universität wien

MASTERARBEIT | MASTER'S THESIS

Titel | Title

Humor als Schutzschild: Eine quantitative Studie zur Rezeption misogyner Hatespeech mit humoristischen Elementen

verfasst von | submitted by

Pauline Pondorf

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 841
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Publizistik- und Kommunikationswissenschaft
Betreut von | Supervisor:	Univ.-Prof. Dr. Jörg Matthes

Die Kommunikationslandschaft hat durch die digitale Ära einen radikalen Wandel vollzogen, wobei sich Social-Media-Plattformen als Schauplätze für öffentliche Diskussionen etabliert haben. Diese Veränderung hat zu einem Anstieg der Verbreitung von Hatespeech, die oft misogyn motiviert ist, geführt. Innerhalb dieses Kontextes ist die subtile Hassrede, die auf einer Kombination aus diskriminierenden Inhalten und humoristischen Elementen beruht, ein problematisches Phänomen. Daher zielt diese Studie darauf ab, die Wirkung von expliziter und offen Hassrede im Vergleich zur subtilen, humorvollen Form der Hatespeech zu untersuchen, wobei das Geschlecht der Teilnehmenden als moderierende Variable berücksichtigt wird. Basierend auf dem Stand der aktuellen Forschung werden die Hypothesen entwickelt, dass explizite Hassrede zu einer negativeren emotionalen Stimmung führt, die Counterspeech-Bereitschaft erhöht und misogyne Einstellungen verstärkt. Es wird ebenfalls angenommen, dass diese Effekte bei Männern tendenziell ausgeprägter sind als bei Frauen. Ein experimentelles 2 x 3 Between-Subject Design wird verwendet, um diese Hypothesen zu prüfen. Die Ergebnisse zeigen, dass die explizite Form von Hatespeech keine signifikanten Unterschiede zur untersuchten Variablen im Vergleich zur humoristischen Form aufweist. Dies könnte darauf hindeuten, dass auch humorvolle Hassreden als Verstoß gegen Diskursnormen wahrgenommen werden und sich nicht von expliziten Formen unterscheidet. Insgesamt liefert diese Studie Erkenntnisse über die Dynamik von Hassrede, insbesondere im Hinblick auf implizite, humorvolle Formen, und trägt zur Weiterentwicklung von Strategien zur Bekämpfung von Diskriminierung im digitalen Raum bei.

Inhaltsverzeichnis

Humor als Schutzschild: Eine quantitative Studie zur Rezeption misogyner Hatespeech mit humoristischen Elementen

1. Einleitung

"*Violence against women happens anywhere, there is no safe place, not even at home. On the contrary. Woman are targeted at home as well as in their workplace, in schools and universities, on the street, in displacement and migration, and increasingly online through cyber violence and hate speech.*" (European Comision, 2019)

Die Kommunikationslandschaft hat sich aufgrund der digitalen Ära und der damit einhergehenden Verbreitung von Social-Media-Plattformen tiefgreifend verändert. Die Plattformen bergen das Potenzial, weltweite Vernetzung zu fördern und Informationen in Echtzeit auszutauschen. Sie können als Raum für offene Diskussionen und den Austausch von Ideen dienen, aber auch als Katalysator für Diskriminierung fungieren. Das Phänomen der Hatespeech (dt. Hassrede) stellt dabei eine Gefahr dar. Laut einer aktuellen Studie der EU-Agentur für Grundrechte (FRA, 2023) sind vor allem weiblich gelesene Personen Ziel von feindlichen und diskriminierenden Äußerungen. Jedoch tragen die Auswirkungen nicht nur die Betroffenen selbst. Hassrede kann gesellschaftliche Veränderungen fördern und einen Einfluss auf Normen und die öffentliche Meinung der Gesellschaft haben (Schmid, 2023). Mitlesende sind ebenso in den Prozess von Hassrede involviert (Zochniak et al., 2023; Hsueh et al., 2015; Schäfer et al., 2022). Die Erkenntnisse machen deutlich, dass eine Notwendigkeit besteht, misogyn motivierte Hatespeech systematisch zu untersuchen.

Innerhalb des Spektrums der Hassrede stellt die subtile, humoristische Form eine besonders komplexe Herausforderung dar. Während explizite Hatespeech eindeutig als diskriminierend gilt, können implizite Formen ein ähnliches Wirkungspotenzial auf Rezipient:innen haben. (Schmid, 2023; Schmid et al., 2024). Hassrede mit humorvollen Elementen treten als scheinbar witzige oder satirische Kommentare auf. Diese können subtil

diskriminierendes Gedankengut verbreiten (Matamoros-Fernández et al., 2023; Munn, 2019). Durch die Verschleierung der Hatespeech durch humorvolle Elemente, ist die Hassrede dabei schwer zu identifizieren und zu bekämpfen (Matamoros-Fernández et al., 2023).

Durch eine unzureichende Forschung des Phänomens der humoristischer Hassrede im Kontext von Misogynie, ergibt sich eine relevante Forschungslücke. Direkte, explizite Hatespeech ist ausreichend dokumentiert, jedoch fehlen Untersuchungen, wie humoristische Hatespeech auf die Rezeption Einfluss nimmt (Wachs et al., 2021; Schmid, 2023). Diese Lücken sind alarmierend, da humoristische Hassrede nicht nur das Wohlbefinden der betroffenen Personen selbst beeinträchtigt, sondern möglicherweise auch zur Festigung von misogynen Einstellungen führt. Was zu diskriminierenden, gesellschaftlichen Normen beitragen kann (Schmid, 2023).

Die Studie zielt darauf ab, diese Forschungslücken zu schließen, indem die Rezeption von humorvoller Hatespeech in Kontrast zur expliziten, misogyn motivierten Hassrede analysiert wird. Im Einzelnen wird untersucht, ob und in welcher Weise explizite Hatespeech die emotionale Stimmung beeinflusst, ob die Form von Hatespeech eine Reaktion in Form von Abwehr, sprich Counterspeech (dt. Gegenrede), provoziert sowie misogyne Einstellungen verstärken kann. Geschlecht wird als Moderator berücksichtigt, um unterschiedliche Reaktionen der Versuchspersonen auf die beiden Formen von Hatespeech zu identifizieren. Zur Erfassung der Kausalzusammenhänge unter kontrollierten Bedingungen ist ein Online-Experiment geeignet, weshalb auf dieses Forschungsdesign zurückgegriffen wird (Koch et al., 2019). Es wird ein 2 x 3 Between-Subject Design verwendet, wobei die unabhängige Variable (UV) die Explizität der Hatespeech in den misogynen Nutzerkommentaren bildet. Die UV umfasst zwei Gruppen: die Experimentalgruppe „humoristische Hatespeech“ und die Kontrollgruppe „explizite Hatespeech“. Untersucht werden dabei drei abhängige Variablen (AV): die „Emotionale Stimmung“ zum Zeitpunkt

der Befragung, die „Bereitschaft zur Counterspeech“ sowie die „Verstärkung misogyner Einstellungen“. Bei der Anwendung des Faktors „Geschlecht“ handelt es sich um ein Moderator, um mögliche Unterschiede der Hatespeech-Effekte nach dem Geschlecht bzw. Betroffenensicht der Teilnehmer:innen analysieren zu können.

Die Studie kann dazu beitragen, die komplexe Dynamik von Hatespeech und ihre Einflüsse auf die Gesellschaft zu verstehen. Die Differenzierung verschiedener Arten von Hassrede hilft, ein besseres Verständnis dieser Phänomene zu entwickeln, wodurch präzisere Instrumente zur Bekämpfung von Diskriminierung und zur Erreichung des respektvollen und inklusiven Dialogs in den sozialen Medien entstehen können.

Der Aufbau der Arbeit folgt einer klaren Struktur, um die Forschungsfragen systematisch zu beantworten und die Ergebnisse nachvollziehbar darzustellen. Zunächst wird der theoretische Rahmen abgesteckt, um relevante Konzepte und Definitionen zu Hatespeech, insbesondere im Kontext von Misogynie in sozialen Medien, zu erläutern. Der Fokus liegt dabei auf der Unterscheidung zwischen expliziter und humorvoller Hassrede und den Auswirkungen auf Rezipient:innen. Es folgt eine Aufstellung der Forschungsfragen und eine präzise Herleitung der Hypothesen. Anschließend wird die Methodik der Studie, einschließlich der Operationalisierung und der Methodenbegründung, beschrieben. Nachdem die Rekrutierung der Proband:innen und der Versuchsablauf erläutert wurden, wird im Ergebnisteil über die Resultate berichtet, die auf der Analyse der erhobenen Daten und der Hypothesenprüfung basieren. Abschließend werden diese im Rahmen einer Diskussion interpretiert und in Beziehung zur Literatur gesetzt, wobei auf Limitationen der Studie eingegangen wird. Die Arbeit endet mit einer Conclusio, die die wichtigsten Ergebnisse zusammenfasst und die Bedeutung für zukünftige Forschungsfragen und Maßnahmen zur Vermeidung von Hatespeech hervorhebt.

2. Theoretischer Rahmen

Das folgende Kapitel erfasst den aktuellen Stand der Forschung sowie die theoretischen Konzepte zum Thema Hatespeech, Online-Misogynie und diskriminierendem Humor. Anschließend werden das Forschungsdefizit und die Relevanz der vorliegenden Studie zusammenfassend erläutert.

2.1 Das Phänomen der Hatespeech

Im Folgenden wird das Phänomen Hatespeech definiert und die Kategorien und Merkmale von Hassrede dargestellt. Darüber hinaus wird der Kontext von Social-Media-Plattformen berücksichtigt sowie die Auswirkungen von Hassrede auf die Rezipient:innen erläutert. Zudem werden Gegenmaßnahmen, insbesondere die Strategie der Counterspeech beschrieben.

2.1.1 Definitorische Annäherung des Begriffs der Hatespeech

Im Allgemeinen lässt sich Hatespeech als Diffamierung oder Verunglimpfung von Gruppen bezeichnen (Brown, 2017). Diese werden durch gemeinsame Eigenschaften und Merkmale konstituiert (Unger, 2013). Nicht jede Gruppe kann dabei zwangsläufig Ziel von Hassrede sein. Die vereinenden Merkmale korrelieren durch eine historisch benachteiligte oder diskriminierende Position innerhalb der Gesellschaft (Cohen-Almagor, 2013; Matsuda, 1989; Stone, 2000). Die Gründe für die Diskriminierung haben oft keinen rationalen Ursprung. Vielmehr resultieren sie aus diskriminierenden, gesellschaftlichen Strukturen und Systemen, die durch sozialisierende Prozesse erworben werden (Schmitt, 2017). Die kollektivistischen Eigenschaften sind vielfältig und intersektional. So können Merkmale, wie sexuelle Orientierung, Religion, Herkunft, der soziokulturelle Status oder das Geschlecht Individuen zu einer dieser Gruppe formen (Erjavec & Kovačič, 2012). Insbesondere

Angehörige von Minderheiten, vor allem aus religiösen oder ethischen Gemeinschaften, sind betroffen. Auch Frauen werden häufig Opfer von Hatespeech (Sponholz, 2018; FRA, 2023; Schmid, 2023).

Im wissenschaftlichen Diskurs wurde eine klare und abschließende Definition von Hatespeech noch nicht erreicht (Mendel, 2012). Es existieren eine Vielzahl unterschiedlicher Definitionen, welche in drei zentralen Merkmalen übereinstimmen: Hatespeech beinhaltet Kommunikation, Öffentlichkeit und Diskriminierung (Brown, 2017; Ben-David & Fernández, 2016; Sponholz, 2018). Diese Wesensmerkmale lassen sich anhand von drei grundlegenden Fragen präzise bestimmen:

Erstens: *Wer ist die Zielgruppe*? (Brown, 2017). In der vorliegenden Forschungsarbeit wird sich auf weiblich gelesene Personen fokussiert, insbesondere im Kontext misogyn motivierter Hatespeech. Die Zielgruppe umfasst demnach weibliche Personen, die durch diskriminierende und abwertende Äußerungen aufgrund ihres Geschlechts betroffen sind. Die Gründe dafür werden in *Kapitel 2.2* erläutert. Zweitens: *Welche Handlungen kennzeichnen die „communication of animosity*“? (Committee on the Elimination of Racial Discrimination, 2013). In dieser Untersuchung wird zwischen expliziter und subtiler, humorvoller Hassrede unterschieden. Explizite Hassrede bezieht sich auf klare, direkte Angriffe, während implizite, humorvolle Hassrede subtilere Formen der Diskriminierung umfassen, die durch Satire oder vermeintlich humorvolle Ausdrucksweisen vermittelt werden (Schmid, 2023). Eine klare Abgrenzung wird im Folgenden noch weiter erläutert. Drittens: *Wo finden diese Handlungen statt*? (Delgado & Stefancic, 2004; Matsuda, 1989). Der Fokus liegt im Weiteren auf Social-Media-Plattformen.

Im Folgenden soll Hatespeech anlehnend an Schmid (2023) und Hawdon (2017) zur Abgrenzung von anderen Konzepten als jegliche Äußerung von Hass oder herabwürdigende

Haltung gegenüber sozialen Gruppen verstanden werden (Hawdon et al., 2016). Diese kann sich in verbalen oder visuellen Elementen von Hass manifestieren (Rieger et al., 2021). In dieser Studie wird sich auf verbale Ausdrücke von Hass beschränkt, da dies für die Forschungslogik eines Online-Experiments zugänglicher ist.

Hatespeech wird in dieser Studie, um das Phänomen von anderen negativen Äußerungen wie Unhöflichkeit und Cyber-Mobbing zu unterscheiden, anhand von vier Merkmalen definiert (Pondorf, 2024): (1) Die am häufigsten genannte Eigenschaft in Definitionen von Hassrede betrifft ein spezifisches Ziel (Bilewicz & Soral, 2020). Hatespeech ist stets ein diskriminierender Ausdruck für eine soziale Gruppe als Ganzes (Rieger et al., 2021). Jede Eigenschaft kann dabei ein Auslöser für Diskriminierung sein, jedoch sind die häufigsten Kategorien Ethnizität, Nationalität, Geschlecht, Religion, sexuelle Orientierung oder Behinderungen (ElSherief et al., 2018; Silva et al., 2016). Zudem (2) ist Hatespeech stets an Einzelpersonen oder soziale Gruppen gerichtet, die der bzw. die Angreifende nicht persönlich kennt, wodurch es sich beispielsweise zu anderen Phänomenen wie Cybermobbing abgrenzt (Obermaier et al., 2018). Der Zweck (3) von Hassrede dient zur Schädigung der Angegriffenen oder zur Unterdrückung der sozialen Gruppe, die diskriminiert wird (Guo et al., 2020). Darüber hinaus kann die Intensität (4) des Angriffs variieren. Nicht nur schwerwiegende Beleidigung und Aufforderung zu physischer Gewalt sind diskriminierende Ausdrucksformen von Hatespeech. Auch die Wiederholung von Stereotypen kann dem Phänomen zugeordnet werden. Es existiert eine Vielzahl von Formen (Sellars, 2016).

Aus kommunikationswissenschaftlicher Sicht wird Hassrede nicht nur als sprachlicher Ausdruck von Hass oder als ausgesprochene Diskriminierung betrachtet, wie es häufig in der linguistischen Analyse der Fall ist (Meibauer, 2013; Marx, 2018). Vielmehr wird Hatespeech als ein komplexes Phänomen verstanden, das in einem spezifischen Handlungskontext

entsteht. Dabei sind in diesem Kontext nicht nur die sprechende Person, sondern auch die Zuhörenden bzw. das Publikum, die vermittelten Botschaften und das genutzte Medium involviert (Ouirdi et al., 2014). Bei Hatespeech handelt es sich im Wesentlichen um eine kommunikative Herstellung von Diskriminierung durch Personen. Dabei werden bewusst oder absichtlich Gegensätze geschaffen, die verschiedene Menschengruppen als ungleichwertig und sich gegenseitig ausschließend darstellen (Sponholz, 2018). Diese Gegensätze entstehen nicht erst durch Hassrede, sondern sind bereits im gesellschaftlichen Denken verankert. Sie sind latent vorhanden und werden durch Kommunikation – in diesem Fall durch Hatespeech – aktiviert und verstärkt (Bergmann, 1997). Der Unterschied zur linguistischen Definition wird besonders bei mediatisierten Formen wie der Online-Hassrede deutlich. In einer Online-Umgebung beeinflussen Medien die Hatespeech durch ihre eigene Logik. Diese Medienlogik umfasst Normen, Strategien, Mechanismen und Prozesse, welche die Dynamik eines Mediums prägen und im Folgenden ausführlich erläutert werden (Klinger & Svensson, 2015; van Dijck & Poell, 2013). Hatespeech kann daher nicht am Inhalt oder der Struktur erkannt werden. Die betroffene Gruppe muss nicht zwingend beschimpft oder explizit angesprochen werden.

Diese Perspektive ist vor allem für die vorliegende Forschung von besonderer Bedeutung, da sie die Notwendigkeit hervorhebt, auch implizite und subtile Formen von Hassrede zu erfassen und zu analysieren, um deren Auswirkungen auf die Rezipient:innen adäquat zu untersuchen.

2.1.2 Elemente und Kategorisierung von Hatespeech

Die Identifizierung und Unterscheidung verschiedener Formen von Hassrede erfordern eine eingehende Analyse ihrer Strategien und Merkmale. Meibauer (2013) betont, dass Hatespeech häufig durch Verschleierungsstrategien verbreitet wird, die ihre Erkennung

erschweren. Dabei unterscheidet er fünf Strategien und einhergehende Elemente, mit denen Hassrede verbreitet werden kann: Direktheit, Offenheit, Intensität, Autorität und Gewalt (Meibauer, 2013). Anstatt direkter Beleidigungen werden Vorurteile oder verallgemeinernde Aussagen genutzt, die denselben Effekt haben (Direktheit). Häufig wird Hatespeech nicht als solche gelesen, sondern erscheint verschleiert in Kommentaren oder Diskussionen, zum Beispiel in Medien oder als vermeintliches Bildungsangebot (Borgeson & Valeri, 2004) (Offenheit). Auch autoritäre Institutionen wie der Staat oder größere Bevölkerungsgruppen können Hassrede verbreiten, ohne dass diese als solche wahrgenommen werden (Autorität). Dabei werden oft nostalgische Rückblicke genutzt, etwa auf die deutsche Teilung oder das Wirtschaftswunder (Ben-David & Matamoros-Fernández, 2016). Auch Gewalt als Begleiterscheinung von Hatespeech wird selten als solche benannt (Gewalt). Ein weiteres Phänomen ist die satirische oder humoristische Darstellung von Hass, etwa in Form von Memes, die den Ernst der Hatespeech verschleiern und sie harmloser erscheinen lassen, als sie tatsächlich ist (Intensität).

Diese Verschleierungsstrategien machen deutlich, dass zwischen expliziter und impliziter Hatespeech unterschieden werden muss. Während explizite Hassrede direkte Angriffe und Drohungen umfassen, sind verdeckte Formen subtiler und können beispielsweise durch humoristische Inhalte oder Fake News soziale Gruppen marginalisieren (Hajok & Selg, 2018). Oftmals taucht Hassrede in Form der impliziten Weise auf, da dieser Inhalt nicht rechtswidrig ist und somit nicht eingeschränkt werden kann (Paasch-Colberg et al., 2021; Rieger et al., 2021). Gerade bei dieser Form ist das Erkennen und die Moderation von Hatespeech schwierig. Algorithmen scheitern oft an der Identifikation solcher Inhalte (Ben-David & Matamoros-Fernández, 2016). Dies gibt den Verfasser:innen die Möglichkeit, ihre Ideologien unbemerkt zu verbreiten und soziale Gruppen herabzusetzen.

Zur unterschiedlichen Kategorisierung von Hatespeech orientiert sich vorliegende Arbeit an der Studie der EU-Agentur für Grundrechte (FRA, 2023). Die Einteilung der Hassrede in fünf Hauptkategorien ist insbesondere relevant, da sie die Identifikation von expliziter und impliziter Hatespeech ermöglicht.

Aufforderung zu Gewalt, Diskriminierung oder Hass stellt die erste Kategorie dar. Diese umfasst Äußerungen, die direkt zu konkreten Handlungen aufrufen, beispielweise zur Vergewaltigung oder Mord. Diese Form wird im Folgenden als explizite Hatespeech betrachtet, da sie klare, direkte Angriffe oder Drohungen enthält. Diese unterscheidet sich von der zweiten Kategorie, *Herabwürdigung*, die sich auf die Diffamierung von Fähigkeiten, dem Charakter oder Ruf von Personen im Zusammenhang mit deren (wahrgenommener) Mitgliedschaft in einer Gruppe bezieht. Hierzu zählen objektivierende und entmenschlichende Sprache. Auch diese Form wird in der vorliegenden Studie als ausschließlich explizite Hatespeech betrachtet, da dabei direkt und unverschleiert verletzende Ausdrücke gegen bestimmte Gruppen geäußert werden. Die dritte Kategorie umfasst verletzende, abwertende und/oder obszöne Sprache, die sich auf eine bestimmte Eigenschaft einer Personengruppe bezieht und die als *Beleidigende Sprache* zusammengefasst werden kann. Dabei ist dies kontextabhängig gemeint und nicht konkret einordbar, da die Einschätzung sich, ob eine Äußerung als beleidigend empfunden wird, je nach Rezipient:in unterscheidet. Im Folgenden können diese Äußerungen als explizite sowie implizite Hassrede klassifiziert werden. Dabei ist vor allem der vorliegende Kontext entscheidend sowie die rezipierende Person. Unter der vierten Kategorie, *Negative Stereotypisierung*, versteht man die Zuweisung negativer Eigenschaften und Merkmale zu einer sozialen Gruppe. Auch diese Form der Hassrede kann sowohl explizit als auch subtil erfolgen, jedoch gilt sie für vorliegende Studie eher für die implizite Form, wie die Literatur zeigt (Schmid, 2023). Als *Andere hasserfüllte Inhalte* kann die letzte Kategorie zusammengefasst werden. Diese

umfasst Unterstützung für hasserfüllte Ideologien oder Holocaustleugnung. Durch die Konzentration auf misogyne Hatespeech, ist diese Kategorie im Nachfolgenden nicht vertreten.

Zu beachten ist, dass sich die Kategorien nicht zwangsläufig gegenseitig ausschließen (FRA, 2023). Ein Ausdruck von Online-Hass kann in mehrere Kategorien gleichzeitig fallen. Es ist außerdem wichtig zu berücksichtigen, dass die Einordnung von Beiträgen anhand dieser Merkmale stets eine subjektive Bewertung darstellt. So kann es vorkommen, dass Personen Kommentare als unproblematisch einstufen, während das betroffene Opfer die darin verwendete Sprache als verletzend empfindet (FRA, 2023). Die Kategorisierung zeigt dabei aber, eine Differenzierung in verschiedene Stufen der Explizität. Die Unterscheidung zwischen expliziter und impliziter Hatespeech, wie sie durch Meibauer (2013) und die EU-Studie (FRA, 2023) hervorgehoben wird, erlaubt es, sowohl offensichtliche als auch subtile Formen der Hassrede zu erkennen. Schmid et al. (2024) unterscheiden ebenso explizite Formen von Hatespeech wie etwa die Androhung von Gewalt von impliziten Formen, die sich beispielsweise in Stereotypisierungen ausdrücken. Der Grad an Feindseligkeit kann zu einer unterschiedlichen Wahrnehmung des Inhalts bei den Rezipierenden führen. Deshalb untersucht die vorliegende Studie, wie solche „humorvollen“ Elemente Angriffe auf bestimmte Gruppen verbergen und analysiert die Auswirkungen dieser impliziten Formen von Hassrede auf die Reaktionen der Rezipient:innen.

2.1.3 Hatespeech als Social-Media-Phänomen

Hatespeech tritt zwar in sämtlichen Interneträumen auf, jedoch sind vor allem Social-Media-Plattformen durch die innehabenden Mechanismen und Merkmale besonders von dem komplexen und multifaktoriellen Problem betroffen (Wachs & Wright, 2018).

Speziell der Online-Disinhibitionseffekt erklärt, warum Hatespeech oftmals in sozialen Medien auftritt. Dieser beschreibt, welche Bedingungen und Ursachen bei Nutzer:innen eher zu enthemmten und aggressiven Verhalten in digitalen Räumen führen, als dies im interpersonellen Kontakt der Fall ist (Wachs & Wright, 2018). Diese Annahme trifft insbesondere auf Social-Media-Plattformen zu, die durch spezifische Strukturen und funktionale Merkmale charakterisiert sind (Matamoros-Fernández & Farkas, 2021; Zhang & Luo, 2019). Die technische Infrastruktur, die eine global umfassende Kommunikation fördert, begünstigt ebenfalls die Bildung und Ausbreitung von Hassgruppen. Diese sogenannten Hass-Cluster profitieren von der globalen Vernetzung und können Hatespeech zirkulieren lassen (Johnson et al., 2019). Zusätzlich tragen Algorithmen und funktionale Merkmale, wie die Kommentarfunktion oder das Teilen von Inhalten, zu diesem Prozess bei, indem sie bestimmte, diskriminierende Beiträge bzw. Kommentare besonders hervorheben und dadurch unbeteiligte Nutzer:innen mit Hassrede in Kontakt kommen (Matamoros-Fernández, 2017; Schulze et al., 2022). Des Weiteren trägt die Anonymität auf Social-Media-Plattformen zur Enthemmung bei. Das anonyme Auftreten und die Unnahbarkeit der Kommunikationspartner:innen senken die Hemmschwelle für diskriminierendes Verhalten. Studien zeigen, dass Menschen mit Identitätsverschleierung und fehlender physischer Gegenwart eher beleidigend und hasserfüllt sind (Suler, 2004; Brown, 2018). Da die Möglichkeit besteht, mehrere Konten zu erstellen und fiktive Benutzernamen zu verwenden, wird die Anonymität weiter verstärkt. User:innen verhalten sich durch die Unsichtbarkeit auf Social-Media-Plattformen oftmals provokanter und aggressiver (Suler, 2004). Die Logik der Plattformen selbst ermöglicht jedoch nicht nur diskriminierende Diskurse, sie erhöht diese auch aktiv. Die spezielle Struktur der Plattformen, einschließlich der algorithmischen Elemente, trägt entscheidend zur Sichtbarkeit und Verbreitung von Hassrede bei (Matamoros-Fernández, 2017). So wird Hatespeech, wie Maarouf et al. (2024) in einer

aktuellen Untersuchung zeigen, nicht nur stärker, sondern auch schneller verbreitet als neutrale oder positive Inhalte.

Zudem kann der Negativitätseffekt, die Verbreitung von Hassrede auf Social-Media-Plattformen erklären (Maarouf et al., 2024). Dieser beschreibt die Tendenz, dass negative Emotionen, wie sie in Hassrede transportiert werden, mehr Reaktionen hervorrufen als bei positiven Inhalten (Maarouf et al., 2024). Als starke negative Emotion besitzt Hass das Potenzial, vermehrte Reaktionen hervorzurufen, wodurch die Interaktionsraten auf Social-Media-Plattformen steigen, und die weitere Verbreitung begünstigt wird.

Durch die genannten spezifischen Ursachen konzentriert sich vorliegende Studie auf Hatespeech in sozialen Medien, da die Rahmenbedingungen eine vertiefte Analyse im Kontext von Social-Media-Plattformen als unerlässlich darstellen. Die Präventionsarbeit ist vor allem dort vorzunehmen und erfordert weitere Betrachtungen des Phänomens in diesem Umfeld.

2.1.4 Wahrnehmung von Hatespeech aus Rezipierendensicht

Die Wahrnehmung von Inhalten auf Social-Media-Plattformen, wird von einer Vielzahl von kontextuellen Faktoren beeinflusst. Analog zu anderen Medienformen kann die Nutzung nicht universell beschrieben werden, sondern hängt maßgeblich von individuellen Eigenschaften der Rezipierenden sowie der Art der Nachrichtenpräsentation und des spezifischen Inhaltes ab (Schmid et al., 2024)

Ohme und Mothes (2020) differenzieren zwischen zwei Ebenen der Wahrnehmung von Hassrede, indem sie die selektive Exposition in sozialen Medien systematisch analysieren. Die erste Ebene bildet die Aufmerksamkeit der Nutzer:innen für einzelne Beiträge mit Hassrede - während des Durchscrollens der Social-Media-Feeds ab. Diese umfasst die Erkennung von Hatespeech und die Entscheidung, das Durchscrollen zu

verlangsamen oder anzuhalten, um bestimmte Beiträge bzw. Textpassagen genauer zu betrachten (Ohme & Mothes, 2020). Die zweite Ebene beinhaltet eine detaillierte Analyse der Gefühle, Einstellungen und Meinungen der Nutzer:innen im Hinblick auf den ihnen begegneten Inhalte von Hassrede (Ohme & Mothes, 2020). Die vorliegende Arbeit analysiert ausschließlich die zweite Wahrnehmungsebene

Die Studie von Schmid et al. (2024) zeigt, dass strafrechtlich relevante Hassrede schwerwiegender wahrgenommen wird als Hatespeech, die nicht strafrechtlich verfolgt wird. Das bestätigt frühere Forschungsarbeiten. Diese zeigen, dass direkte und explizite Hassrede mit beleidigender Sprache oder Gewaltandrohungen als besonders hasserfüllt empfunden wird (Leonhard et al., 2018; Kenski et al., 2020; Stryker et al., 2016). Schmid et al. (2024) können dabei aufgrund ihrer methodischen Ausrichtung nicht abschließend klären, ob indirekte Hassrede eventuell effektiver wird, wenn sie mit hoher Frequenz auftritt, wie es von Paasch-Colberg et al. (2021) vorgeschlagen wird.

In der qualitativen Mixed-Methods-Studie bewerten einige Teilnehmer:innen die hasserfüllten Intentionen der impliziten Hatespeech als besonders negativ, nachdem sie diese identifiziert haben, und führen eine Reflexion über deren subtile Auswirkungen durch (Schmid et al., 2024). Darüber hinaus kommen andere Forschungen zu dem Ergebnis, dass das Publikum antagonistische Stereotypen ähnlich unsachlich und schädlich wahrnimmt, wie direkte hasserfüllte Äußerungen (Weber et al., 2020). So werden in einigen Kontexten die Äußerungen von impliziter Hatespeech sogar eher als Hassrede eingestuft als explizite (Benikova et al., 2018).

Da sich die Studie mit misogyner Hatespeech befasst, sind geschlechterspezifischen Unterschiede in der Wahrnehmung von Hatespeech relevant: Die Studie von Kenski et al. (2017) zeigt, dass Frauen Aussagen tendenziell häufiger als unhöflich bewerten. Studien

belegen, dass sich geschlechtsspezifische Unterschiede bereits im jungen Alter manifestieren (Perry & Weiss, 1989). Zudem geben weiblich gelesene Personen tendenziell weniger negative Kommentare in Online-Foren ab (Moss-Racusin et al., 2015), was möglicherweise auf eine erhöhte Sensibilität für Unhöflichkeit zurückzuführen ist. Diese Ergebnisse verdeutlichen, dass geschlechtsspezifische Normen und somit das Geschlecht eine Rolle dabei spielen, wie Unhöflichkeit ausgedrückt und empfunden wird.

Insgesamt zeigen widersprüchliche Forschungsergebnisse über die Wahrnehmung verschiedener Formen von Hassrede und deren Relevanz hinsichtlich des Meinungsklimas die Komplexität dieses Themas auf. Offene Fragen bleiben bestehen, insbesondere in Bezug auf die Wirkung indirekter Hassrede im Vergleich zu expliziten Formen von Hatespeech.

2.1.5 Auswirkungen von Hatespeech auf die Stimmung und Einstellung der Rezipierenden

Online-Kommunikation in unserer Gesellschaft spielt bei der Gestaltung des Meinungsklimas eine maßgebliche Rolle (Ben-David & Matamoros-Fernández, 2016). Die generelle Stimmungslage wird auch durch Hatespeech beeinflusst. Dabei sind Social-Media-Plattformen besonders geeignet, Räume für Personen mit radikalen Ansichten zu bieten (Costello et al., 2016; Ben-David & Matamoros-Fernández, 2016). Diese Personen und Gruppen sind meistens dem politisch rechten Spektrum zuzuordnen, religiös extremer angesiedelt und propagieren dabei rassistische, nationalistische sowie regierungskritische oder kulturfeindliche Hassrede. Außenstehende Personen, die bereits zu solchen Ansichten tendieren, sind durch Hatespeech anfällig, weiter beeinflusst zu werden (Costello et al., 2016).

Des Weiteren können Online-Gruppen, die extreme Ansichten teilen, ein homogenes Meinungsklima schaffen, in dem Hassrede nicht als störend empfunden wird, sondern als „normal“ gilt und den eigenen Ansichten entspricht. In solchen Umgebungen kann

extremistisches Gedankengut gedeihen und langfristig legitimiert werden (Ben-David & Matamoros-Fernández, 2016). Mit dem Narrativ *„Wir gegen die Anderen*“ kommt es zu einer Identifikation mit der eigenen Gruppe und führt parallel zur Ablehnung von Anderen (Leets, 2001). Dabei kann es zu einer ideologischen Assimilation führen und in diskriminierendem Verhalten oder sogar in Gewalttaten enden (Costello et al., 2016).

Hatespeech fördert nicht nur extreme Einstellungen, sondern wirkt sich auch auf die Meinungen der Rezipierenden aus. In einem homogenen Meinungsklima sind Personen weniger bereit wiedersprechende Ansichten zu äußern. Sie passen sich, um Konflikte zu vermeiden, eher der vorherrschenden Meinung an (Noelle-Neumann, 1974). Individuelle Einstellungen können durch das Umfeld von Social-Media-Plattformen beeinflusst werden. Vor allem durch die starke Verbreitung der Meinung einzelner Individuen, erscheinen diese als gesellschaftliche Norm. Kommentare und Interaktionen erzeugen ein Bild von vorherrschenden Einstellungen, wodurch Hatespeech fälschlicherweise als die Meinung der Mehrheit verstanden werden kann (Kreißel et al., 2018).

Hassrede kann somit Auswirkungen auf die individuelle Ebene von Rezeption haben. Dazu zählen Einstellungs- und Verhaltensänderungen, kognitive Prozesse wie Wissen und Überzeugungen sowie emotionale Reaktionen, welche als sekundäre Wahrnehmungsebene angesehen werden kann (Kümpel & Rieger, 2019; Ohme & Mothes, 2020). Müller und Lopez-Sanchez (2021) zeigen, dass die Konfrontation mit Hass Stress auslösen kann. Rezipierende sind dabei nicht gleichermaßen betroffen. Die individuelle, psychische Belastbarkeit ist ein relevanter Faktor. Personen mit niedriger Belastbarkeit sind anfälliger für emotionale Instabilität und empfinden den Stress stärker als jene mit höherer.

Weitere Studien legen nahe, dass Hatespeech emotionale und psychische Beeinträchtigungen verursachen (Leets & Giles, 1997). Hassrede kann Schmerzgefühle bei

den Angegriffenen auslösen (Eisenberger, 2015), wobei die Lebensqualität für Opfer dadurch gemindert werden kann (Vedeler et al., 2019), da ihnen häufiger Krankheiten wie Depressionen und posttraumatische Belastungsstörungen diagnostiziert werden (Wypych & Bilewicz, 2021). Die Betroffenen unternehmen häufiger Suizidversuche als jene die keine Opfer von Hatespeech sind (Marshall et al., 2011).

Zochniak et al. (2023) stellen fest, dass das Wohlbefinden von Personen aus der LGBT+ Gemeinschaft durch die Konfrontation mit homophober Hatespeech erheblich beeinträchtigt wird. Mit zwei Studien wiesen sie konsistent nach, dass der Kontakt mit homophober Hassrede sowohl das momentane affektive Wohlbefinden als auch die generelle Lebenszufriedenheit mindert. Personen, sie sich besonders stark mit der LGBT + Gemeinschaft identifizieren, sind dabei stärker davon betroffen. Das Ergebnis zeigt, dass Hatespeech insbesondere die Personen belastet, die bereits Ziel von Hatespeech geworden sind und sich mit der betroffenen Gruppe identifizieren.

Für die Bildung von stereotypischen Einstellungen spielen Medienabbildung eine relevante Rolle (Ward et al., 2020; Schemer, 2012). Auch hasserfüllte Benutzerkommentare und Hatespeech sind hiermit inbegriffen. Die Forschung zeigt, dass Personen durch Konfrontation mit Hassrede expliziten, negativen Stereotypen vermehrt zustimmen (Schäfer et al., 2022; Weber et al., 2020; Soral et al., 2018). Die Interaktion mit misogynen Online-Inhalten kann sexistische Einstellungen verstärken (Fox et al., 2015). Williams et al. (2016) zeigen, dass die Sensibilität von Personen, die zuvor Hatespeech erlebt haben, gegenüber ähnlichen Inhalten gesteigert wird. Betroffene Personen empfanden die Hatespeech zudem als beleidigender, im Vergleich zu jenen, die bisher nicht davon betroffen waren (Pondorf, 2024).

Diese Ergebnisse zeigen, dass die Rezeption von Hatespeech einen unmittelbaren Einfluss auf die Nutzer:innen von Social-Media-Plattformen hat. Dabei ist die Forschung zu verschiedenen Formen von Hatespeech noch nicht ausreichend vorhanden.

2.1.6 Counterspeech als Strategie zur Bekämpfung von Hatespeech

Aufgrund der Auswirkungen von Hatespeech auf Betroffene und Rezipierende wird zunehmend über geeignete Maßnahmen zur Bekämpfung diskutiert. Eine rechtliche Regulierung könnte ein Ansatzpunkt sein, um Hassrede einzudämmen. In Österreich wurde dafür das *„Gesetz zum Schutz der Nutzer auf Kommunikationsplattformen im Internet“* eingeführt (Republik Österreich, 2020). Inhaltlich wird die Verantwortung für das Vorgehen gegen Hatespeech vor allem den Betreibenden der Social-Media-Plattformen überlassen. Dabei liegt das Problem darin, dass Inhalte erst nach einer Meldung durch andere Nutzer:innen überprüft werden und im Anschluss gegebenenfalls entfernt werden. Somit liegt die Regulierung von Content bei kommerziellen Unternehmen. Dadurch besteht eine Gefahr für die Meinungsfreiheit, da Inhalte durch kapitalistische Plattformen und nicht etwa durch demokratische Institutionen kontrolliert werden (Alkiviadou, 2019; Howard, 2019).

Eine weitere Strategie zur Bekämpfung von Hatespeech ist das Konzept der Counterspeech. Diese reagiert unmittelbar auf Hassreden und kann schädliche Auswirkungen mindern (Bahador, 2021). Sie stellt eine bedeutende Form der Intervention gegen Hatespeech dar und fungiert als kollektive Reaktion auf hasserfüllte Inhalte. Frühere Untersuchungen zu Counterspeech haben vor allem die Beweggründe von Internetnutzer:innen beleuchtet, sich an dieser Form der Gegenrede zu beteiligen (Kalch & Naab, 2017; Leonhard et al., 2018; Ziegele et al., 2019; Ping et al., 2024).

Hinter Counterspeech steht theoretisch betrachtet das Medienpriming. Negative Stereotype und Eigenschaften sind durch das Konzept zwar nicht verhinderbar, jedoch

können Effekte abgeschwächt werden, indem auch positive Eigenschaften der angegriffenen Gruppe hervorgehoben werden (Schäfer et al., 2023). So zeigen Studien, dass Gegenrede zu einer Veränderung der Wahrnehmung innerhalb von Online-Diskussionen führen kann (Obermaier et al., 2023; Schieb & Preuss, 2018). Unbeteiligte Zuschauer:innen können durch das Äußern gegenteiliger Meinungen dazu beitragen, negative Folgen für die Betroffenen zu mildern (Krasodomski-Jones et al., 2023). Obwohl es nicht immer gelingt, die Täter:innen zu überzeugen oder zum Schweigen zu bringen, trägt Counterspeech das Potenzial, Diskursnormen zu prägen. Dabei wird unbeteiligten Nutzer:innen gezeigt, dass Hassrede nicht die Meinung der Mehrheit ist und sie können, ermutigt werden ebenfalls einzugreifen. Das Phänomen kann zu einem breiteren Spektrum an Gegenargumenten führen, was als effektive Bekämpfungsstrategie von Hassrede angesehen wird (Delgado et al., 2014). Obermaier et al. (2023) erweitern die Forschung zur Bystander-Intervention (basierend auf dem Bystander-Interventionsmodell) im Kontext von Hatespeech. Dieses Modell der Zuschauerintervention (Latané & Darley, 1970) erklärt das Eingreifen von Beobachter:innen in Situationen bei Online-Hass. Die Zuschauerintervention setzt voraus, dass Zuschauer:innen die Situation selbst als bedrohlich wahrnehmen und dadurch ein Gefühl von persönlicher Verantwortung entwickeln. Dies entscheidet, ob sie eingreifen oder nicht. Das Modell wurde bereits vermehrt in der Forschung zu Reaktionen einschließlich der Counterspeech im Kontext der Hatespeech herangezogen (Obermaier et al., 2016; Ziegele et al., 2020; Leonhard et al., 2018). Obermaier et al. (2023) zeigen, dass Hassrede, die als unhöflich und diskriminierend eingestuft wird, den Bystander-Interventionsprozess auslösen kann. Im Einklang mit Leonhard et al. (2018) intervenieren Bystander nicht automatisch nach dem Lesen von Hassrede. Vielmehr steigt das persönliche Verantwortungsgefühl zum Intervenieren, je nachdem, wie stark das Gelesene als Angriff empfunden wird. Diese Verantwortung führt zu Interventionen durch Gegenargumente oder die Ermutigung anderer

zur Intervention. Selbst Hatespeech, die gesetzlich nicht verfolgt, aber als unhöflich wahrgenommen wird, kann das Verantwortungsgefühl und die Intention zum Eingreifen wecken. Also kann die Wahrnehmung, dass Diskursnormen in einer demokratischen Gesellschaft verletzt werden, bereits ausreichen, um den Counterspeech-Prozess zu initiieren. Insbesondere misogyne Hatespeech löst dabei ein hohes Verantwortungsbewusstsein aus und somit auch den Interventionsprozess (Obermaier et al., 2023). Ping et al. (2024) belegen, dass frühere Opfererfahrungen einen entscheidenden Prädiktor für die Teilnahme an Counterspeech darstellen. Personen, die bereits Ziel von Hatespeech waren, beteiligten sich signifikant häufiger an einer Form der Gegenrede. Dies wird durch die Beobachtung unterstützt, dass Opfer verstärkt motiviert sind, hasserfülltes Verhalten direkt zu konfrontieren, während diese Motivation bei Nicht-Opfern weniger ausgeprägt ist.

Die aktuelle Forschung trifft keine Unterscheidung zwischen impliziter und expliziter Hatespeech, noch wurde der Grad der Beteiligung der Rezipierenden berücksichtigt. Diese Lücke wird in der vorliegenden Studie geschlossen.

2.2 Misogynie auf Social-Media-Plattformen

Misogynie im Online-Kontext stellt ein zunehmendes Problem dar, welches weiblich gelesene Personen, unabhängig von ihrem demografischen oder geografischen Hintergrund, betreffen. Dabei kann diese geschlechtsspezifische Gewalt von Online-Misogynie bis zu Vergewaltigungspornografie, Cyberstalking und Cyberbelästigung sowie textbasierte Übergriffe reichen (Barker & Jurasz, 2018). Dies ist kein neues Phänomen. Vielmehr sind diese Gewaltformen im digitalen Raum eine Fortführung bereits bestehender Ungleichheiten, die bereits offline zu finden sind. Die Ursachen, die der Online-Diskriminierung vorausgehen, unterscheiden sich nicht wesentlich von dem Offline-Kontext. Beide Formen sind tief in den bereits bestehenden ungleichen Geschlechterverhältnissen, patriarchalen

Strukturen sowie der gesellschaftlichen Sozialisierung verwurzelt (Barker & Jurasz, 2018). Sexismus, der auf Social-Media-Plattformen erlebt wird, existiert somit nicht isoliert von realen Interaktionen. Tatsächlich gibt es hohe Raten von Cyber-Aggressionen gegenüber Frauen durch ihre romantischen Partner (Martinez-Pecino & Durán, 2019).

Im Allgemeinen kann Sexismus als Vorurteil bzw. Diskriminierung gegenüber einer Person oder einer Gruppe aufgrund ihres Geschlechts bezeichnet werden. Er kann von Individuen, Gruppen oder Institutionen ausgeübt werden (Handbook Of Prejudice, Stereotyping And Discrimination, 2009). Sexismus kann sowohl in feindlicher Form wie beispielsweise in Form einer Beleidigung als auch in der sogenannten wohlwollenden Form, etwa der Überprotektion von Frauen, auftreten (Glick & Fiske, 1996). Beispiele für wohlwollender Sexismus sind Kommentare wie *„Kein Mann hat Erfolg ohne eine gute Frau an seiner Seite“* und *„Sie sind wahrscheinlich überrascht, wie klug Sie sind, für ein Mädchen“* (Jha & Mamidi, 2017). Die Begrifflichkeiten der Geschlechtervoreingenommenheit und des Sexismus lassen offen, welches Geschlecht diskriminiert wird. Jedoch zeigen Forschungen, dass weiblich gelesene Personen sowohl offline (Nielsen, 2002) als auch online häufiger Opfer von sexistischer Voreingenommenheit sind (Wojatzki et al., 2018; Döring & Mohseni, 2018; Mohseni, 2021; FRA, 2023; Pondorf, 2024). Da die vorliegende Studie lediglich Hatespeech gegenüber weiblich gelesenen Personen untersucht, wird der Begriff der misogynen Hassrede gegenüber anderen Begriffen bevorzugt.

Eine Studie der European Union Agency for Fundamental Rights (FRA, 2023) kam zu dem Ergebnis, dass Misogynie die vorherrschende Form von Hatespeech auf Social-Media-Plattformen ist. Die Untersuchung zeigt, dass Hassreden, die Frauen als Ziel haben, dreimal so oft vorkommt wie rassistische Hatespeech. Dabei dienen vor allem herabwürdigende Sprache, Objektifizierung sowie Entmenschlichung als Mittel, um Frauen

zu diskriminieren. Zudem weisen die Beiträge in der Studie höhere Niveaus an Gewaltaufrufen gegen Frauen im Vergleich zu anderen Gruppen auf, wobei Gewalt gegen Frauen online oft auf sexualisierter Gewalt basiert. Umfragedaten aus den USA zeigen einen kritischen Trend: Zwischen 2017 und 2020 hat sich der Prozentsatz der Frauen, die von sexueller Belästigung online berichteten, verdoppelt und war dreimal so hoch wie der von Männern (Vogels, 2021).

Da weiblich gelesene Personen häufiger von Hatespeech betroffen sind, erleben sie auch deren Konsequenzen verstärkt. Forschungen haben ergeben, dass betroffene Personen Social-Media-Plattformen eher meiden und weniger aktiv sind (Molyneaux et al., 2008). Opfer von Hassreden erleben vermehrt Gefühle wie Wut, Schmerz und Depressionen. In extremen Fällen kann der Hass im Online-Umfeld sogar zu suizidalen Gedanken bzw. durchgeführten Selbstmorden führen (Citron, 2009). Darüber hinaus verringert es die Teilnahme von Frauen am Online-Diskurs. Dies kann ein Hindernis für die Geschlechtergleichstellung darstellen (Amnesty Global Insights, 2017; Im et al., 2022; Vitak et al., 2017). Als Zielpersonen sind Frauen oftmals nicht in der Lage, sich gegen Misogynie auszusprechen (Sasse et al., 2021).

In sozialen Medien sind abwertende Witze, welche die Kompetenz von Frauen infrage stellen, weit verbreitet (Drakett et al., 2018; Fox et al., 2015). Sie werden zudem in alarmierendem Maße sexualisiert (Bell et al., 2018; Davis, 2018). Auf Twitter beispielsweise werden Frauen alle dreißig Sekunden verbal missbraucht, wobei Woman of Colour etwa drei Mal häufiger in problematischen oder missbräuchlichen Tweets erwähnt werden, als weiße Frauen. Besonders Schwarze Frauen sind acht Mal häufiger Ziel von diskriminierenden oder missbräuchlichen Tweets (Amnesty International, 2017). Miranda (2023) zeigt, dass Frauen in untersuchten Beiträgen weniger durch direkte Beleidigungen, dafür aber auf ironische und lächerliche Art angegriffen werden. Insbesondere Feministinnen oder Frauen, die nicht den

gängigen Schönheitsidealen entsprechen, sind dabei Opfergruppen. Diese Diskriminierungen manifestieren sich oft durch die Hypersexualisierung von Frauenkörpern und die Verherrlichung traditioneller, untergeordneter Geschlechterrollen. Dies stützt die Annahme, dass technologisch vermittelte Praktiken bestehende hierarchische Strukturen und Privilegien reproduzieren und Phänomene wie toxische Männlichkeit fördern (Ging & Siapera, 2018). Auch Siddiqui (2008) und Eudey (2008) zeigen, dass vor allem Beiträge mit feministischen Inhalten misogyne Hasskommentare anziehen (Pondorf, 2024). Zudem bestärken die Auswertungen des Guardians von siebzig Millionen Online-Artikeln die wissenschaftlichen Erkenntnisse mit dem vermehrten Auffinden von Hasskommentaren unter Artikeln von weiblichen Journalistinnen und/oder über Feminismus oder Vergewaltigung (Gardiner, 2018; Pondorf, 2024). Auf Social-Media-Plattformen sind feindselige misogyne Inhalte zu finden, wobei etwa sieben Prozent der Tweets, die Frauen adressieren, problematisch oder missbräuchlich sind (Amnesty International, 2018).

Insgesamt ist Forschung zu misogyner Hassrede relevant, da sexistische Inhalte weit verbreitet sind und sich negativ auf Betroffene auswirkt. Misogynie auf Social-Media-Plattformen tritt dabei offen und explizit als auch subtil durch ironische, aber abwertende Hatespeech auf. Diese können unkritisch wahrgenommen werden und tragen zur Normalisierung von misogynen Diskursen bei. Diese Studie legt den Fokus deshalb speziell auf diese subtilen Formen misogyner Hassrede. So könnten die Ergebnisse aufzeigen, ob impliziten Formen von Hassrede zur Verfestigung geschlechterdiskriminierender Strukturen beitragen und einen Einfluss auf die Rezeption haben.

2.3 Kritische Betrachtung von diskriminierendem Humor

Diskriminierender Humor gegen marginalisierte Gruppen ist in der Wissenschaft ausführlich erforscht (Ford et al., 2013; Saucier et al., 2016). Hierbei kann die Theorie der

voreingenommenen Norm herangezogen werden. Dabei trägt Humor mit herabwürdigenden Elementen zur Verstärkung von Vorurteilen bei und fördert somit insgesamt diskriminierendes Verhalten (Ford & Ferguson, 2004; Ford et al., 2008). Zudem unterstützt abwertender Humor die Verfestigung von stereotypischen Überzeugungen. Saucier et al. (2018) kamen zu dem Ergebnis, dass Personen, die rassistische Witze verharmlosen, auch eher negative Stereotype gegenüber Schwarzen Personen akzeptieren. Abwertender Humor kann also Stereotype vor allem gegenüber marginalisierten Gruppen verstärken und diese somit unter dem Deckmantel eines "harmlosen Witzes" gesellschaftlich legitimieren.

Die potenziell schädlichen Auswirkungen von abwertendem Humor auf Frauen sind umfassend dokumentiert. Beispielsweise kann die Exposition von misogynem Humor beeinflussen, wie viel Geld Männer hypothetischen Frauenorganisationen entziehen würden (Ford et al., 2008). Zudem bewerten Männer misogynen Bemerkungen, nachdem sie sexistische Witze gehört haben, am Arbeitsplatz als weniger beleidigend (Ford, 2000; Ford et al., 2001). Ford (2013) kam zu dem Ergebnis, dass Männer nach dem Konsum von misogynem Humor, Überzeugungen teilen, welche die Geschlechterungleichheit rechtfertigen. Zudem zeigt Ford (2000), dass Personen mit höherer Ausprägung von feindseligem Sexismus nach der Auseinandersetzung mit misogynen, humorvollen Inhalten signifikant toleranter gegenüber sexistischen Ereignissen sind. Weitere Studien zeigen, dass Frauen nach der Begegnung mit misogynem Humor eher diskriminierende Verhaltensweisen gegenüber weiblich gelesenen Personen aufweisen. Im Gegensatz dazu zeigen Männer dies unter den gleichen Bedingungen nicht (Abrams & Bippus, 2014). Auch extreme Formen von Sexismus wurden in zahlreichen Studien untersucht. Dabei wurde festgestellt, dass misogyner Humor die Fantasien von Vergewaltigungen, Opferbeschuldigungen und Gewalt an Frauen sowie sogar die eigene Bereitschaft zu einer Vergewaltigung erhöhen kann (Romero-Sanchez et al., 2010; Thomae & Viki, 2013; Thomae & Pina, 2015; Saucier et al., 2016).

Humor dient als Strategie, um die Diskriminierung in der Hassrede zu verbergen (Warren et al., 2021). Dabei wird die Hatespeech oftmals implizit kommuniziert und erfolgt somit auf eine legale Weise (Matamoros-Fernández et al., 2023), wodurch eine Entfernung und ein Verbot auf Social-Media-Plattformen erschwert werden (Gillett et al., 2022). Die Integration von humorvollen Elementen zu Hatespeech tolerieren Rezipierende eher als eine explizite herabsetzende Äußerung (Mendiburo-Seguel et al., 2019). Humoristische Hatespeech kann Vorurteile verstärken und die hierarchische Struktur zwischen sozialen Gruppen aufrechterhalten (Hodson et al., 2010).

Schmid (2023) untersucht die Wahrnehmung und Verarbeitung von humoristischer Hatespeech. Dabei kommt sie zu dem Ergebnis, dass Humor die feindselige Natur von Hassrede nicht über alle Inhalte und Zielgruppen hinweg verschleiern kann, jedoch klassifizieren die Teilnehmer:innen die Feindseligkeit niedriger. Der potenzielle Schaden wird zwar erkannt, doch überwiegt die Auffassung, dass die humoristische Hatespeech primär der Unterhaltungszwecke diene. Insbesondere gilt dies für eine implizite Form von humorvoller Hatespeech, in der negative Stereotype dargestellt werden (Pondorf, 2024). Das lässt erkennen, dass das Phänomen der Hassrede mit humoristischen Elementen stark in der Kategorie *Negative Stereotypisierung* anzusiedeln ist.

Besonders Misogynie manifestiert sich vermehrt in Ausdrucksformen wie humorvollen Bemerkungen, Ironie und Witzen (Worth et al., 2015). Diese Formen können dazu dienen, die Bedrohung für Frauen als harmlos erscheinen zu lassen, indem sie humoristische Elemente beinhalten, wie beispielsweise Emoticons oder die Verwendung der Abkürzung "LOL" (laugh out loud) (Cole, 2015; Pondorf, 2024)

In ihrer Untersuchung zur impliziten, humorvollen Hassrede gegenüber weiblichen Personen stellen Schmid et al. (2024) fest, dass diese häufig nicht als unhöflich oder

problematisch wahrgenommen wird. Dabei sind auch Geschlechterunterschiede zu erkennen. Frauen bewerten hasserfüllten Humor abwertender als Männer (Biber et al., 2002).

Die wissenschaftlichen Untersuchungen verdeutlichen, dass misogyner Humor die soziale Stellung von weiblich gelesenen Frauen, aber auch ihr Selbstbild und ihre psychische Gesundheit bedrohen kann. Dabei wirkt abwertender Humor als Legitimierung von diskriminierendem Verhalten. Misogynie mit humoristischen Elementen wirkt erstmal weniger bedrohlich. Durch eine sozial akzeptierte Form der Diskriminierung kann Humor zu einer höheren Akzeptanz von Vorurteilen beitragen. Mit Hilfe der Subtilität werden negative Auswirkungen oft unterschätzt (Ford & Ferguson, 2004; Ford et al., 2008).

2.4 Forschungsdefizit

Die genannten Entwicklungen in der Online-Kommunikation und die zunehmende Komplexität des Phänomens der Hatespeech machen spezifischere Forschung unerlässlich. Vor allem die Integration von Humor als Verschleierungsstrategie von Hatespeech ist wenig erforscht (Schmid, 2023). Jedoch kann gerade diese implizite Form von Hass weiter zirkulieren, da sich die Erkennung und die Moderation dieser Inhalte schwieriger gestalten (Matamoros-Fernández et al., 2023). Bisherige wissenschaftliche Arbeiten, die sich mit der Rezeption von Hatespeech beschäftigen, haben sich vermehrt auf explizite Formen konzentriert (Schmid et al., 2024; Obermaier et al., 2023; Maarouf et al., 2024). Diese Studien bieten wertvolle Einblicke in die Mechanismen der Hassrede, jedoch fehlt es an umfassenden Analysen der impliziten, humorvoll getarnten Formen von Hass. Humoristische Hassrede, die durch vermeintliche Witze, Emojis oder ironische Formulierungen vermittelt wird, kann sowohl auf eine indirekte Weise Hass transportieren als auch die wahrgenommene Tiefe des Hasses vermindern (Warren et al., 2021). Wie in den vorherigen Kapiteln aufgezeigt, werden vor allem Frauen in Online-Räumen Opfer von geschlechtsspezifischer

Hatespeech. Diese wird oftmals implizit und durch humoristische sowie vermeintlich harmlose Kommentare verschleiert (Brown, 2017). Diese Forschungslücke erfordert eine tiefe Analyse der spezifischen Kontexte, in denen humorvolle Hassrede auftaucht, um zu verstehen, wie sich die Rezeption von impliziter von expliziter Hatespeech unterscheidet. Schmid (2023) erwähnt zudem eine konkrete und zentrale Forschungslücke in der Analyse von Hatespeech mit humoristischen Elementen. Hassrede soll in einem kontextualisierten Rahmen untersucht werden. Dabei werden aufbauend auf ihre Studie bestehende Limitationen weiter minimiert.

Diese Studie ist relevant, da sie das Verständnis von subtiler, humoristischer Hassrede erweitert. Die gewonnenen Erkenntnisse bieten eine Grundlage für präventive Maßnahmen, Bildungsprogramme und Aufklärungskampagnen, die das Bewusstsein für schädliche Auswirkungen solcher Inhalte stärken. Dadurch kann die Kommunikationskultur in Online-Räumen verbessert. Eine umfassende Analyse humoristischer Hassrede liefert zudem wertvolle Ansätze, anhand derer Online-Plattformen und politische Entscheidungsträger:innen gezielte Schutzmaßnahmen für marginalisierte Gruppen entwickeln können. Insgesamt schließt diese Forschung Lücken in der Forschung von Hatespeech und stärkt die Handlungsfähigkeit im Umgang mit dem komplexen Phänomen.

3. Forschungsfragen und Hypothesen

In den vorangegangenen Kapiteln wurden die relevanten theoretischen Konzepte sowie die aktuelle Forschungslage zu expliziter und impliziter, speziell der humoristischen Hassrede, umfassend behandelt. Darüber hinaus wurde die besondere Bedeutung misogyner Hatespeech herausgestellt. Im folgenden Kapitel liegt der Fokus auf der Formulierung der Forschungsfrage sowie der Ableitung der entsprechenden Hypothesen. Diese dienen als Grundlage für weitere empirische Untersuchungen.

Die zentrale Forschungsfrage dieser Studie lautet: *„Inwiefern beeinflusst die Explizität von Hatespeech die Nutzerreaktionen der Rezipierenden?“* Drei wesentliche Dimensionen werden dabei entlang der Forschungslücke untersucht. Zunächst wird die emotionale Reaktion durch die Stimmung der Nutzer:innen zum Zeitpunkt der Befragung analysiert, wie in etwa Empörung, Verärgerung oder Belustigung. Die zweite Dimension befasst sich mit einer kommunikativen Handlung, speziell der Bereitschaft zur Counterspeech. Hierbei wird untersucht, ob Nutzer:innen motiviert sind, sich gegen Hatespeech zu positionieren und intervenierende Kommentare zu leisten. Drittens wird analysiert, wie sich die Konfrontation mit expliziter oder subtiler misogyner Hatespeech auf die sexistischen Einstellungen der Nutzer:innen auswirkt. Hierbei zeigt die Studie, ob die Konfrontation mit Hassrede zu einer Veränderung individueller Einstellungen hinsichtlich Geschlechterrollen bzw. -gleichheit führt. Kognitive Aspekte werden damit einbezogen.

Hypothese 1.1

Wie der Forschungsstand belegt, beeinflusst die Explizität der Hatespeech die Wahrnehmung der Rezipierenden maßgeblich. Schmid et al. (2024) zeigen, dass es eine Korrelation zwischen der Direktheit von Hassrede und einer stärkeren, feindseligeren Wahrnehmung gibt. Insbesondere bei misogynen Äußerungen wird das deutlich. Dies weist

darauf hin, dass explizite, misogyne Hatespeech zu stärkeren negativen Gefühlen wie etwa Empörung und Abscheu führt als subtile Formen von Hatespeech. Studien belegen, dass explizite Formen von Hassrede intensivere emotionale Reaktionen hervorrufen als indirekte, subtile (Kenski et al., 2020; Stryker et al., 2016; Leonhard et al., 2018). Schmid (2023) kam zu dem Ergebnis, dass humoristische Hatespeech als weniger feindselig wahrgenommen wird und somit auch emotional weniger belastend ist als explizite Hassrede. Humorvolle Elemente können dementsprechend dazu beitragen, die Botschaften abzuschwächen und emotionale Effekte der Rezipient:innen zu mildern. Daraus kann geschlossen werden, dass humoristische, misogyn geprägte Hassrede nicht zu denselben negativen emotionalen Reaktionen führt wie explizite, direkte Hatespeech. Daher wird in dieser Hypothese postuliert, dass explizite, misogyne Hassrede eine negativere, emotionale Stimmung erzeugt als humoristische, misogyne Hatespeech.

H 1.1: Explizite, misogyne Hatespeech führt zu einer negativeren, emotionalen Stimmung als humoristische, misogyne Hatespeech.

Hypothese 1.2

Strafrechtlich relevante Hassrede wird als besonders schwerwiegend angesehen (Schmid et al., 2024). Dies steht im Einklang mit der Erkenntnis, dass explizite Hatespeech, die in etwa beleidigende Sprache oder Gewaltandrohungen enthält, als hasserfüllt wahrgenommen wird (Kenski et al., 2020; Stryker et al., 2016). Insbesondere gilt das für die Wahrnehmung von misogyner Hatespeech, die häufig als unhöflich und belastend empfunden wird (Obermaier et al., 2023). Subtilere Formen von Hatespeech wie humorvolle Hassrede werden dagegen als weniger feindselig wahrgenommen (Schmid, 2023). Leonhard et al. (2018) ergänzen diese Perspektive, indem sie zeigen, dass Bystander nicht automatisch nach dem Lesen von Hassrede intervenieren. Studien zeigen, dass das Gefühl der persönlichen

Verantwortung stärker wird, abhängig von der Intensität des als Angriff empfundenen Inhalts (Obermaier et al., 2023). Dieses Verantwortungsgefühl motiviert dann zur Intervention in Form von Counterspeech (Obermaier et al, 2023). In Anbetracht dieser Zusammenhänge postuliert *H 1.2*, dass explizite, misogyne Hatespeech aufgrund ihrer Wahrnehmung, als besonders schwerwiegend und bedrohlich, zu einem stärkeren Gefühl der Verantwortung und damit zu einer höheren Bereitschaft zur Counterspeech führt. Dies steht im Gegensatz zu humoristischen Darstellungen von Hatespeech, die oft als weniger ernst und damit weniger dringlich zum Eingreifen empfunden wird.

H 1.2: Explizite, misogyne Hatespeech führt zu einer höheren Bereitschaft von Counterspeech als humoristische, misogyne Hatespeech.

Hypothese 1.3

Forschungen zeigen, dass die Wirkung von Medienbildern stereotypische Einstellungen formen. Es ist dokumentiert, dass Medieninhalte dazu beitragen, dass Stereotype entwickelt werden (Ward et al., 2020; Schemer, 2012). Insbesondere bei der Betrachtung von diskriminierenden Nutzerkommentaren und Hatespeech ist diese Dynamik relevant. Zahlreiche Studien zeigen, dass der Kontakt mit Hassrede die Zustimmung zu negativen, expliziten Stereotypen über die betroffenen Gruppen aktiviert (Schäfer et al., 2022; Weber et al., 2020; Soral et al., 2018). Fox et al. (2015) weisen nach, dass die Auseinandersetzung mit sexistischen Online-Inhalten misogyne Einstellungen verstärken kann. Diese Erkenntnis wird durch weitere Studien unterstützt, die eine Zunahme der Zustimmung zu negativen Stereotypen bei Exposition gegenüber Hassrede dokumentieren (Schäfer et al., 2022; Hsueh et al., 2015; Weber et al., 2020; Soral et al., 2018). Eine besondere Gefahr geht von humoristischen Formen von Hatespeech aus, die oft nicht auf den ersten Blick erkannt werden (Schmid et al., 2024). Da humorvolle Hatespeech nicht so

schädlich wahrgenommen wird wie die explizite Form, wird angenommen, dass misogyne Einstellungen auch nicht in dem Maße verstärkt werden wie bei der Konfrontation mit direkter Hatespeech. Die explizite Ausrichtung der Hassrede gilt als maßgeblich, da sie potenziell eine intensivere Verstärkung sexistischer Einstellungen bewirken kann (Hodson et al., 2010).

H 1.3: Explizite, misogyne Hatespeech führt zu einer verstärkten sexistischen Einstellung als humoristische, misogyne Hatespeech.

Es wird angenommen, dass der Moderator „Geschlecht" entscheidend beeinflusst, wie stark die Explizität von Hassrede die Rezeption der einzelnen Personen steuert. Vor diesem Hintergrund ergibt sich die folgende Forschungsfrage: *„Inwieweit moderiert das Geschlecht die Wirkung von expliziter, misogyner Hatespeech auf die emotionale Stimmung, die Bereitschaft zur Counterspeech und den eigenen misogynen Einstellungen?"*

Hypothese 2.1

Williams et al. (2016) zeigen, dass Personen, die bereits selbst mit Hassrede konfrontiert wurden, eine erhöhte Sensibilität für ähnliche Inhalte entwickeln. Ihre Studie ergibt auch, dass die Betroffenen die Hassrede als beleidigender empfinden als unbeteiligte Personen. Kenski et al. (2017) zeigen, dass Frauen, im Vergleich zu anderen Gruppen, Aussagen häufiger als unhöflich bewerten. Geschlechterspezifische Normen spielen also eine Rolle wie Hatespeech empfunden wird (Perry & Weiss, 1989). Durch eine erhöhte Wahrnehmung von Betroffenen - also hier weiblich gelesenen Personen - wird postuliert, dass diese eine negativere emotionale Stimmung nach der Rezeption empfinden, insbesondere von expliziter Hatespeech, als Unbeteiligte. Dem schließen sich Zochniak et al. (2023) an. Das Ergebnis zeigt, dass Hatespeech das Wohlbefinden von Personen, die sich der betroffenen Gruppe zugehörig fühlen, erheblich mehr beeinflusst. Hatespeech kann also bei

Betroffenen zu emotionalerer Belastung führen als bei Personen, die der Gruppe nicht zugehörig sind (Eisenberger, 2015; Vedeler, 2019; Wypych & Bilewicz, 2021). Da explizite Hatespeech feindseliger wahrgenommen wird (Schmid, 2023) und Frauen hinsichtlich dessen sensibilisierter sind (Williams et al., 2016) erleben sie diese feindseliger als Männer.

H 2.1: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und einer negativen, emotionalen Stimmung. Frauen zeigen tendenziell eine stärkere negative, emotionale Reaktion auf explizite, misogyne Hatespeech im Vergleich zu humorvoller Hatespeech, während bei Männern dieser Unterschied weniger ausgeprägt ist.

Hypothese 2.2

Ping et al. (2024) belegen, dass frühere Opfererfahrungen einen entscheidenden Prädiktor für die Teilnahme an Counterspeech darstellen. Personen, die schon Ziel von Hatespeech waren, beteiligten sich signifikant häufiger an einer Form der Gegenrede. Die Ergebnisse stützen sich auf frühere Forschungen, die belegen, dass Opfer aufgrund ihrer Erfahrungen besser einschätzen können, wann Interventionen wirksam sind (Obermaier et al., 2023). Dies wird durch die Beobachtung unterstützt, dass Opfer verstärkt motiviert sind, hasserfülltes Verhalten direkt zu konfrontieren, während diese Motivation bei Nicht-Opfern weniger ausgeprägt ist (Obermaier et al., 2023). Da Frauen häufiger Opfer von Hatespeech sind als Männer (FRA, 2023), ist ihr Verantwortungsgefühl auch höher und es wird eher der Interventionsprozess ausgelöst. Studien zeigen, dass durch eine feindseligere Wahrnehmung, die durch die Gruppenzugehörigkeit des Geschlechts erfolgt, die Bereitschaft zur Counterspeech größer ist (Obermaier et al., 2023). Explizite Hatespeech wird feindseliger wahrgenommen als humorvolle Hassrede (Schmid, 2023), vor allem Frauen erkennen Diskriminierungen schneller (Williams et al., 2016). Ausgehend davon wird die Hypothese

aufgestellt, dass Frauen eine höhere Bereitschaft zur Counterspeech aufweisen als Männer, insbesondere bei expliziter, misogyner Hatespeech. Dabei wird angenommen, dass Frauen stärker auf die Inhalte reagieren und somit häufiger konstruktive Gegenrede initiieren.

H 2.2: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und der Bereitschaft zur Counterspeech. Frauen zeigen tendenziell eine größere Bereitschaft zur Gegenrede, insbesondere bei expliziter, misogyner Hatespeech, während dieser Effekt bei Männern weniger ausgeprägt ist.

Hypothese 2.3

Hatespeech erhöht die Akzeptanz negativer Stereotype gegenüber der angegriffenen Gruppe (Hsueh et al., 2015; Schäfer et al., 2022; Soral et al., 2018; Weber et al., 2020). Fox et al. (2015) fanden heraus, dass die Konfrontation mit misogynen Inhalten im Internet die Tendenz verstärkt, sexistischen Einstellungen zuzustimmen. Dieser Effekt kann durch die Art und der Direktheit der Hassrede intensiviert oder abgeschwächt werden. Hatespeech kann humorvoll und subtil sein und wird oft nicht als solche erkannt (Schmid et al., 2024), was die Intensivierung sexistischer Einstellungen indirekt unterstützen kann (Hodson et al., 2010). Die Forschungsergebnisse zeigen, dass es Geschlechtsunterschiede bei Wahrnehmung und Reaktion von Hatespeech gibt. Frauen bewerten hasserfüllten Humor häufiger als feindlich und erkennen den diskriminieren Charakter leichter als Männer (Biber et al., 2002). Zochniak et al. (2023) zeigen, dass Betroffene stärkere emotionale Reaktionen zeigen als Personen, die nicht betroffen sind, was sich auf die Verarbeitung und Bewertung solcher Inhalte auswirken kann (Schmid et al, 2024). Auf dieser Basis können wir schließen, dass das Geschlecht eine wichtige Rolle bei der Wirkung explizit misogyn geprägter Hassrede auf die Intensivierung misogyner Einstellungen spielt.

H 2.3: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und der Verstärkung sexistischer Einstellungen. Frauen sind tendenziell stärker von der Hassrede betroffen, insbesondere bei expliziter, misogyner Hatespeech, und zeigen eine geringere Zustimmung zu sexistischen Stereotypen im Vergleich zu Männern.

4. Untersuchungsanlage und Methodologie

Im folgenden Kapitel wird die methodische Vorgehensweise dieser Studie umfassend dargestellt. Dabei werden zunächst das Forschungsdesign und die Methodenkonzeption, einschließlich der ethischen Maßnahmen, erläutert. Daraufhin folgt eine detaillierte Beschreibung der Durchführung der Studie sowie der Rekrutierung der Teilnehmer:innen. Im Anschluss erfolgt eine Beschreibung der Operationalisierung, die sowohl das Stimulusmaterial als auch die Messung der relevanten Konstrukte im Detail umfassen. Daraufhin wird der Aufbau des Fragebogens sowie der Ablauf des Pretests dargestellt.

4.1 Forschungsdesign

In dieser Studie dient ein Online-Experiment zur Untersuchung der Fragestellungen. Mithilfe der Methode soll herausgefunden werden, ob die Explizität von Hatespeech einen Effekt auf die Rezeption der Nutzer:innen hat. Die Proband:innen waren innerhalb des Online-Experiments verschiedenen Stimuli ausgesetzt, welche in Form eines Facebook-Postings präsentiert wurden. Nach der Konfrontation mit dem Stimulus wurde den Teilnehmenden ein Online-Fragebogen vorgelegt. Dabei wurde anhand von drei Konstrukten gemessen, welchen Einfluss das Rezipierte auf die Reaktionen hat. Die Forschungslogik eines Online-Experimentes fragt nach dem Einfluss von UVs auf AVs (Brosius et al., 2022). UVs werden dabei gezielt manipuliert, um zu prüfen, ob und wie sich diese Veränderungen auf die AVs auswirken. In Online-Fragebogen-Experimenten erfolgt die Manipulation der UV in Form von verschiedenen Stimuli, während die AV anhand der Antworten der Teilnehmer:innen gemessen werden. In dieser Studie bilden die Manipulation und somit die UV die Explizität der Hatespeech. Daraus ergaben sich zwei Experimentalgruppen, um die Effekte von verschiedenen Formen von Hassrede zu untersuchen, wobei eine zugleich als Kontrollgruppe dient. Die Experimentalgruppe wurde mit impliziter, humoristischer Hassrede

konfrontiert, die humorvolle Elemente beinhalteten und subtiler in der Ausdrucksweise waren. Die zweite Experimentalgruppe bzw. Kontrollgruppe sah Inhalte mit expliziter Hassrede, die klare und direkte Beleidigungen enthielt. Die AVs wurden mithilfe von Konstrukten durch einen Online-Fragebogen gemessen. Das experimentelle Design dieser Studie basiert auf einem 2 x 3 Between-Subject Design. Dieses umfasst zwei Ausprägungen der UV hinsichtlich der Explizität der Hatespeech von Nutzerkommentaren. Die Variablen werden operationalisiert als "explizite Hatespeech" und "humoristische Hatespeech". Das ermöglicht es, Unterschiede in der Wirkung von Hatespeech, abhängig von ihrem expliziten oder humoristischen Charakter, zu untersuchen. Die drei AVs der Studie umfassen die emotionalen Reaktionen, insbesondere die aktuelle Stimmung, die Bereitschaft zur Counterspeech als kommunikative Handlung sowie den kognitiven Aspekt in Form misogyner Einstellungen. Zusätzlich wird der Faktor „Geschlecht“ als Moderator integriert, um potenzielle Geschlechtsunterschiede in den Effekten zu identifizieren. Dieser Ansatz ermöglicht es, zu untersuchen, ob und wie sich die Reaktionen auf Hassrede in Abhängigkeit vom Geschlecht der Rezipierenden unterscheiden. Dies ist besonders relevant, da eine Gruppe – die weiblichen Teilnehmenden – unmittelbar von der Hassrede betroffen ist.

4.1.1 Methodenbegründung

Für das Forschungsinteresse wurde als Methode das Online-Experiment gewählt. Dieses wurde gewählt, da es – wie von Koch et al. (2019) beschrieben – Ursache-Wirkungs-Beziehungen aufdecken kann. Speziell für die vorliegenden Fragstellungen ist es deshalb besonders geeignet. Kommunikationswissenschaftliche Forschungen nutzen Experimente häufig, da sie dazu beitragen, klare Kausalzusammenhänge aufzudecken und zu überprüfen (Koch et al., 2019). Die Methode bietet die Möglichkeit, das Stimulusmaterial als Bild direkt in den Fragebogen zu integrieren. Das ist entscheidend für die Validität der Studie (Aguinis

& Bradley, 2014). Die interne Validität wird durch eine präzise Kontrolle und Konsistenz der Manipulation zwischen den Gruppen gestärkt (Vanden Abeele & Postma-Nilsenova, 2018). Vor allem durch das Thema, das eine Online-Umgebung vorsieht, ist ein Online-Experiment dem eines Laborexperiments vorzuziehen (Koch et al., 2019). Durch eine authentische Umgebung in dem Social-Media-Kontext wird im Gegensatz zu einer künstlichen Atmosphäre im Labor die externe Validität der Ergebnisse erhöht (Von Berger et al., 2009). Zudem vereinfacht diese Methode die Rekrutierung der Teilnehmer:innen, da keine physische Präsenz im Labor erforderlich ist. Die Kontrolle der Stimulusbearbeitung ist in einem Online-Experiment sicherzustellen. Dies wurde durch einen Manipulationscheck nach der Rezeption des Stimulus gewährleistet. Dieser wird in *Kapitel 5.1* näher erläutert. Es wurde außerdem ein 2 x 3 Between-Subjects Design gewählt, bei dem jeder Versuchsperson nur einer der experimentellen Bedingungen zugewiesen wurde. Das Design bietet wesentliche Vorteile zu Within-Subjects Designs. Bei diesem durchlaufen Proband:innen alle Bedingungen (Hemmerich, 2016). Das kann zu dem sogenannten *„Carry-Over-Effekt“* führen, bei dem eine vorherige Bedingung unbeabsichtigt die Ergebnisse einer späteren Bedingung beeinflussen könnte (Jones & Kenward, 2014). Bei einem Between-Subject Design reduziert sich die Wahrscheinlichkeit, dass Proband:innen falsche Antworten geben oder die Studie vorzeitig abbrechen, da sie nur eine Bedingung erleben und sich nicht durch wiederholte Tests ermüden (Hemmerich, 2016).

4.1.2 Forschungsethische Maßnahmen

In der vorliegenden Studie wurden umfassende forschungsethische Maßnahmen anlehnend an Schmid (2023) und Obermaier et al. (2023) implementiert, um die Teilnehmenden zu schützen und die ethische Integrität zu gewährleisten. Zu Beginn der Befragung wurde eine Triggerwarnung ausgesprochen, welche die Proband:innen auf

potenziell belastende Inhalte in den gezeigten Stimuli vorbereitete. Zwar wurde erwähnt, dass es sich um misogyne Hatespeech handeln kann, jedoch nicht darauf eingegangen, dass diese auch als implizite Hassrede mit humorvollen Elementen auftreten kann, um gezielt Einfluss zu vermeiden. Zudem wurden sie über die Freiwilligkeit ihrer Teilnahme und mögliche emotionale Beeinträchtigungen informiert. Die Proband:innen erhielten den Hinweis, dass sie sie Umfrage, falls diese zu belastend empfunden wird, jederzeit abbrechen können. Am Ende erfolgte ein detailliertes Debriefing, das, neben der Erläuterung des Studienzwecks, auch Kontaktstellen für Unterstützung im Falle emotionaler Belastungen angab. Diese Maßnahmen tragen dazu bei, das Wohlbefinden der Teilnehmenden zu schützen und die Einhaltung forschungsethischer Standards sicherzustellen.

4.2 Teilnehmer:innen und Durchführung

Zur Überprüfung der Hypothesen und Forschungsfragen wurde das Online-Experiment im Zeitraum vom 15. August 2024 bis 29. September 2024 durchgeführt. Die Grundgesamtheit (N) umfasst dabei Personen, die mit Social-Media-Plattformen vertraut sind. Aufgrund dessen wurde zunächst eine Grundgesamtheitsabfrage gestellt, so dass nur jene Teilnehmer:innen für den restlichen Fragebogen zugelassen wurden, die mindestens einen Social-Media-Account besitzen. Die Grundgesamtheit (N) ist unendlich und nicht bekannt und es wurde auf eine Stichprobe (n) zurückgegriffen. Diese wurde mithilfe einer willkürlichen Auswahl getroffen, wodurch es eine anfallende bzw. ad-hoc-Stichprobe ist (Koch et al., 2019). Die willkürliche Auswahl wird insbesondere bei der in dieser Forschung angewendeten Methode des Experimentes verteidigt, da *„es ja nur um den relativen Nachweis eines Effektes gehe, also die Überprüfung, ob dieser überhaupt auftritt. […] Zentrales Argument der Befürworter ist also, dass Experimente nur Kausalitäten nachweisen wolle, was man prinzipiell mit jeder Stichprobe erreichen kann.“* (Koch et al., 2019, S. 119).

Das eingesetzte Erhebungsinstrument zeigt eine Limitation im Hinblick auf den Digital Divide. Die Stichprobe erfasst lediglich die Bevölkerungsgruppen, die Zugang zum Internet haben und diesen aktiv nutzen. Deshalb wurden die Teilnehmer:innen im entfernten Bekanntenkreis und über Universitätsgruppen auf Social-Media-Plattformen durch einen Online-Link rekrutiert. Vor allem durch den Bezug der Studie zu sozialen Medien war die Gewinnung über die Plattformen relevant, da dies die Zielgruppe der Untersuchung darstellt. Nach der Rekrutierungsphase umfasste die bereinigte Stichprobe (n) = 205.

Bei der Datenreinigung wurden insgesamt neun Fälle ausgeschlossen. Zwei Personen haben der Datenschutzbelehrung nicht zugestimmt. Weitere zwei Proband:innen haben angegeben, keinen Social-Media-Account zu besitzen, sodass sie für die weitere Studie nicht mehr von Relevanz waren. Zudem wurde ein Fall ausgeschlossen, da dieser im Vergleich zu den anderen Teilnehmenden eine auffällig kurze Bearbeitungszeit aufwies. Ein einfaches „Durchklicken“ konnte hier nicht ausgeschlossen werden, was die Ergebnisse verfälscht hätte. Zudem blieben bei zwei Durchführungen der Studie einige Teilfragen unbeantwortet, wodurch diese Fälle für die weitere Auswertung nicht mehr relevant sind. Alle anderen Fälle konnten weiterverwendet werden.

An der Studie nahmen Teilnehmer:innen im Alter von 18 bis über 60 Jahren teil, wobei die größte Gruppe mit 53,7 % aus 21 bis 29 - jährigen bestand. Die Geschlechterverteilung ist mit 72,2 % Frauen und 27,3 % Männern deutlich unausgewogen. In Bezug auf die Bildung ergab sich ein klarer Trend für eine höhere Bildung. 56,6 % der Proband:innen haben einen Bachelor- bzw. Masterabschluss.

Die Stichprobe wurde in eine Experimentalgruppe (Implizite Hatespeech mit humorvollgemeinten Elementen) sowie eine Kontrollgruppe (Explizite Hatespeech)

randomisiert aufgeteilt. Die Größe der Experimentalgruppe ergab damit (n) = 101 und die der Kontrollgruppe (n) = 104.

4.3 Operationalisierung der Variablen

Die UVs und AVs werden gemäß der Forschungsfragen und Hypothesen operationalisiert. Die UV umfasst die Explizität der Hatespeech in misogynen Nutzerkommentaren, die in Form von "expliziter Hatespeech" und "Hatespeech mit humorvollen Elementen" gemessen wird. Die Explizität wird durch die direkte bzw. implizite Darstellung der Hassrede erreicht, wobei bei Letzterem humorvolle Elemente in den Kommentaren zu finden sind. Die AVs umfassen die Stimmung, gefolgt von der Bereitschaft zur Counterspeech sowie sexistischen Einstellungen.

4.3.1 Stimulusmaterial

Zur Operationalisierung der UVs wurde das Stimulusmaterial der vorliegenden Studie gezielt ausgewählt, um die Wirkung von Hassrede auf die Rezipient:innen zu untersuchen. Vor diesem Hintergrund fiel die Wahl auf Facebook als Plattform, da sie weiterhin die meistgenutzte Social-Media-Plattform in Österreich darstellt (Statista, 2024a). Somit hat sie eine hohe Relevanz für die Studienteilnehmenden. Die externe Validität der Studie kann dadurch gesteigert werden, da die Teilnehmer:innen die Umgebung der Social-Media-Plattform aus ihrem Alltag kennen und als authentisch wahrnehmen.

Das Stimulusmaterial basiert auf einem authentischen Facebook-Beitrag des offiziellen Kanals der ARD-Tagesschau, einem renommierten und vertrauenswürdigen Nachrichtenanbieter in Deutschland (Statista, 2024b). Der Beitrag thematisiert den Weltfrauentag in Berlin und enthält zwei verschiedene Stimuli von Hatespeech: explizite Kommentare und Kommentare mit humorvollen Elementen. Der Beitrag wurde so konzipiert,

dass er in Form und Struktur den anderen Beiträgen entspricht, um eine möglichst hohe Authentizität zu gewährleisten. Der Text vermeidet bewusst polarisierende Begriffe und wertende Aussagen, um eine neutrale Darstellung zu gewährleisten. Die Manipulation soll sich bewusst auf die Kommentare beschränken und nicht durch andere Störfaktoren beeinflusst werden. Durch die Nutzung eines realen und neutralen Nachrichtenkanals erhöht sich die ökologische Validität, da die dargestellte Situation für die Teilnehmer:innen glaubwürdig und lebensnah erscheint (Statista, 2024b).

Die Kommentare wurden sowohl direkt von Facebook aus verschiedenen Postings und Nachrichtenseiten entnommen als auch von der Autorin selbst verfasst. Es wird ein Beitrag mit feministischen Inhalten analysiert, da frühere Studien von Siddiqui (2008) und Eudey (2008) zeigen, dass solche Beiträge häufig von misogynen Hasskommentaren betroffen sind. Aus den Analysen wurden gezielt Kommentare ausgewählt, die der festgelegten Definition von Hatespeech entsprechen, um die Auswirkungen von Hatespeech in dem spezifischen Kontext zu untersuchen. Zur Erfassung wurde sich ebenso an die Kategorisierung aus *Kapitel 2.1.2* gehalten. Für die Experimentalgruppe wurden Kommentare aus den Kategorien *Beleidigende Sprache* und eine *Negative Stereotypisierung* gewählt. Zusätzlich wurde darauf geachtet, dass die Kommentare humorvolle Elemente wie Emojis und Ausdrücke wie „Lol“, „Haha“ und satirische Wortspiele enthalten. Für die Kontrollgruppe wurden Kommentare der Kategorie *Aufforderung zu Gewalt, Diskriminierung oder Hass, Herabwürdigung* und *Negative Stereotypisierung* ausgewählt. Sobald humorvolle Elemente wie beispielsweise lachende Emojis enthalten waren, wurden diese ausgeschlossen.

Diese Auswahl der Kommentare wurde mithilfe einer Vorstudie validiert. Hierfür erhielten die teilnehmenden Proband:innen fünfzig Kommentare, die sie auf einer 5-stufigen Likert-Skala hinsichtlich ihrer humorvollen Elemente bewerten sollten (Frage: "Inwiefern

stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?"). Die Teilnehmer:innen setzten sich aus sechzehn Personen aus dem persönlichen Umfeld der Autorin zusammen. Sie wurden über soziale Medien kontaktiert, um ihre Perspektiven in den Auswahlprozess einzubeziehen. Auf Grundlage der Bewertungen der Proband:innen wurden die finalen Kommentare für die Studie ausgewählt. Durch die Einbeziehung einer Vorstudie konnte die Objektivität der Stimulusauswahl erhöht werden. Zudem wurde die Reliabilität gestärkt, da die Auswahl nicht allein durch die Autorin bestimmt wurde. Diese methodischen Vorgehensweisen trugen zur internen Validität der Studie bei, indem sichergestellt wurde, dass die beobachteten Effekte auf die Manipulation und nicht auf andere Störvariablen zurückzuführen sind.

Zehn Kommentare wurden durch die Vorstudie für die Experimentalbedingung ausgewählt. Dabei wird der Facebook-Beitrag auf eine weniger ernste, jedoch potenziell ebenso schädliche Weise kommentiert. In dieser Gruppe werden humorvolle Komponenten wie lachende Emojis und Ausdrücke wie *„Lol“* integriert, zudem kommt satirische Sprache zur Anwendung. Beispiele für die Kommentare im Stimulus lauten: *„Der Tag, an dem Frauen uns erzählen, wie hart ihr Leben ist, während sie sich ihre Nägel machen lassen* 😅😅*“* oder *„Weltfrauentag – ich habe bereits einen Blumenstrauß für die Geschirrspülmaschine gekauft, hahaha“* Im Gegensatz dazu beinhalten die expliziten Kommentare direkte, angreifende Aussagen, die sich auf das Thema des Weltfrauentags sowie auf die Rolle der Frauen in der Gesellschaft beziehen. Dabei wurden ebenfalls zehn Kommentare durch die Vorstudie selektiert. Beispiele für diese expliziten Kommentare sind: *„Ich könnte mich nicht zusammenreißen, würde ich vor so einer dummen Feministenfotze stehen; ich weiß nicht, ob mir da die Hand ausrutschen würde...“* oder *„Dümmer geht’s immer, besonders in Berlin; die Frauen sind so lächerlich“* Die Unterschiede zwischen den beiden Stimulusarten liegen somit in der Art der Ausdrucksweise: Während die humorvollen

Kommentare versuchen, das Thema des Weltfrauentags durch Ironie und Satire zu kommentieren, sind die expliziten Kommentare durch direkte und aggressive Angriffe auf Frauen geprägt.

Die Präsentation der Kommentare erfolgte als Bildstimulus, um eine konsistente und kontrollierte Darstellung zu gewährleisten. Diese Methode bietet mehrere Vorteile. Erstens ermöglicht sie den Teilnehmer:innen, eine gewisse Distanz zum Stimulusmaterial einzunehmen, was die ethische Vertretbarkeit der Studie erhöht. Zweitens wurden die Experimentalbedingungen kontrolliert, indem potenzielle Störfaktoren wie die Interaktivität eliminiert wurden. In einer realen Online-Umgebung könnten Nutzer:innen auf Kommentare reagieren, diese liken oder teilen, was die Wahrnehmung und die Reaktionen der Teilnehmenden beeinflussen könnte. Solche interaktiven Elemente könnten dazu führen, dass die Teilnehmenden ihre Bewertungen nicht nur auf den Inhalt der Kommentare stützen, sondern auch auf die Art und Weise, wie andere Nutzer:innen auf diese Kommentare reagieren. Um eine verzerrte Wahrnehmung zu vermeiden, wurde die Interaktivität ausgeschlossen. Dies gewährleistet, dass die Teilnehmenden ihre Einschätzungen ausschließlich aufgrund des Kommentarinhalts vornehmen, ohne von externen Einflussfaktoren abgelenkt zu werden. Alle Namen und Likes der Kommentare wurden entfernt, um weitere potenzielle Störfaktoren zu vermeiden. Das Material wurde mit einem Bildbearbeitungsprogramm erstellt, um sicherzustellen, dass die Präsentation standardisiert ist und die Bedingungen für alle Teilnehmer:innen gleichbleiben. Durch die standardisierte Präsentation des Stimulusmaterials als Bild wird die interne Validität der Studie gestärkt. Dies stellt sicher, dass die beobachteten Effekte tatsächlich auf die experimentelle Manipulation zurückzuführen sind und nicht durch externe Faktoren verfälscht werden. Die Konsistenz der Präsentationsbedingungen minimiert die Variabilität, die durch

unterschiedliche Darstellungsformate oder unkontrollierbare Elemente entstehen könnten, und gewährleistet so die Zuverlässigkeit der Ergebnisse.

4.3.2 Messung der Konstrukte

Im Folgenden werden die Konstrukte der AVs genauer definiert. Die emotionale Stimmung der Rezipient:innen wird erfasst, um die unmittelbaren Auswirkungen der Hassrede auf die Befindlichkeit der Proband:innen zu untersuchen. Die Bereitschaft zur Counterspeech wird als die Neigung der Teilnehmenden definiert, auf die Hassrede mit einer Gegenrede zu reagieren. Diese wird anhand von Items gemessen, welche die Absicht der Teilnehmenden zur direkten Konfrontation mit der Hassrede erfasst. Veränderungen sexistischer Einstellungen beziehen sich auf die Verschiebung der Einstellungen der Teilnehmenden gegenüber Frauen, die durch die Konfrontation mit misogynen Hassreden hervorgerufen wird. Die Reihenfolge ist so gewählt, dass zunächst die Stimmung gemessen wurde, um unmittelbare und unverfälschte emotionale Reaktionen auf den Stimulus zu erfassen. Anschließend folgt die Messung der Counterspeech (kommunikative Handlungen), da diese stark von der zuvor gemessenen emotionalen Reaktion beeinflusst wird und die Bereitschaft zur Handlung widerspiegelt. Abschließend werden die Einstellungen (kognitiven Aspekte) erfasst, da diese auf tiefergehender Reflexion basieren und weniger von spontanen emotionalen Reaktionen beeinflusst werden. Die Reihenfolge gewährleistet, dass die Erfassung der Reaktionen und Handlungen nicht durch vorangegangene Überlegungen verzerrt wird.

Emotionale Stimmung

Zur Messung der emotionalen Stimmung wurde die deutsche Version der *„Positive and Negative Affect Schedule“* (PANAS; Watson et al., 1988) verwendet. Das Messinstrument wurde von Watson et al. (1988) für den englischsprachigen Raum entwickelt

und von Breyer und Bluemke (2016) in den deutschen Sprachraum adaptiert. Es werden zwanzig Adjektive erfasst, wobei zehn positive und zehn negative Empfindungen beschreiben. Die PANAS ermöglicht es, sowohl *„aktuelle, zeitlich begrenzte Affekte als auch überdauernde, habituelle Affektivitätsmerkmale“* (Breyer & Blumke, 2016, S.1) zu erfassen. Für diesen Untersuchungszweck wurde diese für die emotionale Reaktion, d.h. die Stimmung im aktuellen Moment, angewendet. Dabei wurde das Konstrukt, analog zu Breyer und Blumke (2016), mit einer fünfstufigen Likert-Skala gemessen, welche von „gar nicht“ bis „sehr“ reicht. Die negativen Empfindungen wurden im Vorhinein umkodiert, um ein valides Ergebnis zu erhalten. In *Tabelle 1* sind die umkodierten Items zu finden. Das ist notwendig, um eine konsistente Interpretation der Antworten zu gewährleisten und den Einfluss der unterschiedlichen Stimuli auf die emotionalen Reaktionen der Teilnehmenden präzise zu erfassen. Die interne Konsistenz dient ebenfalls zur Auswahl dieses Messinstrumentes. Sie liegt zwischen .86 und .93 (Breyer & Blumke, 2016). Auch für die vorliegende Studie konnte eine zufriedenstellende interne Konsistenz ($\alpha = .787$) nachgewiesen werden. Die zwanzig Items wurden mithilfe des Mittelwertindex für die weitere Auswertung zusammengefasst. Die nachstehende *Tabelle 1* gibt einen Überblick über die Formulierung der Items sowie Reliabilitäten und die deskriptive Statistik.

Tabelle 1 Emotionale Stimmung im Moment der Befragung

	M	SD
Emotionale Stimmung im Moment der Befragung insgesamt (α = .787)	3,12	,501
interessiert	3,07	1,071
verärgert (*umcodiert)*	3,05	1,410
freudig erregt	1,47	,877
wach	2,96	1,135
stark	2,79	1,080
beschämt (*umcodiert)*	3,83	1,336
begeistert	1,50	,889
schuldig (*umcodiert)*	4,66	,760
entschlossen	2,89	1,154
feindselig (*umcodiert)*	3,63	1,290
aufmerksam	3,35	,864
gereizt (*umcodiert)*	3,26	1,364
angeregt	2,19	1,045
ängstlich (*umcodiert)*	4,30	1,040
aktiv	2,46	1,073
bekümmert (*umcodiert)*	3,44	1,285
stolz	1,88	1,132
erschrocken (*umcodiert)*	3,33	1,385
durcheinander (*umcodiert)*	4,05	1,065
nervös (*umcodiert)*	4,33	,974
(1 = gar nicht, 5 = sehr)		

n = 205

Bereitschaft zur Counterspeech

Die Operationalisierung des Konstrukts der Counterspeech orientiert sich an der Methodik von Obermaier et al. (2021). Da diese Methode schon in anderen Studien (Leonhard et al., 2018; Obermaier et al., 2016) erfolgreich angewendet wurde, bietet sie eine fundierte Basis zur Messung des Konstruktes der Bereitschaft zur Counterspeech. Mittels eines Fragebogens wurden sechs Items auf einer fünfstufigen Likert-Skala abgefragt, welche

in *Tabelle 2* zu finden sind. Der Fragebogen enthält Items, welche die Bereitschaft zur Gegenrede und zur konstruktiven Intervention messen. Zwei Items wurden in diesem Kontext umkodiert, um eine präzise Messung zu gewährleisten und die Konsistenz der Antworten zu überprüfen und sicherzustellen, dass die Reaktionen der Teilnehmenden korrekt erfasst werden. Dies ermöglichte eine differenzierte Analyse der Einstellung der Teilnehmenden gegenüber der Counterspeech auf hasserfüllte Kommentare. Analog zu den Arbeiten von Obermaier et al. (2021) und Leonhard et al. (2018) bestätigt der errechnete Cronbach´s Alpha-Wert eine zufriedenstellende Reliabilität von $\alpha = .775$. Dies bedeutet, dass die Items konsistent dasselbe zugrunde liegende Konstrukt messen und das Ergebnis belegt, dass es sich um ein valides Instrument zur Erfassung der Bereitschaft zur Counterspeech handelt.

Tabelle 2 Bereitschaft zur Counterspeech

	M	SD
Bereitschaft zur Counterspeech insgesamt (α = .775)	2,60	,954
Ich fühle mich persönlich verpflichtet, den Kommentaren zu widersprechen.	2,86	1,453
Ich sehe es nicht als meine persönliche Pflicht an, einzugreifen. (*umcodiert)*	3,00	1,304
Ich schreibe einen Kommentar, der den Kommentaren sachlich widerspricht.	2,05	1,257
Ich verfasse keinen Kommentar, um Personen zu unterstützen. (*umcodiert)*	3,11	1,590
Ich widerlege den Kommentar mit sachlichen Argumenten.	2,34	1,354
Ich verfasse einen Kommentar, der die Aussagen klar und direkt kritisiert.	2,27	1,365
(1 = gar nicht, 5 = sehr)		

n = 205

Misogyne Einstellungen

Für die Erfassung von misogynen Einstellungen wurde der Gender-Blind-Sexism-Inventory (GBSI; Stoll et al., 2016) verwendet. Dieser wurde auf der Bais bestehender Modelle zum zeitgenössischen Sexismus wie Modern Sexism (Swim et al., 1995), Neosexism (Tougas et al., 1995) und Ambivalent Sexism (Glick & Fiske, 1996) entwickelt. Er zielt darauf ab, vier zentrale Dimensionen von gender-blindem Sexismus zu messen. Der GBSI

umfasst eine 12-Item Skala, die in vier Subskalen mit je drei Items unterteilt ist, welche die (1) Abstract Liberalism, (2) Naturalization, (3) Cultural Sexism und (4) Minimization of Sexism abdecken. Hinter dem Konzept des „Abstract Liberalism“ steht die Überzeugung, dass Gleichheit durch Prinzipien wie Meritokratie und Individualismus erfolgen kann. Soziale Unterschiede werden als natürliche Konsequenzen von individuellen Leistungen angesehen, wodurch strukturelle Benachteiligungen nicht berücksichtigt werden. Die Dimension „Naturalization“ stützt sich auf die Annahme, dass soziale Unterschiede „natürlich“ bzw. „biologisch“ sind und nicht aus Diskriminierung resultieren. „Cultural Sexism“ interpretiert die Unterschiede zwischen Geschlechtern als Resultat sozialer Prozesse, die stereotype Vorstellungen von Geschlechterrollen rechtfertigen. Die Dimension „Minimization of Sexism“ sieht Geschlechterungleichheiten als Problem der Vergangenheit an und ignoriert institutionellen Sexismus als mögliche Erklärung von Ungleichheiten. Mit einem Cronbach's Alpha von .80 konnten Soll et al. (2016) eine hohe Konsistenz feststellen. Für diese Studie wurden die Reliabilitäten einzeln für jede Dimension überprüft. Die Werte für die Dimensionen „Abstract Liberalism“ sowie „Cultural Sexism“ lagen dabei unter dem Schwellenwert von $\alpha > 0{,}7$. Das Gesamtkonstrukt konnte jedoch einen positiven Cronbach’s Alpha-Wert von $\alpha = 0{,}84$ aufweisen. Dies weist auf eine akzeptable interne Konsistenz hin (Koch et al., 2019). Der GBSI wurde mittels eines Fragebogens operationalisiert, welcher mit einer fünfstufigen Likert-Skala abgefragt wurde. *Tabelle 3* zeigt einen Überblick über die Mittelwerte, Standardabweichungen und Reliabilitäten der einzelnen Dimensionen.

Tabelle 3 Misogyne Einstellungen

	M	SD
Misogyne Einstellungen insgesamt (α = ,84)	2,05	,640
Abstract Liberalism (α = ,685)	1,57	,746
Wenn Frauen weniger verdienen als Männer, liegt das daran, dass sie andere Entscheidungen treffen.	1,67	,948
Gleichstellungsmaßnahmen bevorzugen Frauen auf Kosten der Männer.	1,51	,963
Wenn eine politische Maßnahme speziell Frauen unterstützt, ist das gegenüber Männern unfair.	1,53	,947
(1 = stimme gar nicht zu, 5 = stimme voll zu)		
Naturalization (α = ,753)	2,37	,746
Frauen sind von Natur aus emotionaler als Männer.	2,42	1,284
Männer sind von Natur aus aggressiver als Frauen.	2,84	1,248
Die biologischen Unterschiede zwischen Männern und Frauen sind wichtiger als die Erziehung, um ihre Verhaltensunterschiede zu erklären.	1,84	1,050
(1 = stimme gar nicht zu, 5 = stimme voll zu)		
Cultural Sexism (α = ,402)	1,95	,677
Es ist wichtig, Kindern beizubringen, sich entsprechend ihres Geschlechts zu verhalten.	1,42	,852
Es wird eher akzeptiert, wenn ein Mädchen sich wie ein Wildfang verhält, als wenn ein Junge sich wie ein „Weichei“ benimmt.	3,09	1,312
Es ist besser, Mädchen zu pflegenden und Jungen zu versorgenden Rollen zu erziehen, als umgekehrt.	1,35	,755
Minimization of Sexism (α = .752)	2,31	,887
Sexismus ist in der heutigen Gesellschaft kein großes Problem mehr.	1,59	,944
Heutzutage haben Frauen dieselben Chancen wie Männer.	2,09	1,112
Die Geschlechterungleichheit ist nicht mehr so schlimm wie früher.	3,25	1,186
(1 = stimme gar nicht zu, 5 = stimme voll zu)		

n = 205

4.4 Fragebogen

Der Fragebogen, der im Rahmen dieser Studie verwendet wurde, umfasste neben der Erhebung demografischer Merkmale auch die Messung zentraler Konstrukte. Die

demografischen Fragen erfassten unter anderem Alter, Geschlecht und Bildungsniveau. Besonders das Geschlecht wurde als wichtiger Moderator berücksichtigt, da es potenziell Einfluss auf die Wahrnehmung der Stimuli haben könnte. Nach der Datenschutzbelehrung und der Erhebung der demografischen Informationen wurde den Teilnehmenden zufällig eine von zwei Stimulus-Bedingungen zugewiesen: Die Experimentalgruppe erhielt ein Posting mit humorvoll-impliziten Kommentaren und die Kontrollgruppe mit expliziten Kommentaren. Die Anweisung lautete, das zugewiesene Posting sorgfältig und aufmerksam zu lesen. Im Anschluss wurde ein Manipulationscheck durchgeführt. Dieser dient nicht nur der Verifizierung der Manipulation, sondern auch der Messung der Wahrnehmung der unterschiedlichen Stile der Postings durch die Teilnehmenden. Der Manipulationscheck stellt sicher, dass die Teilnehmenden den Unterschied zwischen expliziten und humorvoll-impliziten Kommentaren korrekt erkennen können. Nach dem Manipulationscheck wurden die Hauptvariablen der Studie erhoben. Zunächst wurden die emotionalen Reaktionen der Teilnehmenden auf die gelesenen Kommentare erfasst. Danach wurde das Konstrukt der Bereitschaft der Counterspeech für die Ermittlung der Interventionstendenz gemessen. Anschließend erfolgte eine Abfrage der sexistischen Einstellungen, um die Haltung über geschlechtsspezifische misogyne Vorurteile beurteilen zu können. Zum Schluss erfolgte eine ausführliches Debriefing, welches detaillierte Beschreibung des Experiments sowie eine Erklärung des Zwecks und der Durchführung beinhaltete. Das Debriefing dient dazu, den Teilnehmenden alle relevanten Informationen über die Studie bereitzustellen, eventuelle Fragen zu klären und Unterstützung anzubieten.

4.5 Pretest

Ein Pretest ist entscheidend für die Sicherstellung der internen Validität und die Konsistenz der Ergebnisse in wissenschaftlichen Studien (Koch et al., 2019). Vier Personen

durchliefen den Pretest vor der eigentlichen Datenerhebung. Das Feedback hilft dabei, Fehler und Unklarheiten im Studienmaterial zu identifizieren und zu beheben (Koch et al., 2019). Die Teilnehmer:innen wurden von der Hauptstudie ausgeschlossen, um die Integrität der Ergebnisse zu wahren. Der Pretest deckte mehrere Rechtschreibfehler und unklare Formulierungen auf. Beispielsweise wurde die Frage „*Geben Sie bitte an, wie Sie sich im Moment allgemein fühlen*“ als zu ungenau empfunden und auf „*Geben Sie bitte an, wie Sie sich in diesem Moment fühlen*“ geändert, um mehr Klarheit zu schaffen. Zudem wurde die Lesbarkeit durch Anpassungen der Schriftgröße und Formatierung verbessert. Diese Änderungen sorgten dafür, dass die Fragen verständlicher waren und das Layout leichter lesbar. Zudem bestätigte der Pretest, dass die Manipulation des Stimulusmaterials erfolgreich ist. Die Teilnehmenden nahmen diese wie beabsichtigt wahr, was die Wirksamkeit des experimentellen Designs bestätigte. Diese Anpassungen und die Bestätigung der Manipulation gewährleisteten, dass das finale Befragungsmaterial präzise und benutzerfreundlich ist und die beabsichtigte Wirkung erzielt wird.

5. Statistische Analyse

In diesem Kapitel werden die statistischen Auswertungen der Studie beschrieben. Dazu zählen der Manipulations- und Randomisierungscheck zur Prüfung der experimentellen Bedingungen sowie die Überprüfung und Auswertung der Hypothesen anhand geeigneter statistischer Verfahren.

5.1 Manipulationscheck

Zur Überprüfung der Frage, ob sich die Stimuli hinreichend voneinander abgrenzen und somit eine Manipulation erfolgreich war, wurde ein Manipulationscheck durchgeführt. Dieser wurde in den Fragebogen integriert und mithilfe eines t-Tests für unabhängige Stichproben in SPSS ausgewertet. Dazu wurde in der Befragung mittels vier Items mit einer Likert-Skala abgefragt, ob die gelesenen Kommentare eindeutig hasserfüllt waren oder humorvoll bzw. satirisch gemeint. Bei der Auswertung wurden die Items des Manipulationstests nach einer bestandenen Reliabilitätsprüfung (α = ,788) zu einer Variable zusammengefasst.

Darauf folgte ein t-Test für unabhängige Gruppen, bei dem die Mittelwerte für den expliziten und den humorvollen Stimulus miteinander verglichen wurden. Die Richtung der Effektstärke ist anhand der Formulierung der Items zu interpretieren. Ein hoher Wert weist auf eine starke Zustimmung hin, dass die Aussage als humorvoll oder sarkastisch interpretiert wird. Da sich die explizite (M = 1.238, SD = 0.3854) und die implizite Hatespeech (M = 2.049, SD = 0.7780) signifikant (t (147.11) = 9.453, p = 0.000) voneinander unterschieden, hat das Stimulus-Material den Manipulationscheck bestanden.

5.2 Randomisierungscheck

Um sicherzustellen, dass die Störvariablen gleichmäßig verteilt sind und strukturell äquivalente Versuchsgruppen entstanden sind, wurde ein Randomisierungscheck durchgeführt (Koch et al., 2019). Dieser dient dazu systematische Verzerrungen und alternative Erklärungen auszuschließen (Koch et al., 2019). Für den Randomisierungscheck wurden bekannte Variablen, hier Geschlecht und Alter, untersucht, um eine gleichmäßige Verteilung zu gewährleisten. Zur Berechnung wurden für die jeweiligen Variablen zwei t-Tests bei unabhängigen Stichproben durchgeführt. Die Experimentalgruppe, welche Hatespeech mit subtilen, humorvollen Kommentaren erhielt (M = 1.32; SD = 0.498) unterscheidet sich von der Kontrollgruppe, die explizite Hatespeech gelesen hatte, in ihrem Geschlecht nicht signifikant (t (204) = 1.090, p = 0.277). Zusätzlich konnten keine signifikanten Unterschiede (t (204) = -0.110, p = 0,913) bei dem Alter der Proband:innen zwischen der Experimentalgruppe (M = 3.90; SD = 1.397) und der Kontrollgruppe (M = 3.92; SD = 1.370) festgestellt werden. Die Randomisierung der Gruppen war somit erfolgreich und sie sind in ihren Merkmalen homogen.

5.3 Überprüfung der Hypothesen 1.1 bis 1.3

Die Überprüfung der Hypothesen *H 1.1, H 1.2* sowie *H 1.3* zielt darauf ab, den Einfluss von Hatespeech mit humorvollen Elementen auf die Rezeption und das Verhalten von Nutzer:innen zu untersuchen. Zur Hypothesenprüfung wird ein t-Test für unabhängige Stichproben als Analysemethode verwendet, wobei zwei UVs (implizite/humorvoll-gemeinte und explizite Hatespeech) hinsichtlich ihrer Mittelwertunterschiede verglichen werden. Eine zentrale Voraussetzung für dieses Verfahren ist die Normalverteilung der Daten (Kubinger et al., 2009). Da jedoch bei einer Stichprobengröße von n > 30 je Gruppe von einer Normalverteilung ausgegangen wird, kann in dieser Analyse aufgrund der ausreichend

großen Stichprobe (n = 205) auf die Überprüfung dieser Voraussetzung verzichtet werden (Kubinger et al., 2009). Eine weitere Voraussetzung für den t-Test ist die Homogenität der Varianzen, welche mittels des Levene-Tests geprüft wird (Kubinger et al., 2009). Varianzhomogenität kann angenommen werden, wenn der Levene-Test nicht signifikant ausfällt ($p > 0,05$). Bei Varianzheterogenität ($p < 0,05$) werden die Ergebnisse des Welch-Tests (t-Test mit Welch-Korrektur) berichtet (Ortmann, 2023).

H 1.1: Explizite, misogyne Hatespeech führt zu einer negativeren Stimmung als humoristische, misogyne Hatespeech.

Zur Überprüfung von *H 1.1* wurde ein t-Test für unabhängige Stichproben durchgeführt, wobei AV die emotionale Stimmung und die UV die Explizität der Hatespeech war. Die deskriptive Analyse zeigt, dass es keine großen Unterschiede in der emotionalen Belastung zwischen der impliziten Hatespeech mit humorvollen Elementen ($M = 3.152$, $SD = 0.502$) und der expliziten Hassrede ($M = 3.094$, $SD = 0.5011$) gibt. Ob der Unterschied dennoch signifikant ist, wird durch den t-Test bestimmt. Da die Varianzhomogenität eine Voraussetzung für den t-Test darstellt, wurde zunächst der Levene-Test herangezogen. Der Levene-Test zeigt, dass die Varianzhomogenität nicht angenommen werden kann ($F = 0.83$, $p = 0.774$). Daher wurde der Welch-Test für ungleiche Varianzen verwendet. Die Ergebnisse des t-Tests zeigten jedoch keinen signifikanten Unterschied zwischen den beiden Gruppen ($t(203) = .831$, $p = 0.407$). Demnach unterscheidet sich die emotionale Stimmung von expliziter und impliziter Hatespeech nicht signifikant. *H 1.1* wird somit verworfen.

H 1.2: Explizite, misogyne Hatespeech führt zu einer höheren Bereitschaft von Counterspeech als humoristische, misogyne Hatespeech.

Zur Überprüfung der *H 1.2* wurde ebenfalls ein t-Test für unabhängige Stichproben durchgeführt, wobei die AV die Counterspeech darstellt. Die deskriptive Analyse zeigt, dass

es ebenfalls keine großen Differenzen in der Bereitschaft von Counterspeech zwischen der impliziten Hatespeech mit humorvollen Elementen (M = 2.498, SD = 0.8529) und der expliziten Hassrede (M = 2.707, SD = 1.0375) gibt. Ob der Unterschied dennoch signifikant ist, wird durch den t-Test bestimmt. Der Levene-Test zeigt, dass die Varianzhomogenität angenommen werden kann (F = 7.895, p = 0.005). Die Ergebnisse des t-Tests verdeutlichen jedoch keinen signifikanten Unterschied zwischen den beiden Gruppen (t(203) = -1.568, p = 0.118). Demnach unterscheidet sich die Bereitschaft zur Counterspeech von expliziter und impliziter Hatespeech nicht signifikant. *H 1.2* kann nicht bestätigt werden.

H 1.3: Explizite, misogyne Hatespeech führt zu einer verstärkten miogynen Einstellung als humoristische, misogyne Hatespeech.

Die misogynen Einstellungen wurden anhand der Dimensionen „Abstract Liberalism“, „Naturalization“, „Cultural Sexism“ und „Minimization of Sexism“ erhoben. Um *H 1.3* zu testen, wurde ein t-Test für unabhängige Stichproben mit der Variable der misogynen Einstellungen (= AV) und Explizität der Hatespeech (= UV) durchgeführt. Dabei lässt sich kein signifikanter Unterschied zwischen den verschiedenen Gruppen erkennen (t(203) = .415, p = .678). Entgegen der Vermutung beeinflusst explizite Hatespeech (M = 2.03, SD = 0.6113) die sexistischen Einstellungen von Rezipierenden nicht signifikant positiv im Vergleich zur impliziten Hatespeech mit humorvollen Elementen (M = 2.07, SD = 0.6713). Somit muss *H 1.3* verworfen werden.

5.4 Überprüfung der Hypothesen 2.1 bis 2.3

Die Überprüfung der Hypothesen *H 2.1*, *H 2.2* und *H 2.3* zielt darauf ab, den Einfluss von Hatespeech mit humorvollen Elementen auf die Rezeption und das Verhalten von Nutzer:innen zu untersuchen. Insbesondere wird analysiert, inwieweit das Geschlecht als Moderator die Wirkung von expliziter, misogyner Hatespeech auf die emotionale Stimmung,

die Bereitschaft zur Counterspeech und die eigenen misogynen Einstellungen beeinflusst. Zur Überprüfung dieser Hypothesen wird eine Moderationsanalyse durchgeführt. Dabei werden die Effekte der expliziten, misogynen Hatespeech auf die genannten abhängigen Variablen in Abhängigkeit vom Geschlecht der Teilnehmenden untersucht. Eine zentrale Voraussetzung für die Durchführung der Moderationsanalyse ist die Normalverteilung der Daten (Kubinger et al., 2009; Hayes, 2013). Bei einer Stichprobengröße von n > 30 je Gruppe kann jedoch in der Regel von einer Normalverteilung ausgegangen werden. In dieser Analyse wird aufgrund der ausreichend großen Stichprobe (n = 205) auf die Überprüfung dieser Voraussetzung verzichtet, sodass die Normalverteilung als gegeben angenommen werden kann (Kubinger et al., 2009; Hayes, 2013). Des Weiteren wird eine lineare Beziehung zwischen den Variablen vorausgesetzt (Hayes, 2013). Diese Linearität wurde durch eine visuelle Inspektion des Streudiagramms mit LOESS-Glättung überprüft, die ein etwa lineares Verhältnis der Variablen zeigte. Die Moderationsanalysen wurden mit dem PROCESS-Makro von Hayes (2013) durchgeführt, das lineare Regressionen mittels der Methode der kleinsten Quadrate berechnet und unstandardisierte Koeffizienten liefert. Zur Berechnung der Konfidenzintervalle wurde Bootstrapping mit 5000 Iterationen eingesetzt, ergänzt durch heteroskedastizitätskonsistente Standardfehler (Chambers et al., 1993).

H 2.1: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und einer negativen, emotionalen Stimmung. Frauen zeigen tendenziell eine stärkere negative, emotionale Reaktion auf explizite, misogyne Hatespeech im Vergleich zu humorvoller Hatespeech, während bei Männern dieser Unterschied weniger ausgeprägt ist.

Eine Moderationsanalyse wurde durchgeführt, um zu untersuchen, ob das Geschlecht die Beziehung zwischen expliziter, misogyner Hatespeech und der negativen Stimmung signifikant moderiert. Die Ergebnisse zeigten keinen signifikanten Moderationseffekt des

Geschlechts auf den Zusammenhang zwischen misogyner Hatespeech und negativer Stimmung, $\Delta R^2 = 0.17$ %, $F(1, 201) = 0.297$, $p = .586$, 95% CI [-0.2458, 0.3771]. Gemäß den Empfehlungen von Hayes (2018) wurde der Interaktionsterm im Modell belassen, da er keine signifikante Veränderung des erklärten Varianzanteils bewirkte. Zudem wiesen sowohl das Geschlecht, $B = 0.059$, $p = .421$, als auch die Hatespeech, $B = -0.054$, $p = 0.446$, keine signifikanten Haupteffekte auf die negative Stimmung auf. *H 2.1* kann somit nicht bestätigt werden.

H 2.2: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und der Bereitschaft zur Counterspeech. Frauen zeigen tendenziell eine größere Bereitschaft zur Gegenrede, insbesondere bei expliziter, misogyner Hatespeech, während dieser Effekt bei Männern weniger ausgeprägt ist.

Um zu untersuchen, ob das Geschlecht den Zusammenhang zwischen expliziter, misogyner Hatespeech und der Bereitschaft zur Counterspeech moderiert, wurde eine Moderationsanalyse durchgeführt. Die Ergebnisse zeigten, dass kein signifikanter Moderationseffekt des Geschlechts auf diesen Zusammenhang besteht, $\Delta R^2 = 0{,}06\%$, $F(1, 201) = 0{,}161$, $p = .689$, 95% CI [-0.4034, 0.5456]. Gemäß den Empfehlungen von Hayes (2018) wurde der Interaktionseffekt im Modell belassen, da er die erklärte Varianz nicht signifikant beeinflusste. Die Haupteffekte des Geschlechts ($B = 0{,}175$, $p = .205$) und der Hatespeech ($B = 0{,}222$, $p = .097$) waren ebenfalls nicht signifikant. Somit kann *H 2.2* nicht nachgewiesen werden.

H 2.3: Das Geschlecht moderiert den Zusammenhang zwischen der Art der Hatespeech (explizit vs. humorvoll) und der Verstärkung sexistischer Einstellungen. Frauen sind tendenziell stärker von der Hassrede betroffen, insbesondere bei expliziter, misogyner

Hatespeech, und zeigen eine geringere Zustimmung zu sexistischen Stereotypen im Vergleich zu Männern.

Eine Moderationsanalyse wurde durchgeführt, um zu prüfen, ob das Geschlecht den Zusammenhang zwischen expliziter misogyner Hatespeech und der Verstärkung misogyner Einstellungen moderiert. Die Analyse ergab, dass das Geschlecht keinen signifikanten Moderationseffekt auf diese Beziehung hat, $\Delta R^2 = 2{,}35\%$, $F(1, 201) = 4{,}53$, $p = .035$, 95% CI [-0.798, -0.0528]. Im Einklang mit den Empfehlungen von Hayes wurde der Interaktionseffekt im Modell belassen, da eine leichte, aber signifikante Veränderung in der erklärten Varianz festgestellt wurde. Die Haupteffekte zeigten, dass weder die Hatespeech ($B = -0{,}021$, $p = .815$) noch das Geschlecht ($B = 0{,}213$, $p = .021$) einzeln signifikante Einflüsse auf die Verstärkung misogyner Einstellungen hatten. *H 2.3* konnte somit nicht verifiziert werden.

6. Diskussion

Im Onlinebereich ist misogyne Hassrede eine Bedrohung für betroffene Personen sowie das gesellschaftliche Stimmungsbild. Insbesondere Frauen sind Ziel von diskriminierender Hatespeech (FRA, 2023). Die Präsenz von Hassrede in digitalen Räumen ist ein zunehmendes Problem. Resultate früherer Experimentaluntersuchungen belegen, dass der Konsum hasserfüllter Kommentare nicht nur das Wohlbefinden der adressierten Personen schadet (Zochniak et al., 2023), sondern auch zu negativen Emotionen bei unbeteiligten Rezipient:innen sowie zur stärkeren Verfestigung stereotyper Muster führen kann (Hsueh et al., 2015; Schäfer et al., 2022). Munn (2019) beobachtet, dass der Konsum ironischer, aber diskriminierender radikaler Inhalte oft den Beginn von Radikalisierungsprozessen darstellt. Dabei sind die Inhalte zwar oft nicht strafrechtlich relevant, können aber ebenso schädliche Auswirkungen haben (Matamoros-Fernández et al., 2023). Da Plattformen solche Äußerungen ungleichmäßig regulieren (Gillett et al., 2022) und sie schwieriger zu identifizieren sind als explizite Hatespeech (Schmid et al., 2024), bergen sie eine besondere Gefahr. Es kann zu Normalisierung, Gewöhnung und Desensibilisierung der Gesellschaft gegenüber diskriminierenden Inhalten führen (Bilewicz & Soral, 2020; Soral et al., 2018).

Ziel der vorliegenden Untersuchung ist es auf Basis des bisherigen Forschungsstandes, festzustellen, wie explizite, misogyne Hatespeech im Vergleich zu humorvoll-subtil, misogyn motivierter Hassrede auf die emotionale Befindlichkeit, die Bereitschaft zur Counterspeech sowie die Unterstützung sexistischer Stereotype wirken. Die anvisierte Frage lautet: Unterscheiden sich explizite und subtile Formen der Hassrede hinsichtlich ihrer Rezeption durch die Nutzer:innen voneinander? Wobei das Geschlecht als Moderator beachtet wird. Obgleich die Unterschiede zwischen den beiden Erscheinungsformen der Hassrede im Hinblick auf die Wirkungen auf die Nutzer:innen

methodisch präzise im Blick behalten wurden, resultierte aus den Ergebnissen kein signifikanter Unterschied. Das Resultat widerspricht einem Großteil der Forschung (Schmid et al., 2024; Obermaier et al., 2023; Kenski et al., 2020; Schmid, 2023; Leonhard et al., 2018; Fox et al., 2015). Die Befunde nähern sich Forschungen an, die unterstreichen, dass implizite Hatespeech analog zu expliziter Hatespeech wahrgenommen wird. Die Ergebnisse von Weber et al. (2020) zeigen, dass sowohl subtile als auch explizite Formen von Hatespeech schädigenden Einfluss auf Rezipierende haben können. Auch Schmid et al. (2024) konnten innerhalb einer qualitativen Forschung Anzeichen feststellen, dass implizite, insbesondere humorvolle misogyne Hatespeech, als negativ wahrgenommen wird. Vor allem, nachdem die Proband:innen die Hassrede hinter dem vermeintlichen Humor entdeckten, bewerteten sie diese kritisch.

Dieses Ergebnis legt nahe, dass beide Formen von Hassrede – unabhängig von ihrer Erscheinungsweise oder dem Geschlecht der Rezipierenden – ähnlich wahrgenommen werden und vergleichbare Auswirkungen haben können. Es unterstreicht die Relevanz, auch subtilen Formen von HassredeAufmerksamkeit zu widmen, da diese möglicherweise ebenso schädlich wie explizite Hassrede sind. Die fehlenden Unterschiede könnten darauf hinweisen, dass implizite Inhalte nicht weniger problematisch sind, sondern lediglich anders verarbeitet werden, was im Kontext der Normalisierung und Desensibilisierung gegenüber Feindseligkeit besonders kritisch betrachtet werden muss. Im Folgenden wird ein Fokus auf die Diskussion der einzelnen Hypothesen gelegt.

H 1.1

Es wurde die Hypothese aufgestellt, dass explizite Hatespeech aufgrund einer bedrohlicheren Wahrnehmung zu einer schlechteren emotionalen Stimmung führt als subtile, humorvolle Hatespeech. Entgegen dieser Erwartung wurden keine signifikanten Unterschiede

hinsichtlich der emotionalen Reaktionen zwischen den Versuchsgruppen festgestellt. Das widerspricht bestehenden Studien (Schmid, 2023; Zochniak et al., 2023). Beide Versuchsgruppen scheinen die Formen von Hassrede ähnlich wahrgenommen zu haben, wodurch es zu keinen Unterschieden in der emotionalen Reaktion kam. Möglicherweise sind die Unterschiede in der Wahrnehmung zwischen der expliziten und der subtilen, humorvollen Form von Hatespeech nicht so ausgeprägt, wie das die Literatur bisher suggeriert. Obwohl Studien wie die von Schmid (2023) festgestellt haben, dass humorvolle Hassrede als weniger diskriminierend erkannt wird, könnten die Proband:innen in dieser Studie beide Formen durch den misogynen Charakter gleich eingestuft haben. Obermaier et al. (2023) zeigen, dass insbesondere misogyne Hatespeech als diskriminierend eingestuft wird im Gegensatz zu anderen Betroffenengruppen. Offenbar waren die Teilnehmer:innen hinsichtlich misogyner Inhalte sensibilisiert, sodass sie den diskriminierenden Kern hinter dem Humor schnell erkannt haben, wodurch eine vergleichbare Reaktion zwischen den Versuchsgruppen gemessen werden konnte. Die Resultate schließen sich der Forschung von Schmid et al. (2024) an. Dabei bewerteten einige Proband:innen misogyne, humorvolle Hassrede negativ, sobald sie die feindseligen Absichten dahinter erkannten. Zusätzlich könnte der fehlende signifikante Unterschied darauf zurückzuführen sein, dass das manipulative Potenzial von Humor überschätzt wird. Die Rezeption von expliziter und humorvoller Hatespeech könnte zwar zu einer unterschiedlichen Wahrnehmung führen, dadurch dass beide aber diskriminierende Aussagen enthalten, könnten die emotionalen Reaktionen für beide Formen gleich sein. Müller und Lopez-Sanchez (2021) kommen zu dem Ergebnis, dass die individuelle psychische Resilienz emotionale Reaktionen beeinflusst. Dabei sind Personen mit höherer Resilienz, welche laut den demografischen Daten in dieser Studie wahrscheinlich vertreten waren, besser in der Lage negative bzw. belastende Inhalte zu bewältigen. Die Personen sind zwar von den Inhalten emotional betroffen, können ihre Gefühle aber besser

regulieren. Das Resultat der Studie zeigt, dass die Inhalte als belastend wahrgenommen wurden. Dies könnte darauf hindeuten, dass spezifische Inhalte – unabhängig davon, ob sie explizit oder subtil-humorvoll präsentiert werden – selbst bei resilienten Personen als stark emotional empfunden wurde, insbesondere wenn sie klare diskriminierende Botschaften enthalten. Die psychische Resilienz könnte dazu beigetragen haben, dass die Teilnehmer:innen in ihrer Gesamtwahrnehmung nicht stärker zwischen den beiden Formen der Hassrede differenzierten. Beide Stimuli wurden inhaltlich möglicherweise als moralisch und emotional belastend wahrgenommen, wodurch Resilienz die emotionalen Reaktionen eher auf einem moderaten Niveau hielt und so die Unterschiede zwischen expliziter und subtiler Hassrede überdeckte.

H 1.2

Die Hypothese, dass explizite Hatespeech zu einer höheren Bereitschaft zur Gegenrede führt, wurde ebenfalls nicht bestätigt. Es gab keinen signifikanten Unterschied in der Counterspeech-Bereitschaft zwischen den beiden Versuchsgruppen. Dies könnte darauf hindeuten, dass die Art und Weise, wie Hatespeech formuliert wird (explizit oder humorvoll), keinen signifikanten Einfluss auf das Aktivwerden von Rezipierenden hat. Menschen reagieren möglicherweise unabhängig vom Tonfall der Hassrede gleich stark oder schwach auf sie, wenn es darum geht, Gegenrede zu leisten. Dies könnte darauf zurückzuführen sein, dass humorvolle Hassrede, ähnlich wie die in Obermaier et al. (2021) beschriebene unhöfliche und diskriminierende Hatespeech, bereits die Diskursnormen in einer demokratischen Gesellschaft verletzt. Solche Verstöße rufen möglicherweise bereits ein Gefühl persönlicher Verantwortung bei den Rezipient:innen hervor, wodurch der Bystander-Interventionsprozess ausgelöst wird. Wie Obermaier et al. (2023) verdeutlichen, hängt das Eingreifen von Bystandern nicht automatisch vom bloßen Lesen von Hassrede ab, sondern vom Ausmaß, in dem das Gelesene als Angriff wahrgenommen wird, was wiederum die

Bereitschaft zur Intervention erhöht. Die Teilnehmer:innen könnten sensibilisiert genug sein, so dass auch die humorvolle Hatespeech diesen Prozess auslöste. Eine eher homogene Zusammensetzung der Versuchspersonen, die insbesondere durch ein hohes Bildungsniveau geprägt war, könnte die Ergebnisse dieser Studie beeinflusst haben. Außerdem könnten viele der Teilnehmer:innen möglicherweise den Nutzen von Counterspeech als individuelle Reaktion auf Hassrede infrage gestellt haben. In Übereinstimmung mit den Kommentaren einiger Proband:innen, die angaben, dass Counterspeech wirkungslos sei, könnte dies darauf hindeuten, dass das Problem der Hassrede von vielen als Aufgabe der Politik oder der Plattformbetreibenden und nicht des Individuums betrachtet wird. Diese Haltung könnte erklären, warum die erwarteten Unterschiede in der Bereitschaft zur Intervention ausblieben. Wie Ping et al. (2024) feststellen, sind frühere Opfererfahrungen ein entscheidender Prädiktor für die Teilnahme an Counterspeech. Personen, die keine direkten Erfahrungen als Opfer von Hassrede gemacht haben, könnten weniger motiviert sein, sich aktiv an Gegenmaßnahmen zu beteiligen. Dies könnte einen weiteren Grund dafür darstellen, warum die Interventionsbereitschaft in dieser Studie gering ausfiel.

H 1.3

Andere Studien konnten feststellen, dass der Kontakt mit diskriminierenden Inhalten, insbesondere expliziter Hatespeech, negative Einstellungen gegenüber der angegriffenen Gruppe fördert (Schäfer et al., 2022; Weber et al., 2020; Soral et al., 2018). An diese Resultate kann die vorliegende Studie nicht anknüpfen. Die Hypothese, dass explizite Hatespeech zu misogyneren Einstellungen als subtile, humorvolle Hassrede führt, kann nicht angenommen werden. Es ergab sich kein signifikanter Unterschied zwischen den Versuchsgruppen, was bedeutet, dass weder explizite noch humorvoll, implizite Hatespeech einen stärkeren Einfluss auf die misogyne Einstellungen der Rezipierenden hatte. Eine Erklärung für das Ergebnis liegt in der generellen niedrigen Zustimmung zu misogynen

Einstellungen durch die Proband:innen. Es ist anzunehmen, dass dies im Zusammenhang mit den demografischen Merkmalen der Teilnehmenden steht. In allem bestanden die Versuchsgruppen aus überwiegend jungen, gebildeten Personen, die dadurch möglicherweise eine hohe Sensibilität gegenüber Phänomenen wie Sexismus und Hatespeech aufweisen. Diese homogene Gruppe ist eventuell diskriminierenden Inhalten kritisch gegenüber eingestellt und dadurch nicht anfällig - ob explizit oder implizite Inhalte - die eigenen Einstellungen zu verändern, da diese stark ausgeprägt sind. So zeigen auch Williams et al. (2016), dass Erfahrungen, welche mit Hatespeech gemacht worden sind, die Sensibilität für ähnliche Inhalte erhöht. Die Versuchsgruppen haben sich möglicherweise in der Vergangenheit mit dem Phänomen der Hatespeech auseinandergesetzt oder sind ihm auf reflektierte Art und Weise begegnet. Dadurch wurden eine Sensibilität und Reflexionsfähigkeit aufgebaut, so dass die humorvollen Elemente den diskriminierenden Kern nicht verschleiern konnten. In heterogeneren Gruppen mit einem breiteren Bildungshintergrund oder politischen Einstellungen, könnten somit die Ergebnisse anders ausfallen. Paasch-Colberg et al. (2021) argumentieren, dass die wiederholte Auseinandersetzung mit indirekter Hatespeech über einen längeren Zeitraum über die Wirkung solcher Inhalte verstärken kann. Durch die experimentellen Rahmenbedingungen hatten die Proband:innen Zeit und Gelegenheit über ihre Inhalte nachzudenken und zu reflektieren, was in einem realen Medienumfeld nicht üblich ist. In der vorliegenden Studie hatten die Teilnehmer möglicherweise mehr Zeit, die Inhalte zu verarbeiten und ihre Emotionen einzuordnen, da sie aktiv, in einem kontrollierten Umfeld und nur einmalig mit diesen Inhalten konfrontiert wurden. Im Gegensatz dazu ist die Rezeption von Hatespeech in sozialen Medien oft flüchtiger und reaktiver, was zu einer weniger intensiven und reflexiven Auseinandersetzung führt (Koch et al, 2019). In den sozialen Medien konsumieren

Nutzer:innen Inhalte schnell und in einem konstanten Strom, was die Möglichkeit zur kritischen Reflexion einschränkt.

H 2.1 bis 2.3

Die zweite Forschungsfrage ging der Fragestellung nach, ob das Geschlecht als Moderator die Wirkung von expliziter bzw. humorvoller Hatespeech auf die Rezeption der Versuchspersonen hat. Analog zu der ersten Forschungsfrage wurden die Dimensionen emotionale Stimmung, die Bereitschaft zur Counterspeech und die Unterstützung misogyner Einstellungen als AVs untersucht. Hingegen anderer Forschungsarbeiten (Kenski et al., 2017; Zochniak et al., 2023; Williams et al., 2016; Obermaier et al., 2023; Hsueh et al., 2015; Fox et al., 2015), konnte in der vorliegenden Arbeit kein moderierender Effekt bei den drei Dimensionen nachgewiesen werden. Die Ergebnisse widersprechen damit früheren Studien und weisen keine Wirkung von dem Geschlecht bzw. der Zugehörigkeit der Betroffenengruppe auf die Rezeption von Hatespeech auf (Kenski et al., 2017; Zochniak et al., 2023; Williams et al., 2016; Obermaier et al., 2023; Hsueh et al., 2015; Fox et al., 2015). Der bisherige Forschungsstand ist bis lang davon ausgegangen, dass Personen, welche sich der diskriminierenden Gruppe nahe oder angehörig fühlen, eine andere Reaktion zeigen als Unbeteiligte (Soral et al; 2018, Fox et al., 2015, Schmid et al., 2024, Zochniak et al., 2023; Williams et al., 2016). Insbesondere misogyne Themen und Beiträge sind davon betroffen (Obermaier at al., 2023). Die Forschung zeigt, dass weiblich gelesene Personen sensibler gegenüber diskriminierenden Inhalten sind und diese unhöflicher bewerten (Kenski et al., 2017). Dabei sind sie Personen mit früheren Opfererfahrungen häufiger dazu bereit sich in Form von Counterspeech gegen Hatespeech auszusprechen (Ping et al, 2024). Entgegen diesen Erwartungen zeigte sich in der vorliegenden Studie kein signifikanter Einfluss des Geschlechts auf die emotionale Reaktion, die Unterstützung misogyner Stereotype oder die Bereitschaft zur Gegenrede. Mögliche Gründe dafür könnte wie bereits diskutiert in der

Homogenität der Stichprobe liegen. Eine vorhandene Sensibilisierung könnte dazu geführt haben, dass alle Geschlechter ähnlich kritisch auf die gezeigten Inhalte reagierten. Die Unterschiede zwischen den Geschlechtern könnten in einer weniger sensibilisierten bzw. homogenen Stichprobe bzw. breiteren ideologischen und kulturellen verankerten Stichprobe ausgeprägter sein. Eine hohe Medienkompetenz und eine kritische Einstellung gegenüber diskriminierenden Inhalten, kann nach Fox et al. (2015) zu einem schwächeren Effekt von Hassrede führen.

Ein weiterer Aspekt, der in Betracht gezogen werden muss, ist die Art der verwendeten Stimuli. Die Stimuli der vorliegenden Studie bestanden aus expliziter und humorvoll misogyn motivierter Hassrede, die möglicherweise von beiden Geschlechtern als gleich verletzend wahrgenommen wurden. Studien wie die von Schmid (2023) zeigen, dass humorvolle Hassrede oft subtiler und weniger als direkt bedrohlich wahrgenommen wird, was die erwarteten geschlechtsbedingten Unterschiede abschwächen könnte. Auch explizite Hassrede könnte in einem experimentellen Setting anders wahrgenommen werden als im realen Leben, da die Teilnehmer:innen in einem kontrollierten Umfeld möglicherweise bewusster und analytischer auf die Inhalte reagierten, was zu weniger emotionalen Unterschieden zwischen den Geschlechtern führte.

6.1 Limitationen und zukünftige Forschung

Entgegen der bestehen Literatur, wurden keine signifikanten Unterschiede zwischen den beiden Formen von Hassrede festgestellt. Im Folgenden werden deshalb Limitationen der Studie ausgehend von der Diskussion ausgezeigt, die fehlende signifikante Effekte erklären könnten.

Die Homogenität der Stichprobe stellt eine methodische Schwachstelle der Studie dar. Die Versuchspersonen waren überwiegend jung und gebildet und stammten so

möglicherweise aus einem ähnlichen sozialen Umfeld. Diese Zusammensetzung kann maßgeblich dafür verantwortlich sein, dass keine signifikanten Unterschiede zwischen den Gruppen auftraten. Ein hoher Bildungsgrad und eine Sensibilisierung innerhalb des Themenbereichs lässt Personen kritischer und reflektiert in Bezug zu Hatespeech handeln (Williams et al., 2016). Zukünftige Forschungen sollen deshalb auf die Heterogenität der Stichprobe achten und verschiedene Bildungshintergründe und politische Einstellungen innerhalb dieser abdecken. Dies könnte zu differenzierten Ergebnissen führen und zeigen, welche Gruppen unterschiedlich auf verschiedene Formen von Hatespeech reagieren. Unterschiedliche Moderatoren wie die soziale Schicht, politischen Einstellungen, Alter oder einer Unterscheidung zwischen Land- und Stadtbewohner:innen könnten wertvolle Erkenntnisse liefern wie Hassrede in verschiedenen Bevölkerungsgruppen wahrgenommen wird und wo Bildungsarbeit erforderlich ist. Zusätzlich könnte qualitative Forschung beispielsweise mit Fokusgruppen oder Interviews individuelle Faktoren besser erfassen.

Das experimentelle Setting kann als weitere Limitation betrachtet werden. Die Teilnehmer:innen reagierten in keinem realen und authentischen Social-Media-Umfeld auf die Hatespeech, sondern unter kontrollierten Bedingungen. Dadurch könnten Reaktionen anders ausgefallen sein, als dies im realen Setting wäre (Koch et al., 2019). In der realen Nutzung sozialer Medien werden Inhalte oft in einem überladenen und dynamischen Umfeld konsumiert, in dem die Rezipient:innen mit einer Vielzahl an Informationen konfrontiert werden und nicht die gleiche Fokussierung auf einen einzelnen Stimulus möglich ist (Koch et al., 2019). Daher sollten zukünftige Forschungen Hassrede in realistischeren Medienkontexten untersuchen. In realen Social-Media-Feeds oder in simulierten digitalen Umgebungen könnten andere Ergebnisse erzielt werden sowie herausgefunden werden, wie die Wahrnehmung und Wirkung von Hatespeech in einem authentischen Kontext ausfällt. Möglicherweise können gerade dann subtile Formen von Hatespeech weniger bewusst

wahrgenommen werden und dadurch andere Reaktionen ausgelöst werden. Die Ergebnisse solcher Studien können dazu beitragen, besser zu verstehen, wie Hatespeech im täglichen Medienkonsum die Einstellungen und Emotionen von Nutzenden beeinflusst.

Der Manipulationscheck könnte einen potenziellen Einfluss auf die Ergebnisse gehabt haben. Dieser wurde unmittelbar nach der Präsentation der Stimuli durchgeführt was ein Messartefakt verursacht haben könnte. Die Teilnehmenden waren somit schon vor der Abfrage der AVs gezwungen, die Stimuli bewusst zu reflektieren und zu kategorisieren. Dies hat möglicherweise dazu geführt, den diskriminierenden Kern hinter dem Humor schneller zu erkennen und deshalb auch darauf ähnlich zu der expliziten Hatespeech zu reagieren. Deshalb sollte in zukünftigen Studien der Manipulationscheck erst nach der Messung der emotionalen und kognitiven Reaktionen durchgeführt werden, um die spontanen und affektiven Reaktionen nicht durch bewusste Reflexion zu überlagern. Zudem könnte der Manipulationscheck in einem weniger direktiven Format erfolgen, um die natürliche Reaktion der Proband:innen nicht zu beeinträchtigen.

Diese Studie untersuchte lediglich kurzfristige Effekte von Hatespeech, was als Limitation betrachtet werden kann. Zwar wurden die unmittelbaren und kurzfristigen Reaktionen untersucht, jedoch könnte die subtile Form von Hatespeech über einen längeren Zeitraum eine andere Wahrnehmung und Reaktion der Versuchspersonen auslösen. Langzeitstudien, könnten das Ergebnis dieser Studie stützen und zeigen, ob diskriminierende Inhalte durch wiederholte Rezeption zu einer Normalisierung führen, wie es Paasch-Colberg et al. (2021) nahelegen. Diese argumentieren, dass indirekte Hassrede zu langfristigen Medieneffekten führen kann und sich so Stereotype verfestigen können (Paasch-Colberg et al., 2021). Die Effekte werden oft erst nach wiederholter Exposition sichtbar und konnten in dieser Studie aufgrund ihres kurzen Untersuchungszeitraums nicht erfasst werden. Es sollten daher in zukünftiger Forschung auch Langzeitstudien durchgeführt werden, um die Rezeption

vor allem der humorvollen, subtilen Form besser zu verstehen. Dies könnte zu wichtigeren Erkenntnissen über die schleichenden Effekte von Hassrede führen, die sich erst über Monate oder Jahre hinweg entfalten.

Die Verwendung von mehrdimensionalen Messinstrumenten könnte für zukünftige Forschungsarbeiten relevant sein. Dabei könnten neben psychologischen Reaktionen auch physiologische Messungen, wie etwa Herzfrequenzvariabilität oder Hautleitfähigkeit, eingesetzt werden, um unbewusste Reaktionen auf die Stimuli zu erfassen. Dadurch könnten die Reaktionen bei der Rezeption differenzierter betrachtet und das Verständnis für die Verarbeitung der unterschiedlichen Formen von Hatespeech ausgebaut werden. Zudem können interaktive Elemente in die Forschung integriert werden, um die Interaktion der Versuchspersonen zu untersuchen. Beispielsweise kann dieser simulierte Kontext durch Funktionen, wie Teilen, Kommentieren oder Liken geschaffen werden. Solche dynamischen Studien könnten Aufschluss darüber geben, wie sich die Bereitschaft zur Gegenrede oder die Unterstützung misogyner Stereotype in Echtzeit entwickeln und ob bestimmte Moderatoren, wie z. B. der Grad der Identifikation mit den betroffenen Gruppen, einen Einfluss auf diese Reaktionen haben.

6.2 Ausblick

Die Ergebnisse zeigen auf wissenschaftlicher Ebene, dass die Rezeption von verschiedenen Direktheitsformen von Hatespeech komplexer ist als bisher angenommen. Im Gegensatz zu früheren Forschungen, deutet das Resultat der Studie darauf hin, dass humorvolle Hatespeech zu ähnlich starken Reaktionen führt wie explizite Hassrede, insbesondere wenn den Rezipient:innen die diskriminierende Absicht bewusst ist.

Daher bieten die Ergebnisse nicht nur wissenschaftliche Erkenntnisse, sondern auch klare Ansatzpunkte für konkrete Anwendungen in verschiedenen gesellschaftlichen

Bereichen. Die Notwendigkeit, subtilere Formen von Hatespeech erst zu nehmen und sowohl auf einer technologischen als auch gesellschaftlichen Ebene darauf zu reagieren, wird deutlich, um langfristige Auswirkungen zu minimieren.

Die Erkenntnisse sind für Social-Media-Plattformen relevant. Die Überarbeitung von bisherigen Moderationsstrategien, könnte subtile Formen von Hassrede miteinschließen. Auch Algorithmen zur Erkennung von diskriminierenden Inhalten können erweitert werden, sodass alle Arten von Hatespeech identifiziert werden können. Der Einsatz von Technologien wie natürlicher Sprachverarbeitung (Natural Language Processing, NLP) könnte helfen, Muster von Ironie oder Elemente von Humor zu erkennen, die in humorvoller Hassrede häufig vorkommen. Ergänzend dazu könnten spezialisierte Teams aus geschulten Fachleuten eingesetzt werden, die problematische Inhalte analysieren und bewerten, wenn Algorithmen an ihre Grenzen stoßen. Die Unterstützung von Künstlicher Intelligenz kann ebenfalls dazu beitragen, diskriminierende Inhalte schneller zu Erkennen und diese zu Entfernen.

Die Ergebnisse betonen zudem die Bedeutung von Medienbildung. Programme zur Förderung von Medienkompetenz sollten nicht nur darauf abzielen offensichtliche, explizite Hatespeech, sondern auch subtilere Formen zu verstehen und kritisch zu hinterfragen. Durch Workshops und Schulungsprogramme kann mit realen Beispielen gearbeitet werden, was die Teilnehmer:innen hilft Inhalte besser einordnen zu können. Insbesondere in der Frühbildung, d.h. in Schulen oder Jugendeinrichtungen können Unterrichtseinheiten mit konkreten Fallbeispielen entwickelt werden. Dabei liegt das Ziel nicht nur in der Erkennung von Hatespeech und Diskriminierung aller Art, sondern auch in dem aktiv werden gegen hasserfüllte Inhalte, wie beispielsweise mit der Strategie der Counterspeech.

Für das politische Interesse können die Ergebnisse genutzt werden, um gesetzliche Rahmenbedingungen anzupassen. Gesetze könnten erweitert werden, um sicherzustellen,

dass auch solche Formen von Hassrede in den Fokus regulatorischer Maßnahmen rücken. Dies könnte durch klare Definitionen von subtilem Hass und durch die Förderung von Maßnahmen zur Aufklärung über dessen Gefahren erfolgen. Solche Regelungen könnten Plattformbetreiber:innen verpflichten, spezifische Richtlinien für humorvolle und indirekte Hassrede zu entwickeln und deren Einhaltung systematisch zu überwachen.

Zusammenfassend kann intensive Forschung innerhalb des Themengebiets Hatespeech, das öffentliche Bewusstsein für mögliche schädlichen Auswirkungen schärfen. Dies gilt insbesondere für verschiedene Darstellungsformen des Phänomens. Der digitale Raum trägt immer mehr zur gesellschaftlichen Meinungsbildung bei, deshalb ist es relevant nicht nur offensichtliche, sondern auch subtile Mechanismen von Diskriminierung zu verstehen und bekämpfen. Durch eine stärkere Sensibilisierung und Bildungsarbeit in dem Bereich kann eine inklusivere und respektvollere Diskussionskultur gefördert werden. Technologische Fortschritte in der Moderation, gezielte Bildungsmaßnahmen und angepasste gesetzliche Rahmenbedingungen könnten zusammenwirken, um langfristig die Verbreitung und Wirkung von Hassrede – sowohl explizit als auch subtil –zu reduzieren.

7. Conclusio

Die vorliegende Arbeit liefert wesentliche Erkenntnisse über die Wahrnehmung von Hatespeech, insbesondere von humorvoller Hassrede, im Vergleich zu expliziten Formen. Es konnte gezeigt werden, dass keine signifikanten Unterschiede in der Rezeption zwischen der expliziten und der humorvollen, subtilen Form der Hatespeech vorhanden sind. Dabei wurden die Dimensionen emotionalen Stimmung, die Bereitschaft zur Counterspeech sowie der Verstärkung von sexistischen Einstellungen abgefragt, als auch der moderierende Faktor „Geschlecht“ beachtet. Das Resultat zeigt die Notwendigkeit auf, dass auch subtile und indirekte Formen von Hatespeech ernst genommen werden müssen. Durch methodische Limitationen der Studie wie die Homogenität der Versuchspersonen oder das experimentelle Setting, könnten die Ergebnisse beeinflusst worden sein. Daher sollten zukünftige Forschungen heterogenere Stichproben und realistischere Medienszenarien untersuchen, um ein tieferes Verständnis der Auswirkungen von Hatespeech zu gewinnen. Insgesamt trägt die Arbeit zur Vertiefung des Verständnisses von Hassrede in digitalen Räumen bei und liefert wichtige Anhaltspunkte für die Entwicklung effektiver Maßnahmen zur Bekämpfung von Diskriminierung auf Social-Media-Plattformen.

8. Literaturverzeichnis

Abrams, J. R., & Bippus, A. M. (2014). *Gendering jokes*. Journal Of Language And Social Psychology, 33(6), 692–702. https://doi.org/10.1177/0261927x14544963

Abrams, J. R., Bippus, A. M., & McGaughey, K. J. (2015). *Gender disparaging jokes: An investigation of sexist-nonstereotypical jokes on funniness typicality and the moderating role of ingroup identification*. Humor - International Journal Of Humor Research, 28(2).

Aguinis, H. & Bradley, K. J. (2014). *Best Practice Recommendations for Designing and Implementing Experimental Vignette Methodology Studies*. Organizational Research Methods, *17*(4), 351–371. https://doi.org/10.1177/1094428114547952

Alesina, A. F., Giuliano, P., & Nunn, N. (2013). *On the origins of gender roles: Women and the plough*. The Quarterly Journal of Economics, 128(2), 469–530. https://doi.org/10.1093/qje/qjt005

Alkiviadou, N. (2019). *Hate speech on social media networks: Towards a regulatory framework?* Information & Communications Technology Law, 28(1), 19–35. https://doi.org/10.1080/13600834.2018.1494417

Amnesty Global Insights. (2017). *Unsocial media: The real toll of online abuse against women*. Medium. https://medium.com/amnesty-insights/unsocial-media-the-real-toll-of-online-abuse-against-women-37134ddab3f4

Amnesty International. (2018). *Troll patrol findings: Using crowdsourcing, data science & machine learning to measure violence and abuse against women on Twitter*. https://decoders.amnesty.org/projects/troll-patrol/findings#introductio

Åkerlund, M. (2021). *Influence without metrics: Analyzing the impact of far-right users in an online discussion forum*. Social Media + Society, 7(2), 205630512110088. https://doi.org/10.1177/20563051211008831

American Time Use Survey (2019). *United States Department of Labor*. https://www.bls.gov/news.release/pdf/atus.pdf

Bahador, B. (2021). *Countering hate speech*. In H. Tumber & S. Waisbord (Eds.), *The Routledge companion to media disinformation and populism* (pp. 507–518). Routledge.

Barker, K. & Jurasz, O. (2018). *Online Misogyny as Hate Crime: A Challenge for Legal Regulation?*

http://oro.open.ac.uk/57063/

Bell, B. T., Cassarly, J. A., & Dunbar, L. (2018). *Selfie-objectification: Self-objectification and positive feedback ("Likes") are associated with frequency of posting sexually objectifying self-images on social media*. Body Image, 26, 83–89. https://doi.org/10.1016/j.bodyim.2018.06.005

Ben-David, A., & Fernandez, A. M. (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, 27. http://eprints.qut.edu.au/101369/

Benesch, S. (2014). *Defining and diminishing hate speech*. In P. Grant (Hrsg.), State of the world's minorities and indigenous peoples 2014 (S. 19–25). Minority Rights Group.

Benikova, D., Wojatzki, M., & Zesch, T. (2018). *What does this imply? Examining the impact of implicitness on the perception of hate speech*. In Lecture Notes in Computer Science (S. 171–179). https://doi.org/10.1007/978-3-319-73706-5_14

Bergmann, W. (1997). *Kommunikationslatenz und Vergangenheitsbewältigung*. In: König, H., Kohlstruck, M., & Wöll, A. (eds), Vergangenheitsbewältigung am Ende des zwanzigsten Jahrhunderts. Leviathan Sonderhefte, vol 18. VS Verlag für Sozialwissenschaften Wiesbaden. https://doi.org/10.1007/978-3-663-11730-8_18

Biber, J. K., Doverspike, D., Baznik, D., Cober, A. B., & Ritter, B. (2002). *Sexual harassment in online communications: Effects of gender and discourse medium*. Cyberpsychology & Behavior, 5(1), 33–42. https://doi.org/10.1089/109493102753685863

Bilewicz, M. & Soral, W. (2020). *Hate Speech Epidemic. The Dynamic Effects of Derogatory Language on Intergroup Relations and Political Radicalization*. Political Psychology, 41(S1), 3–33. https://doi.org/10.1111/pops.12670

Blee, K. M., & Jackman, M. R. (1994). *The velvet glove: Paternalism and conflict in gender, class, and race relations*. Contemporary Sociology, A Journal Of Reviews, 23(6), 792. https://doi.org/10.2307/2076034

Block, K., Croft, A., De Souza, L., & Schmader, T. (2019). *Do people care if men don't care about caring? The asymmetry in support for changing gender roles*. Journal Of Experimental Social Psychology, 83, 112–131. https://doi.org/10.1016/j.jesp.2019.03.013

Boesveld, S. (2020). *What's driving the gender pay gap in medicine?* Canadian Medical Association Journal, 192(1), E19–E20. https://doi.org/10.1503/cmaj.1095831

Borgeson, K., & Valeri, R. (2004). *Faces of hate*. Journal Of Applied Sociology, os-21(2), 99–111. https://doi.org/10.1177/19367244042100205

Breyer, B., & Bluemke, M. (2016). *Deutsche Version der Positive and Negative Affect Schedule PANAS (GESIS Panel)*, 20, 20. https://doi.org/10.6102/zis242

Brosius, H., Haas, A. & Unkel, J. (2022). *Methoden der empirischen Kommunikationsforschung*. In Studienbücher zur Kommunikations- und Medienwissenschaft. https://doi.org/10.1007/978-3-658-34195-4

Brown, A. (2017). *What is hate speech? Part 1: The myth of hate*. Law And Philosophy, 36(4), 419–468. https://doi.org/10.1007/s10982-017-9297-1

Brown, A. (2018). *What is so special about online (as compared to offline) hate speech?* Ethnicities, 18(3), 297–326. https://doi.org/10.1177/1468796817709846

Burke, M., & Kraut, R. E. (2014). *Growing closer on Facebook: Changes in tie strength through social network site use*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/2556288.2557094

Chambers, M. J., Davidson, R. & MacKinnon, J. G. (1993*). Estimation and Inference in Econometrics*. The Economic Journal, *104*(424), 703. https://doi.org/10.2307/2234656

Citron, D. K. (2009). *Law's expressive value in combating cyber gender harassment*. Michigan Law Review, 108, 373–416.

Cohen-Almagor, R. (2013). *Freedom of expression v. social responsibility: Holocaust denial in Canada*. Journal Of Mass Media Ethics, 28(1), 42–56. https://doi.org/10.1080/08900523.2012.746119

Cole, K. (2015). "*It's Like She's Eager to be Verbally Abused": Twitter, Trolls, and (En)Gendering Disciplinary Rhetoric*. Feminist Media Studies, 15(2), 356–358. https://doi.org/10.1080/14680777.2015.1008750

Committee on the Elimination of Racial Discrimination. (2013). *General recommendation No.35*. Retrieved 16.10.2024 from https://tbinternet.ohchr.org/_layouts/treatybodyexternal/

Download.aspx?symbolno=CERD/C/GC/35&Lang=en.

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

Croft, A., Schmader, T., & Block, K. (2015). *An underexamined inequality*. Personality And Social Psychology Review, 19(4), 343–370. https://doi.org/10.1177/1088868314564789

Davies, P. G., Spencer, S. J., & Steele, C. M. (2005). *Clearing the air: Identity safety moderates the effects of stereotype threat on women's leadership aspirations*. Journal Of Personality And Social Psychology, 88(2), 276–287. https://doi.org/10.1037/0022-3514.88.2.276

Davis, S. E. (2018). Objectification, Sexualization, and Misrepresentation: Social Media and the College Experience. *Social Media + Society*, *4*(3), 205630511878672. https://doi.org/10.1177/2056305118786727

Delgado, R., & Stefancic, J. (2004). *Understanding words that wound*. Choice Reviews Online, 42(02), 42–1071. https://doi.org/10.5860/choice.42-1071

Delgado, R., & Stefancic, J. (2009). *Four observations about hate speech*. SSRN Electronic Journal. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2401935

Delgado, R., & Stefancic, J. (2014). *Hate speech in cyberspace*. Wake Forest Law Review, 49, 319. https://scholarship.law.ua.edu/cgi/viewcontent.cgi?article=1559&context=fac_articles

Döring, N., & Mohseni, M. (2018). *Male dominance and sexism on YouTube: Results of three content analyses*. Feminist Media Studies, 19(4), 512–524. https://doi.org/10.1080/14680777.2018.1467945

Döring, N., & Mohseni, M. (2020). *Gendered hate speech in YouTube and YouNow comments: Results of two content analyses*. Studies in Communication Media, 9(1), 62–88. https://doi.org/10.5771/2192-4007-2020-1-62

Drakett, J., Rickett, B., Day, K., & Milnes, K. (2018). *Old jokes, new media – Online sexism and constructions of gender in internet memes*. Feminism & Psychology, 28(1), 109–127. https://doi.org/10.1177/0959353517727560

Eisenberger, N. I. (2015). *Social pain and the brain: Controversies, questions and where to go from here*. Annual Review Of Psychology, 66(1), 601–629. https://doi.org/10.1146/annurev-psych-010213-115146

ElSherief, M., Kulkarni, V., Nguyen, D., Wang, W. Y., & Belding, E. (2018). *Hate lingo: A target-based linguistic analysis of hate speech in social media*. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1804.04257

Erjavec, K., & Kovačič, M. P. (2012). *"You don't understand this is a new war!" Analysis of hate speech in news websites' comments*. Mass Communication & Society, 15(6), 899–920. https://doi.org/10.1080/15205436.2011.619679

Eudey, B. (2008*). Keyword: Feminism: Evaluating representations of feminism in YouTube*. Feminist Collections, 29(1), 28

European Commission (2019). *Stop violence against women: Statement by the European Commission and the High Representative*. https://ec.europa.eu/commission/presscorner/detail/en/statement_19_6300

Falk, E. (2010). *Women for president: Media bias in nine campaigns*. University of Illinois Press. https://doi.org/10.1080/08821127.2010.10678174

Ford, T. E. (1997). *Effects of stereotypical television portrayals of African-Americans on person perception*. Social Psychology Quarterly, 60(3), 266. https://doi.org/10.2307/2787086

Ford, T. E. (2000). *Effects of sexist humor on tolerance of sexist events*. Personality And Social Psychology Bulletin, 26(9), 1094–1107. https://doi.org/10.1177/01461672002611006

Ford, T. E., & Ferguson, M. A. (2004). *Social consequences of disparagement humor: A prejudiced norm theory*. Personality And Social Psychology Review, 8(1), 79–94. https://doi.org/10.1207/s15327957pspr0801_4

Ford, T. E., Wentzel, E. R., & Lorion, J. (2001). *Effects of exposure to sexist humor on perceptions of normative tolerance of sexism*. European Journal Of Social Psychology, 31(6), 677–691. https://doi.org/10.1002/ejsp.56

Ford, T. E., Boxer, C. F., Armstrong, J., & Edel, J. R. (2008). *More than "just a joke": The prejudice-releasing function of sexist humor*. Personality And Social Psychology Bulletin, 34(2), 159–170. https://doi.org/10.1177/0146167207310022

Ford, T. E., Buie, H. S., Mason, S. D., Olah, A. R., Breeden, C. J., & Ferguson, M. A. (2019). *Diminished self-concept and social exclusion: Disparagement humor from the target's perspective*. Self And Identity, 19(6), 698–718. https://doi.org/10.1080/15298868.2019.1653960

Ford, T. E., Woodzicka, J. A., Triplett, S. R. & Kochersberger, A. O. (2013). Sexist Humor and Beliefs that Justify Societal Sexism. *De Gruyter Mouton*. https://www.uiowa.edu/crisp/files/crisp/files/art9.26.13_1.pdf

Ford, T. E., & Woodzicka, J. A. (2015). *Sexist humor as a trigger of state self-objectification in women*. Humor - International Journal Of Humor Research, 28(2). https://doi.org/10.1515/humor-2015-0018

Fosch-Villaronga, E., Poulsen, A., Søraa, R. A., & Custers, B. (2021). *Gendering algorithms in social media*. ACM SIGKDD Explorations Newsletter, 23(1), 24–31. https://doi.org/10.1145/3468507.3468512

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

FRA (2023). *Online Content Moderation*. Abgerufen von: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2023-online-content-moderation_en.pdf

Gardiner, B. (2018). *"It's a terrible way to go to work:" what 70 million readers' comments on the Guardian revealed about hostility to women and minorities online*. Feminist Media Studies, *18*(4), 592–608. https://doi.org/10.1080/14680777.2018.1447334

Gelber, K., & McNamara, L. (2015). *Evidencing the harms of hate speech*. Social Identities, 22(3), 324–341. https://doi.org/10.1080/13504630.2015.1128810

Gillett, R., Stardust, Z., & Burgess, J. (2022). *Safety for whom? Investigating how platforms frame and perform safety and harm interventions*. Social Media + Society, 8(4), 205630512211443. https://doi.org/10.1177/20563051221144315

Ging, D. & Siapera, E. (2018). *Special issue on online misogyny*. Feminist Media Studies,*18*(4), 515–524. https://doi.org/10.1080/14680777.2018.1447345

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. https://doi.org/10.1037/0022-3514.70.3.491

Glick, P., & Fiske, S. T. (1997). *Hostile and benevolent sexism*. Psychology Of Women Quarterly, 21(1), 119–135. https://doi.org/10.1111/j.1471-6402.1997.tb00104.x

Glick, P., & Fiske, S. T. (2001). *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

Guo, L., & Johnson, B. G. (2020). *Third-person effect and hate speech censorship on Facebook*. Social Media + Society, 6(2), 205630512092300. https://doi.org/10.1177/2056305120923003

Hajok, D. & Selg, O. (2018). *Kommunikation auf Abwegen? Fake News und Hate Speech in kritischer Betrachtung*. Jugend Medien Schutz-Report, 41(4), 2–6. https://doi.org/10.5771/0170-5067-2018-4-2

Handbook of Prejudice, Stereotyping, and Discrimination. (2009). *Psychology Press eBooks*. https://doi.org/10.4324/9781841697772

Hawdon, J., Oksanen, A., & Räsänen, P. (2016). *Exposure to online hate in four nations: A cross-national consideration*. Deviant Behavior, 38(3), 254–266. https://doi.org/10.1080/01639625.2016.1196985

Hayes, A. F. (2013). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. https://ci.nii.ac.jp/ncid/BB1323391X

Hemmerich, W. (2016). *StatistikGuru: Between-Groups Design*. Retrieved from https://statistikguru.de/lexikon/between-groups-design.html

Henry, N., & Powell, A. (2016). *Technology-facilitated sexual violence: A literature review of empirical research*. Trauma, Violence, & Abuse, 19(2), 195–208. https://doi.org/10.1177/1524838016650189

Howard, J. W. (2019). *Free speech and hate speech. Annual Review of Political Science*, 22(1), 93–109. https://doi.org/10.1146/annurev-polisci-051517-012343

Hodson, G., Rush, J., & MacInnis, C. C. (2010). *A joke is just a joke (except when it isn't): Cavalier humor beliefs facilitate the expression of group dominance motives*. Journal Of Personality And Social Psychology, 99(4), 660–682. https://doi.org/10.1037/a0019627

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(4), 557–576. https://doi.org/10.1111/hcre.12059

Im, J., Schoenebeck, S., Iriarte, M., Grill, G., Wilkinson, D., Batool, A., Alharbi, R., Funwie, A., Gankhuu, T., Gilbert, E., & Naseem, M. (2022). *Women's perspectives on harm and justice after online harassment*. Proceedings Of The ACM On Human-Computer Interaction, 6(CSCW2), 1–23. https://doi.org/10.1145/3555775

Jha, A., & Mamidi, R. (2017). *When does a compliment become sexist? Analysis and classification of ambivalent sexism using Twitter data*. Proceedings of the second workshop on NLP and computational social science. https://doi.org/10.18653/v1/w17-290

Johnson, N. F., Leahy, R., Restrepo, N. J., Velasquez, N., Zheng, M., Manrique, P., Devkota, P., & Wuchty, S. (2019). *Hidden resilience and adaptive dynamics of the global online hate ecology*. Nature, 573(7773), 261–265. https://doi.org/10.1038/s41586-019-1494-7

Jones, B., & Kenward, M. G. (2014). *Design and analysis of cross-over trials*. Chapman and Hall/CRC eBooks. https://doi.org/10.1201/b17537

Kalch, A., & Naab, T. K. (2017). *Replying, disliking, flagging: How users engage with uncivil and impolite comments on news sites*. Studies in Communication Media, 6(4), 395–419. https://doi.org/10.5771/2192-4007-2017-4-395

Kenski, K., Coe, K., & Rains, S. A. (2017). *Perceptions of uncivil discourse online: An examination of types and predictors*. Communication Research, 47(6), 795–814. https://doi.org/10.1177/0093650217699933

Kenski, K., Coe, K., & Rains, S. A. (2020). *Perceptions of uncivil discourse online: An examination of types and predictors*. Communication Research, 47(6), 795–814.

Klinger, U., & Svensson, J. (2014). *The emergence of network media logic in political communication: A theoretical approach*. New Media & Society, 17(8), 1241–1257. https://doi.org/10.1177/1461444814522952

Koch, T., Peter, C. & Müller, P. (2019). *Das Experiment in der Kommunikations- und Medienwissenschaft*. Springer eBooks. https://doi.org/10.1007/978-3-658-19754-4

Kong, G., Kong, T. D., & Lu, R. (2020). *Political promotion incentives and within-firm pay gap: Evidence from China*. Journal Of Accounting And Public Policy, 39(2), 106715. https://doi.org/10.1016/j.jaccpubpol.2020.106715

Krasodomski-Jones, A. (2023, 24. Februar). *Counter-Speech*. Demos. http://www.demos.co.uk/project/counter-speech/

Kreißel, P., Ebner, J., Urban, A., & Guhl, J. (2018). *Hass auf Knopfdruck: Rechtsextreme Trollfabriken und das Ökosystem koordinierter Hasskampagnen im Netz*. London: Institute for Strategic Dialogue.

Kubinger, K. D., Rasch, D. & Moder, K. (2009). Zur Legende der Voraussetzungen des t-Tests für unabhängige Stichproben. *Psychologische Rundschau*, *60*(1), 26–27. https://doi.org/10.1026/0033-3042.60.1.26

Kümpel, A. S. & Rieger, D. (2019). *Wandel der Sprach- und Debattenkultur in sozialen Online-Medien: Ein Literaturüberblick zu Ursachen und Wirkungen von inziviler Kommunikation.* Konrad Adenauer Stiftung. https://doi.org/10.5282/ubm/epub.68880

Latané, B. & Darley, J. M. (1970). *The unresponsive bystander: Why Doesn't He Help?* Prentice Hall.

Leets, L. (2001). *Responses to internet hate sites: Is speech too free in cyberspace?* Communication Law And Policy, 6(2), 287–317. https://doi.org/10.1207/s15326926clp0602_2

Leets, L., & Giles, H. (1997). *Words as weapons? When do they wound? Investigations of harmful speech*. Human Communication Research, 24(2), 260–301. https://doi.org/10.1111/j.1468-2958.1997.tb00415.x

Leonhard, L., Rueß, C., Obermaier, M., & Reinemann, C. (2018). *Perceiving threat and feeling responsible: How severity of hate speech, number of bystanders, and prior reactions of others affect bystanders' intention to counterargue against hate speech on Facebook*. Studies in Communication Media, 7(4), 555–579. https://doi.org/10.5771/2192-4007-2018-4-555

Maarouf, A., Pröllochs, N., & Feuerriegel, S. (2024). *The virality of hate speech on social media*. Proceedings Of The ACM On Human-Computer Interaction, 8(CSCW1), 1–22. https://doi.org/10.1145/3641025

Martinez-Pecino, R. & Durán, M. (2016*). I Love You but I Cyberbully You: The Role of Hostile Sexism. Journal Of Interpersonal Violence*, 34(4), 812–825. https://doi.org/10.1177/0886260516645817

Marshall, A. D., Robinson, L. R. & Azar, S. T. (2011). *Cognitive and emotional contributors to intimate partner violence perpetration following trauma. Journal Of Traumatic Stress*, 24(5), 586–590. https://doi.org/10.1002/jts.20681

Marx, K. (2018). *Hate speech – ein Thema für die Linguistik*. In *Nomos Verlagsgesellschaft mbH & Co. KG eBooks* (S. 37–58). https://doi.org/10.5771/9783845293288-37

Matamoros-Fernández, A. (2017). *Platformed racism: The mediation and circulation of an Australian race-based controversy on Twitter, Facebook, and YouTube*. Information Communication & Society, 20(6), 930–946. https://doi.org/10.1080/1369118x.2017.1293130

Matamoros-Fernández, A. (2018). *Inciting anger through Facebook reactions in Belgium: The use of emoji and related vernacular expressions in racist discourse*. First Monday. https://doi.org/10.5210/fm.v23i9.9405

Matamoros-Fernández, A., & Farkas, J. (2021). *Racism, hate speech, and social media: A systematic review and critique*. Television & New Media, 22(2), 205–224. https://doi.org/10.1177/1527476420982230

Matamoros-Fernández, A., Bartolo, L., & Troynar, L. (2023). *Humour as an online safety issue: Exploring solutions to help platforms better address this form of expression*. Internet Policy Review, 12(1). https://doi.org/10.14763/2023.1.1677

Matsuda, M. J. (1989). *Public response to racist speech: Considering the victim's story*. Michigan Law Review, 87(8), 2320. https://doi.org/10.2307/1289306

Matsuda, M. J., Lawrence, C. R., Delgado, R., & Crenshaw, K. W. (1993). *Words that wound: Critical race theory, assaultive speech, and the First Amendment*. Choice Reviews Online, 31(04), 31–2366. https://doi.org/10.5860/choice.31-2366

Meibauer, J. (2013). *Hassrede: Interdisziplinäre Beiträge zu einer aktuellen Diskussion*.

Mendel, T. (2012). *Does International Law Provide for Consistent Rules on Hate Speech?* In Cambridge University Press eBooks (S. 417–429). https://doi.org/10.1017/cbo9781139042871.029

Mendiburo-Seguel, A. & Ford, T. E. (2019). *The effect of disparagement humor on the acceptability of prejudice*. Current Psychology, 42(19), 16222–16233. https://doi.org/10.1007/s12144-019-00354-2

Miranda, S. L. (2023). *Analyzing Hate Speech Against Women on Instagram. Open Information Science*, 7(1). https://doi.org/10.1515/opis-2022-0161

Mohseni, M. (2021). *Sexistische Online-Hassrede auf Videoplattformen*. In Springer eBooks, 39–51. https://doi.org/10.1007/978-3-658-31793-5_3

Molyneaux, H., Gibson, K. & Singer, J. (2008). *Exploring the Gender Divide on YouTube: An Analysis of the Creation and Reception of Vlogs*. American Communication Journal. http://ac-journal.org/journal/2008/Spring/3GenderandYoutube.pdf

Moss-Racusin, C. A., Molenda, A. K. & Cramer, C. R. (2015). *Can Evidence Impact Attitudes? Public Reactions to Evidence of Gender Bias in STEM Fields*. Psychology Of Women Quarterly, 39(2), 194–209. https://doi.org/10.1177/0361684314565777

Müller, A., & Lopez-Sanchez, M. (2021). *Countering negative effects of hate speech in a multi-agent society*. In *Frontiers in Artificial Intelligence and Applications*. https://doi.org/10.3233/faia210122

Munn, L. (2019). *Alt-right pipeline: Individual journeys to extremism online*. First Monday. https://doi.org/10.5210/fm.v24i6.10108

Nakamura, L. (2014). *'I WILL DO EVERYTHING That Am Asked': Scambaiting, digital show-space, and the racial violence of social media*. Journal Of Visual Culture, 13(3), 257–274. https://doi.org/10.1177/1470412914546845

Nielsen, L. B. (2002). *Subtle, Pervasive, Harmful: Racist and Sexist Remarks in Public as Hate Speech*. Journal Of Social Issues, 58(2), 265–280. https://doi.org/10.1111/1540-4560.00260

Noelle-Neumann, E. (1974). *The spiral of silence: A theory of public opinion*. Journal Of Communication, 24(2), 43–51. https://doi.org/10.1111/j.1460-2466.1974.tb00367.x

Obermaier, M., Fawzi, N. & Koch, T. (2016). *Bystanding or standing by? How the number of bystanders affects the intention to intervene in cyberbullying*. New Media & Society, 18(8), 1491–1507. https://doi.org/10.1177/1461444814563519

Obermaier, M., Hofbauer, M., & Reinemann, C. (2018). *Journalists as targets of hate speech: How German journalists perceive the consequences for themselves and how they cope with it*. Studies in Communication Media, 7(4), 499–524. https://doi.org/10.5771/2192-4007-2018-4-499

Obermaier, M., Schmuck, D. & Saleem, M. (2021). *I'll be there for you? Effects of Islamophobic online hate speech and counter speech on Muslim in-group bystanders' intention to intervene*. New Media & Society, *25*(9), 2339–2358. https://doi.org/10.1177/14614448211017527

Obermaier, M., Schmid, U. K., & Rieger, D. (2023). *Too civil to care? How online hate speech against different social groups affects bystander intervention*. European Journal Of Criminology, 20(3), 817–833. https://doi.org/10.1177/14773708231156328

Ohme, J., & Mothes, C. (2020). *What affects first- and second-level selective exposure to journalistic news? A social media online experiment*. Journalism Studies, 21(9), 1220–1242.

Ortmann, M. (2023). *Welch´s t-Test | Dr. Magdalene Ortmann | Glossar*. Dr. Magdalene Ortmann. https://ortmann-statistik.de/glossar/welchs-t-test/

Ouirdi, M. E., Ouirdi, A. E., Segers, J., & Henderickx, E. (2014). *Social media conceptualization and taxonomy*. Journal Of Creative Communications, 9(2), 107–126. https://doi.org/10.1177/0973258614528608

Paasch-Colberg, S., & Strippel, C. (2021). *"The boundaries are blurry . . .": How comment moderators in Germany see and respond to hate comments*. Journalism Studies, 23(2), 224–244.

Paasch-Colberg, S., Trebbe, J., Strippel, C., et al. (2021). *Insults, criminalization, and calls for violence: Forms of hate speech and offensive language in German user comments on immigration*. In: Monnier, A., Boursier, A., & Seoane, A. (eds), *Cyberhate in the context of migrations. Postdisciplinary studies in discourse*. Cham: Palgrave Macmillan, pp. 137–163.

Parker, K., & Funk, C. (2017, April 14). *Gender discrimination comes in many forms for today's working women*. Pew Research Center. https://www.pewresearch.org/short-reads/2017/12/14/gender-discrimination-comes-in-many-forms-for-todays-working-women/

Perry, D. G., Perry, L. C., & Weiss, R. J. (1989). *Sex differences in the consequences that children anticipate for aggression*. Developmental Psychology, 25, 312–319.

Ping, K., Kumar, A., Ding, X. & Rho, E. (2024). *Behind the Counter: Exploring the Motivations and Barriers of Online Counterspeech Writing*. arXiv (Cornell University). https://doi.org/10.48550/arxiv.2403.17116

Pondorf, P. (2024). *Konzept der Masterarbeit. Humor als Schutzschild: Eine quantitative Studie zur Rezeption von misogyner Hatespeech mit humoristischen Elementen.*

Power, J. G., Murphy, S. T., & Coover, G. (1996). *Priming prejudice: How stereotypes and counter-stereotypes influence attribution of responsibility and credibility among ingroups and outgroups*. Human Communication Research, 23(1), 36–58. https://doi.org/10.1111/j.1468-2958.1996.tb00386.x

Republik Österreich. (2020). *Bundesgesetz über Maßnahmen zum Schutz der Nutzer auf Kommunikationsplattformen*. https://www.parlament.gv.at/PAKT/VHG/XXVII/ME/ME_00049/index.shtml

Rieger, D., Kümpel, A. S., Wich, M., Kiening, T., & Groh, G. (2021). *Assessing the extent and types of hate speech in fringe communities: A case study of alt-right communities on 8chan, 4chan, and Reddit*. DOAJ (Directory of Open Access Journals). https://doi.org/10.1177/20563051211052906

Romero-Sánchez, M., Durán, M., Carretero-Dios, H., Megías, J. L., & Moya, M. (2010). *Exposure to sexist humor and rape proclivity: The moderator effect of aversiveness ratings*. Journal Of Interpersonal Violence, 25(12), 2339–2350. https://doi.org/10.1177/0886260509354884

Ryan, K.M., Kanjorski, J. (1998). *The Enjoyment of Sexist Humor, Rape Attitudes, and Relationship Aggression in College Students*. Sex Roles 38, 743–756. https://doi.org/10.1023/A:1018868913615

Saleem, H. M., Dillon, K. P., Benesch, S., & Ruths, D. (2017). *A web of hate: Tackling hateful speech in online social spaces*. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1709.10159

Sasse, J., & Grossklags, J. (2023). *Breaking the silence: Investigating which types of moderation reduce negative effects of sexist social media content*. Proceedings Of The ACM On Human-Computer Interaction, 7(CSCW2), 1–26. https://doi.org/10.1145/3610176

Sasse, J., Van Breen, J. A., Spears, R., & Gordijn, E. H. (2021). *The rocky road from experience to expression of emotions: Women's anger about sexism*. Affective Science, 2(4), 414–426. https://doi.org/10.1007/s42761-021-00081-7

Saucier, D. A., O'Dea, C. J. & Strain, M. L. (2016). *The bad, the good, the misunderstood: The social effects of racial humor. Translational Issues in Psychological* Science, 2(1), 75–85. https://doi.org/10.1037/tps0000059

Saucier, D. A., Strain, M. L., Miller, S. S., O'Dea, C. J. & Till, D. F. (2018). "*What do you call a Black guy who flies a plane?": The effects and understanding of disparagement and confrontational racial humor.* Humor - International Journal Of Humor Research, 31(1), 105–128. https://doi.org/10.1515/humor-2017-0107

Schafer, J. R., & Navarro, J. (2003). *The seven-stage hate model: The psychopathology of hate groups*. In PsycEXTRA Dataset. https://doi.org/10.1037/e315272004-001

Schäfer, S., Rebasso, I., Boyer, M. M., & Planitzer, A. M. (2023). *Can we counteract hate? Effects of online hate speech and counter speech on the perception of social groups*. Communication Research. https://doi.org/10.1177/00936502231201091

Schäfer, S., Sülflow, M., & Reiners, L. (2022). *Hate speech as an indicator for the state of the society*. Journal Of Media Psychology, 34(1), 3–15. https://doi.org/10.1027/1864-1105/a000294

Schemer, C. (2012). *The Influence of News Media on Stereotypic Attitudes Toward Immigrants in a Political Campaign*. Journal Of Communication, 62(5), 739–757. https://doi.org/10.1111/j.1460-2466.2012.01672.x

Schieb, C., & Preuss, M. (2018). *Considering the elaboration likelihood model for simulating hate and counter speech on Facebook*. Studies in Communication Media, 7(4), 580–606. https://doi.org/10.5771/2192-4007-2018-4-580

Schieb, C. (2022). *Zwischen Like und Empörung: Theoretische Modellierung und empirische Untersuchung von Rezeptions- und Persuasionswirkungen von sexistischem Hate Speech*. https://doi.org/10.17879/63079514069

Schmid, U. K., Kümpel, A. S. & Rieger, D. (2024). *How social media users perceive different forms of online hate speech: A qualitative multi-method study*. New Media & Society, 146144482210911. https://doi.org/10.1177/14614448221091185

Schmid, U. K. (2023). *Humorous hate speech on social media: A mixed-methods investigation of users' perceptions and processing of hateful memes*. New Media & Society. https://doi.org/10.1177/14614448231198169

Schmitt, J. B. (2017). *Online-hate speech: Definition und Verbreitungsmotivationen aus psychologischer Perspektive*. In K. Kaspar, L. Gräßer, & A. Riff (Hrsg.), *Online Hate Speech: Perspektiven auf eine neue Form des Hasses* (S. 51–56). Grimme-Institut.

Schulze, H., Hohner, J., & Rieger, D. (2022). *Soziale Medien und Radikalisierung*. In Nomos Verlagsgesellschaft mbH & Co. KG eBooks (S. 319–330). https://doi.org/10.5771/9783748904212-319

Schwertberger, U., & Rieger, D. (2021). *Hass und seine vielen Gesichter: Eine sozial- und kommunikationswissenschaftliche Einordnung von Hate Speech*. In Springer eBooks (S. 53–77). https://doi.org/10.1007/978-3-658-31793-5_4

Sellars, A. (2016). *Defining hate speech*. Social Science Research Network. https://doi.org/10.2139/ssrn.2882244

Siddiqui, S. (2008). *YouTube and feminism: A class action project*. Feminist Collections, 29(1), 24–25

Silva, L., Mondal, M., Correa, D., Benevenuto, F. & Weber, I. (2016). *Analyzing the Targets of Hate in Online Social Media*. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1603.07709

Sponholz, L. (2018). *Hate Speech in den Massenmedien: Theoretische Grundlagen und empirische Umsetzung*. In Springer eBooks. http://ci.nii.ac.jp/ncid/BB28183713

Sponholz, L. (2021). *Hass mit Likes: Hate speech als Kommunikationsform in den Social Media*. In Springer eBooks (S. 15–37). https://doi.org/10.1007/978-3-658-31793-5_2

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

Statista (2024a, 6. Mai). *Österreich - meistgenutzte Social-Media Plattformen 2023* | https://de.statista.com/statistik/daten/studie/184936/umfrage/beliebteste-soziale-netzwerke-in-oesterreich-nach-nutzung/

Statista. (2024b, 17. Juni). *Ranking der vertrauenswürdigsten Nachrichtenquellen in Deutschland 2024*. https://de.statista.com/statistik/daten/studie/877238/umfrage/ranking-der-vertrauenswuerdigsten-nachrichtenquellen-in-deutschland/

Stoll, L. C., Lilley, T. G. & Pinter, K. (2016). *Gender-Blind Sexism and Rape Myth Acceptance*. Violence Against Women, *23*(1), 28–45. https://doi.org/10.1177/1077801216636239

Stone, G. R. (2000). *First Amendment*. In L. W. Levy (Hrsg.), *Encyclopedia of the American Constitution* (S. 1055–1057). Macmillan.

Strain, M. L., Martens, A. L., & Saucier, D. A. (2016). *"Rape is the new black": Humor's potential for reinforcing and subverting rape culture*. Translational Issues in Psychological Science, 2(1), 86–95. https://doi.org/10.1037/tps0000057

Stryker, R., Conway, B. A., & Danielson, J. T. (2016). *What is political incivility?* Communication Monographs, 83(4), 535–556.

Suler, J. (2004). *The online disinhibition effect*. CyberPsychology & Behavior, 7(3), 321–326. https://doi.org/10.1089/1094931041291295

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. https://doi.org/10.1037/0022-3514.68.2.199

Tajfel, H., & Turner, J. C. (2004). *The social identity theory of intergroup behavior*. In *Psychology Press eBooks* (S. 276–293). https://doi.org/10.4324/9780203505984-16

Thomae, M., & Pina, A. (2015). *Sexist humor and social identity: The role of sexist humor in men's in-group cohesion, sexual harassment, rape proclivity, and victim blame*. Humor - International Journal Of Humor Research, 28(2). https://doi.org/10.1515/humor-2015-0023

Thomae, M., & Viki, G. T. (2013). *Why did the woman cross the road? The effect of sexist humor on men's rape proclivity*. Journal of Social Evolutionary and Cultural Psychology, 7(3), 250–269. https://doi.org/10.1037/h0099198

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. https://doi.org/10.1177/0146167295218007

Tynes, B. M., Del Toro, J., & Lozada, F. T. (2015). *An unwelcomed digital visitor in the classroom: The longitudinal impact of online racial discrimination on academic motivation*. School Psychology Review, 44(4), 407–424. https://doi.org/10.17105/spr-15-0095.1

Unger, D. (2013). *Kriterien zur Einschränkung von Hate Speech: Inhalt, Kosten oder Wertigkeit von Äußerungen?* In J. Meibauer (Hrsg.), Hassrede/Hate Speech: Interdisziplinäre Beiträge zu einer aktuellen Diskussion (S. 257–285). Gießener Elektronische Bibliothek.

Van Dijck, J., & Poell, T. (2013). *Understanding social media logic*. Media And Communication, 1(1), 2–14. https://doi.org/10.17645/mac.v1i1.70

Vanden Abeele, M. M. P. & Postma-Nilsenova, M. (2018). *More Than Just Gaze: An Experimental Vignette Study Examining How Phone-Gazing and Newspaper-Gazing and Phubbing-*

While-Speaking and Phubbing-While-Listening Compare in Their Effect on Affiliation. Communication Research Reports, *35*(4), 303–313. https://doi.org/10.1080/08824096.2018.1492911

Vedeler, J. S., Olsen, T., & Eriksen, J. (2019). *Hate speech harms: A social justice discussion of disabled Norwegians' experiences*. Disability & Society, 34(3), 368–383. https://doi.org/10.1080/09687599.2018.1515723

Vitak, J., Chadha, K., Steiner, L., & Ashktorab, Z. (2017). *Identifying women's experiences with and strategies for mitigating negative effects of online harassment*. Proceedings Of The 2017 ACM Conference On Computer Supported Cooperative Work And Social Computing, 1231–1245. https://doi.org/10.1145/2998181.2998337

Vogels, E. (2021). *The state of online harassment: Survey report*. Pew Research Center. https://www.pewresearch.org/internet/2021/01/13/the-state-of-online-harassment/

Von Berger, R., Burek, M. & Saller, C. (2009). *Online-Vignettenexperimente. Methode und Anwendung auf spieltheoretische Analysen.* In VS Verlag für Sozialwissenschaften eBooks (S. 305–319). https://doi.org/10.1007/978-3-531-91791-7_19

Wachs, S., & Wright, M. (2018). *Associations between bystanders and perpetrators of online hate: The moderating role of toxic online disinhibition*. International Journal Of Environmental Research And Public Health, 15(9), 2030. https://doi.org/10.3390/ijerph15092030

Wachs, S., Koch-Priewe, B. & Zick, A. (2021). *Wenn Hass redet und schädigt. Einleitung in den Sammelband*. In Springer eBooks (S. 3–12). https://doi.org/10.1007/978-3-658-31793-5_1

Waldron, J. (2010). *Dignity and defamation: The visibility of hate*. Harvard Law Review, 123(7), 1596–1657. https://dialnet.unirioja.es/servlet/articulo?codigo=3216568

Ward, L. M., & Grower, P. (2020). *Media and the development of gender role stereotypes*. Annual Review Of Developmental Psychology, 2(1), 177–199. https://doi.org/10.1146/annurev-devpsych-051120-010630

Warren, C., Barsky, A. & McGraw, A. P. (2020). *What Makes Things Funny? An Integrative Review of the Antecedents of Laughter and Amusement*. Personality And Social Psychology Review, 25(1), 41–65. https://doi.org/10.1177/1088868320961909lle

Watson, D., Clark, L. A., & Tellegen, A. (1988). *Development and validation of brief measures of positive and negative affect: The PANAS scales*. Journal Of Personality And Social Psychology, 54(6), 1063–1070. https://doi.org/10.1037/0022-3514.54.6.1063

Weber, A. (2009). *Manual on hate speech*. http://icm.sk/subory/Manual_on_hate_speech.pdf

Weber, M., Viehmann, C., Ziegele, M., & Schemer, C. (2020). *Online hate does not stay online – How implicit and explicit attitudes mediate the effect of civil negativity and hate in user comments on prosocial behavior*. Computers in Human Behavior, 104, 106192. https://doi.org/10.1016/j.chb.2019.106192

Weston-Scheuber, K. (2013). *Gender and the prohibition of hate speech*. QUT Law Review, 12(2). https://doi.org/10.5204/qutlr.v12i2.504

Williams, A., Oliver, C., Aumer, K., & Meyers, C. (2016). *Racial microaggressions and perceptions of internet memes*. Computers in Human Behavior, 63, 424–432. https://doi.org/10.1016/j.chb.2016.05.067

Wojatzki, M., Horsmann, T., Gold, D., et al. (2018). *Do women perceive hate differently? Examining the relationship between hate speech, gender, and agreement judgments*. Presented at the 14th Conference on Natural Language Processing (KONVENS 2018), Vienna, Austria.

Worth, A., Augoustinos, M., & Hastie, B. (2015). *"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

Wypych, M., & Bilewicz, M. (2021). *Psychological toll of hate speech: The role of acculturation stress in the effects of exposure to ethnic slurs on mental health among Ukrainian immigrants in Poland*. Cultural Diversity & Ethnic Minority Psychology, 30(1), 35–44. https://doi.org/10.1037/cdp0000522

Zhang, Z., & Luo, L. (2019). *Hate speech detection: A solved problem? The challenging case of long tail on Twitter*. Semantic Web, 10(5), 925–945. https://doi.org/10.3233/sw-180338

Ziegele, M., Naab, T. K., & Jost, P. (2019). *Lonely together? Identifying the determinants of collective corrective action against uncivil comments*. New Media & Society, 22(5), 731–751. https://doi.org/10.1177/1461444819870130

Zochniak, K., Lewicka, O., Wybrańska, Z., & Bilewicz, M. (2023). *Homophobic hate speech affects well-being of highly identified LGBT people*. Journal Of Language And Social Psychology, 42(4), 453–463. https://doi.org/10.1177/0261927x231174569

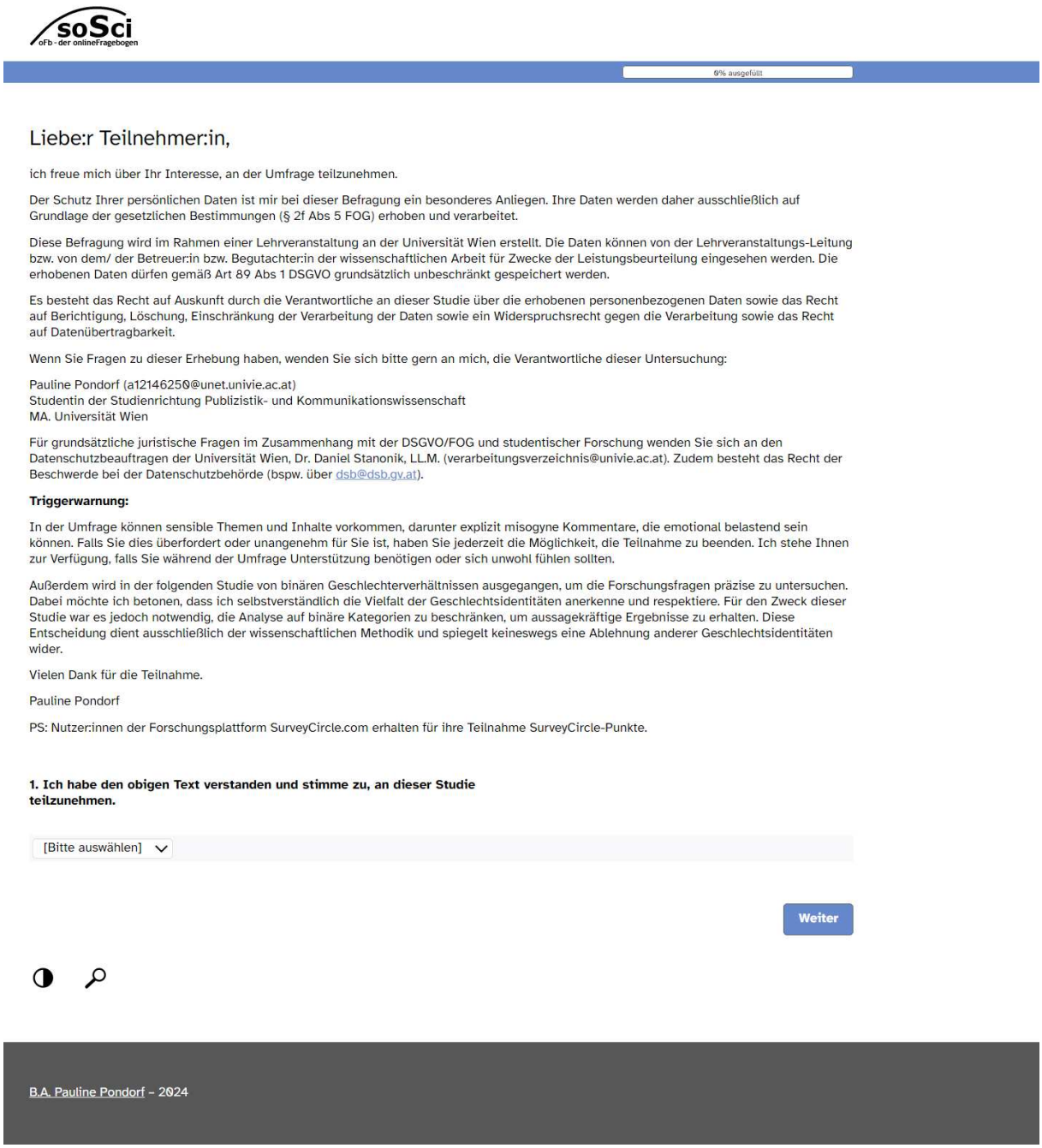
9. Tabellenverzeichnis

10. Anhang

A.1 Abstract

The digital era has fundamentally changed ways of communication and turned social media platforms into central places for public discussion. This development has led to the spread of hate speech, especially misogynist discrimination. A complex phenomenon in this context is the subtle, humorous form of hate speech, which is a combination of offensive content and humorous elements. The aim of the present study is to investigate the effects of explicit, misogynistic versus subtle, humorous forms of hate speech in terms of their reception, with the gender of the participants included as a moderating variable. Based on existing research, it is hypothesised that explicit hate speech leads to a more negative emotional mood, an increased willingness to use counterspeech and a reinforcement of sexist attitudes. It is also assumed that these effects tend to be more pronounced in men than in women. An experimental 2 x 3 between-subject design is used to test these assumptions. The results of the study show no significant differences between the explicit and the subtle, humorous form of hate speech in relation to the variables analysed. This could indicate that humorous hate speech is also perceived as a violation of discourse norms and does not differ from explicit forms. Overall, this study provides insights into the dynamics of hate speech, especially with regard to implicit, humorous forms, and contributes to the further development of strategies to combat discrimination in the digital space.

A.2 Fragebogen der Hauptstudie



soSci
oFb - der onlineFragebogen

0% ausgefüllt

Liebe:r Teilnehmer:in,

ich freue mich über Ihr Interesse, an der Umfrage teilzunehmen.

Der Schutz Ihrer persönlichen Daten ist mir bei dieser Befragung ein besonderes Anliegen. Ihre Daten werden daher ausschließlich auf Grundlage der gesetzlichen Bestimmungen (§ 2f Abs 5 FOG) erhoben und verarbeitet.

Diese Befragung wird im Rahmen einer Lehrveranstaltung an der Universität Wien erstellt. Die Daten können von der Lehrveranstaltungs-Leitung bzw. von dem/ der Betreuer:in bzw. Begutachter:in der wissenschaftlichen Arbeit für Zwecke der Leistungsbeurteilung eingesehen werden. Die erhobenen Daten dürfen gemäß Art 89 Abs 1 DSGVO grundsätzlich unbeschränkt gespeichert werden.

Es besteht das Recht auf Auskunft durch die Verantwortliche an dieser Studie über die erhobenen personenbezogenen Daten sowie das Recht auf Berichtigung, Löschung, Einschränkung der Verarbeitung der Daten sowie ein Widerspruchsrecht gegen die Verarbeitung sowie das Recht auf Datenübertragbarkeit.

Wenn Sie Fragen zu dieser Erhebung haben, wenden Sie sich bitte gern an mich, die Verantwortliche dieser Untersuchung:

Pauline Pondorf (a1214625@@unet.univie.ac.at)
Studentin der Studienrichtung Publizistik- und Kommunikationswissenschaft
MA. Universität Wien

Für grundsätzliche juristische Fragen im Zusammenhang mit der DSGVO/FOG und studentischer Forschung wenden Sie sich an den Datenschutzbeauftragen der Universität Wien, Dr. Daniel Stanonik, LL.M. (verarbeitungsverzeichnis@univie.ac.at). Zudem besteht das Recht der Beschwerde bei der Datenschutzbehörde (bspw. über dsb@dsb.gv.at).

Triggerwarnung:

In der Umfrage können sensible Themen und Inhalte vorkommen, darunter explizit misogyne Kommentare, die emotional belastend sein können. Falls Sie dies überfordert oder unangenehm für Sie ist, haben Sie jederzeit die Möglichkeit, die Teilnahme zu beenden. Ich stehe Ihnen zur Verfügung, falls Sie während der Umfrage Unterstützung benötigen oder sich unwohl fühlen sollten.

Außerdem wird in der folgenden Studie von binären Geschlechterverhältnissen ausgegangen, um die Forschungsfragen präzise zu untersuchen. Dabei möchte ich betonen, dass ich selbstverständlich die Vielfalt der Geschlechtsidentitäten anerkenne und respektiere. Für den Zweck dieser Studie war es jedoch notwendig, die Analyse auf binäre Kategorien zu beschränken, um aussagekräftige Ergebnisse zu erhalten. Diese Entscheidung dient ausschließlich der wissenschaftlichen Methodik und spiegelt keineswegs eine Ablehnung anderer Geschlechtsidentitäten wider.

Vielen Dank für die Teilnahme.

Pauline Pondorf

PS: Nutzer:innen der Forschungsplattform SurveyCircle.com erhalten für ihre Teilnahme SurveyCircle-Punkte.

1. Ich habe den obigen Text verstanden und stimme zu, an dieser Studie teilzunehmen.

[Bitte auswählen]

Weiter

B.A. Pauline Pondorf – 2024

soSci
oFb - der onlineFragebogen

16% ausgefüllt

2. Bitte beantworten Sie folgende Frage

Besitzen Sie einen Account auf mindestens einer Social Media Plattform (z.B. WhatsApp, Facebook, Instagram, Snapchat, YouTube, etc.)?

- ◯ ja
- ◯ nein

Weiter

soSci
oFb - der onlineFragebogen

26% ausgefüllt

3. Welchem Geschlecht fühlen Sie sich zugehörig?

- ◯ weiblich
- ◯ männlich
- ◯ divers
- ◯ keine Angabe

4. Wie alt sind Sie?

- ◯ 17 oder jünger
- ◯ 18-20
- ◯ 21-29
- ◯ 30-39
- ◯ 40-49
- ◯ 50-59
- ◯ 60 oder älter

5. Was ist der höchste Schul- oder Bildungsabschluss, den Sie erreicht haben?

- ◯ Ich habe keine Schule besucht.
- ◯ Ich habe keine Schule abgeschlossen.
- ◯ Volks- Grundschule oder weniger
- ◯ Hauptschule oder AHS-Unterstufe
- ◯ Sonder- Förderschule
- ◯ Polytechnikum
- ◯ Lehre/Berufsschule
- ◯ BMS [Fachschule, z.B. HASCH]
- ◯ AHS mit Matura
- ◯ BHS mit Matura [z.B. HTL, HAK, HBLA]
- ◯ Hochschulverwandte Lehranstalt
- ◯ Kolleg
- ◯ Bachelor/Bakkalaureat
- ◯ Magister/Master/DI/FH
- ◯ Doktor/PhD
- ◯ Andere

Weiter

soSci
oFb - der onlineFragebogen

39% ausgefüllt

Es wird Ihnen im Nachfolgenden ein Posting präsentiert. Bitte lesen Sie sich die Kommentare konzentriert durch. Wenn Sie alles aufmerksam gelesen haben, können Sie am Ende der Seite mit einem Klick auf "weiter" mit der Befragung fortfahren.

Weiter

Weiter

soSci
oFb - der onlineFragebogen

59% ausgefüllt

6. Nun geht es darum, wie Sie die Kommentare, die Sie gerade gelesen haben, bewerten. Inwiefern treffen die folgenden Aussagen in Bezug auf die Kommentare zu?

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Die Kommentare waren humorvoll gemeint.	○	○	○	○	○
Die Kommentare waren eindeutig hasserfüllt.	○	○	○	○	○
Die Kommentare zeigten eine klare Absicht, zu verletzen und zu bedrohen.	○	○	○	○	○
Die Kommentare waren als Scherze oder satirisch gemeint.	○	○	○	○	○

Weiter

soSci
oFb - der onlinefragebogen

69% ausgefüllt

7. Auf den nächsten Fragebogen-Seiten bekommen Sie mehrere Aussagen zu sehen.
Zunächst möchte ich gerne von Ihnen wissen, wie Sie sich fühlen. Die folgenden Wörter beschreiben unterschiedliche Gefühle und Empfindungen. Lesen Sie jedes Wort und tragen Sie dann in die Skala neben jedem Wort die Intensität ein. Sie haben die Möglichkeit, zwischen fünf Abstufungen zu wählen. Geben Sie bitte an, wie Sie sich in diesem Moment fühlen.

Wie sehr fühlen Sie sich im Moment ...?	gar nicht	ein wenig	mäßig	ziemlich	sehr
interessiert?	○	○	○	○	○
verärgert?	○	○	○	○	○
freudig erregt?	○	○	○	○	○
wach?	○	○	○	○	○
stark?	○	○	○	○	○

Wie sehr fühlen Sie sich im Moment ...?	gar nicht	ein wenig	mäßig	ziemlich	sehr
beschämt?	○	○	○	○	○
begeistert?	○	○	○	○	○
schuldig?	○	○	○	○	○
entschlossen?	○	○	○	○	○
feindselig?	○	○	○	○	○

Wie sehr fühlen Sie sich im Moment ...?	gar nicht	ein wenig	mäßig	ziemlich	sehr
aufmerksam?	○	○	○	○	○
gereizt?	○	○	○	○	○
angeregt?	○	○	○	○	○
ängstlich?	○	○	○	○	○
aktiv?	○	○	○	○	○

Wie sehr fühlen Sie sich im Moment ...?	gar nicht	ein wenig	mäßig	ziemlich	sehr
bekümmert?	○	○	○	○	○
stolz?	○	○	○	○	○
erschrocken?	○	○	○	○	○
durcheinander?	○	○	○	○	○
nervös?	○	○	○	○	○

Weiter

soSci
oFb - der onlineFragebogen

79% ausgefüllt

Inwiefern stimmen Sie folgenden Aussagen zu? Bitte kreuzen Sie Zutreffendes an.

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Ich fühle mich persönlich verpflichtet, den Kommentaren zu widersprechen.	○	○	○	○	○
Ich sehe es nicht als meine persönliche Pflicht an, einzugreifen.	○	○	○	○	○
Ich schreibe einen Kommentar, der den Kommentaren sachlich widerspricht.	○	○	○	○	○
Ich verfasse keinen Kommentar, um Personen zu unterstützen.	○	○	○	○	○
Ich widerlege den Kommentar mit sachlichen Argumenten.	○	○	○	○	○
Ich verfasse einen Kommentar, der die Aussagen klar und direkt kritisiert.	○	○	○	○	○

Weiter

soSci
oFb - der onlineFragebogen

8. Zuletzt geht es um Aussagen über Männer und Frauen, sowie deren Beziehungen zueinander.

Inwiefern stimmen Sie folgenden Aussagen zu? Bitte kreuzen Sie Zutreffendes an.

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Wenn Frauen weniger verdienen als Männer, liegt das daran, dass sie andere Entscheidungen treffen.	○	○	○	○	○
Gleichstellungsmaßnahmen bevorzugen Frauen auf Kosten der Männer.	○	○	○	○	○
Wenn eine politische Maßnahme speziell Frauen unterstützt, ist das gegenüber Männern unfair.	○	○	○	○	○

Inwiefern stimmen Sie folgenden Aussagen zu? Bitte kreuzen Sie Zutreffendes an.

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Frauen sind von Natur aus emotionaler als Männer.	○	○	○	○	○
Männer sind von Natur aus aggressiver als Frauen.	○	○	○	○	○
Die biologischen Unterschiede zwischen Männern und Frauen sind wichtiger als die Erziehung, um ihre Verhaltensunterschiede zu erklären.	○	○	○	○	○

Inwiefern stimmen Sie folgenden Aussagen zu? Bitte kreuzen Sie Zutreffendes an.

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Es ist wichtig, Kindern beizubringen, sich entsprechend ihres Geschlechts zu verhalten.	○	○	○	○	○
Es wird eher akzeptiert, wenn ein Mädchen sich wie ein Wildfang verhält, als wenn ein Junge sich wie ein „Weichei" benimmt.	○	○	○	○	○
Es ist besser, Mädchen zu pflegenden und Jungen zu versorgenden Rollen zu erziehen, als umgekehrt.	○	○	○	○	○

Inwiefern stimmen Sie folgenden Aussagen zu? Bitte kreuzen Sie Zutreffendes an.

	stimme gar nicht zu 0	1	2	3	stimme voll zu 4
Sexismus ist in der heutigen Gesellschaft kein großes Problem mehr.	○	○	○	○	○
Heutzutage haben Frauen dieselben Chancen wie Männer.	○	○	○	○	○
Die Geschlechterungleichheit ist nicht mehr so schlimm wie früher.	○	○	○	○	○

Weiter

soSci
oFb - der onlineFragebogen

99% ausgefüllt

Vielen Dank für Ihre Teilnahme an dieser Studie!

Das Ziel der Studie war es, herauszufinden, wie verschiedene Arten von Kommentaren zum Weltfrauentag auf Social Media wahrgenommen werden und welche Reaktionen und Effekte sie auslösen. Dabei habe ich zwischen explizit hasserfüllten Kommentaren und solchen mit humorvollen Elementen unterschieden. Diese Untersuchung soll helfen, das Verständnis für die Dynamik und Rezeption von Hatespeech, vor allem mit humorvollen Elementen, zu vertiefen und Grundlagen für Gegenmaßnahmen zu schaffen.

Humorvolle Hatespeech stellt dabei ein besonderes Problem dar, weil sie Vorurteile und negative Stereotype subtil verstärken und gesellschaftlich akzeptabler erscheinen lassen kann. Durch die Einbettung von Hassrede in humoristische Kontexte wird die feindselige Natur oft verschleiert, was die Wahrnehmung und Reaktion der Rezipient:innen womöglich beeinflussen kann.

Einige der Kommentare, die Sie gesehen haben, sind echt und stammen von Facebook-Postings. Andere wurden für das Experiment manipuliert. Diese Manipulation war notwendig, um meine Forschungsfragen zu untersuchen.

Falls Sie Fragen oder Bedenken zu Ihrer Studienteilnahme haben oder über die Ergebnisse dieser Forschung informiert werden möchten, können Sie dies im untenstehenden Feld (mit Angabe Ihrer E-Mail-Adresse) äußern.

Wenn das Thema belastend für Sie war, Sie Hilfe benötigen oder mehr Informationen über Hatespeech haben wollen, finden Sie hier Strategien und weitere Informationen:

Für allgemeine Unterstützung, Hilfe und Fragen zum Umgang mit Hassrede empfehle ich den Helpdesk der Neuen Deutschen Medienmacher:

https://helpdesk.neuemedienmacher.de/

Weitere Informationen und Strategien im Umgang mit Hassrede finden Sie hier:

https://no-hate-speech.de/

Zum Abschluss möchte ich mich nochmals herzlich für Ihre Teilnahme bedanken. Ihre Unterstützung ist von großer Bedeutung für diese Forschung.

Für Nutzer von SurveyCircle: Der Survey Code lautet: WHV4-T8NG-SGUS-PYUX

Freundliche Grüße,

Pauline Pondorf

9. Falls Sie Anregungen und Fragen haben, nutzen Sie bitte nachfolgende Kommentarfunktion.

Weiter

A.3 Fragebogen Vorstudie

0% ausgefüllt

Sehr geehrte Teilnehmer:innen,

herzlich willkommen zum Expert:innencheck im Rahmen meiner Vorbereitungen für eine umfassende Studie von Kommentaren zum Weltfrauentag in den sozialen Medien. Ihre Teilnahme ist von großem Wert, um erste Einblicke und Perspektiven zu gewinnen, die mich bei der späteren Analyse unterstützen werden.

Ziel und Zweck der Vorstudie:

Diese Vorstudie zielt darauf ab, grundlegende Themen und Tonlagen zu identifizieren, die in den Kommentaren zum Weltfrauentag auf Plattformen wie Twitter, Facebook und Instagram zu finden sind.

Ablauf:

Während dieses Expert:innenchecks bitte ich Sie, Ihre Einschätzungen und Interpretationen der Kommentare offen zu teilen. Es gibt keine richtigen oder falschen Antworten; ich interessiere mich für Ihre persönliche Perspektive. Sie haben jederzeit die Möglichkeit, die Teilnahme abzubrechen, insbesondere falls Ihnen die Inhalte unangenehm werden sollten.

Triggerwarnung:

In der Vorstudie können explizite misogynistische und gewaltverherrlichende Kommentare vorkommen. Falls Sie dies belastet oder überfordert, haben Sie jederzeit die Möglichkeit, die Teilnahme zu beenden. Ich stehe Ihnen jederzeit zur Verfügung, falls Sie während des Expert:innen-Checks Unterstützung benötigen oder sich unwohl fühlen sollten.

Nochmals vielen Dank für Ihre Teilnahme und Ihre wertvollen Beiträge zu diesem Expert:innencheck.

Sollten Sie weitere Fragen haben, stehe ich Ihnen gerne zur Verfügung.

Mit freundlichen Grüßen,

Pauline Pondorf

Weiter


1. Lesen Sie sich folgende Aussagen genau durch und kreuzen Sie nach Ihrer Einschätzung an.

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Wo bleibt meine popcorn Tüte bei so einem Emanzen haufen? ready to battle every guy 😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Heute feiern wir Frauen für ihre unglaubliche Fähigkeit, stundenlang zu shoppen und trotzdem nichts Passendes zu finden	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Frauen seh ich da aber relativ wenige eher so hysterische Kampfemanzen die gegen jeden Kerl in den Krieg ziehen ... meine Güte.	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Frauen sollen in die Küche gehen... Abwaschen saugen putzen essen anrichten. Und nicht ständig wegen gleichberechtigung rumjammern	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Sollen die dich mal die Arbeit machen, was wir Männer machen denn wir wissen alle, dass es eine olympische Disziplin ist, den Kühlschrank zu plündern 😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Haha ungelogen dieses Frauen Tamtam...Vielleicht sollten sie eine eigene Band gründen – die Tam-Tam schwestern	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Und eigentlich wollen die bitches doch alle mal nur wieder gevögelt werden...	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Frauen sind genauso gleichberichtg wie Männer nur eben auf falscher weiße ... eigentlüch müssten mütter gleichberechtigt werden und nicht die frauen an sich	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Ich könnte mich nicht zusammenreißen würde ich vor so einer dummen feministenfotze stehen, weiß nicht ob mir da die hand ausrutschen würde...	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Wenn meine Frau da mitlaufen würde würde sie im keller landen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Das heißt wohl, dass ich heute extra freundlich sein muss wenn meine alte mir erklärt, wie schwer es ist, ein kissen auf die couch zu legen??!?!?	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
„die Frauen sind doch die ersten die heulen wenn was wehtut vorallem bei der Arbeit und brauchen dann Sonderrechte.	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Ja die armen frauen... die sollen doch mal welche Frauen Europa „verenden“ wollten...	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
viele Frauen beweisen doch wie verblödet sie sind indem sie sich operieren lassen und aufblasen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Bei dem Theater schäme ich mich schön langsam eine Frau zu sein	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Aber nackte Ti**en euer Recht und Ausdruck unser so einer fortschrittlichen Gesellschaft	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Das Problem ist das Frauen sich aussuchen wann sie emanzipiert sind und wann nicht ! Geht es um Führungspositionen gleiches Gehalt wie die Männer usw usw sind sie emanzipiert ! Geht es darum 2x 25 Kilo Säcke Zement zu schleppen sind es wieder die armen zarten Mädels	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Die müssen einfach alle mal vergewaltigt werden damit die aus ihrem Feminismuswahn aufwachen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Ein großteil der Frauen können doch mittlerweile nichtmal mehr kochen...	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Frauen in Österreich protestieren dass es ihnen schlecht ergeht.. die armen, Den Frauen geht's schon zu gut in Österreich alles nur hysterische Weiber	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Dümmer geht's immer, besonders in berlin dumm mit Auszeichnung bestanden haben die Frauen 😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Haben die eigentlich flexible arbeitszeiten wie oft die sich aussuchen wann sie emanzipiert sein wollen und wann nicht haha?	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Sollen mal die Arbeit machen was wir männer machen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Da müsste mal einer mit der pistole durchlaufen und bisschen aufräumen aber dann müssten sie ja wieder einen starken mann zu hilfe rufen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Jetzt fehlt aber noch eine Liste mit den Frauen, die Europa „verenden" wollen... Vielleicht wird es ja sogar eine Bestseller-Liste	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Die müssen einfach mal wieder die beine breit machen und schon sind ihre „probleme gelöst"	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Der Tag, an dem Frauen uns erzählen, wie hart ihr Leben ist, während sie sich ihre Nägel machen lassen 💅😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Hä Frauen dürfen sich doch eh stark und unabhängig fühlen – solange sie den Müll rausbringen, bevor ich nach Hause komme, verstehe das problem nicht	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
mir egal hauptsache die frau kann kochen und putzen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
würden die Frauen endlich damit aufhörn der Chef sein zu wollen, ist doch klar, dass männer intelligenter sind	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Heute feiere ich frauen für ihre wirklich unglaubliche Fähigkeit, stundenlang zu reden und trotzdem nichts Wichtiges zu sagen 😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Weltfrauentag – habe schon einen Blumenstrauß für die Geschirrspülmaschine gekauft hahaha	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Die sind so arm und immer benachteiligt... Mimmimimimi	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Dieses Frauen Tam – Tam nervt mich gewaltig!	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Was wollen die feministischen Berufsempörten jetzt schon wieder? Mich langweilts	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Diese mädels sollte man mal wieder gscheid ficken dass sie auf klare Gedanken kommen	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Klingt wie ein guter Tag, um meiner Frau zu erklären, dass sie genauso gut Mathe kann wie ich... solange es um kochrezepte geht	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Ja ja wo Geld da ist Aber manche Frauen kann man dort nicht brauchen weil sie zu blöd sind.	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Witz des Jahres: Frauen in Österreich protestieren, dass es ihnen schlecht ergeht.... Wollen die nen oscar für hysterische darstellung?	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Wann ist eigentlich nochmal weltmännertag?.. damit wir endlich auch mal einen Grund haben uns über ungebügelte Hemden zu beschweren.🤷‍♀️😂	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Dümmer gehts immer besonders in berlin die Frauen sind so lächerlich	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Vielleicht gibts bald ne Petition für ‚Betatsch-Rechte' 🤣	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Einfach nur peinlich....vielleicht nennen sie sich bald peinlich emanzen und sind bei rtl 2 zu sehen..	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Weltfrauentag? Das hieß doch früher mal frühjahrsputz – nur dass heute staubsauger und putzmittel endlich genderneutral sind höhöhö	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Die sind soooooo immer benachteiligt. Mimimimimi fast wie der Refrain eines neuen Pop-Hits	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Inwiefern stimmen Sie zu, dass folgende Aussagen humorvolle Elemente enthalten?

Die Sorgen dieser geistesgestörten Frauen möchte ich haben.... Alles eine Ausgeburt des Feminismus	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Ja fühlt euch heute alle mal stark und unabhängig, bevor ihr wieder fragt, wie man den Fernseher bedient höhöh	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Einfach nur peinlich kann man nicht mehr als Frauen bezeichne	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu
Zeit für ne neue kochsendung -Kochen für Emanzen, damit die auch mal wieder kochen lernen hahaha	Ich stimme gar nicht zu	○○○○○	Ich stimme sehr zu

Weiter

soSci
oFb - der onlineFragebogen

67% ausgefüllt

Vielen Dank für Ihre Teilnahme!

Die Vorstudie zielt darauf ab, Kommentare zum Weltfrauentag, die in sozialen Medien verbreitet werden, zu analysieren und zu klassifizieren. Besonderes Augenmerk liegt dabei darauf festzustellen, ob diese Kommentare explizite Hassrede oder humorvolle Hatespeech darstellen. Diese Untersuchung dient als Vorbereitung für eine umfassendere Hauptstudie, die die Auswirkungen solcher Kommentare auf Rezipient:innen differenziert nach Geschlecht untersuchen wird. Ziel ist es, ein vertieftes Verständnis für die Dynamik und die Rezeption von Hatespeech zu gewinnen und Grundlagen für wirkungsvolle Gegenmaßnahmen zu erarbeiten.

Falls Sie Fragen oder Anmerkungen haben, können Sie mich gerne kontaktieren unter:

a12146250@unet.univie.ac.at

oder einen Kommentar unten hinterlassen.

Mit freundlichen Grüßen
Pauline Pondorf

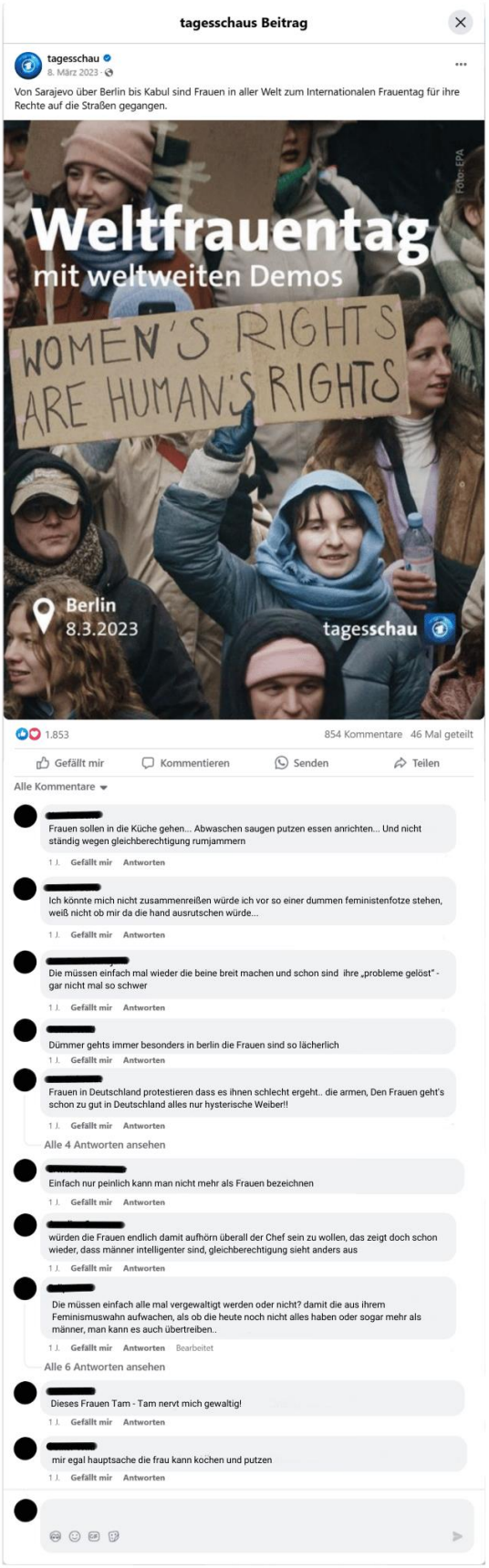
2. Haben Sie noch Anmerkungen?

Weiter

A.4 Stimulusmaterial

Experimentalgruppe 1: Subtile Hatespeech mit humorvollen Elementen

Experimentalgruppe 2 bzw. Kontrollgruppe: Explizite Hatespeech



A.5 SPSS-Output der Ergebnisse

T-Test H 1.1

Gruppenstatistiken

	ZH01	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
SM_insg	1	101	3,1520	,50159	,04991
	2	104	3,0938	,50105	,04913

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
									95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
SM_insg	Varianzen sind gleich	,083	,774	,831	203	,407	,05823	,07003	-,07986	,19632
	Varianzen sind nicht gleich			,831	202,811	,407	,05823	,07004	-,07986	,19632

T-Test H 1.2

Gruppenstatistiken

	ZH01	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
CS_insg	1	101	2,4983	,85293	,08487
	2	104	2,7067	1,03746	,10173

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
									95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
CS_insg	Varianzen sind gleich	7,895	,005	-1,568	203	,118	-,20838	,13286	-,47035	,05359
	Varianzen sind nicht gleich			-1,573	197,652	,117	-,20838	,13248	-,46965	,05288

T-Test H 1.3

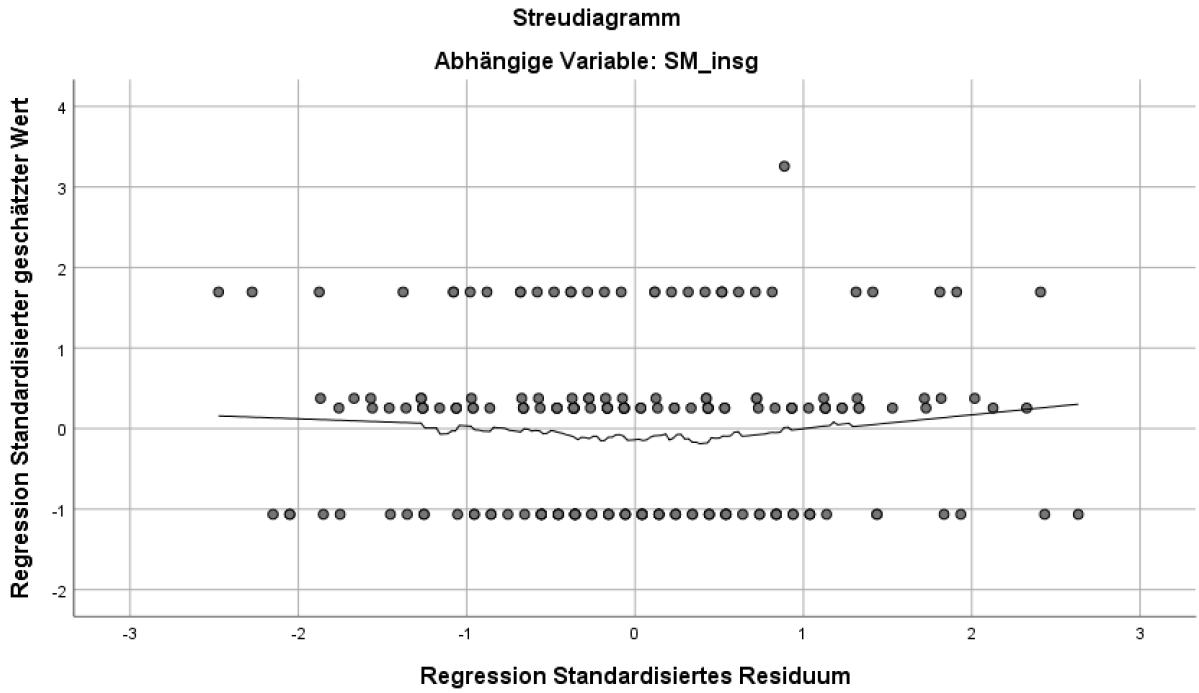
Gruppenstatistiken

	ZH01	N	Mittelwert	Std.-Abweichung	Standardfehler des Mittelwertes
sex_ing	1	101	2,0685	,67128	,06679
	2	104	2,0312	,61129	,05994

Test bei unabhängigen Stichproben

		Levene-Test der Varianzgleichheit		T-Test für die Mittelwertgleichheit						
									95% Konfidenzintervall der Differenz	
		F	Signifikanz	T	df	Sig. (2-seitig)	Mittlere Differenz	Standardfehler der Differenz	Untere	Obere
sex_ing	Varianzen sind gleich	,682	,410	,415	203	,678	,03723	,08962	-,13948	,21395
	Varianzen sind nicht gleich			,415	199,993	,679	,03723	,08975	-,13974	,21420

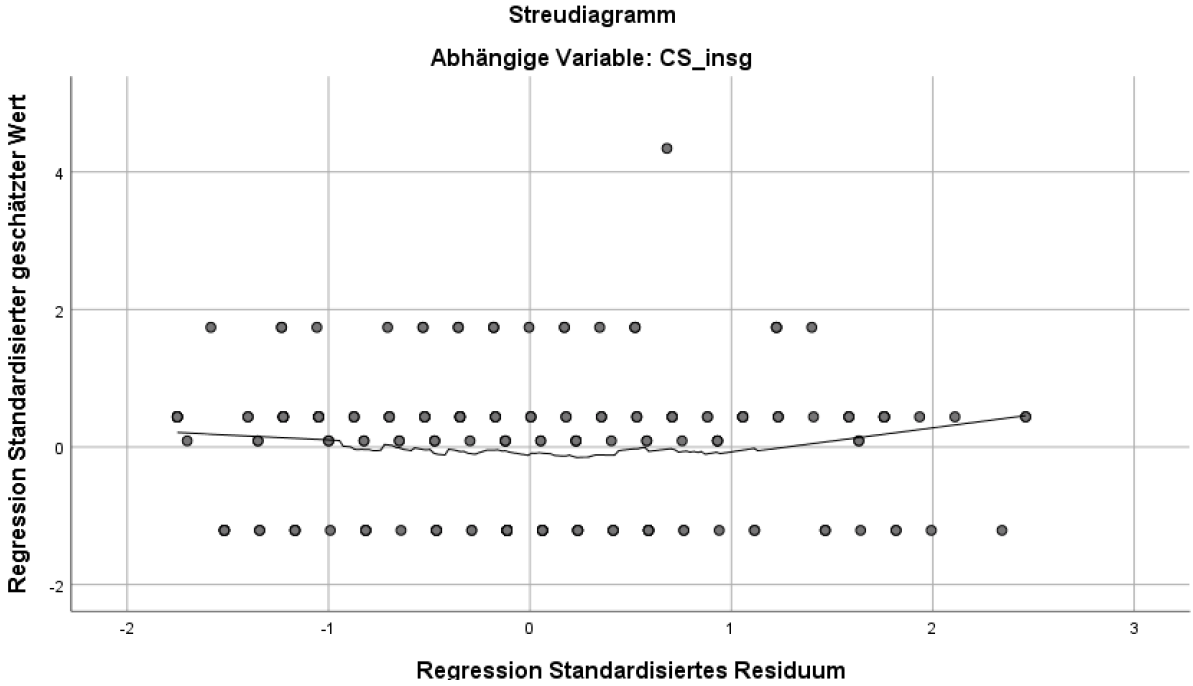
Moderationsanalyse H 2.1



```
Run MATRIX procedure:

***************** PROCESS Procedure for SPSS Version 4.2 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : SM_insg
    X  : ZH01
    W  : DM01

Sample
Size:  205

**************************************************************************
OUTCOME VARIABLE:
 SM_insg

Model Summary
          R       R-sq        MSE     F(HC3)        df1        df2
 p
      ,0911      ,0083      ,2526      ,5631     3,0000   201,0000      ,64
00

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     3,1241      ,0355    88,0116      ,0000     3,0541     3,1941
ZH01         -,0540      ,0710     -,7609      ,4476     -,1940      ,0860
DM01          ,0563      ,0786      ,7166      ,4744     -,0986      ,2112
Int_1         ,0857      ,1573      ,5449      ,5864     -,2245      ,3959

Product terms key:
 Int_1    :        ZH01     x        DM01

Test(s) of highest order unconditional interaction(s):
       R2-chng     F(HC3)        df1        df2          p
X*W      ,0017      ,2969     1,0000   201,0000      ,5864
----------
    Focal predict: ZH01     (X)
          Mod var: DM01     (W)

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.
```

```
DATA LIST FREE/
   ZH01       DM01       SM_insg    .
BEGIN DATA.
     -,5073     -,2878     3,1478
      ,4927     -,2878     3,0691
     -,5073      ,0000     3,1515
      ,4927      ,0000     3,0975
     -,5073      ,4852     3,1577
      ,4927      ,4852     3,1453
END DATA.
GRAPH/SCATTERPLOT=
 DM01     WITH     SM_insg  BY       ZH01     .

*********** BOOTSTRAP RESULTS FOR REGRESSION MODEL PARAMETERS ************

OUTCOME VARIABLE:
 SM_insg

              Coeff    BootMean     BootSE    BootLLCI    BootULCI
constant     3,1241     3,1231      ,0354     3,0552     3,1941
ZH01         -,0540     -,0551      ,0704     -,1930      ,0841
DM01          ,0563      ,0523      ,0783     -,1082      ,2017
Int_1         ,0857      ,0771      ,1585     -,2470      ,3799

************************ ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000

NOTE: One SD below the mean is below the minimum observed in the data for W
,
      so the minimum measurement on W is used for conditioning instead.

NOTE: A heteroscedasticity consistent standard error and covariance matrix
estimator was used.

NOTE: The following variables were mean centered prior to analysis:
          DM01     ZH01

------ END MATRIX -----
```

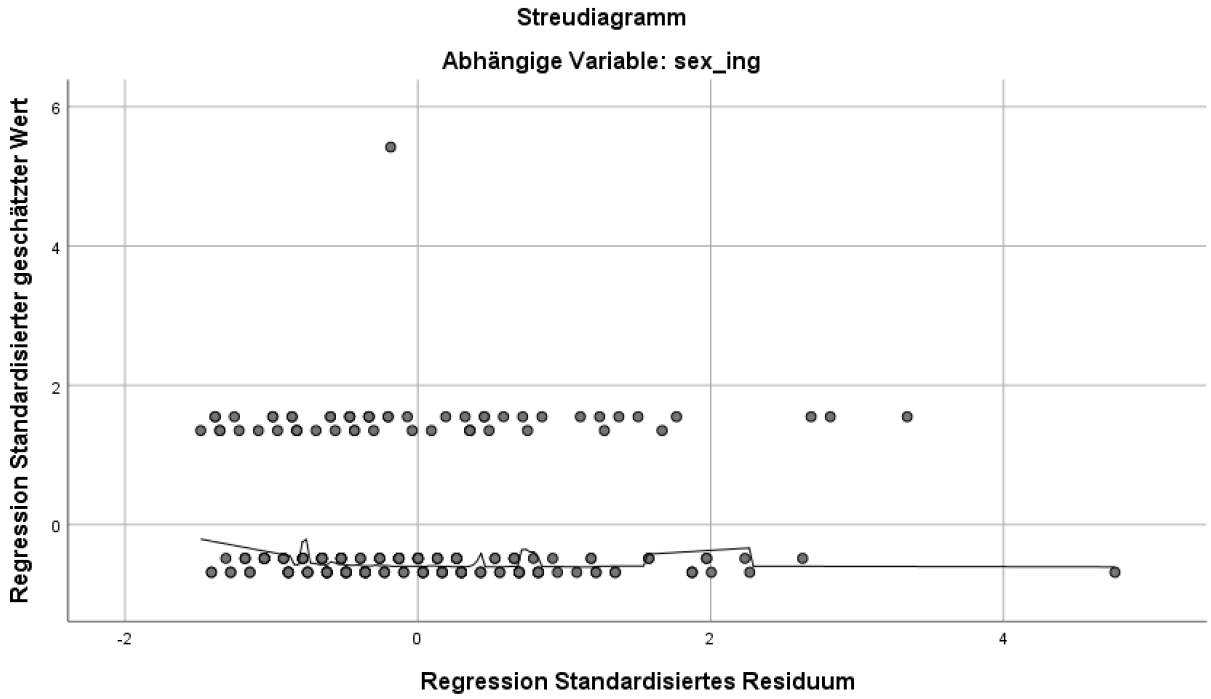
Moderationsanalyse H 2.2



```
Run MATRIX procedure:

***************** PROCESS Procedure for SPSS Version 4.2 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : CS_insg
    X  : ZH01
    W  : DM01

Sample
Size:  205

**************************************************************************
OUTCOME VARIABLE:
 CS_insg

Model Summary
          R       R-sq        MSE     F(HC3)        df1        df2
 p
      ,1430      ,0204      ,9057     1,8030     3,0000   201,0000      ,14
78

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     2,6059      ,0667    39,0604      ,0000     2,4744     2,7375
ZH01          ,2215      ,1331     1,6637      ,0977     -,0410      ,4840
DM01          ,1722      ,1224     1,4075      ,1608     -,0691      ,4136
Int_1         ,0981      ,2449      ,4007      ,6891     -,3847      ,5809

Product terms key:
 Int_1    :        ZH01     x        DM01

Test(s) of highest order unconditional interaction(s):
       R2-chng     F(HC3)        df1        df2          p
X*W      ,0006      ,1606     1,0000   201,0000      ,6891
----------
    Focal predict: ZH01     (X)
          Mod var: DM01     (W)

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.
```

```
DATA LIST FREE/
   ZH01       DM01       CS_insg    .
BEGIN DATA.
     -,5073     -,2878     2,4583
      ,4927     -,2878     2,6516
     -,5073      ,0000     2,4936
      ,4927      ,0000     2,7151
     -,5073      ,4852     2,5530
      ,4927      ,4852     2,8221
END DATA.
GRAPH/SCATTERPLOT=
 DM01     WITH     CS_insg  BY       ZH01     .

*********** BOOTSTRAP RESULTS FOR REGRESSION MODEL PARAMETERS ************

OUTCOME VARIABLE:
 CS_insg

              Coeff    BootMean     BootSE   BootLLCI   BootULCI
constant     2,6059     2,6048      ,0659     2,4765     2,7378
ZH01          ,2215      ,2182      ,1328     -,0423      ,4776
DM01          ,1722      ,1653      ,1217     -,0845      ,3999
Int_1         ,0981      ,0807      ,2438     -,4149      ,5451

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000

NOTE: One SD below the mean is below the minimum observed in the data for W
,
      so the minimum measurement on W is used for conditioning instead.

NOTE: A heteroscedasticity consistent standard error and covariance matrix
estimator was used.

NOTE: The following variables were mean centered prior to analysis:
          DM01     ZH01

------ END MATRIX -----
```

Moderationsanalyse H 2.3



```
Run MATRIX procedure:

***************** PROCESS Procedure for SPSS Version 4.2 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2022). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : sex_ing
    X  : ZH01
    W  : DM01

Sample
Size:  205

**************************************************************************
OUTCOME VARIABLE:
 sex_ing

Model Summary
          R       R-sq        MSE     F(HC3)        df1        df2
 p
      ,2244      ,0504      ,3951     2,6587     3,0000   201,0000      ,04
94

Model
              coeff    se(HC3)          t          p       LLCI       ULCI
constant     2,0418      ,0437    46,6768      ,0000     1,9555     2,1280
ZH01         -,0196      ,0875     -,2235      ,8234     -,1922      ,1530
DM01          ,2243      ,0953     2,3534      ,0196      ,0364      ,4123
Int_1        -,4072      ,1914    -2,1275      ,0346     -,7846     -,0298

Product terms key:
 Int_1    :        ZH01     x        DM01

Test(s) of highest order unconditional interaction(s):
       R2-chng     F(HC3)        df1        df2          p
X*W      ,0235     4,5262     1,0000   201,0000      ,0346
----------
    Focal predict: ZH01     (X)
          Mod var: DM01     (W)

Conditional effects of the focal predictor at values of the moderator(s):
```

```
      DM01     Effect    se(HC3)          t          p       LLCI       UL
CI
    -,2878      ,0976      ,0987      ,9892      ,3237     -,0970      ,29
22
     ,0000     -,0196      ,0875     -,2235      ,8234     -,1922      ,15
30
     ,4852     -,2171      ,1338    -1,6229      ,1062     -,4809      ,04
67

Moderator value(s) defining Johnson-Neyman significance region(s):
     Value    % below    % above
    1,0593    99,5122      ,4878

Conditional effect of focal predictor at values of the moderator:
      DM01     Effect    se(HC3)          t          p       LLCI       UL
CI
    -,2878      ,0976      ,0987      ,9892      ,3237     -,0970      ,29
22
    -,1378      ,0366      ,0889      ,4112      ,6813     -,1387      ,21
18
     ,0122     -,0245      ,0878     -,2794      ,7803     -,1977      ,14
86
     ,1622     -,0856      ,0957     -,8942      ,3723     -,2744      ,10
32
     ,3122     -,1467      ,1108    -1,3242      ,1869     -,3651      ,07
17
     ,4622     -,2078      ,1305    -1,5923      ,1129     -,4651      ,04
95
     ,6122     -,2688      ,1531    -1,7565      ,0805     -,5707      ,03
30
     ,7622     -,3299      ,1774    -1,8596      ,0644     -,6798      ,01
99
     ,9122     -,3910      ,2029    -1,9269      ,0554     -,7911      ,00
91
    1,0593     -,4509      ,2287    -1,9718      ,0500     -,9018      ,00
00
    1,0622     -,4521      ,2292    -1,9726      ,0499     -,9040     -,00
02
    1,2122     -,5132      ,2560    -2,0048      ,0463    -1,0179     -,00
84
    1,3622     -,5742      ,2831    -2,0281      ,0439    -1,1325     -,01
59
    1,5122     -,6353      ,3106    -2,0456      ,0421    -1,2477     -,02
29
```

```
     1,6622     -,6964      ,3382    -2,0589      ,0408    -1,3633     -,02
95
     1,8122     -,7575      ,3660    -2,0693      ,0398    -1,4793     -,03
57
     1,9622     -,8186      ,3940    -2,0776      ,0390    -1,5955     -,04
17
     2,1122     -,8796      ,4220    -2,0843      ,0384    -1,7118     -,04
74
     2,2622     -,9407      ,4502    -2,0897      ,0379    -1,8284     -,05
31
     2,4122    -1,0018      ,4784    -2,0942      ,0375    -1,9451     -,05
85
     2,5622    -1,0629      ,5066    -2,0980      ,0372    -2,0619     -,06
39
     2,7122    -1,1240      ,5349    -2,1011      ,0369    -2,1788     -,06
92

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.

DATA LIST FREE/
   ZH01       DM01       sex_ing    .
BEGIN DATA.
     -,5073     -,2878     1,9277
      ,4927     -,2878     2,0253
     -,5073      ,0000     2,0517
      ,4927      ,0000     2,0321
     -,5073      ,4852     2,2608
      ,4927      ,4852     2,0436
END DATA.
GRAPH/SCATTERPLOT=
 DM01     WITH     sex_ing  BY       ZH01     .

*********** BOOTSTRAP RESULTS FOR REGRESSION MODEL PARAMETERS ************

OUTCOME VARIABLE:
 sex_ing

              Coeff    BootMean     BootSE   BootLLCI   BootULCI
constant     2,0418     2,0406      ,0435     1,9579     2,1263
ZH01         -,0196     -,0212      ,0875     -,1946      ,1487
DM01          ,2243      ,2206      ,0941      ,0329      ,4032
Int_1        -,4072     -,4167      ,1889     -,8049     -,0589

*********************** ANALYSIS NOTES AND ERRORS ************************
```

```
Level of confidence for all confidence intervals in output:
  95,0000

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000

W values in conditional tables are the minimum, the mean, and 1 SD above th
e mean.

NOTE: One SD below the mean is below the minimum observed in the data for W
,
      so the minimum measurement on W is used for conditioning instead.

NOTE: A heteroscedasticity consistent standard error and covariance matrix
estimator was used.

NOTE: The following variables were mean centered prior to analysis:
          DM01     ZH01

------ END MATRIX -----
```