>
- Connor, R. A., Glick, P. & Fiske, S. T. (2016). Ambivalent sexism in the twenty-first century. In C. G. Sibley & F. K. Barlow (Hrsg.), *Cambridge Handbooks in Psychology. The Cambridge handbook of the psychology of prejudice* (S. 295–320). Cambridge University Press.
- Correll, J., Judd, C. M., Park, B. & Wittenbrink, B. (2010). Measuring prejudice, stereotypes and discrimination. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Hrsg.), *The SAGE handbook of prejudice, stereotyping and discrimination* (S. 45–62). Sage.
- Correll, S. J., Benard, S. & Paik, I. (2007). Getting a job: Is there a motherhood penalty? *American Journal of Sociology, 112*(5), 1297–1338.

- Costello, M. & Hawdon, J. (2018). Who are the online extremists among us? Sociodemographic characteristics, social networking, and online experiences of those who produce online hate materials. *Violence and Gender*, 5(1), 55–60. <https://doi.org/10.1089/vio.2017.0048>
- Costello, M., Hawdon, J. & Ratliff, T. N. (2017). Confronting online extremism: The effect of self-help, collective efficacy, and guardianship on being a target for hate speech. *Social Science Computer Review*, 35(5), 587–605. <https://doi.org/10.1177/0894439316666272>
- Costello, M., Hawdon, J., Ratliff, T. & Grantham, T. (2016). Who views online extremism? Individual attributes leading to exposure. *Computers in Human Behavior*, 63, 311–320. <https://doi.org/10.1016/j.chb.2016.05.033>
- Cottrell, C. A. & Neuberg, S. L. (2005). Different emotional reactions to different groups: a sociofunctional threat-based approach to "prejudice". *Journal of Personality and Social Psychology*, 88(5), 770–789. <https://doi.org/10.1037/0022-3514.88.5.770>
- Craig, K. M. (2012). Examining hate-motivated aggression: A review of the social psychological literature on hate crimes as a distinct form of aggression. In B. Perry (Hrsg.), *Hate and bias crime: A reader* (S. 117–130). Taylor and Francis.
- Craker, N. & March, E. (2016). The dark side of Facebook®: The dark tetrad, negative social potency, and trolling behaviours. *Personality and Individual Differences*, 102, 79–84. <https://doi.org/10.1016/j.paid.2016.06.043>
- Crisp, R. J., Heuston, S., Farr, M. J. & Turner, R. N. (2007). Seeing red or feeling blue: Differentiated intergroup emotions and ingroup identification in soccer fans. *Group Processes & Intergroup Relations*, 10(1), 9–26. <https://doi.org/10.1177/1368430207071337>
- Crowley, J. P. (2014). Expressive writing to cope with hate speech: Assessing psychobiological stress recovery and forgiveness promotion for lesbian, gay, bisexual, or queer victims of hate speech. *Human Communication Research*, 40, 238–261. <https://doi.org/10.1111/hcre.12020>
- Crown Prosecution Service. (2018). *Hate crime annual report 2017-18*. <https://www.cps.gov.uk/sites/default/files/documents/publications/cps-hate-crime-report-2018.pdf>
- Cuddy, A. J., 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., Fiske, S. T. & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, 92(4), 631–648. <https://doi.org/10.1037/0022-3514.92.4.631>
- Cuddy, A. J., Fiske, S. T. & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the bias map. In M. P. Zanna & J. M. Olson (Hrsg.), *Advances in experimental social psychology* (S. 61–149). Academic Press. [https://doi.org/10.1016/S0065-2601\(07\)00002-0](https://doi.org/10.1016/S0065-2601(07)00002-0)
- Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S [Stéphanie], Leyens, J.-P [Jaques-Philippe], Bond, M. H., Croizet, J.-C., Ellemers, N., Sleebos, E., Htun, T. T., Kim, H.-J., Maio, G. R., Perry, J., Petkova, K., Todorov, V., Rodriguez-Bailon, R., Morales, E., Moya, M., . . . Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology*, 48, 1–33. <https://doi.org/10.1348/014466608X31493>

- Cunningham, G. B. (2005). The importance of a common in-group identity in ethnically diverse groups. *Group Dynamics: Theory, Research, and Practice*, 9(4), 251–260. <https://doi.org/10.1037/1089-2699.9.4.251>
- Daft, R. L. & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, 32(5), 554–571. <https://doi.org/10.1287/mnsc.32.5.554>
- Davidson, R. & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. Oxford University Press. <https://doi.org/10.2307/2554780>
- Davidson, T., Warmesley, D., Macy, M. & Weber, I. (2017). *Automated hate speech detection and the problem of offensive language* (Proceedings of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017)). <https://www.aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/viewPaper/15665>
- Davison, W. P. (1983). The third-person effect in communication. *Public Opinion Quarterly*, 47(1), 1–15. <https://doi.org/10.1086/268763>
- Deaux, K. & Kite, M. (1993). Gender stereotypes. In F. Denmark & M. A. Paludi (Hrsg.), *Psychology of women: A handbook of issues and theories* (S. 107–139). Greenwood Press.
- Delgado, R. (1982). Words that wound. A tort action for racial insults, epithets, and name-calling. *Harvard Civil Rights-Civil Liberties Law Review*, 17, 133–181.
- Delgado, R. & Stefancic, J. (2009). Four observations about hate speech. *Wake Forest Law Review*, 44, 353–370.
- Delgado, R. & Stefancic, J. (2014). Hate speech in cyberspace. *Wake Forest Law Review*, 49(319), 1–21.
- Delgado, R. & Yun, D. (1995). "The speech we hate": First Amendment totalism, the ACLU, and the principle of dialogic politics. *Arizona State Law Journal*, 27, 1281–1300.
- Di Blasio, P. & Milani, L. (2008). Computer-mediated communication and persuasion: Peripheral vs. central route to opinion shift. *Computers in Human Behavior*, 24, 798–815. <https://doi.org/10.1016/j.chb.2007.02.011>
- Didden, R., Scholte, R. H. J., Korzilius, H., Moor, J. M. H. de, Vermeulen, A., O'Reilly, M., Lang, R. & Lancioni, G. E. (2009). Cyberbullying among students with intellectual and developmental disability in special education settings. *Developmental Neuropsychology*, 12(3), 146–151. <https://doi.org/10.1080/17518420902971356>
- Diekman, A. B., Brown, E. R., Johnston, A. M. & Clark, E. K. (2010). Seeking congruity between goals and roles: A new look at why women opt out of science, technology, engineering, and mathematics careers. *Psychological Science*, 21(8), 1051–1057. <https://doi.org/10.1177/0956797610377342>
- Diekmann, A. (2007). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen* (12. Aufl.). Rororo Rowohlt's Enzyklopädie: Bd. 55678. Rowohlt Taschenbuch Verlag.
- Dinoff, B. L. & Kowalski, R. M. (1999). Reducing AIDS risk behavior: The combined efficacy of protection motivation theory and the elaboration likelihood model. *Journal of Social and Clinical Psychology*, 18(2), 223–239. <https://doi.org/10.1521/jscp.1999.18.2.223>
- Disha, I., Cavendish, J. C. & King, R. D. (2011). Historical events and spaces of hate: Hate crimes against Arabs and Muslims in Post-9/11 America. *Social Problems*, 58(1), 21–46. <https://doi.org/10.1525/sp.2011.58.1.21>

- Donath, J. S. (1999). Identity and deception in the virtual community. In M. A. Smith & P. Kollock (Hrsg.), *Communities in cyberspace* (S. 29–59). Routledge.
- Döring, N. & Mohseni, M. R. (2020). Gendered hate speech in YouTube and YouNow comments: Results of two content analyses. *Studies in Communication and Media*, 9(1), 62–88. <https://doi.org/10.5771/2192-4007-2020-1-62>
- Douglas, K. M., McGarty, C., Bliuc, A.-M. & Lala, G. (2005). Understanding cyberhate: Social competition and social creativity in online white supremacist groups. *Social Science Computer Review*, 23(1), 68–76. <https://doi.org/10.1177/0894439304271538>
- Douglas, S. C., Kiewitz, C., Martinko, M. J., Harvey, P., Kim, Y. & Chun, J. U. (2008). Cognitions, emotions, and evaluations: An elaboration likelihood model for workplace aggression. *Academy of Management Review*, 33(2), 425–451.
- Dovidio, J. F. & Gaertner, S. L. (Hrsg.). (1986a). *Prejudice, discrimination, and racism*. Academic Press.
- Dovidio, J. F. & Gaertner, S. L. (1986b). Prejudice, discrimination, and racism: Historical trends and contemporary approaches. In J. F. Dovidio & S. L. Gaertner (Hrsg.), *Prejudice, discrimination, and racism* (S. 1–34). Academic Press.
- Dovidio, J. F., Hewstone, M., Glick, P. & Esses, V. M. (2010). Prejudice, stereotyping and discrimination: Theoretical and empirical overview. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Hrsg.), *The SAGE handbook of prejudice, stereotyping and discrimination* (S. 3–28). Sage. <https://doi.org/10.4135/9781446200919.n1>
- Downs, D. A. (1985). Skokie revisited: Hate group speech and the First Amendment. *Notre Dame Law Review*, 60(4), 629–685.
- Duckitt, J. & Sibley, C. G. (2010). Personality, ideology, prejudice, and politics: A dual-process motivational model. *Journal of Personality*, 78(6), 1861–1893. <https://doi.org/10.1111/j.1467-6494.2010.00672.x>
- Dunbar, N. E., Connelly, S., Jensen, M. L., Adame, B. J., Rozzell, B., Griffith, J. A. & O'Hair, H. D. (2014). Fear appeals, message processing cues, and credibility in the websites of violent, ideological, and nonideological groups. *Journal of Computer-Mediated Communication*, 19, 871–889. <https://doi.org/10.1111/jcc4.12083>
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum Associates.
- Eagly, A. H. (1997). Sex differences in social behavior: Comparing social role theory and evolutionary psychology. *American Psychologist*(December), 1380–1383.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.
- Eagly, A. H. & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037//0033-295X.109.3.573>
- Eagly, A. H., Karau, S. J. & Makhijani, M. G. (1995). Gender and the effectiveness of leaders: A meta-analysis. *Psychological Bulletin*, 117(1), 125–145. <https://doi.org/10.1037/0033-2909.117.1.125>
- Eagly, A. H., Makhijani, M. G. & Klonsky, B. G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111(1), 3–22. <https://doi.org/10.1037/0033-2909.111.1.3>

- Eagly, A. H. & Steffen, V. J. (1984). Gender stereotypes stem from the distribution of women and men into social roles. *Journal of Personality and Social Psychology*, 46(4), 735–754.
- Eagly, A. H. & Wood, W. (2012). Social role theory. In P. A. van Lange, A. Kruglanski & E. T. Higgins (Hrsg.), *Handbook of theories of social psychology* (1. Aufl., S. 458–476). Sage. <https://doi.org/10.4135/9781446249222.n49>
- Eagly, A. H., Wood, W. & Chaiken, S. (1978). Causal inferences about communicators and their effect on opinion change. *Journal of Personality and Social Psychology*, 36(4), 424–435.
- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, 47(3/4), 99–114.
- Eckes, T. (2010). Geschlechterstereotype: Von Rollen, Identitäten und Vorurteilen. In R. Becker & B. Kortendiek (Hrsg.), *Handbuch Frauen- und Geschlechterforschung* (3. Aufl., S. 178–189). 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), 221–228. <https://doi.org/10.1024//0044-3514.30.4.211>
- Eifler, S. & Leitgöb, H. (2019). Experiment. In N. Baur & J. Blasius (Hrsg.), *Handbuch Methoden der empirischen Sozialforschung* (2. Aufl., S. 203–218). Springer Fachmedien Wiesbaden. [https://doi.org/10.1007/978-3-658-21308-4\\_13](https://doi.org/10.1007/978-3-658-21308-4_13)
- Ellison, N. B. & Boyd, D. M. (2013). Sociality through social network sites. In W. H. Dutton (Hrsg.), *The Oxford handbook of internet studies* (S. 151–172). Oxford University Press.
- Erb, H.-P., Pierro, A., Mannetti, L., Spiegel, S. & Kruglanski, A. W. (2007). Biassed processing of persuasive information: On the functional equivalence of cues and message arguments. *European Journal of Social Psychology*, 37(5), 1057–1075. <https://doi.org/10.1002/ejsp.420>
- Erjavec, K. & Kovacic, M. P. (2012). “You don’t understand, this is a new war!” Analysis of hate speech in news web sites’ comments. *Mass Communication & Society*, 15, 899–920. <https://doi.org/10.1080/15205436.2011.619679>
- Esses, V. M., Medianu, S. & Lawson, A. S. (2013). Uncertainty, threat, and the role of the media in promoting the dehumanization of immigrants and refugees. *Journal of Social Issues*, 69(3), 518–536.
- Farley, S., Coyne, I., Sprigg, C., Axtell, C. & Subramanian, G. (2015). Exploring the impact of workplace cyberbullying on trainee doctors. *Medical Education*, 49(4), 436–443. <https://doi.org/10.1111/medu.12666>
- Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Hrsg.), *Handbook of motivation and cognition: Foundations of social behavior* (S. 204–243). Guilford Press.
- Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, 23, 75–109. [https://doi.org/10.1016/S0065-2601\(08\)60318-4](https://doi.org/10.1016/S0065-2601(08)60318-4)

- Fazio, R. H., Ledbetter, J. E. & Towles-Schwen, T. (2000). On the costs of accessible attitudes: Detecting that the attitude object has changed. *Journal of Personality and Social Psychology*, 78(2), 197–210. <https://doi.org/10.1037//0022-3514.78.2.197>
- Fazio, R. H. & Roskos-Ewoldsen, D. R. (2005). Acting as we feel: When and how attitudes guide behavior. In T. C. Brock & M. C. Green (Hrsg.), *Persuasion: Psychological insights and perspectives* (2. Aufl., S. 41–62). SAGE Publications.
- Fekkes, M., Pijpers, F. I. M., Fredriks, A. M., Vogels, T. & Verloove-Vanhorick, S. P. (2006). Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. *Pediatrics*, 117(5), 1568–1574. <https://doi.org/10.1542/peds.2005-0187>
- Felser, G. (2015). *Werbe- und Konsumentenpsychologie*. Springer Verlag. [https://doi.org/10.1007/978-3-642-37645-0\\_14](https://doi.org/10.1007/978-3-642-37645-0_14)
- Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson.
- Festinger, L. & Maccoby, N. (1964). On resistance to persuasive communications. *Journal of Abnormal Psychology*, 68, 359–366. <https://doi.org/10.1037/h0049073>
- Festl, R. (2015). *Täter im Internet: Eine Analyse individueller und struktureller Erklärungsfaktoren von Cybermobbing im Schulkontext*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-09239-9>
- Festl, R. (2016). Perpetrators on the internet: Analyzing individual and structural explanation factors of cyberbullying in school context. *Computers in Human Behavior*, 59, 237–248. <https://doi.org/10.1016/j.chb.2016.02.017>
- Festl, R., Scharkow, M. & Quandt, T. (2015). The individual or the group: A multilevel analysis of cyberbullying in school classes. *Human Communication Research*, 41(4), 535–556. <https://doi.org/10.1111/hcre.12056>
- Feuchter, A. & Funke, J. (2004). Positive Effekte sozialen Faulenzens beim Lösen komplexer Probleme. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 56(2), 304–325.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th edition). Sage.
- Fink, C. (2018). Dangerous speech, anti-Muslim violence, and Facebook in Myanmar. *Journal of International Affairs*, 71(1.5), 43–52.
- Fiske, S. T. (2000). Stereotyping, prejudice, and discrimination at the seam between the centuries: evolution, culture, mind, and brain. *European Journal of Social Psychology*, 30, 299–322.
- Fiske, S. T. (2015). Intergroup biases: A focus on stereotype content. *Current Opinion in Behavioral Sciences*, 3, 45–50. <https://doi.org/10.1016/j.cobeha.2015.01.010>.
- Fiske, S. T., Cuddy, A. J. & Glick, P. (2006). Universal dimensions of social cognition: warmth and competence. *TRENDS in Cognitive Sciences*, 11(2). <https://doi.org/10.1016/j.tics.2006.11.005>
- Fiske, S. T., Cuddy, A. J., 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. <https://doi.org/10.1037//0022-3514.82.6.87>
- Fiske, S. T. & North, M. S. (2015). Measures of stereotyping and prejudice: Barometers of bias. In G. J. Boyle, D. H. Saklofske & G. Matthews (Hrsg.), *Measures of personality and social psychological constructs* (S. 684–718). Elsevier; Academic Press. <https://doi.org/10.1016/B978-0-12-386915-9.00024-3>

- Fiske, S. T. & Russell, A. M. (2010). Cognitive processes. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Hrsg.), *The SAGE handbook of prejudice, stereotyping and discrimination* (S. 115–130). Sage. <https://doi.org/10.4135/9781446200919.n7>
- Fleischhack, J. (2017). Der "Hass" der vielen Formen. In K. Kaspar, L. Gräßer & A. Riffi (Hrsg.), *Schriftenreihe zur digitalen Gesellschaft NRW: Band 4. Online Hate Speech* (S. 23–28). kopaed.
- Flynn, B. S., Worden, J. K., Bunn, J. Y., Connolly, S. W. & Dorwaldt, A. L. (2011). Evaluation of smoking prevention television messages based on the elaboration likelihood model. *Health Education Research*, 26(6), 976–987. <https://doi.org/10.1093/her/cyr082>
- Foschi, M. & Valenzuela, J. (2012). Who is the better applicant? Effects from gender, academic record, and type of decision. *Social Science Research*, 41(4), 949–964. <https://doi.org/10.1016/j.ssresearch.2012.02.001>
- Fox, J., Cruz, C. & Lee, J. Y. (2015). Perpetuating online sexism offline: Anonymity, interactivity, and the effects of sexist hashtags on social media. *Computers in Human Behavior*, 52, 436–442. <https://doi.org/10.1016/j.chb.2015.06.024>
- Fox, J. & Rooney, M. C. (2015). The Dark Triad and trait self-objectification as predictors of men's use and self-presentation behaviors on social networking sites. *Personality and Individual Differences*, 76, 161–165. <https://doi.org/10.1016/j.paid.2014.12.017>
- Fraas, C., Meier, S. & Pentzold, C. (2012). *Online-Kommunikation: Grundlagen, Praxisfelder und Methoden*. Oldenbourg.
- Frees, B. & Koch, W. (2018). ARD/ZDF-Onlinestudie 2018: Zuwachs bei medialer Internetnutzung und Kommunikation. *Media Perspektiven*(9), 398–413.
- Friedman, R., Anderson, C., Brett, J., Olekalns, M., Goates, N. & Lisco, C. C. (2004). The positive and negative effects of anger on dispute resolution: evidence from electronically mediated disputes. *The Journal of Applied Psychology*, 89(2), 369–376. <https://doi.org/10.1037/0021-9010.89.2.369>
- Friedman, R. & Currall, S. C. (2003). Conflict escalation: Dispute exacerbating elements of e-mail communication. *Human Relations*, 56(11), 1325–1347. <https://doi.org/10.1177/00187267035611003>
- Friesel, E. (2013). Juden-Hass gestern und heute: Ein historischer Blick auf 130 Jahre judeo-phobische Feindseligkeit. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 17–27). Gießener Elektronische Bibliothek.
- Frischlich, L., Boberg, S. & Quandt, T. (2017). Unmenschlicher Hass: Die Rolle von Empfehlungsalgorithmen und Social Bots für die Verbreitung von Cyberhate. In K. Kaspar, L. Gräßer & A. Riffi (Hrsg.), *Schriftenreihe zur digitalen Gesellschaft NRW: Band 4. Online Hate Speech* (S. 71–79). kopaed.
- Frischlich, L., Boberg, S. & Quandt, T. (2019). Comment sections as targets of dark participation? Journalists' evaluation and moderation of deviant user comments. *Journalism Studies*, 19(2), 1–20. <https://doi.org/10.1080/1461670X.2018.1556320>
- Früh, W. (2007). *Inhaltsanalyse: Theorie und Praxis* (6., überarb. Aufl.). UVK.
- Fuchs, T. & Schäfer, F. (2020). Normalizing misogyny: hate speech and verbal abuse of female politicians on Japanese Twitter. *Japan Forum*, 1–27. <https://doi.org/10.1080/09555803.2019.1687564>

- Gagliardone, I. (2019). Defining online hate and its "public lives": What is the place for "extreme speech"? *International Journal of Communication*, 13, 3068–3087.
- Gagliardone, I., Gal, D., Alves, T. & Martinez, G. (2015). *Countering online hate speech*. United Nations Educational, Scientific and Cultural Organization.
- Gagliardone, I., Pohjonen, M., Beyene, Z., Zerai, A., Aynekulu, G., Bekalu, M., Bright, J., Moges, M., Seifu, M., Stremlau, N., Taflan, P., Gebrewolde, T. M. & Teferra, Z. (2016). *Mechachal: Online debates and elections in Ethiopia - From hate speech to engagement in social media*.
- Ganz, K. & Meßmer, A.-K. (2015). Anti-Genderismus im Internet. Digitale Öffentlichkeiten als Labor eines neuen Kulturkampfes. In S. Hark & P.-I. Villa (Hrsg.), *(Anti-)Genderismus: Sexualität und Geschlecht als Schauplätze aktueller politischer Auseinandersetzungen* (S. 59–77). transcript.
- Garcia, L., McDevitt, J., Gu, J. & Balboni, J. (1999). *Psychological and behavioral effects of bias and non-bias motivated assault*. Washington, DC. U.S. Department of Justice, National Institute of Justice.
- Garcia-Moreno, C., Jansen, H. A. F. M., Ellsberg, M., Heise, L. & Watts, C. (2005). *WHO multi-country study on women's health and domestic violence against women. Initial results on prevalence, health outcomes and women's responses*. Geneva.
- Gardner, D., O'Driscoll, M., Cooper-Thomas, H. D., Roche, M., Bentley, T., Catley, B., Teo, S. T. T. & Trenberth, L. (2016). Predictors of Workplace Bullying and Cyber-Bullying in New Zealand. *International journal of environmental research and public health*, 13(5), 448. <https://doi.org/10.3390/ijerph13050448>
- Gawronski, B., Hofmann, W. & Wilbur, C. J. (2006). Are "implicit" attitudes unconscious? *Consciousness and Cognition*, 15(3), 485–499. <https://doi.org/10.1016/j.con-cog.2005.11.007>
- Gee, G. C. (2002). A multilevel analysis of the relationship between institutional and individual racial discrimination and health status. *American Journal of Public Health*, 92(4), 615–623.
- Geissler, B. & Oechsle, M. (2000). Die Modernisierung weiblicher Lebenslagen. *Aus Politik und Zeitgeschichte*(B 31-32), 11–17.
- Gelber, K. & McNamara, L. (2015). Evidencing the harms of hate speech. *Social Identities*, 1–14. <https://doi.org/10.1080/13504630.2015.1128810>
- Gensch, K. & Kliegl, C. (2011). *Studienabbruch. Was können Hochschulen dagegen tun? Bewertung der Maßnahmen aus der Initiative „Wege zu mehr MINT-Absolventen“*. München.
- George, C. (2015). Hate speech law and policy. In R. Mansell & P. H. Ang (Hrsg.), *The international encyclopedia of digital communication and society* (S. 305–314). John Wiley & Sons.
- Gerhards, J. & Neidhardt, F. (1991). Strukturen und Funktionen moderner Öffentlichkeit. In S. Müller-Doohm & K. Neumann-Braun (Hrsg.), *Öffentlichkeit, Kultur, Massenkommunikation* (S. 31–89). Bibliotheks- und Informationssystem der Universität Oldenburg.
- Gerstenfeld, P. B., Grant, D. R. & Chiang, C.-P. (2003). Hate online: A content analysis of extremist internet sites. *Analysis of Social Issues and Public Policy*, 3(1), 29–44.

- Gervais, B. T. (2015). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics*, 12(2), 167–185. <https://doi.org/10.1080/19331681.2014.997416>
- Gervais, S. J. & Hillard, A. L. (2011). A role congruity perspective on prejudice toward Hillary Clinton and Sarah Palin. *Analyses of Social Issues and Public Policy*, 11(1), 221–240. <https://doi.org/10.1111/j.1530-2415.2011.01263.x>
- Geschke, D., Klaßen, A., Quent, M. & Richter, C. (2019). *Hass im Netz: Der schleichende Angriff auf unsere Demokratie. Eine bundesweite repräsentative Untersuchung*. [https://www.idz-jena.de/fileadmin/user\\_upload/\\_Hass\\_im\\_Netz\\_-\\_Der\\_schleichende\\_Angriff.pdf](https://www.idz-jena.de/fileadmin/user_upload/_Hass_im_Netz_-_Der_schleichende_Angriff.pdf)
- Gilmore, D. D. (2001). *Misogyny: The male malady*. University of Pennsylvania Press.
- Glaser, J., Dixit, J. & Green, D. P. (2002). Studying hate crime with the internet: What makes racists advocate racial violence? *Journal of Social Issues*, 58(1), 177–193.
- Glick, P., Diebold, J., Bailey-Werner, B. & Zhu, L. (1997). The two faces of Adam: Ambivalent sexism and polarized attitudes towards 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. (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. *American Psychologist*, 56(2), 109–118. <https://doi.org/10.1037/0003-066X.56.2.109>
- Glick, P. & Fiske, S. T. (2001b). Ambivalent sexism. In M. P. Zanna (Hrsg.), *Advances in experimental social psychology* (33. Aufl., S. 115–188). Academic Press.
- Glick, P., Fiske, S. T., Mladinic, A., Saiz, J. L., Abrams, D., Masser, B., Adetoun, B., Osagie, J. E., Akande, A., Alao, A., Annetje, B., Willemsen, T. M., Chipeta, K., Dardenne, B., Dijksterhuis, A., Wigboldus, D., Eckes, T., Six-Materna, I., Expósito, F., . . . López, W. L. (2000). Beyond prejudice as simple antipathy: Hostile and benevolent sexism across cultures. *Journal of Personality and Social Psychology*, 79(5), 763–775. <https://doi.org/10.1037/0022-3514.79.5.763>
- Glick, P., Lameiras, M., Fiske, S. T., Eckes, T., Masser, B., Volpato, C., Manganelli, A. M., Pek, J. C. X., Huang, L.-L., Sakalli-Ugurlu, N., Rodríguez Castro, Y., Pereira, M. L. D., Willemsen, T. M., Brunner, A., Six-Materna, I. & Wells, R. (2004). Bad but bold: Ambivalent attitudes toward men predict gender inequality in 16 nations. *Journal of Personality and Social Psychology*, 86(5), 713–728. <https://doi.org/10.1037/0022-3514.86.5.713>
- Glick, P. & Rudman, L. A. (2010). Sexism. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Hrsg.), *The SAGE handbook of prejudice, stereotyping and discrimination* (S. 328–344). Sage. <https://doi.org/10.4135/9781446200919.n20>
- Goldberg, D. T. (1995). Afterword: Hate, or power? In R. K. Whillock & D. Slayden (Hrsg.), *Hate speech* (S. 267–275). SAGE Publications.
- Golf-Papez, M. & Veer, E. (2017). Don't feed the trolling: rethinking how online trolling is being defined and combated. *Journal of Marketing Management*, 33(15-16), 1336–1354. <https://doi.org/10.1080/0267257X.2017.1383298>

- Görz y Moratalla, M., Kolvenbach, M. & Schader, N. (2021). *Baerbock am häufigsten Ziel von Fake News*. <https://www.tagesschau.de/investigativ/swr/fakenews-bundestagswahl-baerbock-laschet-101.html>
- Grässle, K. (2009). *Frau Dr. Ing: Wege ebnen für Frauen in technische Studiengänge*. Budrich.
- Green, D. P., Abelson, R. P. & Garnett, M. (1999). The distinctive political views of hate-crime perpetrators and white supremacists. In D. A. Prentice & D. T. Miller (Hrsg.), *Cultural divides: Understanding and overcoming group conflict* (S. 429–464). Russell Sage Foundation.
- Green, D. P., Glaser, J. & Rich, A. (1998). From lynching to gay bashing: The elusive connection between economic conditions and hate crime. *Journal of Personality and Social Psychology*, 75(1), 82–92.
- Greenwald, A. G. (1968). Cognitive learning, cognitive response to persuasion, and attitude change. In A. G. Greenwald, T. C. Brock & T. M. Ostrom (Hrsg.), *Psychological foundations of attitudes* (S. 147–170). Academic Press.
- Greenwald, A. G. & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>
- Greenwald, A. G., McGhee, D. E. & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480.
- Groshek, J. & Cutino, C. (2016). Meaner on mobile: Incivility and impoliteness in communicating contentious politics on sociotechnical networks. *Social Media + Society*(October-December), 1–10. <https://doi.org/10.1177/2056305116677137>
- Guadagno, R. E. & Cialdini, R. B. (2002). Online persuasion: An examination of gender differences in computer-mediated interpersonal influence. *Group Dynamics: Theory, Research, and Practice*, 6(1), 38–51. <https://doi.org/10.1037/1089-2699.6.1.38>
- Guadagno, R. E. & Cialdini, R. B. (2007). Persuade him by email, but see her in person: Online persuasion revisited. *Computers in Human Behavior*, 23, 999–1015. <https://doi.org/10.1016/j.chb.2005.08.006>
- Gunitsky, S. (2015). Corrupting the cyber-commons: Social media as a tool of autocratic stability. *Perspectives on Politics*, 13(1), 42–54. <https://doi.org/10.1017/S1537592714003120>
- Gunther, A. C. (1991). What We Think Others Think. *Communication Research*, 18(3), 355–372. <https://doi.org/10.1177/009365091018003004>
- Gupta, V. K., Han, S., Mortal, S. C., Silveri, S. & Turban, D. B. (2018). Do women CEOs face greater threat of shareholder activism compared to male CEOs? A role congruity perspective. *Journal of Applied Psychology*, 103(2), 228–236. <https://doi.org/10.1037/apl0000269>
- Habermas, J. (1992). *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats*. Suhrkamp.
- Hammond, M. D. & Overall, N. C. (2016). Sexism in intimate contexts: How romantic relationships help explain the origins, functions, and consequences of sexist attitudes. In C. G. Sibley & F. K. Barlow (Hrsg.), *Cambridge Handbooks in Psychology. The Cambridge handbook of the psychology of prejudice* (S. 321–343). Cambridge University Press.

- Hampel, T. (2002). Wofür schämt sich Harald Ewert? <https://www.tagesspiegel.de/gesellschaft/panorama/hitlergruss-wofuer-schaemt-sich-harald-ewert/1375334.html>
- Han, J. Y., Hou, J., Kim, E [Eunkyung] & Gustafson, D. H. (2014). Lurking as an active participation process: a longitudinal investigation of engagement with an online cancer support group. *Health Communication, 29*(9), 911–923. <https://doi.org/10.1080/10410236.2013.816911>
- Hanes, E. & Machin, S. (2014). Hate crime in the wake of terror attacks. *Journal of Contemporary Criminal Justice, 30*(3), 247–267. <https://doi.org/10.1177/1043986214536665>
- Hardaker, C. (2010). Trolling in asynchronous computer-mediated communication: From user discussions to academic definitions. *Journal of Politeness Research, 6*, 215–242. <https://doi.org/10.1515/JPLR.2010.011>
- Hardaker, C. (2013). “Uh...not to be nitpicky,,,,,but...the past tense of drag is dragged, not drug.”. *Journal of Language Aggression and Conflict, 1*(1), 58–86. <https://doi.org/10.1075/jlac.1.1.04har>
- Harkins, S. G. & Petty, R. E. (1982). Effects of task difficulty and task uniqueness on social loafing. *Journal of Personality and Social Psychology, 43*(6), 1214–1229.
- Haslam, A. S. & Dovidio, J. F. (2010). Prejudice. In J. M. Levine & M. A. Hogg (Hrsg.), *Encyclopedia of group processes & intergroup relations* (S. 656–660). SAGE Publications.
- Haslam, N. (2006). Dehumanization: An integrative review. *Personality and Social Psychology Review, 10*(3), 252–264. [https://doi.org/10.1207/s15327957pspr1003\\_4](https://doi.org/10.1207/s15327957pspr1003_4)
- Haslam, N. & Stratemeyer, M. (2016). Recent research on dehumanization. *Current Opinion in Psychology, 11*, 25–29. <https://doi.org/10.1016/j.copsyc.2016.03.009>
- Haugtvedt, C. P. & Petty, R. E. (1992). Personality and persuasion: Need for cognition moderates the persistence and resistance of attitude changes. *Journal of Personality and Social Psychology, 63*(2), 308–319. <https://doi.org/10.1037/0022-3514.63.2.308>
- Hauser, D. J., Ellsworth, P. C. & Gonzalez, R. (2018). Are manipulation checks necessary? *Frontiers in Psychology, 9*, 998. <https://doi.org/10.3389/fpsyg.2018.00998>
- Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach* (2. Aufl.). The Guilford Press.
- Hayes, A. F. & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior Research Methods, 39*(4), 709–722. <https://doi.org/10.3758/bf03192961>
- Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research, 21*(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>
- Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women’s ascent up the organizational ladder. *Journal of Social Issues, 57*(4), 657–674.
- Heilman, M. E. (2012). Gender stereotypes and workplace bias. *Research in Organizational Behavior, 32*, 113–135. <https://doi.org/10.1016/j.riob.2012.11.003>
- Heilman, M. E. & Okimoto, T. G. (2007). Why are women penalized for success at male tasks? The implied communality deficit. *Journal of Applied Psychology, 92*(1), 81–92. <https://doi.org/10.1037/0021-9010.92.1.81>

- Heilman, M. E., Wallen, A. S., Fuchs, D. & Tamkins, M. M. (2004). Penalties for success: Reactions to women who succeed at male gender-typed tasks. *Journal of Applied Psychology, 89*(3), 416–427. <https://doi.org/10.1037/0021-9010.89.3.416>
- Heinbach, D., Ziegele, M. & Quiring, O. (2018). Sleeper effect from below: Long-term effects of source credibility and user comments on the persuasiveness of news articles. *New Media & Society, 20*(12), 4765–4786. <https://doi.org/10.1177/1461444818784472>
- Hemmer, J. J. (2000). Hate speech codes: A narrow perspective. *North Dakota Journal of Speech & Theatre, 13*, 21–27.
- Henson, B., Reyns, B. W. & Fisher, B. S. (2013). Does gender matter in the virtual world? Examining the effect of gender on the link between online social network activity, security and interpersonal victimization. *Security Journal, 26*(4), 315–330. <https://doi.org/10.1057/sj.2013.21>
- Herek, G. M., Cogan, J. C. & Gillis, J. R. (2002). Victim experiences in hate crimes based on sexual orientation. *Journal of Social Issues, 58*(2), 319–339.
- Herring, S., Job-Sluder, K., Scheckler, R. & Barab, S. (2002). Searching for safety online: Managing "trolling" in a feminist forum. *The Information Society, 18*(5), 371–384. <https://doi.org/10.1080/01972240290108186>
- Hewstone, M., Rubin, M. & Willis, H. (2002). Intergroup bias. *Annual Review of Psychology, 53*, 575–604.
- Hinduja, S. & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research, 14*(3), 206–221. <https://doi.org/10.1080/13811118.2010.494133>
- Hinduja, S. & Patchin, J. W. (2019). Connecting adolescent suicide to the severity of bullying and cyberbullying. *Journal of School Violence, 18*(3), 333–346. <https://doi.org/10.1080/15388220.2018.1492417>
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H. & Schmitt, M. (2005). A meta-analysis on the correlation between the implicit association test and explicit self-report measures. *Personality and Social Psychology Bulletin, 31*(10), 1369–1385. <https://doi.org/10.1177/0146167205275613>
- Holbert, R. L. (2012). Media influence as persuasion. In J. P. Dillard & S. Lijiang (Hrsg.), *The SAGE handbook of persuasion: Developments in theory and practice* (S. 36–52). SAGE Publications.
- Hölig, S. & Hasebrink, U. (2018). Nachrichtennutzung und soziale Medien: Befunde aus dem Reuters Institute Digital News Survey 2018. *Media Perspektiven*(12), 574–582.
- Holland, J. (2007). *Misogynie: Die Geschichte des Frauenhasses*. Zweitausendeins.
- Holling, H. & Gediga, G. (2016). *Statistik - Testverfahren*. Hogrefe.
- Honeypot. (2018). *Frauen in der IT-Branche 2018*. Honeypot. <https://www.honeypot.io/de/women-in-tech-2018/>
- Hornscheidt, L. (2013). Der Hate Speech-Diskurs als Hate Speech: Pejorisation als konstruktivistisches Modell zur Analyse diskriminierender Sprachhandlungen. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 28–58). Gießener Elektronische Bibliothek.
- Howard, J. W. (2019). Dangerous speech. *Philosophy & Public Affairs, 47*(2), 208–254. <https://doi.org/10.1111/papa.12145>

- Hsueh, M., Yogeeswaran, K. & Malinen, S. (2015). "Leave your comment below": Can biased online comments influence our own prejudicial attitudes and behaviors? *Human Communication Research*, 41, 557–576. <https://doi.org/10.1111/hcre.12059>
- Hu, M.-C., Pavlicova, M. & Nunes, E. V. (2011). Zero-inflated and hurdle models of count data with extra zeros: examples from an HIV-risk reduction intervention trial. *The American Journal of Drug and Alcohol Abuse*, 37(5), 367–375. <https://doi.org/10.3109/00952990.2011.597280>
- Huang, F. L. & Cornell, D. G. (2012). Pick your Poisson: A tutorial on analyzing counts of student victimization data. *Journal of School Violence*, 11(3), 187–206. <https://doi.org/10.1080/15388220.2012.682010>
- Huck, I. & Brosius, H.-B. (2007). Der Third-Person-Effekt — Über den vermuteten Einfluss der Massenmedien. *Publizistik*, 52(3), 355–374. <https://doi.org/10.1007/s11616-007-0171-2>
- Hughes, B. (2019). Invalidating emotions in the non-disabled imaginary. Fear, pity and disgust. In N. Watson & S. Vehmas (Hrsg.), *Routledge handbook of disability studies* (S. 89–101). Routledge.
- Hughey, M. W. & Daniels, J. (2013). Racist comments at online news sites: A methodological dilemma for discourse analysis. *Media, Culture & Society*, 35(3), 332–347. <https://doi.org/10.1177/0163443712472089>
- Human Rights Watch. (2002). *World Report 2002*. <https://www.hrw.org/legacy/wr2k2/us.html#Hate%20Crimes>
- Hutchens, M. J., Cicchirillo, V. J. & Hmielowski, J. D. (2014). How could you think that?!?! Understanding intentions to engage in political flaming. *New Media & Society*, 17(8), 1201–1219. <https://doi.org/10.1177/1461444814522947>
- Iganski, P. (2008). *Hate crime and the city*. Policy Press. <https://doi.org/10.2307/j.ctt9qgq3n>
- Infratest dimap. (2021). *Blitz TV-Triell 12. September 2021*. <https://www.infratest-dimap.de/umfragen-analysen/bundesweit/ard-deutschlandtrend/2021/blitz-tv-triell/>
- Inter-Parliamentary Union. (2018). *Sexism, harassment and violence against women in parliaments in Europe*. <https://www.ipu.org/resources/publications/issue-briefs/2018-10/sexism-harassment-and-violence-against-women-in-parliaments-in-europe>
- Jaccard, J. & Blanton, H. (2005). The origins and structure of behavior: Conceptualizing behavior in attitude research. In D. Albarracín, B. T. Johnson & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 125–171). Psychology Press.
- Jackman, M. R. (1994). *Paternalism and conflict: Ideology and coercion in gender, class, and race relations*. University of California Press.
- Jacks, W. & Adler, J. R. (2015). A proposed typology of online hate crime. *Open Access Journal of Forensic Psychology*, 7, 64–89.
- Jane, E. A. (2014). "YOUR A UGLY, WHORISH, SLUT": Understanding E-bile. *Feminist Media Studies*, 14(4), 531–546. <https://doi.org/10.1080/14680777.2012741073>
- Jane, E. A. (2015). Flaming? What flaming? The pitfalls and potentials of researching online hostility. *Ethics and Information Technology*, 17(1), 65–87. <https://doi.org/10.1007/s10676-015-9362-0>
- Jane, E. A. (2017). *Misogyny online: A short (and brutish) history*. Sage. <https://doi.org/10.4135/9781473916029>
- Janiszewski, C. (1993). Preattentive mere exposure effects. *Journal of Consumer Research*, 20(3), 376. <https://doi.org/10.1086/209356>

- Janshen, D. & Rudolph, H. (1987). *Ingenieurinnen: Frauen für die Zukunft*. Walter de Gruyter.
- Jeong, S.-H. & Hwang, Y. (2012). Does multitasking increase or decrease persuasion? Effects of multitasking on comprehension and counterarguing. *Journal of Communication*, 62(4), 571–587. <https://doi.org/10.1111/j.1460-2466.2012.01659.x>
- Jewkes, Y. (Hrsg.). (2005). *Dot.cons: Crime, deviance and identity on the internet*. Willan.
- John, N. A. (2013). Sharing and the web 2.0: The emergence of a key-word. *New Media & Society*, 15(2), 167–182. <https://doi.org/10.1177/1461444812450684>
- Johnen, M., Jungblut, M. & Ziegele, M. (2018). The digital outcry: What incites participation behavior in an online firestorm? *New Media & Society*, 20(9), 3140–3160. <https://doi.org/10.1177/1461444817741883>
- Johnson, B. T. & Eagly, A. H. (1989). Effects of involvement on persuasion: A meta-analysis. *Psychological Bulletin*, 106(2), 290–314.
- Johnson, K. L., Freeman, J. B. & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, 102(1), 116–131. <https://doi.org/10.1037/a0025335>
- Johnson, N. A., Cooper, R. B. & Chin, W. W. (2009). Anger and flaming in computer-mediated negotiation among strangers. *Decision Support Systems*, 46, 660–672. <https://doi.org/10.1016/j.dss.2008.10.008>
- Johnston, L. & Coolen, P. (1995). A dual processing approach to stereotype change. *Personality and Social Psychology Bulletin*, 21(7), 660–673. <https://doi.org/10.1177/0146167295217001>
- Joinson, A. N. (2007). Disinhibition and the internet. In J. Gackenbach (Hrsg.), *Psychology and the Internet: Intrapersonal, interpersonal and transpersonal implications* (2. Aufl., S. 75–92). Elsevier. <https://doi.org/10.1016/B978-012369425-6/50023-0>
- Josey, C. S. (2010). Hate speech and identity: An analysis of neo racism and the indexing of identity. *Discourse & Society*, 21(1), 27–39. <https://doi.org/10.1177/0957926509345071>
- Jost, J. T. & Kay, A. C. (2005). Exposure to benevolent sexism and complementary gender stereotypes: consequences for specific and diffuse forms of system justification. *Journal of Personality and Social Psychology*, 88(3), 498–509. <https://doi.org/10.1037/0022-3514.88.3.498>
- Juvonen, J. & Gross, E. F. (2008). Extending the school grounds? Bullying experiences in cyberspace. *Journal of School Health*, 78(9), 496–505. <https://doi.org/10.1111/j.1746-1561.2008.00335.x>
- Kaakinen, M., Oksanen, A. & Räsänen, P. (2018). Did the risk of exposure to online hate increase after the November 2015 Paris attacks? A group relations approach. *Computers in Human Behavior*, 78, 90–97. <https://doi.org/10.1016/j.chb.2017.09.022>
- Kaakinen, M., Räsänen, P., Näsi, M., Minkkinen, J., Keipi, T. & Oksanen, A. (2018). Social capital and online hate production: A four country survey. *Crime, Law, and Social Change*, 69, 25–39. <https://doi.org/10.1007/s10611-017-9764-5>
- Kahn, K. B., Spencer, K. & Glaser, J. (2013). Online prejudice and discrimination: From dating to hating. In Y. Amichai-Hamburger (Hrsg.), *The social net: Understanding our online behavior* (2. Aufl., S. 201–219). Oxford University Press.
- Kahneman, D. (1973). *Attention and effort*. Prentice Hall.

- Kaigo, M. & Watanabe, I. (2007). Ethos in chaos? Reaction to video files depicting socially harmful images in the channel 2 Japanese internet forum. *Journal of Computer-Mediated Communication*, 12(4), 1248–1268. <https://doi.org/10.1111/j.1083-6101.2007.00372.x>
- Kalch, A. & Naab, T. K. (2017). Replying, disliking, flagging: How users engage with uncivil and impolite comments on news sites. *Studies in Communication and Media*, 6(4), 395–419. <https://doi.org/10.5771/2192-4007-2017-4-395>
- Karau, S. J. & Williams, K. D. (1993). Social loafing: A meta-analytic review and theoretical integration. *Journal of Personality and Social Psychology*, 65(4), 681–706.
- Karpinski, A. & Hilton, J. L. (2001). Attitudes and the implicit association test. *Journal of Personality and Social Psychology*, 81(5), 774–788. <https://doi.org/10.1037//0022-3514.81.5.774>
- Kaspar, K. (2017). Hassreden im Internet – Ein besonderes Phänomen computervermittelter Kommunikation? In K. Kaspar, L. Gräßer & A. Riffi (Hrsg.), *Schriftenreihe zur digitalen Gesellschaft NRW: Band 4. Online Hate Speech* (S. 63–70). kopaed.
- Kaspar, K., Gräßer, L. & Riffi, A. (Hrsg.). (2017). *Schriftenreihe zur digitalen Gesellschaft NRW: Band 4. Online Hate Speech*. kopaed.
- Kayany, J. M. (1998). Contexts of uninhibited online behavior: Flaming in social newsgroups on usenet. *Journal of the American Society for Information Science*, 49(12), 1135–1141.
- Keipi, T., Näsi, M., Oksanen, A. & Räsänen, P. (2017). *Online hate and harmful content: Cross-national perspectives*. Routledge.
- Kellow, C. L. & Steeves, H. L. (1998). The role of radio in the Rwandan genocide. *Journal of Communication*, 48(3), 107–128. <https://doi.org/10.1111/j.1460-2466.1998.tb02762.x>
- Kemp, S. (2018). *Digital 2018: Germany*. We are social. <https://datareportal.com/reports/digital-2018-germany?rq=digital%202018%3A%20germany>
- Kepplinger, H. M. (2001). Der Ereignisbegriff in der Publizistikwissenschaft. *Publizistik*, 46(2), 117–139. <https://doi.org/10.1007/s11616-001-0032-3>
- Kerr, G., Schultz, D. E., Kitchen, P. J., Mulhern, F. J. & Beede, P. (2015). Does traditional advertising theory apply to the digital world? *Journal of Advertising Research*, 55(4), 390–400. <https://doi.org/10.2501/JAR-2015-001>
- Kessler, T. & Fritsche, I. (2018). *Sozialpsychologie*. Springer Fachmedien Wiesbaden.
- Kiesler, S. & Sproull, L. (1992). Group decision making and communication technology. *Organizational Behavior and Human Decision Processes*, 52, 96–123.
- Kim, A. J. (2000). *Community building on the Web*. Peachpit Press.
- Kim, H.-S. & Sundar, S. S. (2014). Can online buddies and bandwagon cues enhance user participation in online health communities? *Computers in Human Behavior*, 37, 319–333. <https://doi.org/10.1016/j.chb.2014.04.039>
- Kimotho, S. G. & Nyaga, R. N. (2016). Digitized ethnic hate speech: Understanding effects of digital media hate speech on citizen journalism in Kenya. *Advances in Language and Literary Studies*, 7(3). <https://doi.org/10.7575/aiac.all.v.7n.3p.189>
- King, R. D. & Sutton, G. M. (2013). High times for hate crimes: Explaining the temporal clustering of hate-motivated offending. *Criminology*, 51(4), 871–894. <https://doi.org/10.1111/1745-9125.12022>

- Kitchen, P. J., Kerr, G., E. Schultz, D., McColl, R. & Pals, H. (2014). The elaboration likelihood model: Review, critique and research agenda. *European Journal of Marketing*, 48(11/12), 2033–2050. <https://doi.org/10.1108/EJM-12-2011-0776>
- Klapper, J. T. (1960). *The effects of mass communication*. Free Press.
- Klaus, E. & Lünenborg, M. (2013). Zwischen (Post-)Feminismus und Antifeminismus: Reflexionen zu gegenwärtigen Geschlechterdiskursen in den Medien. *Gender: Zeitschrift für Geschlecht, Kultur und Gesellschaft*, 5(2), 78–93.
- Klimmt, C. (2011). *Das Elaboration-Likelihood-Modell*. Nomos.
- Klimpel, L. (2021). *Wie Politikerinnen im Netz diskreditiert werden*. <https://www.tageschau.de/faktenfinder/geschlechtsspezifische-desinformation-101.html>
- Knoll, J. (2015). *Persuasion in sozialen Medien: Der Einfluss nutzergenerierter Inhalte auf die Rezeption und Wirkung von Onlinewerbung*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-10577-8>
- Koehler, C. & Weber, M [Mathias] (2018). "Do I really need to help?!" Perceived severity of cyberbullying, victim blaming, and bystanders' willingness to help the victim. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 12(4). <https://doi.org/10.5817/CP2018-4-4>
- König, M. & König, W. (2016). *#MythosTwitter: Chancen und Grenzen eines sozialen Mediums*. [https://www.otto-brenner-stiftung.de/fileadmin/user\\_data/stiftung/02\\_Wissenschaftsportal/03\\_Publikationen/AP24\\_Twitter\\_Koenig\\_2016\\_10\\_05.pdf](https://www.otto-brenner-stiftung.de/fileadmin/user_data/stiftung/02_Wissenschaftsportal/03_Publikationen/AP24_Twitter_Koenig_2016_10_05.pdf)
- Kowalski, R. M. & Limber, S. P. (2013). Psychological, physical, and academic correlates of cyberbullying and traditional bullying. *The Journal of Adolescent Health*, 53(1), 13–20. <https://doi.org/10.1016/j.jadohealth.2012.09.018>
- Kowalski, R. M., Limber, S. & Agatston, P. W. (2012). *Cyberbullying: Bullying in the digital age* (2. Aufl.). Wiley Blackwell.
- Kowalski, R. M., Morgan, C. A., Drake-Lavelle, K. & Allison, B. (2016). Cyberbullying among college students with disabilities. *Computers in Human Behavior*, 57, 416–427. <https://doi.org/10.1016/j.chb.2015.12.044>
- Kreißel, P., Ebner, J., Urban, A. & Guhl, J. (2018). *Hass auf Knopfdruck. Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz*. London u.a. Institute for Strategic Dialogue (ISD). [http://www.isdglobal.org/wp-content/uploads/2018/07/ISD\\_Ich\\_Bin\\_Hier\\_2.pdf](http://www.isdglobal.org/wp-content/uploads/2018/07/ISD_Ich_Bin_Hier_2.pdf)
- Krosnick, J. A., Judd, C. M. & Wittenbrink, B. (2005). The measurement of attitudes. In D. Albarracín, B. T. Johnson & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 21–76). Psychology Press.
- Kteily, N. & Bruneau, E. (2017). Darker demons of our nature: The need to (re)focus attention on blatant forms of dehumanization. *Current Directions in Psychological Science*, 26(6), 487–494. <https://doi.org/10.1177/0963721417708230>
- Kteily, N., Bruneau, E., Waytz, A. & Cotterill, S. (2015). The ascent of man: Theoretical and empirical evidence for blatant dehumanization. *Journal of Personality and Social Psychology*, 109(5), 901–931. <https://doi.org/10.1037/pspp0000048>
- Kunczik, M. & Zipfel, A. (2005). *Publizistik: Ein Studienhandbuch* (2., durchgesehene und aktualisierte Auflage). Böhlau.
- Küppers, C. (2012). Soziologische Dimensionen von Geschlecht. *Aus Politik und Zeitgeschichte*(20-21). <http://www.bpb.de/apuz/135431/soziologische-dimensionen-von-geschlecht?p=all>

- Kwak, H., Lee, C., Park, H. & Moon, S. (2010). What is Twitter, a social network or a news media? *WWW '10: Proceedings of the 19th international conference on world wide web*, 591–600. <https://doi.org/10.1145/1772690.1772751>
- Laczniak, R. N. & Muehling, D. D. (1993). The relationship between experimental manipulations and tests of theory in an advertising message involvement context. *Journal of Advertising*, 22(3), 59–74. <https://doi.org/10.1080/00913367.1993.10673411>
- Lambe, J. L. (2004). Who wants to censor pornography and hate speech? *Mass Communication & Society*, 7(3), 279–299. [https://doi.org/10.1207/s15327825mcs0703\\_2](https://doi.org/10.1207/s15327825mcs0703_2)
- Landesanstalt für Medien NRW. (2021). *Ergebnisbericht: forsa-Befragung zu: Hate Speech 2021*. [https://www.medienanstalt-nrw.de/fileadmin/user\\_upload/NeueWebsite\\_0120/Themen/Hass/forsa\\_LFMNRW\\_Hassrede2021\\_Ergebnisbericht.pdf](https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Themen/Hass/forsa_LFMNRW_Hassrede2021_Ergebnisbericht.pdf)
- Landrine, H. & Klonoff, E. A. (1996). The schedule of racist events: A measure of racial discrimination and a study of its negative physical and mental health consequences. *Journal of Black Psychology*, 22(2), 144–168. <https://doi.org/10.1177/00957984960222002>
- Lange, P. G. (2007). Searching for the 'you' in 'YouTube': An analysis of online response ability. *Ethnographic Praxis in Industry Conference Proceedings*(1), 36–50. <https://doi.org/10.1111/j.1559-8918.2007.tb00061.x>
- Lange, P. G. (2014). Commenting on YouTube rants: Perceptions of inappropriateness or civic engagement. *Journal of Pragmatics*, 73, 53–65. <https://doi.org/10.1016/j.pragma.2014.07.004>
- Lapidot-Lefler, N. & Barak, A. (2012). Effects of anonymity, invisibility, and lack of eye-contact on toxic online disinhibition. *Computers in Human Behavior*, 28, 434–443.
- Lavigne, C. (2013). *Cyberpunk women, feminism and science fiction: A critical study*. McFarland.
- Lawrence, C. R., III (1990). If he hollers let him go: Regulating racist speech on campus. *Duke Law Journal*, 39(3), 431–483.
- Lazarsfeld, P. F., Berelson, B. & Gaudet, H. (1944). *The people's choice. How the voter makes up his mind in a presidential campaign*. Columbia University Press.
- Lea, M., O'Shea, T., Fung, P. & Spears, R. (1992). 'Flaming' in Computer-Mediated Communication: Observations, Explanations, Implications. In M. Lea (Hrsg.), *Contexts of Computer-Mediated Communication* (S. 89–112). Harvester Wheatsheaf.
- Leader, T., Mullen, B. & Rice, D. (2009). Complexity and valence in ethnophobias and exclusion of ethnic out-groups: What puts the "hate" into hate speech? *Journal of Personality and Social Psychology*, 96(1), 170–182. <https://doi.org/10.1037/a0013066>
- Lee, E. & Leets, L. (2002). Persuasive storytelling by hate groups online: Examining its effects on adolescents. *American Behavioral Scientist*, 45(6), 927–957.
- Lee, E.-J. & Jang, Y. J. (2010). What do others' reactions to news on internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research*, 37(6), 825–846. <https://doi.org/10.1177/0093650210376189>
- Leets, L. (2001). Responses to internet hate sites: Is speech too free in cyberspace? *Journal of Communications Law and Policy*, 6, 287–317.
- Leets, L. (2002). Experiencing hate speech: Perceptions and responses to anti-semitism and antigay speech. *Journal of Social Issues*, 58(2), 341–361.

- Leets, L. & Bowers, P. J. (1999). Loud and angry voices: The insidious influence. *Communication Monographs*, 66(4), 325–340. <https://doi.org/10.1080/03637759909376483>
- Leets, L. & Giles, H. (1997). Words as weapons - When do they wound? Investigations of harmful speech. *Human Communication Research*, 24(2), 260–301.
- Legewie, J. (2013). Terrorist events and attitudes toward immigrants: A natural experiment. *American Journal of Sociology*, 118(5), 1199–1245.
- Legewie, J. (2016). Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. *American Journal of Sociology*, 122(2), 379–424. <https://doi.org/10.1086/687518>
- Leiner, D. J. (2016). Our research's breadth lives on convenience samples: A case study of the online respondent pool "SoSci Panel". *Studies in Communication and Media*, 5(4), 367–396. <https://doi.org/10.5771/2192-4007-2016-4-367>
- Lenz, K. (2006). *Soziologie der Zweierbeziehung: Eine Einführung* (3., überarbeitete Aufl.). VS Verlag für Sozialwissenschaften / GWV Fachverlage.
- Lerner, J. S. & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition & Emotion*, 14(4), 473–493. <https://doi.org/10.1080/026999300402763>
- Lerner, J. S. & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81(1), 146–159. <https://doi.org/10.1037/0022-3514.81.1.146>
- Levin, B. (2002). Cyberhate: A legal and historical analysis of extremists' use of computer networks in America. *American Behavioral Scientist*, 45(6), 958–988.
- Levin, J. & McDevitt, J. (1993). *Hate crimes: The rising tide of bigotry and bloodshed*. Springer. <https://doi.org/10.1007/978-1-4899-6108-2>
- Levin, J. & McDevitt, J. (2002). *Hate crimes revisited: America's war against those who are different*. Westview Press.
- Lev-On, A. (2018). The anti-social network? Framing social media in wartime. *Social Media + Society*, 4(3), 205630511880031. <https://doi.org/10.1177/2056305118800311>
- Leyens, J.-P [Jacques-Philippe], Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S [Stéphanie], Rodriguez-Perez, A. & Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, 4(2), 186–197. [https://doi.org/10.1207/S15327957PSPR0402\\_06](https://doi.org/10.1207/S15327957PSPR0402_06)
- Li, Q. (2007). New bottle but old wine: A research of cyberbullying in schools. *Computers in Human Behavior*, 23(4), 1777–1791. <https://doi.org/10.1016/j.chb.2005.10.005>
- Liberman, A. & Chaiken, S. (1992). Defensive processing of personally relevant health messages. *Personality and Social Psychology Bulletin*, 18(6), 669–679. <https://doi.org/10.1177/0146167292186002>
- Lickel, B., Miller, N., Stenstrom, D. M., Denson, T. F. & Schmader, T. (2006). Vicarious retribution: the role of collective blame in intergroup aggression. *Personality and Social Psychology Review*, 10(4), 372–390. [https://doi.org/10.1207/s15327957pspr1004\\_6](https://doi.org/10.1207/s15327957pspr1004_6)
- Light, B. & Cassidy, E. (2014). Strategies for the suspension and prevention of connection: Rendering disconnection as socioeconomic lubricant with Facebook. *New Media & Society*, 16(7), 1169–1184. <https://doi.org/10.1177/1461444814544002>
- Livingstone, S., Stoilova, M. & Kelly, A. (2016). Cyberbullying: Incidence, trends and consequences. In United Nations (Hrsg.), *Ending the torment: Tackling bullying from the schoolyard to cyberspace* (S. 115–120). United Nations Publications.

- Llorent, V. J., Ortega-Ruiz, R. & Zych, I. (2016). Bullying and cyberbullying in minorities: Are they more vulnerable than the majority group? *Frontiers in Psychology*, 7, 1507. <https://doi.org/10.3389/fpsyg.2016.01507>
- Loch, D. (2019). Rechtspopulistische Parteien in Europa. In H. U. Brinkmann & I.-C. Panreck (Hrsg.), *Rechtspopulismus in Einwanderungsgesellschaften: Die politische Auseinandersetzung um Migration und Integration* (S. 43–74). Springer. [https://doi.org/10.1007/978-3-658-23401-0\\_3](https://doi.org/10.1007/978-3-658-23401-0_3)
- Locke, V. & Johnston, L. (2001). Stereotyping and prejudice: A social cognitive approach. In M. Augoustinos & K. J. Reynolds (Hrsg.), *Understanding prejudice, racism, and social conflict* (S. 107–125). Sage.
- Loosen, W., Häring, M., Kurtanović, Z., Merten, L., Reimer, J., van Roessel, L. & Maalej, W. (2017). Making sense of user comments: Identifying journalists' requirements for a comment analysis framework. *Studies in Communication and Media*, 6(4), 333–364. <https://doi.org/10.5771/2192-4007-2017-4-333>
- Lopes, B. & Yu, H. (2017). Who do you troll and Why: An investigation into the relationship between the dark triad personalities and online trolling behaviours towards popular and less popular Facebook profiles. *Computers in Human Behavior*, 77, 69–76. <https://doi.org/10.1016/j.chb.2017.08.036>
- Mackie, D. M., Devos, T. & Smith, E. R. (2000). Intergroup emotions: Explaining offensive action tendencies in an intergroup context. *Journal of Personality and Social Psychology*, 79(4), 602–616. <https://doi.org/10.1037//0022-3514.79.4.602>
- Maihofer, A. & Schutzbach, F. (2015). Vom Antifeminismus zum ›Anti-Genderismus‹. Eine zeitdiagnostische Betrachtung am Beispiel Schweiz. In S. Hark & P.-I. Villa (Hrsg.), *(Anti-)Genderismus: Sexualität und Geschlecht als Schauplätze aktueller politischer Auseinandersetzungen* (S. 201–217). transcript.
- Maio, G. R. & Haddock, G. (2010). *The psychology of attitudes and attitude change*. Sage. <https://doi.org/10.4135/9781446214299>
- Maletzke, G. (1998). *Kommunikationswissenschaft im Überblick: Grundlagen, Probleme, Perspektiven*. Westdeutscher Verlag.
- Mancera, A. M. C. (2018). The perception of the expression of taboos: A sociolinguistic study. In A. Pizarro Pedraza (Hrsg.), *Linguistic taboo revisited: Novel insights from cognitive perspectives* (S. 269–290). De Gruyter Mouton. <https://doi.org/10.1515/9783110582758-014>
- Mansbridge, J. & Flaster, K. (2007). The cultural politics of everyday discourse: The case of “male chauvinist”. *Critical Sociology*, 33(4), 627–660. <https://doi.org/10.1163/156916307X210973>
- March, E., Grieve, R., Marrington, J. & Jonason, P. K. (2017). Trolling on Tinder® (and other dating apps): Examining the role of the dark tetrad and impulsivity. *Personality and Individual Differences*, 110, 139–143. <https://doi.org/10.1016/j.paid.2017.01.025>
- March, E. & Marrington, J. (2019). A qualitative analysis of internet trolling. *Cyberpsychology, Behavior and Social Networking*, 22(3), 192–197. <https://doi.org/10.1089/cyber.2018.0210>
- Marder, B., Slade, E., Houghton, D. & Archer-Brown, C. (2016). “I like them, but won't ‘like’ them”: An examination of impression management associated with visible political party affiliation on Facebook. *Computers in Human Behavior*, 61, 280–287. <https://doi.org/10.1016/j.chb.2016.03.047>

- Margolis, J. & Fisher, A. (2001). *Unlocking the Clubhouse: Women in Computing*. MIT Press.
- Marker, K. (2013). Know Your Enemy. Zur Funktionalität der Hassrede für wehrhafte Demokratien. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 59–94). Gießener Elektronische Bibliothek.
- Matsuda, M. J., Lawrence, C. R., Delgado, R. & Crenshaw, K. (Hrsg.). (1993). *Words that wound: Critical race theory, assaultive speech, and the First Amendment*. Routledge.
- Maurer, M. (2016). *Nonverbale politische Kommunikation*. Springer VS.
- Maurer, S. (2018). Hedwig Dohms "Die Antifeministen". *Aus Politik und Zeitgeschichte*, 68(17), 40–46.
- Maynard, J. L. & Benesch, S. (2016). Dangerous speech and dangerous ideology: An integrated model for monitoring and prevention. *Genocide Studies and Prevention: An International Journal*, 9(3), 70–95. <https://doi.org/10.5038/1911-9933.9.3.1317>
- McCluskey, M. & Hmielowski, J. D. (2011). Opinion expression during social conflict: Comparing online reader comments and letters to the editor. *Journalism*, 13(3), 303–319. <https://doi.org/10.1177/1464884911421696>
- McDevitt, J., Levin, J. & Bennett, S. (2002). Hate crime offenders: An expanded typology. *Journal of Social Issues*, 58(2), 303–317.
- McKee, H. (2002). "YOUR VIEWS SHOWED TRUE IGNORANCE!!!": (Mis)Communication in an online interracial discussion forum. *Computers and Composition*, 19, 411–434.
- Meddaugh, P. M. & Kay, J. (2009). Hate speech or "reasonable racism?" The other in Stormfront. *Journal of Mass Media Ethics*, 24(4), 251–268. <https://doi.org/10.1080/08900520903320936>
- Megarry, J. (2014). Online incivility or sexual harassment? Conceptualising women's experiences in the digital age. *Women's Studies International Forum*, 47, 46–55. <https://doi.org/10.1016/j.wsif.2014.07.012>
- Meibauer, J. (2013). Hassrede - von der Sprache zur Politik. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 1–16). Gießener Elektronische Bibliothek.
- Meier, A. & Allen, G. (2009). Romantic relationships from adolescence to young adulthood: Evidence from the national longitudinal study of adolescent health. *The Sociological Quarterly*, 50(2), 308–335. <https://doi.org/10.1111/j.1533-8525.2009.01142.x>
- Mesch, S. (15. März 2017). Von Nacktheit, Selbstzensur und Opferumkehr. [https://www.deutschlandfunk.de/hasskommentare-von-nacktheit-selbstzensur-und-opferumkehr.700.de.html?dram:article\\_id=381307](https://www.deutschlandfunk.de/hasskommentare-von-nacktheit-selbstzensur-und-opferumkehr.700.de.html?dram:article_id=381307)
- Meyer, I. H. (2003). Prejudice, social stress, and mental health in lesbian, gay, and bisexual populations: conceptual issues and research evidence. *Psychological Bulletin*, 129(5), 674–697. <https://doi.org/10.1037/0033-2909.129.5.674>
- Miller, G. R. (2002). On being persuaded: Some basic distinctions. In J. P. Dillard & M. Pfau (Hrsg.), *The persuasion handbook: Developments in theory and practice* (S. 3–16). SAGE Publications.
- Miller, J. B. (2012). *Toward a new psychology of women* (Second Edition). Beacon Press.
- Millett, K. (2016 [1969]). *Sexual politics*. Columbia University Press.
- Mischke, J. & Wingerter, C. (2012). *Frauen und Männer auf dem Arbeitsmarkt: Deutschland und Europa*. Wiesbaden.

- Mohan, S., Guha, A., Harris, M., Popowich, F., Schuster, A. & Priebe, C. (2017). The impact of toxic language on the health of Reddit communities. In M. Mouhoub & P. Langlais (Hrsg.), *Advances in artificial intelligence* (S. 51–56). Springer International Publishing.
- Möller, S. (1999). *Sexual Correctness: Die Modernisierung antifeministischer Debatten in den Medien*. VS Verlag für Sozialwissenschaften. <https://doi.org/10.1007/978-3-322-95179-3>
- Mongeau, P. A. & Stiff, J. B. (1993). Specifying causal relationships in the elaboration likelihood model. *Communication Theory*, 3(1), 65–72. <https://doi.org/10.1111/j.1468-2885.1993.tb00057.x>
- Moor, P. J., Heuvelman, A. & Verleur, R. (2010). Flaming on YouTube. *Computers in Human Behavior*, 26, 1536–1546. <https://doi.org/10.1016/j.chb.2010.05.023>
- Morf, C. C. & Koole, S. L. (2014). Das Selbst. In K. Jonas, W. Stroebe & M. Hewstone (Hrsg.), *Lehrbuch. Sozialpsychologie* (6. Aufl., S. 141–195). Springer.
- Morgan, A., Quesenberry, J. & Trauth, E. M. (2004). Exploring the importance of social networks in the IT workforce: Experiences with the "Boy's Club", 1313–1320 (AMCIS 2004 Proceedings).
- Mortensen, T. E. (2018). Anger, fear, and games: The long event of #GamerGate. *Games and Culture*, 13(8), 787–806. <https://doi.org/10.1177/1555412016640408>
- Moss-Racusin, C., Dovidio, J. F., Brescoll, V. L., Graham, M. J. & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America*, 109(41), 16474–16479. <https://doi.org/10.1073/pnas.1211286109>
- Mounk, Y. (2018). *Der Zerfall der Demokratie: Wie der Populismus den Rechtsstaat bedroht*. Droemer.
- Muddiman, A. & Stroud, N. J. (2017). News values, cognitive biases, and partisan incivility in comment sections. *Journal of Communication*, 67(4), 586–609. <https://doi.org/10.1111/jcom.12312>
- Müller, F., van Zoonen, L. & Hirzalla, F. (2014). Anti-Islam propaganda and its effects: Fitna, fear-based communication and the moderating role of public debate. *Middle East Journal of Culture and Communication*, 7, 82–100.
- Munn, L. (2020). Angry by design: Toxic communication and technical architectures. *Humanities and Social Sciences Communications*, 7(1). <https://doi.org/10.1057/s41599-020-00550-7>
- Murphy, E. (1999). 'Breast is best': Infant feeding decisions and maternal deviance. *Sociology of Health and Illness*, 21(2), 187–208. <https://doi.org/10.1111/1467-9566.00149>
- Neuberg, S. L. & Cottrell, C. A. (2002). Intergroup emotions: A biocultural approach. In D. M. Mackie & E. R. Smith (Hrsg.), *From prejudice to intergroup emotions: Differentiated reactions to social groups* (S. 265–283). Routledge.
- Neumann, K. & Arendt, F. (2016). "Der Pranger der Schande": Eine inhaltsanalytische Untersuchung der Wirkung des Bild-Prangers auf das Postingverhalten von Facebook Nutzern zur Flüchtlingsdebatte. *Publizistik*, 61, 247–265. <https://doi.org/10.1007/s11616-016-0283-7>

- Ng, T. W. H., Eby, L. T., Sorensen, K. L. & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58(2), 367–408. <https://doi.org/10.1111/j.1744-6570.2005.00515.x>.
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175–220.
- Nielsen, L. B. (2002). Subtle, pervasive, harmful: Racist and sexist remarks in public as hate speech. *Journal of Social Issues*, 58(2), 265–280.
- Niemann, J. (2016). *Risiken und Nutzen der Kommunikation auf Social Networking Sites: Theoretische Modellierung und empirische Befunde auf Basis der "Theory of Reasoned Action"*. Herbert von Halem.
- Nocentini, A., Calmaestra, J., Schultze-Krumbholz, A., Scheithauer, H., Ortega, R. & Mene-sini, E. (2010). Cyberbullying: Labels, behaviours and definition in three European countries. *Australian Journal of Guidance and Counselling*, 20(2), 129–142. <https://doi.org/10.1375/ajgc.20.2.129>
- Oberschall, A. (2000). The manipulation of ethnicity: From ethnic cooperation to violence and war in Yugoslavia. *Ethnic and Racial Studies*, 23(6), 982–1001. <https://doi.org/10.1080/014198700750018388>
- O'Keefe, D. J. (1999). Variability of persuasive message effects. *Document Design*, 1(2), 87–97. <https://doi.org/10.1075/dd.1.2.02oke>
- O'Keefe, D. J. (2003). Message properties, mediating states, and manipulation checks: Claims, evidence, and data analysis in experimental persuasive message effects research. *Communication Theory*, 13(3), 251–274. <https://doi.org/10.1111/j.1468-2885.2003.tb00292.x>
- O'Keefe, D. J. (2004). Trends and prospects in persuasion theory and research. In J. S. Seiter & R. H. Gass (Hrsg.), *Perspective on persuasion, social influence, and compliance gaining* (S. 31–43). Allyn & Bacon.
- O'Keefe, D. J. (2008). The elaboration likelihood model. In W. Donsbach (Hrsg.), *The international encyclopedia of communication* (1475–1480). Blackwell Publishing. <https://doi.org/10.4135/9781452218410.n9>
- O'Keefe, D. J. (2012). The elaboration likelihood model. In J. P. Dillard & S. Lijiang (Hrsg.), *The SAGE handbook of persuasion: Developments in theory and practice* (S. 137–149). SAGE Publications. <https://doi.org/10.4135/9781452218410.n9>
- Okimoto, T. G. & Brescoll, V. L. (2010). The price of power: Power seeking and backlash against female politicians. *Personality and Social Psychology Bulletin*, 36(7), 923–936. <https://doi.org/10.1177/0146167210371949>
- Olian, J. D., Schwab, D. P. & Haberfeld, Y. (1988). The impact of applicant gender compared to qualifications on hiring recommendations. *Organizational Behavior and Human Decision Processes*, 41(2), 180–195. [https://doi.org/10.1016/0749-5978\(88\)90025-8](https://doi.org/10.1016/0749-5978(88)90025-8)
- Olweus, D. (1993). *Bullying at School: What we know and what we can do*. John Wiley & Sons.
- Olweus, D. (2012). Cyberbullying: An overrated phenomenon? *European Journal of Developmental Psychology*, 9(5), 520–538. <https://doi.org/10.1080/17405629.2012.682358>
- Olweus, D. & Limber, S. P. (2018). Some problems with cyberbullying research. *Current Opinion in Psychology*, 19, 139–143. <https://doi.org/10.1016/j.copsyc.2017.04.012>

- Orengo Castellá, V., Zornoza Abad, A. M., Prieto Alonso, F. & Peiró Silla, J. M. (2000). The influence of familiarity among group members, group atmosphere and assertiveness on uninhibited behavior through three different communication media. *Computers in Human Behavior*, 16(2), 141–159. [https://doi.org/10.1016/S0747-5632\(00\)00012-1](https://doi.org/10.1016/S0747-5632(00)00012-1)
- Ortak, N. (2004). *Persuasion: Zur textlinguistischen Beschreibung eines dialogischen Strategiemusters*. Max Niemeyer. <https://doi.org/10.1515/9783110933246>
- O'Sullivan, P. B. & Flanagin, A. J. (2003). Reconceptualizing 'flaming' and other problematic messages. *New Media & Society*, 5(1), 69–94.
- OSZE. (2018). *#SOFJO: Why online attacks on female journalists are an affront to plurality, media freedom and democracy*. <https://www.osce.org/representative-on-freedom-of-media/405026>
- Papacharissi, Z. (2002). The virtual sphere: The internet as a public sphere. *New Media & Society*, 4(1), 9–27. <https://doi.org/10.1177/14614440222226244>
- Papacharissi, Z. (2009). The virtual sphere 2.0: The internet, the public sphere and beyond. In A. Chadwick & P. N. Howard (Hrsg.), *Routledge handbook of internet politics* (S. 230–245). Routledge.
- Parekh, B. (2006). Hate speech: Is there a case for banning? *Public Policy Research*, 12(4), 213–223. <https://doi.org/10.1111/j.1070-3535.2005.00405.x>
- Patchin, J. W. & Hinduja, S. (2010). Cyberbullying and self-esteem. *The Journal of School Health*, 80(12), 614–621. <https://doi.org/10.1111/j.1746-1561.2010.00548.x>
- Paulhus, D. L. & Dutton, D. G. (2016). Everyday sadism. In V. Zeigler-Hill & D. K. Marcus (Hrsg.), *The dark side of personality: Science and practice in social, personality, and clinical psychology* (S. 109–120). American Psychological Association. <https://doi.org/10.1037/14854-006>
- Paulhus, D. L. & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavelianism, and psychopathy. *Journal of Research in Personality*, 36(6), 556–563.
- Paustian-Underdahl, S. C., Walker, L. S. & Woehr, D. J. (2014). Gender and perceptions of leadership effectiveness: A meta-analysis of contextual moderators. *Journal of Applied Psychology*, 99(6), 1129–1145. <https://doi.org/10.1037/a0036751>
- Pells, K., Ogando Portela, M. J. & Espinoza, P. (2016). Poverty and inequity: Multi-country evidence on the structural drivers of bullying. In United Nations (Hrsg.), *Ending the torment: Tackling bullying from the schoolyard to cyberspace* (S. 41–47). United Nations Publications.
- Pempek, T. A., Yermolayeva, Y. A. & Calvert, S. L. (2009). College students' social networking experiences on Facebook. *Journal of Applied Developmental Psychology*, 30, 227–238. <https://doi.org/10.1016/j.appdev.2008.12.010>
- Perdue, B. C. & Summers, J. O. (1986). Checking the success of manipulations in marketing experiments. *Journal of Marketing Research*, 23(4), 317–326. <https://doi.org/10.1177/002224378602300401>
- Pérez-Escobar, M. & Noguera-Vivo, J. M. (Hrsg.). (2021). *Hate speech and polarization in participatory society*. Routledge, Taylor & Francis.
- Perloff, R. M. (2003). *The dynamics of persuasion. Communication and attitudes in the 21st century*. Lawrence Erlbaum Associates.

- Petty, R. E. & Briñol, P. (2012). The elaboration likelihood model. In P. A. van Lange, A. Kruglanski & E. T. Higgins (Hrsg.), *Handbook of theories of social psychology* (1. Aufl., S. 224–245). Sage.
- Petty, R. E. & Briñol, P. (2015). Emotion and persuasion: Cognitive and meta-cognitive processes impact attitudes. *Cognition and Emotion*, *29*(1), 1–26.  
<https://doi.org/10.1080/02699931.2014.967183>
- Petty, R. E., Briñol, P., Loersch, C. & McCaslin, M. (2009). The need for cognition. In M. R. Leary & R. H. Hoyle (Hrsg.), *Handbook of individual differences in social behavior* (S. 318–329). Guilford Press.
- Petty, R. E. & Cacioppo, J. T. (1981). Issue involvement as a moderator of the effects on attitude of advertising content and context. In K. B. Monroe & A. Abor (Hrsg.), *Advances in Consumer Research* (S. 20–24). Association for Consumer Research.
- Petty, R. E. & Cacioppo, J. T. (1984a). The effects of involvement on responses to argument quantity and quality: Central and peripheral routes to persuasion. *Journal of Personality and Social Psychology*, *46*(1), 69–81. <https://doi.org/10.1037/0022-3514.46.1.69>
- Petty, R. E. & Cacioppo, J. T. (1984b). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research*, *11*(1), 668–672.
- Petty, R. E. & Cacioppo, J. T. (1986a). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer.
- Petty, R. E. & Cacioppo, J. T. (1986b). The elaboration likelihood model of persuasion. *Advances in Experimental Social Psychology*, *19*, 123–205.
- Petty, R. E., Cacioppo, J. T. & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*(5), 847–855. <https://doi.org/10.1037/0022-3514.41.5.847>
- Petty, R. E., Cacioppo, J. T. & Schumann, D. (1983). Central and peripheral routes to advertising effectiveness: The moderating role of involvement. *Journal of Consumer Research*, *10*(2), 135–146.
- Petty, R. E., Harkins, S. G. & Williams, K. D. (1980). The effects of group diffusion of cognitive effort on attitudes: An information-processing view. *Journal of Personality and Social Psychology*, *38*(1), 81–92.
- Petty, R. E., Haugtvedt, C. P. & Smith, S. M. (1995). Elaboration as a determinant of attitude strength: Creating attitudes that are persistent, resistant, and predictive of behavior. In R. E. Petty & J. A. Krosnick (Hrsg.), *Attitude strength: Antecedents and consequences* (S. 93–130). Erlbaum.
- Petty, R. E. & Wegener, D. T. (1999). The elaboration likelihood model: Current status and controversies. In S. Chaiken & Y. Trope (Hrsg.), *Dual-process theories in social psychology* (S. 41–72). Guilford Press.
- Petty, R. E., Wegener, D. T. & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual Review of Psychology*, *48*, 609–647.
- Petty, R. E., Wegener, D. T., Fabrigar, L. R., Priester, J. R. & Cacioppo, J. T. (1993). Conceptual and methodological issues in the elaboration likelihood model of persuasion: A reply to the Michigan State critics. *Communication Theory*, *3*(4), 336–342.  
<https://doi.org/10.1111/j.1468-2885.1993.tb00078.x>

- Petty, R. E., Wells, G. L. & Brock, T. C. (1976). Distraction can enhance or reduce yielding to propaganda: Thought disruption versus effort justification. *Journal of Personality and Social Psychology*, 34(5), 874–884.
- Petzel, T. (2009). *Die Autoritäre Persönlichkeit: Eine Integration traditioneller moderner Sichtweisen*. Cuvillier.
- Piorkowski, C. D. (2018). *Klar denken in der Empörungsdemokratie*. <https://www.tagesspiegel.de/wissen/philosophen-und-hatespeech-klar-denken-in-der-empowerungsdemokratie/22733856.html>
- Planert, U. (2000). Mannweiber, Urniden und sterile Jungfern. Die Frauenbewegung und ihre Gegner im Kaiserreich. *Feministische Studien*, 18(1), 22–35.
- Pohjonen, M. (2019). A comparative approach to social media extreme speech: Online hate speech as media commentary. *International Journal of Communication*, 13, 3088–3103.
- Pohjonen, M. & Udupa, S. (2017). Extreme speech online: An anthropological critique of hate speech debates. *International Journal of Communication*, 11, 1173–1191.
- Pollock, E. (2009). Researching white supremacists online: Methodological concerns of researching hate 'speech'. *Internet Journal of Criminology*. [https://docs.wix-static.com/ugd/b93dd4\\_3686f65909044639a07b17e644b64f92.pdf](https://docs.wix-static.com/ugd/b93dd4_3686f65909044639a07b17e644b64f92.pdf)
- Porst, R. (2008). *Fragebogen: Ein Arbeitsbuch. Lehrbuch*. VS Verlag für Sozialwissenschaften.
- Postmes, T., Spears, R. & Lea, M. (1998). Breaching or building social boundaries? SIDE-Effects of computer-mediated communication. *Communication Research*, 25(6), 689–715. <https://doi.org/10.1177/009365098025006006>
- Postmes, T., Spears, R. & Lea, M. (2000). The formation of group norms in computer-mediated communication. *Human Communication Research*, 26(3), 341–371.
- Postmes, T., Spears, R., Sakhel, K. & Groot, D. de (2001). Social influence in computer-mediated communication: The effects of anonymity on group behavior. *Personality and Social Psychology Bulletin*, 27(10), 1243–1254. <https://doi.org/10.1177/01461672012710001>
- Poteat, V. P. & Birkett, M. (2016). Sexual prejudice: Advances in conceptual and empirical models. In C. G. Sibley & F. K. Barlow (Hrsg.), *Cambridge Handbooks in Psychology. The Cambridge handbook of the psychology of prejudice* (S. 371–391). Cambridge University Press.
- Pratto, F., Sidanius, J., Stallworth, L. M. & Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67(4), 741–763. <https://doi.org/10.1037/0022-3514.67.4.741>
- Preece, J., Nonnecke, B. & Andrews, D. (2004). The top five reasons for lurking: Improving community experiences for everyone. *Computers in Human Behavior*, 20(2), 201–223. <https://doi.org/10.1016/j.chb.2003.10.015>
- Priester, J. R. & Petty, R. E. (1995). Source attributions and persuasion: Perceived honesty as a determinant of message scrutiny. *Personality and Social Psychology Bulletin*, 21(6), 637–654. <https://doi.org/10.1177/0146167295216010>
- Privitera, C. & Campbell, M. A. (2009). Cyberbullying: The new face of workplace bullying? *Cyberpsychology & Behavior*, 12(4), 395–400. <https://doi.org/10.1089/cpb.2009.0025>

- Pürer, H. (2014). Grundbegriffe der Kommunikationswissenschaft. In H. Pürer (Hrsg.), *Publizistik- und Kommunikationswissenschaft* (2. Aufl., S. 63–107). UVK.
- Quade, D. (1967). Rank analysis of covariance. *Journal of the American Statistical Association*, 62(320), 1187–1200.
- Quandt, T. (2018). Dark participation. *Media and Communication*, 6(4), 36–48. <https://doi.org/10.17645/mac.v6i4.1519>
- Quandt, T. & Festl, R. (2017). Cyberhate. In P. Rössler (Hrsg.), *The international encyclopedia of media effects* (S. 336–344). Wiley-Blackwell. <https://doi.org/10.1002/9781118783764.wbieme0215>
- Quandt, T. & Scheufele, B. (2011). Die Herausforderung einer Modellierung von Mikro-Meso-Makro-Links in der Kommunikationswissenschaft. In T. Quandt & B. Scheufele (Hrsg.), *Ebenen der Kommunikation: Mikro-Meso-Makro-Links in der Kommunikationswissenschaft* (S. 9–22). VS Verlag.
- Quesenberry, J. & Trauth, E. M. (2012). The (dis)placement of women in the IT workforce: an investigation of individual career values and organisational interventions. *Information Systems Journal*, 22(6), 457–473. <https://doi.org/10.1111/j.1365-2575.2012.00416.x>
- Quesenberry, J., Trauth, E. M. & Morgan, A. (2006). Understanding the “mommy tracks”: A framework for analyzing work-family balance in the IT workforce. *Information Resources Management Journal*, 19(2), 37–53.
- Räsänen, P., Hawdon, J., Holkeri, E., Keipi, T., Näsi, M. & Oksanen, A. (2016). Targets of online hate: Examining determinants of victimization among young Finnish Facebook users. *Violence Victims*, 31(4), 708–726. <https://doi.org/10.1891/0886-6708.VV-D-14-00079>
- Ratliff, K. A., Redford, L., Conway, J. & Smith, C. T. (2017). Engendering support: Hostile sexism predicts voting for Donald Trump over Hillary Clinton in the 2016 U.S. presidential election. *Group Processes & Intergroup Relations*, 1, 136843021774120. <https://doi.org/10.1177/1368430217741203>
- Rauchut, F. (2018). „Keine Angst vorm bösen Gender“ – Interventionen in die Antifeminismusdebatte im deutschen Print- und Fernsehjournalismus. *Feministische Studien*, 36(1), 188–196. <https://doi.org/10.1515/fs-2018-0015>
- Reader, B. (2012). Free press vs. free speech? The rhetoric of “civility” in regard to anonymous online comments. *Journalism & Mass Communication Quarterly*, 89(3), 495–513. <https://doi.org/10.1177/1077699012447923>
- Regan, S. & Brown, A. (2019). Experiences of online breastfeeding support: Support and reassurance versus judgement and misinformation. *Maternal & Child Nutrition*, 15(4), o. S. <https://doi.org/10.1111/mcn.12874>
- Reicher, S. D., Spears, R. & Postmes, T. (1995). A social identity model of deindividuation phenomena. *European Review of Social Psychology*, 6(1), 161–198. <https://doi.org/10.1080/14792779443000049>
- Reilly, E. D., Rackley, K. R. & Awad, G. H. (2017). Perceptions of male and female STEM aptitude: The moderating effect of benevolent and hostile sexism. *Journal of Career Development*, 44(2), 159–173. <https://doi.org/10.1177/0894845316641514>
- Reips, U.-D. (2000). The web experiment method: Advantages, disadvantages, and solutions. In M. H. Birnbaum (Hrsg.), *Psychological experiments on the internet* (S. 89–117). Academic Press.

- Retelsdorf, J., Schwartz, K. & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, 107(1), 186–194. <https://doi.org/10.1037/a0037107>
- Reynolds, K. J. & Turner, J. C. (2001). Prejudice as a group process: The role of social identity. In M. Augoustinos & K. J. Reynolds (Hrsg.), *Understanding prejudice, racism, and social conflict* (S. 159–178). Sage.
- Rhodes, N. (2012). Outcomes of persuasion: Behavioral, cognitive, and social. In J. P. Dillard & S. Lijiang (Hrsg.), *The SAGE handbook of persuasion: Developments in theory and practice* (S. 53–69). SAGE Publications.
- Richard, F. D., Bond, C. F. & Stokes-Zoota, J. J. (2003). One hundred years of social psychology quantitatively described. *Review of General Psychology*, 7(4), 331–363. <https://doi.org/10.1037/1089-2680.7.4.331>
- Rim, H. & Song, D. (2016). "How negative becomes less negative": Understanding the effects of comment valence and response sidedness in social media. *Journal of Communication*, 66(3), 475–495. <https://doi.org/10.1111/jcom.12205>
- Roback, A. A. (1944). *A dictionary of international slurs (Ethnophaulisms)*. Sci-Art.
- Rohrlack, C. (2009). Multidimensionale Skalierung. In S. Albers, D. Klapper, U. Konradt, A. Walter & J. Wolf (Hrsg.), *Methodik der empirischen Forschung* (3. Aufl., S. 153–173). Gabler.
- Rohrman, B. (1978). Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung. *Zeitschrift für Sozialpsychologie*, 9, 222–245.
- Rosenfeld, M. (2003). Hate speech in constitutional jurisprudence: A comparative analysis. *Cardozo Law Review*, 24(4), 1523–1567.
- Rosenfeld, M. (2012). Hate speech in constitutional jurisprudence: A comparative analysis. In M. E. Herz & P. Molnár (Hrsg.), *The content and context of hate speech: Rethinking regulation and responses* (S. 242–289). Cambridge University Press.
- Rösner, L., Winter, S. & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior*, 58, 461–470. <https://doi.org/10.1016/j.chb.2016.01.022>
- Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. *Communication Research*, 1–27. <https://doi.org/10.1177/0093650220921314>
- Rudman, L. A. (1998). Self-promotion as a risk factor for women: The costs and benefits of counterstereotypical impression management. *Journal of Personality and Social Psychology*, 74(3), 629–645. <https://doi.org/10.1037/0022-3514.74.3.629>
- Rudman, L. A. & Fairchild, K. (2004). Reactions to counterstereotypic behavior: The role of backlash in cultural stereotype maintenance. *Journal of Personality and Social Psychology*, 87(2), 157–176. <https://doi.org/10.1037/0022-3514.87.2.157>
- Rudman, L. A. & Glick, P. (1999). Feminized management and backlash toward agentic women: The hidden costs to women of a kinder, gentler image of middle managers. *Journal of Personality and Social Psychology*, 77(5), 1004–1010. <https://doi.org/10.1037/0022-3514.77.5.1004>
- Rudman, L. A. & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, 57(4), 743–762.
- Rudman, L. A., Moss-Racusin, C., Phelan, J. E. & Nauts, S. (2012). Status incongruity and backlash effects: Defending the gender hierarchy motivates prejudice against

- female leaders. *Journal of Experimental Social Psychology*, 48(1), 165–179.  
<https://doi.org/10.1016/j.jesp.2011.10.008>
- Sabella, R. A., Patchin, J. W. & Hinduja, S. (2013). Cyberbullying myths and realities. *Computers in Human Behavior*, 29(6), 2703–2711.  
<https://doi.org/10.1016/j.chb.2013.06.040>
- Sager, S. F. (2018). Nonverbale Kommunikation. In E. Neuland & P. Schlobinski (Hrsg.), *Handbuch Sprache in sozialen Gruppen* (S. 168–192). De Gruyter Mouton.
- Salzborn, C. (2017). *Phänomen Shitstorm: Herausforderung für die Onlinekrisenkommunikation von Unternehmen*. Tectum Verlag.
- SanJosé-Cabezudo, R., Gutiérrez-Arranz, A. M. & Gutiérrez-Cillán, J. (2009). The combined influence of central and peripheral routes in the online persuasion process. *Cyberpsychology & Behavior*, 12(3), 299–308. <https://doi.org/10.1089/cpb.2008.0188>
- Santana, A. D. (2014). Virtuous or vitriolic: The effect of anonymity on civility in online newspaper reader comment boards. *Journalism Practice*, 8(1), 18–33.  
<https://doi.org/10.1080/17512786.2013.813194>
- Santos, S., Amaral, I. & Basílio Simões, R. (Hrsg.) (2020). *Hate speech in social media: Perceptions and attitudes of higher education students in Portugal*. Proceedings of INTED2020 Conference. 2nd-4th March 2020, Valencia, Spain.
- Sarup, G., Suchner, R. W. & Gaylord, G. (1991). Contrast effects and attitude change: A test of the two-stage hypothesis of social judgment theory. *Social Psychology Quarterly*, 54(4), 364. <https://doi.org/10.2307/2786848>
- Satow, L. (2012). *SEA - Skala zur Erfassung von Testverfälschung durch positive Selbstdarstellung und sozial erwünschte Antworttendenzen*. <https://doi.org/10.23668/psycharchives.417>
- Savage, M. W. & Tokunaga, R. S. (2017). Moving toward a theory: Testing an integrated model of cyberbullying perpetration, aggression, social skills, and Internet self-efficacy. *Computers in Human Behavior*, 71, 353–361.  
<https://doi.org/10.1016/j.chb.2017.02.016>
- Sax, L. J., Lehman, K. J., Jacobs, J. A., Kanny, M. A., Lim, G., Monje-Paulson, L. & Zimmerman, H. B. (2017). Anatomy of an enduring gender gap: The evolution of women's participation in computer science. *The Journal of Higher Education*, 88(2), 258–293.  
<https://doi.org/10.1080/00221546.2016.1257306>
- Schemer, C., Matthes, J. & Wirth, W. (2007). Werbewirkung ohne Erinnerungseffekte? *Zeitschrift für Medienpsychologie*, 19(1), 2–13. <https://doi.org/10.1026/1617-6383.19.1.2>
- Schenk, M. (2007). *Medienwirkungsforschung* (3., vollst. überarb. Aufl.). Mohr Siebeck.
- Scherr, A. (2010). Diskriminierung und soziale Ungleichheiten: Erfordernisse und Perspektiven einer ungleichheitsanalytischen Fundierung von Diskriminierungsforschung und Antidiskriminierungsstrategien. In U. Hormel & A. Scherr (Hrsg.), *Diskriminierung: Grundlagen und Forschungsergebnisse* (S. 35–60). VS Verlag für Sozialwissenschaften.
- Scherr, A. (2011). Was meint Diskriminierung? *Sozial Extra*, 35(11-12), 34–38.  
<https://doi.org/10.1007/s12054-011-0411-6>
- Scherr, A. (2016). *Diskriminierung: Wie Unterschiede und Benachteiligungen gesellschaftlich hergestellt werden* (2. Aufl.). Springer Fachmedien. <https://doi.org/10.1007/978-3-658-10067-4>

- Scherr, A., El-Mafaalani, A. & Yüksel, G. (2017). Einleitung: Interdisziplinäre Diskriminierungsforschung. In A. Scherr, A. El-Mafaalani & G. Yüksel (Hrsg.), *Handbuch Diskriminierung* (S. v–x). Springer.
- Scheufele, B. (2011). Medienwirkungen jenseits des einzelnen Rezipienten. Die Herausforderungen für die Medienwirkungsforschung durch die Mehr-Ebenen-Problematik am Beispiel von fünf exemplarischen Wirkungsbereichen. In T. Quandt & B. Scheufele (Hrsg.), *Ebenen der Kommunikation: Mikro-Meso-Makro-Links in der Kommunikationswissenschaft* (S. 113–142). VS Verlag.
- Scheufele, B. & Schieb, C. (2014). "Wenn wir Trotz-war nicht hätten": Eine Untersuchung zum Verkauf von Strassenzeitungen zwischen Job und Empowerment. LIT.
- Schieb, C. & Preuss, M. (2016). *Governing hate speech by means of counter speech on Facebook*. Paper presented at the 66th ICA Annual Conference, Fukuoka, Japan.
- Schieb, C. & Preuss, M. (2018). Considering the elaboration likelihood model for simulating hate and counter speech on Facebook. *Studies in Communication and Media*, 7(4), 580–606. <https://doi.org/10.5771/2192-4007-2018-4-580>
- Schmader, T. (2002). Gender identification moderates stereotype threat effects on women's math performance. *Journal of Experimental Social Psychology*, 38(2), 194–201. <https://doi.org/10.1006/jesp.2001.1500>
- Schmerl, C. & Steinbach, D. (1973). Ritterlichkeit - Eine indirekte Form der Misogynie? In H. D. Schmidt, C. Schmerl, A. Krameyer, A. Wagner, D. Steinbach & A. Schmidt-Mummendey (Hrsg.), *Frauenfeindlichkeit: Sozialpsychologische Aspekte der Misogynie* (S. 56–79). Juventa.
- Schmider, E., Ziegler, M., Danay, E., Beyer, L. & Bühner, M. (2010). Is it really robust? *Methodology*, 6(4), 147–151. <https://doi.org/10.1027/1614-2241/a000016>
- Schmidt, J. (2006). *Weblogs: Eine kommunikationssoziologische Studie*. UVK.
- Schmincke, I. (2018). Frauenfeindlich, sexistisch, antifeministisch? Begriffe und Phänomene bis zum aktuellen Antigenderismus. *Aus Politik und Zeitgeschichte*, 68(17), 28–33.
- Schmitt, R. (8. März 2017). Saufen und kiffen auf Kosten der Steuerzahler. *Krone*. <https://www.krone.at/557951>
- Schneider, M. C. & Bos, A. L. (2014). Measuring stereotypes of female politicians. *Political Psychology*, 35(2), 245–266. <https://doi.org/10.1111/pops.12040>
- Scholz, S. (2007). "Männer reden Merkel klein." Männlichkeitskritiken im Bundestagswahlkampf 2005. In S. Scholz (Hrsg.), *"Kann die das?": Angela Merkels Kampf um die Macht : Geschlechterbilder und Geschlechterpolitiken im Bundestagswahlkampf 2005* (S. 103–116). Dietz.
- Schönbach, K. (2009). *Verkaufen, Flirten, Führen: Persuasive Kommunikation - ein Überblick* (1. Aufl.). VS Verlag für Sozialwissenschaften.
- Schwarz-Friesel, M. (2013). "Dies ist kein Hassbrief - sondern meine eigene Meinung über Euch!": Zur kognitiven und emotionalen Basis der aktuellen antisemitischen Hassrede. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 143–164). Gießener Elektronische Bibliothek.
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139173438>

- Sears, D. O. (1988). Symbolic racism. In Katz P. A. & Taylor D. A. (Hrsg.), *Eliminating racism. Perspectives in social psychology* (S. 53–84). Springer. [https://doi.org/10.1007/978-1-4899-0818-6\\_4](https://doi.org/10.1007/978-1-4899-0818-6_4)
- Seibert, T. (2016). US-Muslime fürchten den 11. September. *Der Tagesspiegel*. <https://www.tagesspiegel.de/politik/15-jahrestag-us-muslime-fuerchten-den-11-september/14508388.html>
- Sest, N. & March, E. (2017). Constructing the cyber-troll: Psychopathy, sadism, and empathy. *Personality and Individual Differences*, 119, 69–72. <https://doi.org/10.1016/j.paid.2017.06.038>
- Shepherd, T., Harvey, A., Jordan, T., Srauy, S. & Miltner, K. (2015). Histories of hating. *Social Media + Society*, 1–10. <https://doi.org/10.1177/2056305115603997>
- Sherif, M. & Sherif, C. W. (1969). Ingroup and intergroup relations: Experimental analysis. In M. Sherif & C. W. Sherif (Hrsg.), *Social psychology* (S. 221–266). Harper & Row.
- Shiue, Y.-C., Chiu, C.-M. & Chang, C.-C. (2010). Exploring and mitigating social loafing in online communities. *Computers in Human Behavior*, 26(4), 768–777. <https://doi.org/10.1016/j.chb.2010.01.014>
- Sidanius, J. & Pratto, F. (1999). *Social dominance*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139175043>
- Sidanius, J., Pratto, F. & Bobo, L. (1994). Social dominance orientation and the political psychology of gender: A case of invariance? *Journal of Personality and Social Psychology*, 67(6), 998–1011. <https://doi.org/10.1037/0022-3514.67.6.998>
- Siegel, J., Dubrovsky, V., Kiesler, S. & McGuire, T. W. (1986). Group processes in computer-mediated communication. *Organizational Behavior and Human Decision Processes*, 37(2), 157–187.
- Sielke, S. (2009). *Reading rape: The rhetoric of sexual violence in American literature and culture, 1790-1990*. Princeton University Press.
- Sirsch, J. (2013). Die Regulierung von Hassrede in liberalen Demokratien. In J. Meibauer (Hrsg.), *Linguistische Untersuchungen: Bd. 6. Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion* (S. 165–193). Gießener Elektronische Bibliothek.
- Slonje, R. & Smith, P. K. (2008). Cyberbullying: Another main type of bullying? *Scandinavian Journal of Psychology*, 49(2), 147–154. <https://doi.org/10.1111/j.1467-9450.2007.00611.x>
- Smith, E. R. & Mackie, D. M. (2010). Affective processes. In J. F. Dovidio, M. Hewstone, P. Glick & V. M. Esses (Hrsg.), *The SAGE handbook of prejudice, stereotyping and discrimination* (S. 131–145). Sage.
- Smith, L. R., Smith, K. D. & Blazka, M. (2017). Follow me, what's the harm? Considerations of catfishing and utilizing fake online personas on social media. *Journal of Legal Aspects of Sport*, 27(1), 32–45. <https://doi.org/10.1123/jlas.2016-0020>
- Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S. & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 49(4), 376–385. <https://doi.org/10.1111/j.1469-7610.2007.01846.x>
- Smith, P. K., Mahdavi, J., Carvalho, M. & Tippett, N. (2006). *An investigation into cyberbullying, its forms, awareness and impact, and the relationship between age and gender*

- in cyberbullying* (RBX03-06). London. <https://static.lgfl.net/LgflNet/downloads/online-safety/LGfL-OS-Research-Archive-2006-Goldsmiths-Cyberbullying.pdf>
- Somerville, K. (2011). Violence, hate speech and inflammatory broadcasting in Kenya: The problems of definition and identification. *Ecquid Novi: African Journalism Studies*, 32(1), 82–101. <https://doi.org/10.1080/02560054.2011.545568>
- Soral, W., Bilewicz, M. & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, 44(2), 136–146. <https://doi.org/10.1002/ab.21737>
- Sorial, S. (2015). Hate Speech and distorted communication: Rethinking the limits of incitement. *Law & Philosophy*, 34, 299–324. <https://doi.org/10.1007/s10982-014-9214-9>
- Sorrentino, A., Baldry, A. C [Anna Costanza], Farrington, D. P. & Blaya, C. (2019). Epidemiology of cyberbullying across Europe: Differences between countries and genders. *Educational Sciences: Theory & Practice*, 19(2), 74–91. <https://doi.org/10.12738/estp.2019.2.005>
- Spears, R. & Lea, M. (1994). Panacea or panopticon? The hidden power in computer-mediated communication. *Communication Research*, 21(4), 427–459. <https://doi.org/10.1177/009365094021004001>
- 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. (1972). The attitudes toward women scale: An objective instrument to measure attitudes toward the rights and roles of women in contemporary society. *JSAS Catalog of Selected Documents in Psychology*, 2, 66–67.
- Spence, J. T., Helmreich, R. & Stapp, J. (1973). A short version of the attitudes toward women scale (AWS). *Bulletin of the Psychonomic Society*, 2(4), 219–220.
- Spencer, S. J., Steele, C. M. & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology*, 35, 4–28.
- Sponholz, L. (2018). *Hate Speech in den Massenmedien: Theoretische Grundlagen und empirische Umsetzung*. Springer VS.
- Sponholz, L. (2020). Der Begriff »Hate Speech« in der deutschsprachigen Forschung. Eine empirische Begriffsanalyse. *SWS-Rundschau*, 60(1), 43–65.
- Springer, N. (2014). *Beschmutzte Öffentlichkeit? Warum Menschen die Kommentarfunktion auf Online-Nachrichtenseiten als öffentliche Toilettenwand benutzen, warum Besucher ihre Hinterlassenschaften trotzdem lesen, und wie die Wände im Anschluss aussehen*. LIT.
- Springer, N., Pürer, H. & Eichhorn, W. (2014). Computervermittelte Kommunikation. In H. Pürer (Hrsg.), *Publizistik- und Kommunikationswissenschaft* (2. Aufl., S. 88–107). UVK.
- Springer, N. & Troger, F. (2021). „Du stehst unter genauer Beobachtung, unangenehmer Beobachtung“. *Publizistik*, 66(1), 43–65. <https://doi.org/10.1007/s11616-020-00637-w>
- Sproull, L. & Kiesler, S. (1986). Reducing social context cues: Electronic mail in organizational communication. *Management Science*, 32(11), 1492–1512.
- Stangor, C. & Ford, T. E. (1992). Accuracy and Expectancy-confirming Processing Orientations and the Development of Stereotypes and Prejudice. *European Review of Social Psychology*, 3(1), 57–89. <https://doi.org/10.1080/14792779243000023>

- Stangor, C., Lynch, L., Duan, C. & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology*, 62(2), 207–218.
- Statista. (2020). *Soziale Netzwerke*. <https://de.statista.com/statistik/studie/id/11852/dokument/soziale-netzwerke-statista-dossier/>
- Statista. (2021a). *Statista-Dossier zum Social Network Facebook und dem Konzern: Facebook Inc.* <https://de.statista.com/statistik/studie/id/3180/dokument/facebook-statista-dossier/>
- Statista. (2021b). *Studienabschlüsse: Anzahl der bestandenen Prüfungen an Hochschulen in Deutschland in den Prüfungsjahren von 2010 bis 2020 im Fach Informatik nach Geschlecht*. <https://de.statista.com/statistik/daten/studie/739636/umfrage/bestandene-pruefungen-an-hochschulen-in-deutschland-im-fach-informatik-nach-geschlecht/>
- Statistisches Bundesamt. (2014). *Auf dem Weg zur Gleichstellung? Bildung, Arbeit und Soziales – Unterschiede zwischen Frauen und Männern*. [https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/Arbeitsmarkt/WegZurGleichstellung\\_82014.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Publikationen/WirtschaftStatistik/Arbeitsmarkt/WegZurGleichstellung_82014.pdf?__blob=publicationFile)
- Statistisches Bundesamt. (2021). *Bildung und Kultur. Studierende an Hochschulen*. [https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/studierende-hochschulen-endg-2110410217004.pdf?\\_\\_blob=publicationFile](https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/studierende-hochschulen-endg-2110410217004.pdf?__blob=publicationFile)
- Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist*, 52(6), 613–629. <https://doi.org/10.1037/0003-066X.52.6.613>
- Steffens, M. C. & Ebert, I. D. (2016). *Frauen - Männer - Karrieren: Eine sozialpsychologische Perspektive auf Frauen in männlich geprägten Arbeitskontexten*. Springer Fachmedien. <https://doi.org/10.1007/978-3-658-10750-5>
- Stephan, C. W., Stephan, W. G., Demitrakis, K. M., Yamada, A. M. & Clason, D. L. (2000). Women's Attitudes Toward Men: An Integrated Threat Theory Approach. *Psychology of Women Quarterly*, 24(1), 63–73. <https://doi.org/10.1111/j.1471-6402.2000.tb01022.x>
- Stephan, W. G. (2014). Intergroup anxiety: Theory, research, and practice. *Personality and Social Psychology Review*, 18(3), 239–255. <https://doi.org/10.1177/1088868314530518>
- Stephan, W. G. & Stephan, C. W. (1985). Intergroup anxiety. *Journal of Social Issues*, 41(3), 157–175. <https://doi.org/10.1111/j.1540-4560.1985.tb01134.x>
- Stewart, G. (2003). *Die Motivation von Frauen für ein Studium der Ingenieur- und Naturwissenschaften*. München.
- Stieglitz, S. & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—Sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, 29(4), 217–248. <https://doi.org/10.2753/MIS0742-1222290408>
- Stiff, J. B. & Boster, F. J. (1987). Cognitive processing: Additional thoughts and a reply to Petty, Kasmer, Haugtvedt, and Cacioppo. *Communication Monographs*, 54(3), 250–256. <https://doi.org/10.1080/03637758709390230>

- Stocké, V. (2004). Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit. *Zeitschrift für Soziologie*, 33(4). <https://doi.org/10.1515/zfsoz-2004-0403>
- Stocké, V. (2005). *Einstellung zu Bildung: Zusammenstellung sozialwissenschaftlicher Items und Skalen (ZIS)*. <https://zis.gesis.org/skala/Stocké-Einstellung-zu-Bildung>  
<https://doi.org/10.6102/ZIS136>
- Straus, S. (2007). What Is the relationship between hate radio and violence? Rethinking Rwanda's "Radio Machete". *Politics & Society*, 35(4), 609–637.  
<https://doi.org/10.1177/0032329207308181>
- Stroud, N. J. (2008). Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, 30(3), 341–366. <https://doi.org/10.1007/s11109-007-9050-9>
- Su, L. Y.-F., Xenos, M. A., Rose, K. M., Wirz, C., Scheufele, D. A. & Brossard, D. (2018). Uncivil and personal? Comparing patterns of incivility in comments on the Facebook pages of news outlets. *New Media & Society*, 20(10), 3678–3699.  
<https://doi.org/10.1177/1461444818757205>
- Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior*, 7(3), 321–326.
- Sun, N., Rau, P. P.-L. & Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior*, 38, 110–117.  
<https://doi.org/10.1016/j.chb.2014.05.022>
- 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. & Cohen, L. L. (1997). Overt, covert, and subtle sexism: A comparison between the attitudes toward women and modern sexism scales. *Psychology of Women Quarterly*, 21, 103–118.
- Tajfel, H. (1981). *Human groups and social categories: Studies in social psychology*. Cambridge University Press.
- Tajfel, H. & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Hrsg.), *The social psychology of intergroup relations* (S. 33–47). Brooks Cole.
- Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Hrsg.), *Cognitive processes in stereotyping and intergroup behavior* (S. 83–114). Lawrence Erlbaum.
- Tellis, G. J., MacInnis, D. J., Tirunillai, S. & Zhang, Y. (2019). What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence. *Journal of Marketing*, 83(4), 1–20.  
<https://doi.org/10.1177/0022242919841034>
- Thompson, P. A. (1993). *A social influence model of flaming in computer-mediated communication*. Paper presented at the Annual Meeting of the Western States Communication Association, Albuquerque, NM.
- Thomsen, L., Green, E. G. & Sidanius, J. (2008). We will hunt them down: How social dominance orientation and right-wing authoritarianism fuel ethnic persecution of immigrants in fundamentally different ways. *Journal of Experimental Social Psychology*, 44(6), 1455–1464. <https://doi.org/10.1016/j.jesp.2008.06.011>

- Tiedens, L. Z. & Linton, S. (2001). Judgment under emotional certainty and uncertainty: The effects of specific emotions on information processing. *Journal of Personality and Social Psychology*, 81(6), 973–988. <https://doi.org/10.1037/0022-3514.81.6.973>
- Tobien, J. (2017). *Wo sind die Frauen? Die IT-Branche hat ein Problem*. <https://www.heise.de/newsticker/meldung/Wo-sind-die-Frauen-Die-IT-Branche-hat-ein-Gender-Problem-3639941.html>
- Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior*, 26(3), 277–287. <https://doi.org/10.1016/j.chb.2009.11.014>
- Topczewska, U. (2017). Was sind aggressive Sprechakte? Zu Theorie und Methodologie von pragmlinguistischen Untersuchungen zur verbalen Aggression. In S. Bonacchi (Hrsg.), *Verbale Aggression: Multidisziplinäre Zugänge zur verletzenden Macht der Sprache* (S. 35–50). De Gruyter Mouton. <https://doi.org/10.1515/9783110522976-002>
- 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.
- Trauth, E. M. (2002). Odd girl out: An individual differences perspective on women in the IT profession. *Information Technology & People*, 15(2), 98–118. <https://doi.org/10.1108/09593840210430552>
- Trauth, E. M. (2017). A research agenda for social inclusion in information systems. *The DATA BASE for Advances in Information Systems*, 48(2), 9–20. <https://doi.org/10.1145/3084179.3084182>
- Trauth, E. M., Quesenberry, J. L. & Huang, H. (2008). A multicultural analysis of factors influencing career choice for women in the information technology workforce. *Journal of Global Information Management*, 16(4), 1–23. <https://doi.org/10.4018/jgim.2008100101>
- Trauth, E. M., Quesenberry, J. L. & Huang, H. (2009). Retaining women in the U.S. IT workforce: Theorizing the influence of organizational factors. *European Journal of Information Systems*, 18(5), 476–497. <https://doi.org/10.1057/ejis.2009.31>
- Tsesis, A. (2001). Hate in cyberspace: Regulating hate speech on the internet. *San Diego Law Review*, 38, 817–874.
- Turnage, A. K. (2008). Email flaming behaviors and organizational conflict. *Journal of Computer-Mediated Communication*, 13, 43–59. <https://doi.org/10.1111/j.1083-6101.2007.00385.x>
- Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D. & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Basil Blackwell.
- Tutz, G. (2010). Regression für Zählvariablen. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (1. Aufl., S. 887–904). VS Verlag für Sozialwissenschaften.
- Twenge, J. M. (1997). Attitudes toward women, 1970–1995. A meta-analysis. *Psychology of Women Quarterly*, 21, 35–51.
- Tynes, B., Rose, C. A. & Williams, D. R. (2010). The development and validation of the online victimization scale for adolescents. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 4(2).
- United Nations (Hrsg.). (2016). *Ending the torment: Tackling bullying from the schoolyard to cyberspace*. United Nations Publications.

- Valcour, P. M. & Tolbert, P. (2003). Gender, family and career in the era of boundarylessness: Determinants and effects of intra- and inter-organizational mobility. *The International Journal of Human Resource Management*, 14(5), 768–787.  
<https://doi.org/10.1080/0958519032000080794>
- Valentino, N. A., Wayne, C. & Oceno, M. (2018). Mobilizing sexism: The interaction of emotion and gender attitudes in the 2016 US presidential election. *Public Opinion Quarterly*, 82(S1), 213–235. <https://doi.org/10.1093/poq/nfy003>
- van der Heijden, H. (2003). Factors influencing the usage of websites: The case of a generic portal in The Netherlands. *Information & Management*, 40(6), 541–549.
- van Dijk, T. A. (1995). Elite discourse and the reproduction of racism. In R. K. Whillock & D. Slayden (Hrsg.), *Hate speech* (S. 1–27). SAGE Publications.
- van Mierlo, T. (2014). The 1% rule in four digital health social networks: An observational study. *Journal of Medical Internet Research*, 16(2).  
<https://doi.org/10.2196/jmir.2966>
- van Spanje, J. H. P. & de Vreese, C. H. (2014). The way democracy works: The impact of hate speech prosecution of a politician on citizens' satisfaction with democratic performance. *International Journal of Public Opinion Research*, 26(4), 501–516.  
<https://doi.org/10.1093/ijpor/edt039>
- Vazire, S. & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, 87(1), 123–132.  
<https://doi.org/10.1037/0022-3514.87.1.123>
- Viki, G. & Abrams, D. (2003). Infra-humanization: Ambivalent sexism and the attribution of primary and secondary emotions to women. *Journal of Experimental Social Psychology*, 39(5), 492–499. [https://doi.org/10.1016/S0022-1031\(03\)00031-3](https://doi.org/10.1016/S0022-1031(03)00031-3)
- Vollhardt, J., Coutin, M., Staub, E., Weiss, G. & Deflander, J. (2006/2007). Deconstructing hate speech in the DRC: A psychological media sensitization campaign. *Journal of Hate Studies*, 5(15), 15–35.
- Vorländer, H., Herold, M. & Schäller, S. (2016). *PEGIDA: Entwicklungen, Zusammensetzung und Deutung einer Empörungsbewegung*. Springer VS.  
<https://doi.org/10.1007/978-3-658-10982-0>
- Vrooman, S. S. (2002). The art of invective: Performing identity in cyberspace. *New Media & Society*, 4(1), 51–70.
- Wachs, S., Wright, M. F. & Vazsonyi, A. T. (2019). Understanding the overlap between cyberbullying and cyberhate perpetration: Moderating effects of toxic online disinhibition. *Criminal Behaviour and Mental Health*, 29(3), 179–188.  
<https://doi.org/10.1002/cbm.2116>
- Waldron, J. (2010). Dignity and defamation: The visibility of hate. *Harvard Law Review*, 123(7), 1596–1657.
- Walther, J. B. (2011). Theories of computer-mediated communication and interpersonal relations. In M. L. Knapp & J. A. Daly (Hrsg.), *The SAGE handbook of interpersonal communication* (4. Aufl., S. 443–479). SAGE Publications.
- Wang, L., Fan, L. & Bae, S. (2019). How to persuade an online gamer to give up cheating? Uniting elaboration likelihood model and signaling theory. *Computers in Human Behavior*, 96, 149–162. <https://doi.org/10.1016/j.chb.2019.02.024>
- Watterson, K. M. (1991). The power of words: The power of advocacy challenging the power of hate speech. *University of Pittsburgh Law Review*, 52(4), 955–987.

- Webber, G. & Williams, C. (2008). Mothers in “good” and “bad” part-time jobs. *Gender & Society*, 22(6), 752–777. <https://doi.org/10.1177/0891243208325698>
- Weber, M [Max]. (1980 [1922]). *Wirtschaft und Gesellschaft: Grundriß der verstehenden Soziologie*. J. C. B. Mohr.
- Wegener, D. T. & Carlston, D. E. (2005). Cognitive processes in attitude formation and change. In D. Albarracín, B. T. Johnson & M. P. Zanna (Hrsg.), *The handbook of attitudes* (S. 493–542). Psychology Press.
- Weisel, O. & Böhm, R. (2015). “Ingroup love” and “outgroup hate” in intergroup conflict between natural groups. *Journal of Experimental Social Psychology*, 60, 110–120. <https://doi.org/10.1016/j.jesp.2015.04.008>
- Whitty, M. T. (2002). Liar, liar! An examination of how open, supportive and honest people are in chat rooms. *Computers in Human Behavior*, 18(4), 343–352. [https://doi.org/10.1016/S0747-5632\(01\)00059-0](https://doi.org/10.1016/S0747-5632(01)00059-0)
- Williams, J. E. & Best, D. L. (1990). *Measuring sex stereotypes: A multination study*. Sage.
- Williams, M. J. & Tiedens, L. Z. (2016). The subtle suspension of backlash: A meta-analysis of penalties for women's implicit and explicit dominance behavior. *Psychological Bulletin*, 142(2), 165–197. <https://doi.org/10.1037/bul0000039>
- Wilson, T. D., Lindsey, S. & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review*, 107(1), 101–126. <https://doi.org/10.1037/0033-295X.107.1.101>
- Winter, S., Brückner, C. & Krämer, N. C. (2015). They came, they liked, they commented: Social influence on Facebook news channels. *Cyberpsychology, Behavior, and Social Networking*, 18(8), 431–460. <https://doi.org/10.1089/cyber.2015.0005>
- Winter, S. & Krämer, N. C. (2014). A question of credibility – Effects of source cues and recommendations on information selection on news sites and blogs. *Communications*, 39(4), 434–456. <https://doi.org/10.1515/commun-2014-0020>
- Wirth, W. (2006). Involvement. In J. Bryant & P. Vorderer (Hrsg.), *Psychology of entertainment* (S. 199–213). Lawrence Erlbaum Associates.
- Wirth, W. & Kühne, R. (2013). Grundlagen der Persuasionsforschung. In W. Schweiger & A. Fahr (Hrsg.), *Handbuch Medienwirkungsforschung* (S. 313–332). Springer VS. [https://doi.org/10.1007/978-3-531-18967-3\\_16](https://doi.org/10.1007/978-3-531-18967-3_16)
- Wood, W. (1982). Retrieval of attitude-relevant information from memory: Effects on susceptibility to persuasion and on intrinsic motivation. *Journal of Personality and Social Psychology*, 42(5), 798–810.
- Wood, W. (2000). Attitude change: persuasion and social influence. *Annual Review of Psychology*, 51, 539–570. <https://doi.org/10.1146/annurev.psych.51.1.539>
- Wood, W. & Eagly, A. H. (2002). A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin*, 128(5), 699–727. <https://doi.org/10.1037//0033-2909.128.5.699>
- Wood, W. & Eagly, A. H. (2009). Gender identity. In M. R. Leary & R. H. Hoyle (Hrsg.), *Handbook of individual differences in social behavior* (S. 109–128). Guilford Press.
- Wood, W. & Eagly, A. H. (2012). Biosocial construction of sex differences and similarities in behavior. *Advances in Experimental Social Psychology*, 46, 55–123. <https://doi.org/10.1016/B978-0-12-394281-4.00002-7>
- Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S. & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, 115(3), 323–345. <https://doi.org/10.1037/0033-2909.115.3.323>

- Woods, S. M. (1976). Some dynamics of male chauvinism. *Archives of General Psychiatry*, 33(1), 63–65. <https://doi.org/10.1001/archpsyc.1976.01770010037007>
- World Health Organization. (2013). *Global and regional estimates of violence against women: Prevalence and health effects of intimate partner violence and nonpartner sexual violence*. Geneva.
- Worth, A., Augoustinos, M. & Hastie, B. (2016). “Playing the gender card”: Media representations of Julia Gillard’s sexism and misogyny speech. *Feminism & Psychology*, 26(1), 52–72. <https://doi.org/10.1177/0959353515605544>
- Yanagizawa-Drott, D. (2014). Propaganda and conflict: Evidence from the Rwandan genocide. *The Quarterly Journal of Economics*, 129(4), 1947–1994. <https://doi.org/10.1093/qje/qju020>
- Ybarra, M. L. & Mitchell, K. J. (2004). Online aggressor/targets, aggressors, and targets: A comparison of associated youth characteristics. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 45(7), 1308–1316. <https://doi.org/10.1111/j.1469-7610.2004.00328.x>
- Zárate, M. A., Garcia, B., Garza, A. A. & Hitlan, R. T. (2004). Cultural threat and perceived realistic group conflict as dual predictors of prejudice. *Journal of Experimental Social Psychology*, 40(1), 99–105. [https://doi.org/10.1016/S0022-1031\(03\)00067-2](https://doi.org/10.1016/S0022-1031(03)00067-2)
- Zeileis, A., Kleiber, C. & Jackman, S. (2008). Regression models for count data in R. *Journal of Statistical Software*, 27(8). <https://doi.org/10.18637/jss.v027.i08>
- Zhao, X., Lynch, J. G. & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37(2), 197–206. <https://doi.org/10.1086/651257>
- Zick, A. (2017). Sozialpsychologische Diskriminierungsforschung. In A. Scherr, A. El-Mafaalani & G. Yüksel (Hrsg.), *Handbuch Diskriminierung* (Bd. 32, S. 59–80). Springer. [https://doi.org/10.1007/978-3-658-10976-9\\_4](https://doi.org/10.1007/978-3-658-10976-9_4)
- Ziegele, M. (2016). *Nutzerkommentare als Anschlusskommunikation: Theorie und qualitative Analyse des Diskussionswerts von Online-Nachrichten*. Springer VS. <https://doi.org/10.1007/978-3-658-12822-7>
- Ziegele, M. & Jost, P. B. (2016). Not funny? The effects of factual versus sarcastic journalistic responses to uncivil user comments. *Communication Research*, 1–30. <https://doi.org/10.1177/0093650216671854>
- Ziegele, M., Koehler, C. & Weber, M [Mathias] (2018). Socially destructive? Effects of negative and hateful user comments on readers’ donation behavior toward refugees and homeless persons. *Journal of Broadcasting & Electronic Media*, 62(4), 636–653. <https://doi.org/10.1080/08838151.2018.1532430>
- Ziegele, M., Springer, N., Jost, P. & Wright, S. (2017). Online user comments across news and other content formats: Multidisciplinary perspectives, new directions. *Studies in Communication and Media*, 6(4), 315–332. <https://doi.org/10.5771/2192-4007-2017-4-315>
- Zych, I., Baldry, A. C [Anna C.], Farrington, D. P. & Llorent, V. J. (2019). Are children involved in cyberbullying low on empathy? A systematic review and meta-analysis of research on empathy versus different cyberbullying roles. *Aggression and Violent Behavior*, 45, 83–97. <https://doi.org/10.1016/j.avb.2018.03.004>

## Anhang

## A1: FRAGEBOGEN STUDIE I

**Herzlich willkommen bei unserer Studie zum Thema „Karrierechancen von Frauen in der IT-Branche“.**

Wir danken Ihnen für Ihr Interesse an dieser Befragung. Die Befragung dauert ca. 10-15 Minuten. Am Ende der Befragung haben Sie die Möglichkeit an einem Gewinnspiel teilzunehmen bei dem zehn Thalia-Gutscheine zu je 15,- Euro verlost werden.

**Die Richtlinien guter ethischer Forschung sehen vor, dass sich die Teilnehmer/innen an empirischen Studien explizit und nachvollziehbar mit der Teilnahme einverstanden erklären.**

**Freiwilligkeit.** Ihre Teilnahme an dieser Untersuchung ist freiwillig. Es steht Ihnen zu jedem Zeitpunkt dieser Studie frei, Ihre Teilnahme ohne Angabe von Gründen abzubrechen, ohne dass Ihnen daraus Nachteile entstehen.

**Anonymität.** Ihre Daten sind selbstverständlich vertraulich, werden nur in anonymisierter Form ausgewertet und ausschließlich für wissenschaftliche Zwecke genutzt. Ihre Angaben lassen keinen eindeutigen Schluss auf Ihre Person bzw. Ihre Identität zu.

**Risiken.** Es ist nicht zu erwarten, dass Sie gesundheitlichen Schaden durch die Teilnahme an dieser Studie nehmen. Da es sich um eine medienpsychologische Studie handelt, kann es jedoch vorkommen, dass Ihre Stimmung kurzzeitig beeinflusst wird. Sie werden Facebook-Kommentare lesen, die für manche Menschen problematisch sind.

**Fragen.** Falls Sie noch Fragen zu unserer Studie haben sollten, finden Sie hier die E-Mail-Adresse der [Projektleiter](mailto:online-diskussion.ifk@uni-muenster.de) [online-diskussion.ifk@uni-muenster.de](mailto:online-diskussion.ifk@uni-muenster.de)

Am Ende der Befragung geben wir unsere E-Mail-Adresse erneut für Sie bekannt.

**Hiermit bestätige ich, dass ich mindestens 18 Jahre alt bin sowie die Einverständniserklärung gelesen und verstanden habe.**

- Nein (nicht an der Studie teilnehmen)
- Ja

---

Viele Menschen nutzen soziale Netzwerke im Internet. Sie möchten den Kontakt zu Gleichgesinnten, Freunden und Verwandten pflegen oder sich zu aktuellen Themen informieren.

Im Folgenden möchten wir von Ihnen erfahren, ob und falls ja, wie Sie das aktuell größte soziale Netzwerk Facebook nutzen.

**1. Ich nutze das soziale Netzwerk Facebook regelmäßig (also mindestens 1x im Monat).**

- Ja
- Nein
-







5. Die folgenden Aussagen können mehr oder weniger auf Sie zutreffen. Bitte geben Sie bei jeder Aussage an, inwieweit diese im Allgemeinen auf Sie zutrifft.

	Trifft überhaupt nicht zu					Trifft voll und ganz zu
Ich informiere mich regelmäßig über das Thema „Karrierechancen von Frauen“.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich kompetent an einer Diskussion zum Thema „Karrierechancen von Frauen“ teilzunehmen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

6. Woher haben Sie Ihre Informationen zum Thema „Karrierechancen von Frauen“ erhalten?

Klicken Sie bitte sämtliche Informationsquellen an, die auf Sie zutreffen.

- Ich habe Vorträge dazu besucht.
- Ich habe mit anderen darüber gesprochen.
- Ich habe Bücher oder Artikel dazu gelesen (Print oder deren Onlineausgaben).
- Ich habe mir Fernsehsendungen (auch Online-Mediatheken) dazu angeschaut.
- Ich habe mich dazu auf sonstigen Seiten im Internet (Blogs, Social Media etc.) informiert.
- Ich habe eine andere Quelle genutzt, und zwar

Auf der nächsten Seite sehen Sie ein Facebook-Posting von heise.de, eine der meistgenutzten IT-Nachrichtenseiten in Deutschland. Unter dem Posting ist eine Diskussion zum Thema Karrierechancen von Frauen entbrannt.

**Bitte lesen Sie sich sowohl das Posting als auch die Kommentarbeiträge der Nutzer aufmerksam durch.** Für den weiteren Verlauf der Befragung ist es wichtig, dass Sie diese Seite nicht nur überfliegen.

Für den Fall, dass das Posting **nicht** angezeigt wird, können Sie die Seite in Ihrem Browser neu laden und bestätigen.

Klicken Sie bitte auf "**Weiter**" um zum Artikel zu gelangen.



1 Std · 🌐

Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



### Wo sind die Frauen? Die IT-Branche hat ein Problem

Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.

HEISE.DE

👍 Gefällt mir    💬 Kommentieren    ➦ Teilen

👍 😄 🤔 107

Top-Kommentare ▾

**Robert** Lasst uns Mädchen und Jungen früh fordern, denn gemischte Teams sind kreativer, produktiver und finanziell erfolgreicher, wie unzählige wissenschaftliche Studien zeigen.

Gefällt mir · Antworten · 👍 3 · 6 Min

**Mark** Ich finde, man sollte schon früh mit der IT-Ausbildung beginnen, am besten bereits in den Grundschulen. Je früher, desto besser! Wie sollen wir sonst in einer globalisierten Welt mithalten?

Gefällt mir · Antworten · 👍 2 · 34 Min

**Christian**



Gefällt mir · Antworten · 👍 5 · 27 Min

**Ralf Hofmann** Ich arbeite in der IT-Branche und ich kenne einige sehr fähige, bestens ausgebildete Frauen, die ihren Kollegen in nichts nachstehen. Bekanntere Beispiele: Frauen wie Marie Curie, Lise Meitner oder Grace Hopper haben sich in extrem männlich dominierten Zeiten durchgesetzt und würden das heute mit links schaffen.

Gefällt mir · Antworten · 👍 1 · 1 Std

**Freddie** Wenn ich Arbeitgeber wäre, wäre es mir wichtig, gute Leute zu bekommen, egal welchen Geschlechts. Wer bestens qualifiziert ist, soll Karriere machen. Das Geschlecht ist absolut nebensächlich.

Gefällt mir · Antworten · 27 Min

👤 Kommentieren ...



## 7. Finden Sie die gelesenen Kommentare für sich selbst relevant?

Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.

	Stimme überhaupt nicht zu					Stimme voll und ganz zu
Die Kommentare sind wichtig für mich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Es lohnt sich, die Kommentare im Kopf zu behalten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Kommentare sind es wert, ihnen Aufmerksamkeit zu schenken.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Die Kommentare sind interessant für mich.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

## 8. Beim Lesen der Kommentare haben Sie sich vermutlich einige Gedanken gemacht. In den Textfeldern haben Sie nun die Möglichkeit bis zu zehn Gedanken einzutragen. Die Reihenfolge der Gedanken spielt keine Rolle.

Bitte tragen Sie in jedes Textfeld jeweils nur einen Gedanken ein. Es ist nicht erforderlich zehn Gedanken einzutragen, sondern nur so viele Ihnen einfallen.

Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.

1. Gedanke	<input type="text"/>
2. Gedanke	<input type="text"/>
3. Gedanke	<input type="text"/>
4. Gedanke	<input type="text"/>
5. Gedanke	<input type="text"/>
6. Gedanke	<input type="text"/>
7. Gedanke	<input type="text"/>
8. Gedanke	<input type="text"/>
9. Gedanke	<input type="text"/>
10. Gedanke	<input type="text"/>

**9. Im Folgenden geht es darum, wie wahrscheinlich Sie diese Reaktionen auf die soeben gelesenen Kommentare zeigen würden. Manche Leute würden kommentieren, andere würden die Kommentare lesen, ohne etwas zu tun.**

Es geht um die generelle Tendenz, die Sie nach dem Lesen zeigen würden. Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.

	Sehr unwahr- scheinlich						Sehr wahr- scheinlich					
Ich reagiere mit einem missbilligenden Emoji („Traurig“ oder „Wütend“).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich melde Facebook die Kommentare.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich widerspreche meinen Vorrednern entschieden mit einem beleidigenden Kommentar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich unterstütze meine Vorredner mit einem sachlichen Kommentar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich unterstütze meine Vorredner und schreibe einen beleidigenden Kommentar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich lese die Kommentare, ohne etwas zu tun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich widerspreche meinen Vorrednern entschieden mit einem sachlichen Kommentar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich reagiere mit einem Like („Gefällt mir“) oder einem zustimmenden Emoji („Herz“).	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Ihre Antworten sind für unsere Studie sehr wichtig. Vielen Dank, dass Sie uns unterstützen!

Nun folgen noch einige allgemeine Fragen.

**10. Es folgen noch drei kurze Fragen zum soeben gelesenen Facebook-Beitrag.**

Bitte kreuzen Sie die passende Antwort an.

**Die Kommentare wurden geschrieben von...**

- Männern
- Frauen
- Ich kann mich nicht erinnern.

**11. Erinnern Sie sich an die Anzahl der „Gefällt-mir“-Angaben („Likes“), die der Kommentar mit dem Bild erzielt hatte?**

Tragen Sie bitte in das leere Feld die Zahl der erinnerten „Gefällt-mir“-Angaben ein. Falls Sie sich nicht erinnern, lassen Sie das Feld einfach leer.

Der Bilder-Kommentar hatte  „Gefällt-mir“-Angaben.

---

**Haben Sie eine Vermutung, warum es in der Untersuchung ging?**

Sie können in das Textfeld Ihre Gedanken kurz notieren. Sollten Sie keine Vorstellung haben, dann lassen Sie das Feld einfach frei.

---

**12. Wir möchten Sie noch um einige Angaben zu Ihrer Person bitten.**

Ihr Geschlecht...

Ihr Alter...  Jahre

---

### 13. Welchen Bildungsabschluss haben Sie?

Bitte wählen Sie den höchsten Bildungsabschluss, den Sie bisher erreicht haben.

- Schule beendet ohne Abschluss
  - Noch Schüler
  - Volks- oder Hauptschulabschluss
  - Mittlere Reife, Realschul- oder gleichwertiger Abschluss
  - Abgeschlossene Lehre
  - Fachabitur, Fachhochschulreife
  - Abitur, Hochschulreife
  - Fachhochschul-/Hochschulabschluss
  - Anderer Abschluss, und zwar:
- 

### 14. Was machen Sie beruflich?

- In Ausbildung
  - Student/in
  - Angestellte/r Teilzeit
  - Angestellte/r Vollzeit
  - Beamte/r
  - Selbstständig
  - Arbeitslos/Arbeit suchend
  - Sonstiges:
-

## **Vielen Dank für Ihre Mithilfe!**

Sie haben uns durch Ihre Bereitschaft zur Mitarbeit an unserer Studie sehr geholfen.

In der Studie wird die **Reaktion auf frauenfeindliche Hasskommentare in sozialen Netzwerken** untersucht. Dafür haben wir das Beispiel Karrierechancen von Frauen in der IT-Branche gewählt. Die Meinungen zu diesem Thema sind sehr kontrovers. Mit Ihrer Teilnahme haben Sie einen wichtigen Beitrag zur Erforschung des Themas geleistet. Sie wurden im Verlauf der Studie einer von 16 Bedingungen zugewiesen, bei der die Kommentare von einem eher sachlichen, über einen aggressiven bis hin zu einem angreifenden und offen diskriminierenden Stil mit und ohne Aufforderungen zu Gewalt konstruiert waren. Die Studie untersucht, inwiefern das Geschlecht der Kommentatoren sowie die Anzahl der Kommentarlikes die Reaktion von Lesern beeinflusst.

Der Artikel von Heise.de ist vom 01. März 2017, wurde jedoch nicht bei Facebook veröffentlicht. Wir haben das heise.de-Posting und die darunter liegende Diskussion mit schriftlicher Erlaubnis von heise.de konstruiert. Alle Namen der kommentierenden Personen sind frei erfunden. Die Kommentare basieren aber auf echten Kommentaren in sozialen Netzwerken und wurden für die vorliegende Studie sprachlich und inhaltlich angepasst.

Falls Sie weitere Fragen zu diesem Projekt haben, wenden Sie sich bitte per E-Mail an die verantwortlichen [Projektleiter](mailto:online-diskussion.ifk@uni-muenster.de) [online-diskussion.ifk@uni-muenster.de](mailto:online-diskussion.ifk@uni-muenster.de)

Falls Sie Fragen zu Hasskommentaren haben oder selbst Opfer von (frauenfeindlichen) Hasskommentaren geworden sind, haben Sie die Möglichkeit hier Antwort und Hilfe zu finden: [www.hilfetelefon.de](http://www.hilfetelefon.de) oder [www.no-hate-speech.de](http://www.no-hate-speech.de)

Auf der nächsten Seite haben Sie die Möglichkeit am Gewinnspiel teilzunehmen.

---

## **Um am Gewinnspiel teilzunehmen, tragen Sie bitte Ihre E-Mail-Adresse ein.**

- Ja, ich will am Gewinnspiel teilnehmen. Verlost werden zehn Thalia-Gutscheine im Wert von je 15,- Euro. Ich bin damit einverstanden, dass meine E-Mail-Adresse bis zur Ziehung der Gewinner gespeichert wird. Meine Angaben in dieser Befragung bleiben weiterhin anonym, denn die E-Mail-Adresse wird getrennt von meinen Antworten erhoben. Meine E-Mail-Adresse wird nicht an Dritte weitergegeben.
- 

## **Vielen Dank für Ihre Teilnahme!**

**Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.**

## A2: FRAGEBOGEN STUDIE II

**Herzlich willkommen bei unserer Studie zum Thema „Diskussionen in sozialen Medien“.**

Wir danken Ihnen für Ihr Interesse an dieser Befragung. Die Befragung dauert ca. 10-15 Minuten.

**Die Richtlinien guter ethischer Forschung sehen vor, dass sich die Teilnehmer/innen an empirischen Studien explizit und nachvollziehbar mit der Teilnahme einverstanden erklären.**

**Freiwilligkeit.** Ihre Teilnahme an dieser Untersuchung ist freiwillig. Es steht Ihnen zu jedem Zeitpunkt dieser Studie frei, Ihre Teilnahme ohne Angabe von Gründen abzubrechen, ohne dass Ihnen daraus Nachteile entstehen.

**Risiken.** Es ist nicht zu erwarten, dass Sie gesundheitlichen Schaden durch die Teilnahme an dieser Studie nehmen. Da es sich um eine medienpsychologische Studie handelt, kann es jedoch vorkommen, dass Ihre Stimmung kurzzeitig beeinflusst wird. Sie werden Facebook-Kommentare lesen, die für manche Menschen problematisch sind.

**Informationen zum Datenschutz.** Gemäß der Datenschutzverordnung (DSVGO) wird darauf hingewiesen, dass die von Ihnen zur Verfügung gestellten Daten ausschließlich für Forschungszwecke verwendet und nicht an Dritte weitergeleitet werden. Die Auswertung der Daten erfolgt anonymisiert und Rückschlüsse auf Ihre Person sind nicht möglich. Bitte bestätigen Sie, dass Sie mit den Bedingungen dieser Befragung einverstanden sind.

**Fragen.** Falls Sie noch Fragen zu unserer Studie haben sollten, finden Sie hier die E-Mail-Adresse der [Projektleitung](mailto:online-diskussion.ifk@uni-muenster.de) [online-diskussion.ifk@uni-muenster.de](mailto:online-diskussion.ifk@uni-muenster.de)

Am Ende der Befragung geben wir unsere E-Mail-Adresse erneut für Sie bekannt.

**Hiermit bestätige ich, dass ich mindestens 18 Jahre alt bin sowie die Einverständniserklärung gelesen und verstanden habe.**

- Nein (nicht an der Studie teilnehmen)
- Ja
- 

Viele Menschen nutzen soziale Netzwerke im Internet. Sie möchten den Kontakt zu Gleichgesinnten, Freunden und Verwandten pflegen oder sich zu aktuellen Themen informieren.

Im Folgenden möchten wir von Ihnen erfahren, ob und falls ja, wie Sie das aktuell größte soziale Netzwerk Facebook nutzen.

**1. Ich nutze das soziale Netzwerk Facebook regelmäßig (also mindestens 1x im Monat).**

- Ja
- Nein
-

**2. Wie nutzen Sie Facebook? Manche Leute kommentieren und teilen Links, private Fotos und Erlebnisse, andere lesen gerne mit, posten aber kaum oder sehr selten.**

Nachfolgend lesen Sie eine Reihe von Aussagen, wie man Facebook nutzen kann. Mit den Auswahloptionen können Sie Ihre Beurteilung abstimmen.

Im Folgenden interessiert uns Ihr normales Nutzungsverhalten – nicht etwaige Ausnahmesituationen.

	nie						immer	Ich kenne diese Funktion nicht.
Ich teile Inhalte in meiner Chronik oder als Story.	<input type="radio"/>							
Ich drücke den Gefällt-mir-Button oder Reaktionsbutton für Beiträge oder Kommentare.	<input type="radio"/>							
Ich ignoriere die meisten Beiträge und scrolle nur durch meinen Newsfeed.	<input type="radio"/>							
Ich kommentiere Beiträge oder antworte auf Kommentare.	<input type="radio"/>							
Ich lese Beiträge oder Kommentare.	<input type="radio"/>							
Ich melde unangemessene Beiträge oder Kommentare.	<input type="radio"/>							

**3. Wir möchten Sie um einige Angaben zu Ihrer Person bitten.**

**Ihr Geschlecht...**

- weiblich
- männlich
- anderes

Ich bin  Jahre alt.

#### 4. Welchen Bildungsabschluss haben Sie?

Bitte wählen Sie den höchsten Bildungsabschluss, den Sie bisher erreicht haben.

- Schule beendet ohne Abschluss
- Noch Schüler
- Volks- oder Hauptschulabschluss
- Mittlere Reife, Realschul- oder gleichwertiger Abschluss
- Abgeschlossene Lehre
- Fachabitur, Fachhochschulreife
- Abitur, Hochschulreife
- Fachhochschul-/Hochschulabschluss
- Anderer Abschluss, und zwar:

#### 5. Was machen Sie beruflich?

- In Ausbildung
  - Student/in
  - Angestellte/r Teilzeit
  - Angestellte/r Vollzeit
  - Beamte/r
  - Selbstständig
  - Arbeitslos/Arbeit suchend
  - Sonstiges:
-





Frauen sollten von Männern umsorgt und beschützt werden.	<input type="radio"/>					
Viele Frauen haben Spaß daran, mit Männern zu „spielen“, indem sie sich zuerst verführerisch geben, dann aber die Annäherungsversuche der Männer zurückweisen.	<input type="radio"/>					

**8. Ich bin in der IT-Branche tätig.**

- nein
- ja

**9. Die folgenden Aussagen können mehr oder weniger auf Sie zutreffen. Bitte geben Sie bei jeder Aussage an, inwieweit diese im Allgemeinen auf Sie zutrifft.**

	Trifft überhaupt nicht zu					Trifft voll und ganz zu
Ich informiere mich regelmäßig über das Thema „Karrierechancen von Frauen“.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich fühle mich kompetent an einer Diskussion zum Thema „Karrierechancen von Frauen“ teilzunehmen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**10. Woher haben Sie Ihre Informationen zum Thema „Karrierechancen von Frauen“ erhalten?**

Klicken Sie bitte sämtliche Informationsquellen an, die auf Sie zutreffen.

- Ich habe Vorträge dazu besucht.
- Ich habe mir Fernsehsendungen (auch Online-Mediatheken) dazu angeschaut.
- Ich habe Bücher oder Zeitungsartikel dazu gelesen (Print oder deren Onlineausgaben).
- Ich habe mit anderen darüber gesprochen.
- Ich habe mich dazu auf sonstigen Seiten im Internet (Blogs, Social Media etc.) informiert.
- Ich habe eine andere Quelle genutzt, und zwar

Auf der nächsten Seite sehen Sie ein Facebook-Posting von heise.de, eine der meistgenutzten IT-Nachrichtenseiten in Deutschland. Unter dem Posting ist eine Diskussion entstanden.

**Bitte lesen Sie sich sowohl das Posting als auch die Kommentarbeiträge aufmerksam durch.** Für den weiteren Verlauf der Befragung ist es wichtig, dass Sie diese Seite nicht nur überfliegen.

Für den Fall, dass das Posting **nicht** angezeigt wird, können Sie die Seite in Ihrem Browser **neu laden** und bestätigen.

Klicken Sie bitte auf "**Weiter**" um zum Artikel zu gelangen.

---

Guten Morgen ☺️ (schorlil)



HEISE.DE

### Perseverance landet erfolgreich auf dem Mars

Der NASA-Rover Perseverance hat den Mars erreicht und ist erfolgreich gelandet. Das war sch...

   138

17 Kommentare

 Gefällt mir

 Kommentieren

 Teilen

Relevanteste zuerst ▼



Gib einen Kommentar ein ...



**Harald**  
Die Abstiegsstufe, wurde die, wie beim letzten Skycrane-Manöver, auch gezielt entfernt zerschellen gelassen? Da wurde dieses mal gar nichts von erwähnt. Dürften sie aber auch so gemacht haben, denke ich mal.

Gefällt mir · Antworten · 6 Std.



 Top-Fan

**Jan**  
**Harald** Soweit ich es mitbekommen habe, driftete diese geplant danach kurz ab und zerschellte

Gefällt mir · Antworten · 6 Std.



**Harald**  
**Jan**, ah okay, danke. Wäre auch ungewöhnlich, hätten sie es dieses mal anders gemacht.

Gefällt mir · Antworten · 6 Std.

 1



Verfasse eine Antwort ...





**13. Beim Lesen der Kommentare haben Sie sich vermutlich einige Gedanken gemacht. In den Textfeldern haben Sie nun die Möglichkeit bis zu zehn Gedanken einzutragen. Die Reihenfolge der Gedanken spielt keine Rolle.**

Bitte tragen Sie in jedes Textfeld jeweils nur einen Gedanken ein. Es ist nicht erforderlich zehn Gedanken einzutragen, sondern nur so viele Ihnen einfallen.

Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.

1. Gedanke	<input type="text"/>
2. Gedanke	<input type="text"/>
3. Gedanke	<input type="text"/>
4. Gedanke	<input type="text"/>
5. Gedanke	<input type="text"/>
6. Gedanke	<input type="text"/>
7. Gedanke	<input type="text"/>
8. Gedanke	<input type="text"/>
9. Gedanke	<input type="text"/>
10. Gedanke	<input type="text"/>

---



**16. Im Folgenden geht es darum, wie wahrscheinlich Sie die aufgelisteten Reaktionen auf die soeben gelesenen Kommentare zeigen würden. Manche Leute würden kommentieren, andere würden die Kommentare lesen, ohne etwas zu tun.**

Es geht um die generelle Tendenz, die Sie nach dem Lesen zeigen würden. Antworten Sie möglichst spontan. Es gibt keine richtigen oder falschen Antworten.

	Sehr unwahrscheinlich			Sehr wahrscheinlich		
Ich melde Facebook die Kommentare.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich lese die Kommentare, ohne etwas zu tun.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich reagiere ablehnend mit einem Kommentar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Ich reagiere zustimmend mit einem Kommentar oder einem Like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

---

**Haben Sie alle Aufgaben so ausgeführt, wie in den jeweiligen Instruktionen gebeten? Diese Antwort hat für Sie keine Konsequenzen.**

- Ich habe alle Aufgaben, wie in den Instruktionen verlangt, bewältigt.
- Manchmal habe ich irgendetwas geklickt, weil ich unmotiviert war oder mich einfach nicht ausgekannt habe.
- Ich habe häufig irgendetwas angeklickt, damit ich schnell fertig werde.

---

**Vielen Dank für Ihre Mithilfe!**

Sie haben uns durch Ihre Bereitschaft zur Mitarbeit an unserer Studie sehr geholfen.

In der Studie wird die **Reaktion auf frauenfeindliche Hasskommentare in sozialen Netzwerken** untersucht. Dafür haben wir das Beispiel Karrierechancen von Frauen in der IT-Branche gewählt. Die Meinungen zu diesem Thema sind sehr kontrovers. Mit Ihrer Teilnahme haben Sie einen wichtigen Beitrag zur Erforschung des Themas geleistet. Sie wurden im Verlauf der Studie einer von 3 Bedingungen zugewiesen, bei der Sie entweder frauenfeindliche Hasskommentare oder einen neutralen Beitrag gelesen haben. Die Studie untersucht, inwiefern das Geschlecht der Kommentierenden die Reaktion von Lesern und Leserinnen beeinflusst.

Der Artikel mit den Hasskommentaren wurde jedoch nicht bei Facebook veröffentlicht. Wir haben das heise.de-Posting und die darunter liegende Diskussion mit schriftlicher Erlaubnis von heise.de konstruiert. Alle Namen der kommentierenden Personen sind frei erfunden. Die

Kommentare basieren aber auf echten Kommentaren in sozialen Netzwerken und wurden für die vorliegende Studie sprachlich und inhaltlich angepasst.

Falls Sie weitere Fragen zu diesem Projekt haben, wenden Sie sich bitte per E-Mail an die verantwortliche [Projektleitung](mailto:online-diskussion.ifk@uni-muenster.de) [online-diskussion.ifk@uni-muenster.de](mailto:online-diskussion.ifk@uni-muenster.de)

Falls Sie Fragen zu Hasskommentaren haben oder selbst Opfer von (frauenfeindlichen) Hasskommentaren geworden sind, haben Sie die Möglichkeit hier Antwort und Hilfe zu finden: [www.hilfetelefon.de](http://www.hilfetelefon.de) oder [www.no-hate-speech.de](http://www.no-hate-speech.de)

---

**Vielen Dank für Ihre Teilnahme!**

**Ihre Antworten wurden gespeichert, Sie können das Browser-Fenster nun schließen.**

### A3: STIMULUS SACHLICHER KOMMENTAR – MÄNNLICHE KOMMENTIERENDE – VIELE LIKES



Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



Wo sind die Frauen? Die IT-Branche hat ein Problem  
Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.  
HEISE.DE

Gefällt mir · Kommentieren · Teilen

107

Top-Kommentare ▾

**Robert** Lasst uns Mädchen und Jungen früh fordern, denn gemischte Teams sind kreativer, produktiver und finanziell erfolgreicher, wie unzählige wissenschaftliche Studien zeigen.  
Gefällt mir · Antworten · 112 · 6 Min

**Mark** Ich finde, man sollte schon früh mit der IT-Ausbildung beginnen, am besten bereits in den Grundschulen. Je früher, desto besser! Wie sollen wir sonst in einer globalisierten Welt mithalten?  
Gefällt mir · Antworten · 79 · 34 Min



Gefällt mir · Antworten · 130 · 27 Min

**Ralf** Ich arbeite in der IT-Branche und ich kenne einige sehr fähige, bestens ausgebildete Frauen, die ihren Kollegen in nichts nachstehen. Bekanntere Beispiele: Frauen wie Marie Curie, Lise Meitner oder Grace Hopper haben sich in extrem männlich dominierten Zeiten durchgesetzt und würden das heute mit links schaffen.  
Gefällt mir · Antworten · 35 · 1 Std

**Freddie** Wenn ich Arbeitgeber wäre, wäre es mir wichtig, gute Leute zu bekommen, egal welchen Geschlechts. Wer bestens qualifiziert ist, soll Karriere machen. Das Geschlecht ist absolut nebensächlich.  
Gefällt mir · Antworten · 18 · 58 Min

Kommentieren ...

#### A4: STIMULUS OFFENSIVE SPEECH – MÄNNLICHE KOMMENTIERENDE – VIELE LIKES



Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



### Wo sind die Frauen? Die IT-Branche hat ein Problem

Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.

HEISE.DE

Gefällt mir · Kommentieren · Teilen

107

Top-Kommentare

**Robert Hirsch** Das ist doch genetisch bedingt, dass Frauen besser mit Kindern umgehen können. Und ausserdem hat die Floristik-Branche auch ein Gender-Problem. Keine Männer 😄😄

Gefällt mir · Antworten · 112 · 6 Min

**Mark Us** Jeden IT-Affinen den ich kenne würde vor Freude fast ausflippen wenn er eine Frau träfe mit der er hemmungslos über Technik quatschen könnte.

Gefällt mir · Antworten · 79 · 34 Min

**Christian**



Gefällt mir · Antworten · 130 · 27 Min

**Ralf** Mit erzwungenen Feministinnenquoten die besser gebildeten und wesentlich fähigeren männer in die Arbeitslosigkeit schicken. Schade das mir so eine nicht im dunkeln begegnet, echt schade

Gefällt mir · Antworten · 35 · 1 Std

**Freddie** Eine feste Quote für Frauen in Chef-Positionen ist Unsinn, und befördert zwangsläufig unfähige Personen in die Chef-Etage.

Gefällt mir · Antworten · 18 · 58 Min

Kommentieren ...



## A5: STIMULUS HATE SPEECH – WEIBLICHE KOMMENTIERENDE – WENIG LIKES



1 Std · 🌐

Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



### Wo sind die Frauen? Die IT-Branche hat ein Problem

Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.

HEISE.DE

👍 Gefällt mir    💬 Kommentieren    ➦ Teilen

👍 😂 😡 107

Top-Kommentare ▾

**Steffi** **Wolch** Das ist doch genetisch bedingt, dass eine Trulla besser mit Kindern umgehen kann. Wie kann man nur so dumm sein und es nicht einsehen...

Gefällt mir · Antworten · 🗳️ 3 · 6 Min

**Mela** **Nie** Jeden IT-Affinen den ich kenne würde vor Freude fast ausflippen wenn er eine trafe mit der er hemmungslos über Technik quatschen könnte. Statt Mathe und IT zu machen heulen die Dummbrazten und fordern Sonderprivilegien.

Gefällt mir · Antworten · 🗳️ 2 · 34 Min

**Sabine** **Wolch**



Gefällt mir · Antworten · 🗳️ 5 · 27 Min

**Monika** **Wolch** Mit erzwungenen FeminiSStinnenquoten die besser gebildeten und wesentlich fähigeren männer in die Arbeitslosigkeit schicken. Tolle Geisteshaltung haben diese Bitterfregatten.

Gefällt mir · Antworten · 🗳️ 1 · 1 Std

**Andrea** **Wolch** Eine feste Quote für Frauen in Chef-Positionen kann nur einem Fieberwahn entspringen. Lasst die unfähigen Personen ruhig in die Chef-Etagen...oh weh oh weh.

Gefällt mir · Antworten · 27 Min



Kommentieren ...



## A6: STIMULUS DANGEROUS SPEECH – WEIBLICHE KOMMENTIERENDE – WENIG LIKES



1 Std · 🌐

Das Problem ist nicht neu, aber hochaktuell: In der IT-Branche herrscht Frauenmangel. In Deutschland liegt der weibliche Anteil laut einer Bitkom-Umfrage von 2015 bei 24 Prozent.



### Wo sind die Frauen? Die IT-Branche hat ein Problem

Frauen sind in der Tech-Branche in der Minderheit und müssen oftmals gegen Probleme ankämpfen. Dabei kann die IT-Welt weibliche Fachkräfte dringend gebrauchen.

HEISE.DE

👍 Gefällt mir    💬 Kommentieren    ➦ Teilen

👍 🤔 😬 107

Top-Kommentare ▾

**Steffi Heison** Das ist doch genetisch bedingt, dass eine Muschi besser mit Kindern umgehen kann. Mimimi, arbeiten, Selbstverwirklichung. Angeleint an den Herd gehören sie.

Gefällt mir · Antworten · 🗳️ 3 · 6 Min

**Mela** Jeden IT-Affinen den ich kenne würde vor Freude fast ausflippen wenn er eine trafe mit der er hemmungslos über Technik quatschen könnte. Wahrscheinlich muss man den dummen Schlampen ordentlich eine auf die Nuss geben. Soll das Denkvermögen erhöhen...

Gefällt mir · Antworten · 🗳️ 2 · 34 Min

**Sabine**



Gefällt mir · Antworten · 🗳️ 5 · 27 Min

**Monika** Mit erzwungenen FeminiSStinnenquoten die besser gebildeten und wesentlich fähigeren männer in die Arbeitslosigkeit schicken. Schade das mir so eine nicht im dunkeln begegnet, echt schade

Gefällt mir · Antworten · 🗳️ 1 · 1 Std

**Andrea** Eine feste Quote für Frauen in Chef-Positionen...Sowas entspringt auch nur einem kranken Asselhirn. Zertreten wie eine Assel sollte man sie!!!

Gefällt mir · Antworten · 27 Min

👤 Kommentieren ...



## A7: INHALTSANALYSE DER REZEPTIONSGEDANKEN

**Codiereinheit** Als Codiereinheit wird jeder Gedanke einer Versuchsperson unter TL01\_01 bis TL01\_10 verstanden. Als Gedanke zählen:

- einzelne Stichworte
- ganze Sätze
- Sinn gebende Zeichenfolgen (z.B. Emoticons, Auslassungspunkte, Fragezeichen). Keine Sinn gebenden Zeichenfolgen sind Raumfüller wie Gedankenstriche oder ein einzelner Punkt, Schrägstriche.

**Kontexteinheit** Um einen Gedanken eindeutig codieren zu können (v.a. bei einzelnen Stichworten notwendig), sollen alle Kommentare einer Versuchsperson als Kontexteinheit herangezogen werden.

Gedanken, die sich über mehrere Spalten ziehen (Kette), werden gemeinsam codiert. Der erste Gedanke wird wie üblich codiert. Dieser vererbt sodann seine Eigenschaften an alle folgenden Gedanken, die sinnhaft zu einer Codiereinheit zusammengefasst werden können. Gedanken, die vollständige Sätze darstellen (Subjekt, Prädikat, evtl. Objekt) gelten nicht als Kette und werden einzeln codiert.

Beispiel: *Spalte 1: „viele Kranke im Internet“ Spalte 2: „denen im realen Leben niemand zuhört“*  
→ Codierung der beiden Gedanken identisch mit Valenz der Codierung: negativ, Ursprung des Gedankens: intrinsischer Ursprung

**Vorgehensweise:** Angeraten ist bei der Codierung wie folgt vorzugehen:

- Kategorienschema und Regeln vergegenwärtigen
- Alle Kommentare einer Versuchsperson lesen
- Tendenz der Kommentare einer Versuchsperson feststellen
- Weichen Kommentare einer Versuchsperson vom Tenor ab?
- Codieren nach dem vorliegenden Kategoriensystem

TL02	Summe aller Gedanken aus den Variablen TL01_01 bis TL01_10
<p>Als jeweils ein Gedanke wird ein Eintrag unter TL01_01 bis TL01_10 gezählt, unerheblich davon, wie viele Gedanken tatsächlich in einer Variablen vorkommen (s. Hinweise zur Codiereinheit). TL02 nimmt also einen Wert zwischen 1 und 10 ein, fehlt ein Eintrag unter TL01_01 bis TL01_10 wird das als fehlender Wert mit dem Code -9 versehen</p>	

TL03	Valenz der Gedanken
<p>Anteil der Differenz positiver bzw. neutraler und negativer Gedanken an allen Gedanken unter TL03_01 bis TL03_10. Bei einem Überschuss positiver bzw. neutraler Gedanken, ist der Wert positiv. Werden mehr negative Gedanken genannt, trägt der eingetragene Wert ein negatives Vorzeichen. Fehlende Werte (kein einziger Gedanke wurde genannt) werden in diesem Fall <u>manuell mit -9 codiert</u>.</p>	
<p>Die allgemeine Formel lautet: <math display="block">\frac{\sum \text{positive Gedanken} + \sum \text{neutrale Gedanken} - \sum \text{negative Gedanken}}{(\sum \text{positive Gedanken} + \sum \text{neutrale Gedanken} + \sum \text{negative Gedanken})}</math></p>	
<p>Die Valenz der Gedanken nimmt einen Wert zwischen -1 und +1 ein.</p>	
<p>Beispiel 1: 6 positive Gedanken, 0 neutrale Gedanken, 1 negativer Gedanke.  <math>(6-1)/(6+1)=0,71</math></p> <p>In diesem Fall wird 0,71 eingetragen.</p>	
<p>Beispiel 2: 1 positiver Gedanke, 1 neutraler Gedanke, 3 negative Gedanken.  <math>(1+1-3)/(1+1+3)=-0,2</math></p> <p>In diesem Fall wird -0,2 eingetragen.</p>	
TL03_1	Positive Gedanken
<p>Codiert wird die Summe <u>eindeutig positiver</u> Gedanken (Argumente, Eigenschaften, Assoziationen) unter TL01_01 bis TL01_10, die entweder mit dem Stimulus oder der Thematik zusammenhängen. Es wird <u>naiv codiert</u>, d.h. wörtliche Codierung; weitergehende, psychologische Interpretationen oder Schlussfolgerungen (was gemeint sein <u>könnte</u>) werden vermieden. <b>Ausnahme:</b> <u>Redewendungen, feststehende Ausdrücke und Sprichwörter</u>. Rufen diese positive Assoziationen hervor, erhalten sie diesen Code.</p>	
<p>Bei Uneindeutigkeit werden die übrigen Gedanken einer Versuchsperson als Kontexteinheit herangezogen.</p>	
<p><b>Codierleitende Frage: Handelt es sich um eine eindeutig positive Bewertung?</b></p>	

Beispiel: „Der letzte Kommentar von Freddie... hat mir am besten gefallen und genau ausgedrückt, was ich denke: es geht um die Kompetenz der Leute und nicht um das Geschlecht!“

Positive Gedanken werden mit einem Zahlencode zwischen 0 und 10 versehen: Wird ein positiver Gedanke genannt, wird 1 codiert, werden fünf positive Gedanken genannt, wird dies mit 5 codiert.

Wenn eine Versuchsperson keinen einzigen Gedanken nennt, wird dies mit -9 für fehlende Werte codiert.

### **TL03\_2 Negative Gedanken**

Codiert wird die Summe eindeutig negativer Gedanken (Argumente, Eigenschaften, Assoziationen) unter TL01\_01 bis TL01\_10, die entweder mit dem Stimulus oder der Thematik zusammenhängen. Es wird naiv codiert, d.h. wörtliche Codierung; weitergehende, psychologische Interpretationen oder Schlussfolgerungen (was gemeint sein könnte) werden vermieden. **Ausnahme:** Redewendungen, feststehende Ausdrücke und Sprichwörter. Rufen diese negative Assoziationen hervor, erhalten sie diesen Code.

Bei Uneindeutigkeit werden die übrigen Gedanken einer Versuchsperson als Kontexteinheit herangezogen.

#### **Codierleitende Frage: Handelt es sich um eine eindeutig negative Bewertung?**

Beispiel: „unqualifizierte Aussagen“, „Ich habe das Bild nicht verstanden: „Programmiererinnen sind billig!“ „Sind die alle arbeitslos?“ „bildungsfern“

Negative Gedanken werden mit einem Zahlencode zwischen 0 und 10 versehen: Wird ein negativer Gedanke genannt, wird 1 codiert, werden fünf negative Gedanken genannt, wird dies mit 5 codiert.

Wenn eine Versuchsperson keinen einzigen Gedanken nennt, wird dies mit -9 für fehlende Werte codiert.

### **TL03\_3 Neutrale oder ambivalente Gedanken**

Codiert wird die Summe der neutralen, irrelevanten oder ambivalenten Gedanken (Argumente, Eigenschaften, Assoziationen) unter TL01\_01 bis TL01\_10. Neutrale Gedanken sind v.a. sachliche Feststellungen oder Fakten, die keine Anzeichen einer Bewertung aufweisen. Fällt ein Gedanke ambivalent aus und eben nicht eindeutig positiv oder negativ, wird er als neutraler Gedanke codiert, v.a. wenn eine eigentlich positive/negative Aussage durch eine Relativierung abgeschwächt wird (z.B. „halbwegs themenrelevant“, „etwas unsachgemäß“, „eigentlich ganz angenehm“ etc.).

Es wird naiv codiert, d.h. wörtliche Codierung; weitergehende, psychologische Interpretationen oder Schlussfolgerungen (was gemeint sein könnte) werden vermieden.

Bei Uneindeutigkeit werden die übrigen Gedanken einer Versuchsperson als Kontexteinheit herangezogen.

**Codierleitende Frage: Handelt es sich um eine einfache sachliche Feststellung?**

Beispiel: „Zur Kenntnis genommen“, „Offensichtlich haben viele Personen Schwierigkeiten mit Rechtschreibung und Grammatik.“

Die Variable TL03\_3 kann einen Wert zwischen 0 und 10 einnehmen.

Nennt eine Versuchsperson keinen einzigen Gedanken, wird dies mit -9 für fehlende Werte codiert.

**TL03\_01 bis TL03\_10 Valenz der einzelnen Gedanken**

Jeder einzelne Gedanke einer Versuchsperson unter TL01\_01 bis TL01\_10 wird mit dem Wert 1 (positiver Gedanke), 2 (negativer Gedanke) und 3 (neutraler/irrelevanter Gedanke) codiert. Erläuterungen dazu, was unter einem positiven/negativen/neutralen Gedanken zu verstehen ist, werden den Erläuterungen unter TL03\_1 bis TL03\_3 entnommen. Fehlende Werte werden mit -9 für fehlende Werte codiert.

**TL04 Ursprung der Gedanken**

Welchen Hintergrund bzw. Ursprung lässt ein Gedanke erkennen?

Jeder Gedanke wird einzeln unter TL04\_01 bis TL04\_10 codiert. Der Ursprung eines Gedankens kann entweder durch den Stimulus ausgelöst sein, als Reaktion darauf verstanden werden oder zwar thematisch mit dem Stimulus zusammenhängen, aber auf eigenen Überlegungen beruhen (s.u. für die einzelnen Codes).

Es wird naiv codiert, d.h. wörtliche Codierung; weitergehende, psychologische Interpretationen oder Schlussfolgerungen (was gemeint sein könnte) werden vermieden.

Bei Uneindeutigkeit werden die übrigen Gedanken einer Versuchsperson als Kontexteinheit herangezogen.

## 1 Externer Ursprung

Die Gedanken sind erkennbar durch den Stimulus induziert, da sie sich ausdrücklich auf ihn beziehen, indem sie z.B. Anweisungen wiederholen oder Teile des Stimulus zitieren bzw. paraphrasieren. Der Gedanke thematisiert die Kommentare an sich (z.B. einzelne Aussagen, deren Sachlichkeit oder Aggressivität, auch mögliche Gefühle, die den Kommentatoren unterstellt werden), die Anzahl der Likes, das Geschlecht der Kommentatoren.

**Mögliche Indikatoren:** „Diese/r/s Kommentare/Argumente/Männer/Bild/“ „Hier“ etc.

**Codierleitende Frage: Wie wird der Stimulus beschrieben? → DESKRIPTION**

Es ist keine eindeutige Reaktion (z.B. Gefühlsregung, Bewertung) auf den Kommentar erkennbar.

Beispiel: „*Ich soll jetzt zehn Gedanken nennen.*“ „*Viele Rechtschreibfehler*“ „*Frauenstimmen hört man nicht*“ „*Diese Männer haben das männliche und weibliche Prinzip verinnerlicht.*“ „*Die haben Angst vor Frauen.*“

## 2 Modifizierter externer Ursprung

Die Gedanken sind erkennbare Reaktionen auf den Stimulus, z.B. ausführliche Antworten auf den Stimulus. Es findet eine erkennbare affektive Reaktion oder eine Verhaltensreaktion (z.B. eine Bewertung, eine Gefühlsregung, eine Erwartung) auf den Stimulus durch die Versuchsperson statt.

**Mögliche Indikatoren:** „Das ist traurig, dass...“ „Kann ich nicht nachvollziehen, dass...“

**Codierleitende Frage: Was löst der Stimulus für affektive Reaktionen bei den Versuchspersonen aus? → EMOTION oder VERHALTEN**

Beispiel: „*Ich bin sauer, dass diese Männer so frauenfeindlich kommentieren.*“

„*Das Argument im letzten Kommentar hat mich am meisten überzeugt.*“

„*Wer ernsthaft glaubt Florist zu sein oder in der IT zu arbeiten sei durch die Gene bestimmt, ist ganz schön arm dran.*“

### 3 Intrinsicischer Ursprung

Die Gedanken lassen sich nicht als Reaktion auf den Stimulus interpretieren. Sie können zwar thematisch mit dem Stimulus zusammenhängen, lassen sich aber eher auf einen Erfahrungsabgleich, auf Vorwissen oder eigene Überlegungen zurückführen. Es werden neue Gedanken und Aspekte eingeführt, die eine kognitive Auseinandersetzung (auch kognitive Bewertungen) erkennen lassen.

Indikator: Besonders elaborierte Gedanken, die nicht explizit beim Stimulus ansetzen.

**Mögliche Indikatoren:** „Frauen, die in der IT arbeiten, sollen männlichen IT-Spezialisten die Arbeit wegnehmen? Es gibt ohnehin viel zu wenig Fachkräfte in dieser Branche, da nimmt niemand jemand anderem etwas weg.“

**Codierleitende Frage: Welche zusätzlichen Aspekte zeigen sich im Gedanken? Was weiß die Versuchsperson zusätzlich zum Thema? KOGNITION**

### 4 Sonstiger Ursprung

Nur anwenden, wenn sich der Gedanke nicht eindeutig einem der drei vorhergehenden Merkmalsausprägungen zuordnen lässt. Naive Codierung beachten und bei Uneindeutigkeit die restlichen Gedanken der Versuchsperson heranziehen.

Beispiel: „*nichts*“ „...“ „*puh*“