



*„Hate Speech auf österreichischen und deutschen
Nachrichtenportalen“*

Forschungsprojekt KOA 8.064/2021-001

Kurzbericht

Februar 2025

gefördert aus Mitteln der KommAustria/RTR, gemäß § 11 Abs. 3 PresseFG 2004

Projektleitung: Univ.-Prof. Clemens Hutzinger, Ph.D., Privatuniversität Schloss Seeburg, Seeburgstraße 8, 5201 Seekirchen am Wallersee, Österreich

Forschungspartner: Priv.-Doz. Wolfgang Weitzl, Ph.D., FH Oberösterreich, Wehrgrabengasse 1, 4400 Steyr, Österreich

Inhalt

| | |
|---|----|
| Hintergrund und Problemstellung..... | 1 |
| Forschungsfragen und Begründung der methodischen Vorgehensweise..... | 4 |
| Ausgewählte Empirische Studien und Ergebnisse..... | 6 |
| Studie 1 „Eigenschaften der Online-Hater:innen in Österreich“ (n = 554)..... | 6 |
| Studie 2 „Motive und Reaktionen der Online-Hater:innen in Österreich und Deutschland“ (n = 1420)..... | 9 |
| Studie 3 „Auswirkungen auf Online-Hate-Beobachter:innen auf österreichischen Online-Nachrichtenportalen“ (n = 968)..... | 15 |
| Diskussion..... | 17 |
| Berichterstattung und Dissemination der Ergebnisse und Dialog mit der Gesellschaft..... | 20 |

Hintergrund und Problemstellung

Die österreichischen und deutschen Online-Nachrichtenportale sind eine zentrale Informationsquelle der Bevölkerung. Täglich nutzen hunderttausende Bürger:innen die Onlineausgaben von Zeitungen wie Die Presse, Der Standard, Die Krone oder Der Spiegel, um sich über Politik, Wirtschaft, Sport und Gesellschaft zu informieren. Neben journalistisch aufbereiteten Inhalten schätzen viele die Möglichkeit, sich in Kommentarbereichen, Foren und Social Media über aktuelle Themen auszutauschen. Damit sind diese Portale nicht nur Nachrichtenquellen, sondern auch wichtige Interaktionsplattformen der Gesellschaft.

Nutzer:innenkommentare können die Haltung der Rezipientinnen maßgeblich beeinflussen, da sie als electronic word-of-mouth gelten – öffentliche positive oder negative Stellungnahmen im Internet, die viele erreichen (Hennig-Thurau et al., 2004). Diese Kommunikationsform besitzt eine hohe persuasive Wirkung, da Sender:innen als vertrauenswürdiger wahrgenommen werden als andere Quellen (Nieto et al., 2014) und die Inhalte als unverfälscht und hilfreich gelten (Cheng & Ho, 2015).

Forschung zeigt, dass negative Kommentare über Organisationen, Personen oder Ideen oft stärkeren Einfluss auf Rezipient:innen haben als positive. Dieses Muster, bekannt als Negativity Bias, ist gut belegt (Herr et al., 1991; Salehan & Kim, 2016). Negative Kommentare gelten als informativer, analytischer und überzeugender, da sie von der gesellschaftlichen Norm abweichen, meist Positives zu äußern (Sen & Lerman, 2007). Sie beeinflussen Kaufentscheidungen, können aber Unternehmen schaden (Hutzinger & Weitzl, 2021; Weitzl & Hutzinger, 2017; Weitzl & Hutzinger, 2019; Weitzl et al., 2018). Studien belegen den starken Effekt negativer Aussagen (Cheung & Thadani, 2012), besonders problematisch bei ungerechtfertigten oder hasserfüllten Kommentaren.

Die österreichischen und deutschen Online-Nachrichtenportale haben sich zur zentralen Drehscheibe des öffentlichen Diskurses entwickelt. Eine aktuelle Erscheinung ist die zunehmende Anzahl von aggressiven Nutzer:innen-Kommentaren. Menschen nutzen heute diese Portale intensiv, um ihre Meinungen in der Form von Negativkommentaren bzw. Hasspostings zu verbreiten. Dieses Phänomen wird als „Hate Speech“ (dt. „Hassrede“) bezeichnet. Hate Speech wird in verschiedenen Wissenschaftsdisziplinen untersucht, wie etwa Psychologie (Dellagiacoma et al., 2024), Kommunikationswissenschaften (Schäfer et al., 2024), Mensch-Maschinen-Interaktion (Wachs et al., 2024), oder Kriminologie (Obermaier et al., 2023). Wissenschaftler:innen liefern zahlreiche Definitionen, die die Natur und Eigenschaften von Hate Speech

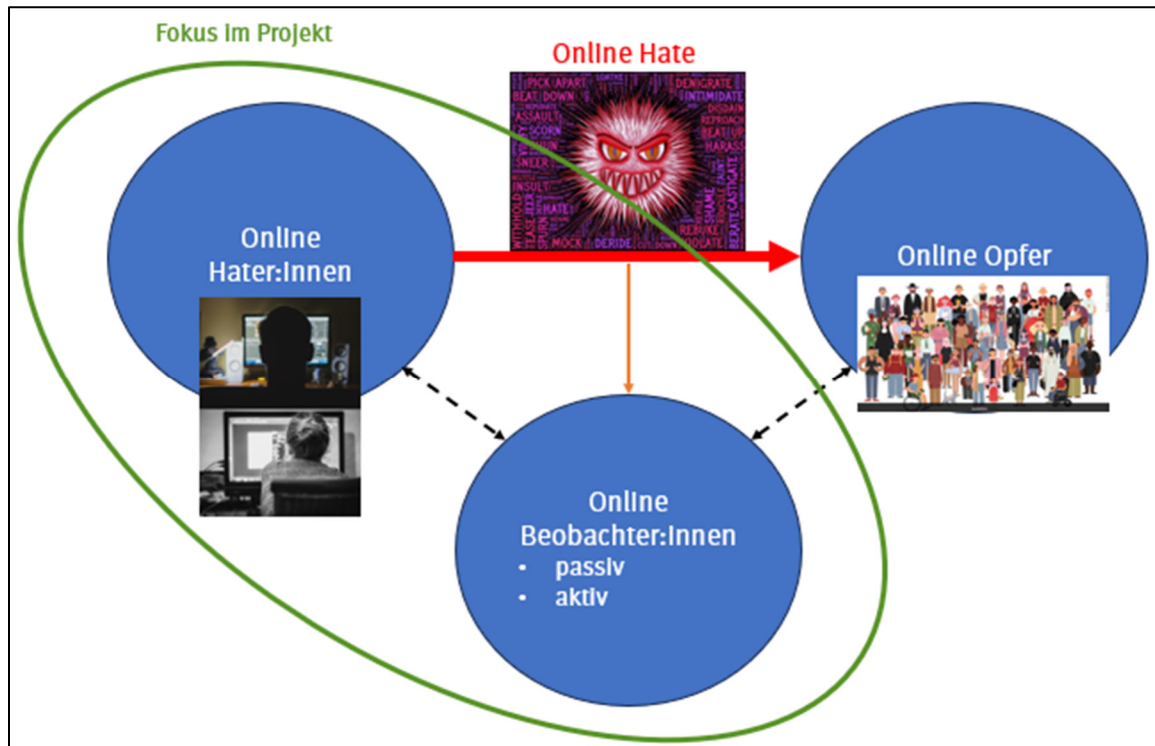
erklären sollen. Eine Gruppe von Definitionen stellt die Feindseligkeit gegenüber anderen aufgrund derer immanenten Eigenschaften (z.B. Geschlecht, Ethnie) in den Vordergrund und definiert Hate Speech als „a bias-motivated, hostile, malicious speech aimed at a person or a group of people because of some of their actual or perceived innate characteristics“ (Cohen-Almagor, 2015, S. 148). Eine weitere Gruppe von Definitionen sehen Hate Speech als einen Mechanismus der gewaltsamen Herabsetzung, welcher sich Angst, Belästigung, Einschüchterung und Diskriminierung bedient (Lederer & Delgado, 1995). Eine wiederum andere Gruppe von Definitionen betont das Potential von Hate Speech andere zu Brutalität und Gewalt aufzuhetzen (Lederer & Delgado, 1995; Meddaugh & Kay, 2009). Diese drei Gruppen von Definitionen von Hate Speech verbindet, dass diese Äußerungen umfassen, welche gegenüber den angegriffenen Personen (Organisationen etc.) eine Geringschätzung auszudrücken (Chetty & Alathur, 2018).

Diese Geringschätzung spiegelt sich auch in einer aktuellen Sichtweise von Wachs und Wright (2018) wider, welche gefragt haben „How often in the past 12 months have you posted hateful or degrading writing or speech online...?“ (S. 3). Typische, jedoch in ihrem Ausmaß variierende, Elemente von Hassnachrichten sind Beleidigung, Bedrohung, Erniedrigung, Aggression und Hasserfülltheit (Kaakinen et al., 2020 Soral et al., 2018; Pace et al., 2021; Wachs & Wright, 2018). Die Verbreitung von online Hass wird von Malecki et al. (2021) als Posten von negativen Bewertungen einer Person gesehen. Aufbauend auf diesen Ideen, wird in diesem Forschungsprojekt hinsichtlich zweier Dimensionen unterschieden:

- Online Hass-Verbreitung: *Wie oft in den letzten 12 Monaten haben Sie öffentlich auf social media Mitteilungen gepostet, die eine negative Einstellung über andere Personen ausgedrückt haben?* (zu beantworten auf einer 7-stufigen Skala: (1) nie; (2) sehr selten; (3) selten; (4) manchmal; (5) einige Mal; (6) oft; (7) sehr oft)
- Durchschnittliche Hasserfülltheit: *Die von mir in den letzten 12 Monaten öffentlich auf social media geposteten Mitteilungen, die eine negative Einstellung über andere Personen ausgedrückt haben, waren üblicherweise ...* (zu beantworten auf fünf 7-stufigen Skalen)
 - *(1) gar nicht beleidigend --- (7) sehr beleidigend*
 - *(1) gar nicht bedrohlich --- (7) sehr bedrohlich*
 - *(1) gar nicht hasserfüllt --- (7) sehr hasserfüllt*
 - *(1) gar nicht erniedrigend --- (7) sehr erniedrigend*
 - *(1) gar nicht aggressiv --- (7) sehr aggressiv*

Wenn Hass im Internet verbreitet wird, gibt es zumindest drei Personengruppen, die Online Hater:innen, die Online Opfer und die Online Beobachter:innen. Diese Projekt fokussiert sich auf die die Online Hater:innen und die Online Beobachter:innen (siehe Abbildung 1).

Abbildung 1: Personengruppen bei Hass im Internet um Fokus im Projekt (Quelle: eigene Darstellung)



Es lässt sich feststellen, dass Online-Nachrichtenportale häufig Schauplätze von Hassverbreitung sind (Mossie & Wang, 2020). Dies kann das Wohlergehen der Betroffenen stark beeinträchtigen, psychisches Leid verursachen (Calvert, 1997; Soral et al., 2018), das Selbstwertgefühl verletzen und Ängste hervorrufen. Die langfristige öffentliche Verfügbarkeit von Hasskommentaren verstärkt diese Auswirkungen und ermöglicht größeren Schaden als offline (Gagliardone et al., 2015). Zudem zeigen erste Studien einen Zusammenhang zwischen Online Hate Speech und Offline Hate Crimes – aus Hass im Netz wird Gewalt auf der Straße (Williams et al., 2020).

Nachrichtenmedien tragen Verantwortung, da sie sowohl Meinungsfreiheit als auch Hate Speech ermöglichen. Das Thema findet regelmäßig Beachtung in der Berichterstattung (z.B. Die Presse, 18.02.2025; Der Standard, 29.11.2024; Kurier, 25.02.2021; Der Standard, 16.02.2020; Kronen Zeitung, 18.11.2020; Der Spiegel, 28.12.2020; Die Zeit, 05.07.2018) und motiviert Medien, sich aktiv an Forschung und Bekämpfung zu beteiligen (z.B. Kurier, 02.04.2019; Wiener Zeitung, 15.01.2021). Auch staatliche Institutionen und der Europarat engagieren sich

(Gagliardone et al., 2015). Die Forschung steckt jedoch noch in den Anfängen und konzentriert sich meist auf Hassforen und spezifische Gruppen (MacAvaney et al., 2019; Saleem et al., 2017). Es fehlt an Studien zu den Auswirkungen von Hate Speech auf prominenten Plattformen wie Online-Nachrichtenportalen. Neue Forschung könnte Medien und Institutionen helfen, Mechanismen hinter Hasskommentaren besser zu verstehen und effektiver gegenzusteuern, um Schaden zu reduzieren und Onlineportale als konstruktive Kommunikationsräume zu stärken.

Forschungsfragen und Begründung der methodischen Vorgehensweise

Forschungsfragen. Die bestehenden Forschungslücken sollen mit den folgenden fünf Hauptforschungsfragen geschlossen werden:

Forschungsfrage 1: Welche Eigenschaften haben Personen, die Hate Speech betreiben?

Forschungsfrage 2: Was sind die Motive für die Beteiligung an Hate Speech?

Die Forschungsfragen eins und zwei untersuchen, wer auf österreichischen und deutschen Online-Nachrichtenportalen Hasskommentare verfasst und warum. Dabei werden psychologische Merkmale (z. B. Persönlichkeit, Motivation, Einstellungen) sowie weitere Eigenschaften (z. B. Verhalten, soziodemografisches Profil) analysiert. Ziel ist die Identifikation und Kategorisierung von Hate-Speech-Verfasser:innen, um zur Ursachenforschung beizutragen.

Forschungsfrage 3: Welche Kontext-Faktoren stimulieren Personen, vermehrt Hasskommentare zu posten?

Forschungsfrage 4: Wie reagieren Personen, die Hasskommentare posten, wenn ihre Kommentare gelöscht werden?

Hate Speech hat vor allem dann drastische negative Konsequenzen, wenn Hasskommentare überhandnehmen und Personen bzw. Personengruppen von mehreren Nutzer:innen angegriffen und diskreditiert werden. Im schlimmsten Fall entfaltet sich ein sogenannter „Shitstorm“ bei welchem sich eine große Masse von Individuen an einer Negativkampagne aktiv beteiligt. Die dritte Forschungsfrage soll aus der Sender:innenperspektive Klarheit darüber bringen, wie und wann eine Entrüstungswelle entsteht. Dabei sollen unterschiedliche Aktivitätslevels von Hate Speech Beteiligung (z.B. das Verfassen von Hasskommentaren, das „Liken“ von Hasskommentaren anderer) im Vordergrund stehen. Dem Thema der Viralität von Social Media Inhalten wird in der wissenschaftlichen Forschung zunehmende Aufmerksamkeit geschenkt (Borges-Tiago et al., 2019). Trotzdem ist über die Viralität von Hate Speech bzw. deren Ausbreitung bis dato wenig bekannt. In Anbetracht der schädlichen Wirkung von Hasskommentaren ist dies eine

bedauerliche Situation. Neues Wissen darüber wann Hate Speech massentauglich wird ist daher dringend nötig. Die vierte Forschungsfrage widmet sich einer bislang noch kaum empirisch untersuchten höchst relevanten Situation: Reaktionen von Nutzer:innen von Onlinekommunikationsplattformen wenn ihre negativen Kommentare gelöscht werden. Die bisherige Forschung zur Zensur von Hate Speech konzentrierte sich vor allem auf die Beobachter:innen von Hate Speech und die Frage, welche von ihnen die Zensur von Hate Speech befürworten (Zhang, 2019). Die wissenschaftliche Fachliteratur diskutiert angeregt inwieweit politische Akteur:innen die Löschung von Hasskommentaren befürworten sollen (Kalsnes & Ihlebæk, 2021). Die Literatur schweigt jedoch größtenteils zur Frage, wie sich das Löschen von Hate Speech auf die Verfasser:innen selbst auswirkt.

Hate Speech hat besonders drastische Folgen, wenn Betroffene massenhaft angegriffen werden, etwa in einem „Shitstorm“. Die dritte Forschungsfrage untersucht aus Sender:innenperspektive, wie und wann solche Entrüstungswellen entstehen, wobei unterschiedliche Aktivitätslevel (z. B. Verfassen oder „Liken“ von Hasskommentaren) im Fokus stehen. Obwohl die Viralität von Social-Media-Inhalten zunehmend erforscht wird (Borges-Tiago et al., 2019), ist über die Verbreitung von Hate Speech wenig bekannt.

Die vierte Forschungsfrage widmet sich der kaum untersuchten Reaktion von Nutzerinnen auf das Löschen ihrer Hasskommentare. Bisherige Forschung konzentrierte sich auf Beobachterinnen und deren Einstellung zur Zensur (Zhang, 2019) sowie politische Akteurinnen (Kalsnes & Ihlebæk, 2021). Unklar bleibt, wie sich Löschungen auf Verfasser:innen auswirken – ob dies zu einer Beruhigung führt oder Shitstorms verstärkt werden.

Forschungsfrage 5: Welche Verhaltensabsichten werden bei Beobachter:innen von Hasskommentaren ausgelöst?

Die fünfte Forschungsfrage untersucht, wie Hasskommentare auf unbeteiligte Leser:innen wirken. Studien zeigen, dass Hate Speech unterschiedlich wahrgenommen wird (Leets, 2002). Dabei spielen individuelle Faktoren (z. B. eigene Hate-Speech-Aktivität, Persönlichkeit) und interpersonelle Faktoren (z. B. Ähnlichkeit zwischen Verfasser:in und Beobachter:in) eine Rolle.

Die sechste Forschungsfrage analysiert die Wahrnehmung von Hasskommentaren in Online-Nachrichtenportalen. Community-Forschung zeigt, dass Online-Subkulturen besondere Eigenschaften besitzen und für Organisationen relevant sind (Kumar & Nayak, 2019). Untersucht wird, ob Nutzer:innen sich an Hate Speech beteiligen, gegensteuern oder passiv bleiben. Zudem wird erforscht, wie Hate Speech das Image und die Wahrnehmung österreichischer und

deutscher Online-Nachrichtenportale beeinflusst, um Medien ein besseres Verständnis der Auswirkungen negativer Inhalte zu ermöglichen.

Begründung der methodischen Vorgehensweise. Zum einen hat sich bisherige Forschung insbesondere auf die automatisierte Identifikation von Hate Speech mittels Algorithmen und künstlicher Intelligenz konzentriert (MacAvaney et al., 2019). Zum anderen lässt sich bei bestehender Primärdatenforschung hinsichtlich Sample und Forschungsdesign feststellen, dass der Großteil der bisherigen Studien zum Thema Hate Speech Studierende als Teilnehmer:innen (Cowan & Hodge, 1996; Leets, 2002; Soral et al., 2018, study 3) hatte, welchen entweder hypothetische Hate Speech Szenarien vorgelegt wurden (Cowan & Hodge, 1996; Leets, 2002) oder von welchen die Intensität, in der sie Hate Speech im Internet wahrgenommen haben (Soral et al., 2018, study 3), erhoben wurde. Im Gegensatz dazu beziehen sich die Forschungsfragen 1 bis 4 jedoch auf die Sender:innenperspektive von Hate Speech: Um ein besseres Verständnis der Hate Speech Betreiber:innen hinsichtlich ihrer Eigenschaften, Motive, Kontext-Faktoren und Reaktionen auf Zensur zu erlangen, wird ein nicht experimentelles Untersuchungsdesign gewählt. In Form einer Onlinebefragung soll ein repräsentatives Sample von Verfasser:innen von Hate Speech auf österreichischen und deutschen Online-Nachrichtenportalen herangezogen werden. Bisher wurden für ein Land repräsentative Samples zum Thema Hate Speech hauptsächlich mit Beobachter:innen von Hate Speech erhoben (Barnidge et al., 2019, Bevölkerung der USA; Soral et al., 2018, study 1, Bevölkerung von Polen). Zusätzlich werden in diesem Projekt auch noch die Beobachter:innen von Hate Speech auf österreichischen und deutschen Online-Nachrichtenportalen untersucht. Die Forschungsfrage 5 bezieht sich auf die Erforschung der Wahrnehmung und Verhaltensabsichten von Beobachter:innen von Hate Speech. Diese Forschungsfrage wird mittels eines Onlineexperiments beantwortet. Durch den gewählten quantitativen Forschungsansatz lassen sich aussagekräftige Ergebnisse erzielen, aus welchen sich konkrete Handlungsempfehlungen ableiten lassen.

Ausgewählte Empirische Studien und Ergebnisse

Studie 1 „Eigenschaften der Online-Hater:innen in Österreich“ (n = 554)

Die Studie 1 hat das Ziel einen Beitrag zur Beantwortung der Forschungsfrage 1 „Welche Eigenschaften haben Personen, die Hate Speech betreiben?“ zu leisten.

Sample und deskriptive Ergebnisse

Für Studie 1, wurde im Oktober 2022 eine Stichprobe von n = 554 Menschen, die in Österreich leben und zwischen 16 und 75 Jahren alt sind, erhoben. Diese Stichprobe entspricht der nach

EUROSTAT 2021 für Geschlecht, Alter und Region Bundesland¹. 50.9% der befragten Personen waren Frauen; 49.1% Männer (M = 45,60 Jahre; SD = 14,97). Studie 1 ist eine online Umfrage und es wurden etablierte wissenschaftliche Messinstrumente verwendet.² Die Charakteristika der Stichprobe lassen sich in Tabellen 1 ablesen.

Tabelle 1: Charakteristika der Stichprobe von Studie 1

| Bundesland | Anzahl | Gültige % |
|------------------|------------|------------|
| Burgenland | 18 | 3,20 |
| Kärnten | 34 | 6,10 |
| Niederösterreich | 107 | 19,30 |
| Oberösterreich | 92 | 16,60 |
| Salzburg | 32 | 5,80 |
| Steiermark | 78 | 14,10 |
| Tirol | 52 | 9,40 |
| Vorarlberg | 23 | 4,20 |
| Wien | 118 | 21,30 |
| <i>Gesamt</i> | <i>554</i> | <i>100</i> |

| Höchster Bildungsabschluss | Anzahl | Gültige % |
|---|------------|------------|
| kein Pflichtschul- oder Hauptschulabschluss | 1 | 0,20 |
| Pflichtschul- oder Hauptschulabschluss | 25 | 4,50 |
| Lehre, Berufsschule, Fachschule | 225 | 40,60 |
| Mittlere Reife | 23 | 4,20 |
| Höhere Schule mit Matura/Abitur | 132 | 23,80 |
| Hochschule (Bakkalaureat oder vergleichbar) | 50 | 9,00 |
| Hochschule (Master oder vergleichbar) | 72 | 13,00 |
| Hochschule (Doktorat oder vergleichbar) | 24 | 4,30 |
| Sonstiges | 2 | 0,40 |
| <i>Gesamt</i> | <i>554</i> | <i>100</i> |

| Politische Orientierung | |
|-------------------------|------------|
| M | 3,95 |
| SD | 1,27 |
| | Gültige % |
| stark links (= 1) | 3,60 |
| 2 | 8,70 |
| 3 | 18,60 |
| 4 | 40,40 |
| 5 | 19,90 |
| 6 | 4,70 |
| stark rechts (= 7) | 4,20 |
| <i>Gesamt</i> | <i>100</i> |

| Beruf | Anzahl | Gültige % |
|---------------------------|------------|------------|
| In Ausbildung | 27 | 4,90 |
| Angestellte/r, Beamter/in | 274 | 49,50 |
| Arbeiter/in | 50 | 9,00 |
| Selbstständig | 32 | 5,80 |
| Pensionist/in | 124 | 22,40 |
| Nicht berufstätig | 35 | 6,30 |
| Sonstiges | 12 | 2,20 |
| <i>Gesamt</i> | <i>554</i> | <i>100</i> |

Ergebnisse

Hier drei erste deskriptive Ergebnisse:

- 8% der Menschen in Österreich zwischen 16 und 75 Jahren wurde aufgrund ihres Verhaltens schonmal einer ihrer social media Accounts gesperrt
- 18% der Menschen in Österreich zwischen 16 und 75 Jahren haben in den letzten 12 Monaten öffentlich auf social media Hass-Mitteilungen gepostet

¹ Die Erhebung wurde vom Online-Marktforschungsdienstleister Bilendi <https://www.bilendi.de/> durchgeführt.

² Bei Fragen zur Befragung und zu den verwendeten Messinstrumenten wenden Sie sich bitte an clemens.hutzinger@uni-seeburg.at

- 37% der Menschen in Österreich zwischen 16 und 75 Jahren, die in den letzten 12 Monaten öffentlich auf social media Mitteilungen gepostet haben, haben in den letzten 12 Monaten öffentlich auf social media Hass-Mitteilungen gepostet

Erklärung für die Häufigkeit Hass-Mitteilungen zu posten

Gibt es einen Unterschied zwischen Männern und Frauen? → *Nein*. Geschlecht hat gar keinen Effekt auf die Häufigkeit Hass-Mitteilungen zu posten. Männer ($M = 3,05$; $SD = 1,43$) posten gleich häufig Hasskommentare wie Frauen ($M = 3,00$; $SD = 1,45$; $t = -0,19$; $Sig. = 0,85$).

Hat Alter einen Effekt? → *Nein*. Alter hat keinen signifikanten Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($r = 0,01$; $Sig. = 0,95$).

Hat die politische Orientierung einen Effekt? → *Nein*. Die eigene politische Orientierung hat keinen signifikanten Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($r = -0,07$; $Sig. = 0,46$).

Hat die Bildung einen Effekt? → *Nein*. Der eigene höchste Bildungsabschluss hat keinen signifikanten Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($F(6, 106) = 0,56$; $Sig. = 0,76$). Das bedeutet beispielsweise, dass Personen mit Pflichtschulabschluss genauso häufig Hasskommentare posten, wie Personen mit Lehrabschluss oder Personen mit Masterstudienabschluss.

Hat die eigene Persönlichkeit einen Effekt? → *Ja*. Narzissmus hat einen signifikanten positiven Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($r = 0,21$; $p < 0,05$). Das bedeutet, dass je mehr Personen bewundert und beachtet werden wollen, desto öfter posten sie Hass-Mitteilungen auf social media. Machiavellismus hat keinen signifikanten Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($r = 0,18$; $Sig. = 0,06$). Das bedeutet, dass sich das Ausmaß, in dem Personen ihren Willen durchsetzen und andere auszunutzen, nicht darauf auswirkt, wie oft sie Hass-Mitteilungen auf social media posten. Psychopathie hat einen signifikanten positiven Effekt auf die Häufigkeit Hass-Mitteilungen zu posten ($r = 0,35$; $p < 0,001$). Das bedeutet, dass je mehr Personen gefühllos und zynisch sind, desto öfter posten sie Hass-Mitteilungen auf social media.

Erklärung für die Hasserrülltheit der Posts

Gibt es einen Unterschied zwischen Männern und Frauen? → *Nein*. Geschlecht hat gar keinen Effekt auf die Hasserrülltheit der Posts. Die Posts der Männer ($M = 2,19$; $SD = 1,32$) unterscheiden sich bezüglich Hasserrülltheit nicht von denen der Frauen ($M = 1,97$; $SD = 1,24$; $t = -0,85$; $Sig. = 0,40$).

Hat Alter einen Effekt? → *Ja*. Alter hat einen signifikanten negativen Effekt auf die Hasserrülltheit der Posts ($r = -0,35$; $p < 0,001$). Das bedeutet, dass mit zunehmendem Alter die Hasserrülltheit der erstellten Nachrichten auf sozial media geringer wird.

Hat die politische Orientierung einen Effekt? → *Nein*. Die eigene politische Orientierung (von 1 = „stark links“ bis 7 = „stark rechts“) hat keinen signifikanten Effekt auf die Hasserrülltheit der Posts ($r = -0,07$; $Sig. = 0,53$).

Hat die Bildung einen Effekt? → *Nein*. Der eigene höchste Bildungsabschluss hat keinen signifikanten Effekt auf die signifikanten Effekt auf die Hasserrülltheit der Posts ($F(6, 91) = 1,08$; $Sig. = 0,38$). Das bedeutet beispielsweise, dass Personen mit Doktoratsabschluss genauso hasserrüllte Kommentare schreiben, wie Personen mit mittlerer Reife oder Personen mit Maturaabschluss.

Hat die eigene Persönlichkeit einen Effekt? → *Ja*. Narzissmus hat einen signifikanten positiven Effekt auf die die Hasserrülltheit der Posts ($r = 0,32$; $p < 0,01$). Das bedeutet, dass je mehr Personen bewundert und beachtet werden wollen, desto mehr Hass steckt in Ihren Mitteilungen auf social media. Machiavellismus hat einen signifikanten positiven Effekt auf die die Hasserrülltheit der Posts ($r = 0,46$; $p < 0,001$). Das bedeutet, dass je mehr Personen ihren Willen durchsetzen und andere auszunutzen, desto mehr Hass steckt in Ihren Mitteilungen auf social media. Psychopathie hat einen signifikanten positiven Effekt auf die die Hasserrülltheit der Posts ($r = 0,55$; $p < 0,001$). Das bedeutet, dass je mehr Personen gefühllos und zynisch sind, desto mehr Hass steckt in Ihren Mitteilungen auf social media.

Studie 2 „Motive und Reaktionen der Online-Hater:innen in Österreich und Deutschland“ (n = 1420)

Die Studie 2 hat das Ziel einen Beitrag zur Beantwortung der Forschungsfrage 1 „Welche Eigenschaften haben Personen, die Hate Speech betreiben?“, der Forschungsfrage 2 „Was sind die Motive für die Beteiligung an Hate Speech?“, der Forschungsfrage 3 „Welche Kontext-

Faktoren stimulieren Personen, vermehrt Hasskommentare zu posten?“ und der Forschungsfrage 4 „Wie reagieren Personen, die Hasskommentare posten, wenn ihre Kommentare gelöscht werden?“ zu leisten.

Sample und deskriptive Ergebnisse

Für Studie 2, wurde im Februar 2023 eine Stichprobe von $n = 1830$ Menschen erhoben³, die in Österreich ($n_1 = 410$) oder Deutschland ($n_2 = 1420$) leben, zwischen 16 und 75 Jahren alt sind und in den letzten 12 Monaten zumindest einmal öffentlich auf social media eine Mitteilung gepostet haben, die eine negative Einstellung über andere Personen ausgedrückt hat. 40.1% der befragten Personen waren Frauen, 59.9% Männer ($M = 50,53$ Jahre; $SD = 14,07$). Studie 2 ist ein online Umfrage und es wurden etablierte wissenschaftliche Messinstrumente verwendet.⁴ Die Charakteristika der Stichprobe lassen sich in Tabellen 2 ablesen.

Tabelle 2: Charakteristika der Stichprobe von Studie 2

| Politische Orientierung | |
|-------------------------|------------|
| M | 3,96 |
| SD | 1,28 |
| | Gültige % |
| stark links (= 1) | 3,10 |
| 2 | 10,10 |
| 3 | 19,00 |
| 4 | 36,80 |
| 5 | 20,70 |
| 6 | 7,40 |
| stark rechts (= 7) | 3,00 |
| <i>Gesamt</i> | <i>100</i> |

| Beruf | Anzahl | Gültige % |
|----------------------------|-------------|------------|
| Lehrling / Auszubildende:r | 9 | 0,80 |
| Arbeiter:in | 155 | 13,50 |
| Angestellte:r | 770 | 67,30 |
| Beamter/in | 48 | 4,20 |
| Selbstständige:r | 138 | 12,10 |
| Sonstiges | 24 | 2,10 |
| <i>Gesamt</i> | <i>1144</i> | <i>100</i> |

| Höchster Bildungsabschluss | Anzahl | Gültige % |
|---|-------------|------------|
| kein Pflichtschul- oder Hauptschulabschluss | 4 | 0,20 |
| Pflichtschul- oder Hauptschulabschluss | 54 | 3,00 |
| Lehre, Berufsschule, Fachschule | 558 | 30,50 |
| Mittlere Reife | 346 | 18,90 |
| Höhere Schule mit Matura/Abitur | 337 | 18,40 |
| Hochschule (Bakkalaureat oder vergleichbar) | 165 | 9,00 |
| Hochschule (Master oder vergleichbar) | 313 | 17,10 |
| Hochschule (Doktorat oder vergleichbar) | 46 | 2,50 |
| Sonstiges | 7 | 0,40 |
| <i>Gesamt</i> | <i>1830</i> | <i>100</i> |

³ Die Erhebung wurde vom Online-Marktforschungsdienstleister Bilendi <https://www.bilendi.de/> durchgeführt.

⁴ Bei Fragen zur Befragung und zu den verwendeten Messinstrumenten, wenden Sie sich bitte an clemens.hutzinger@uni-seeburg.at.

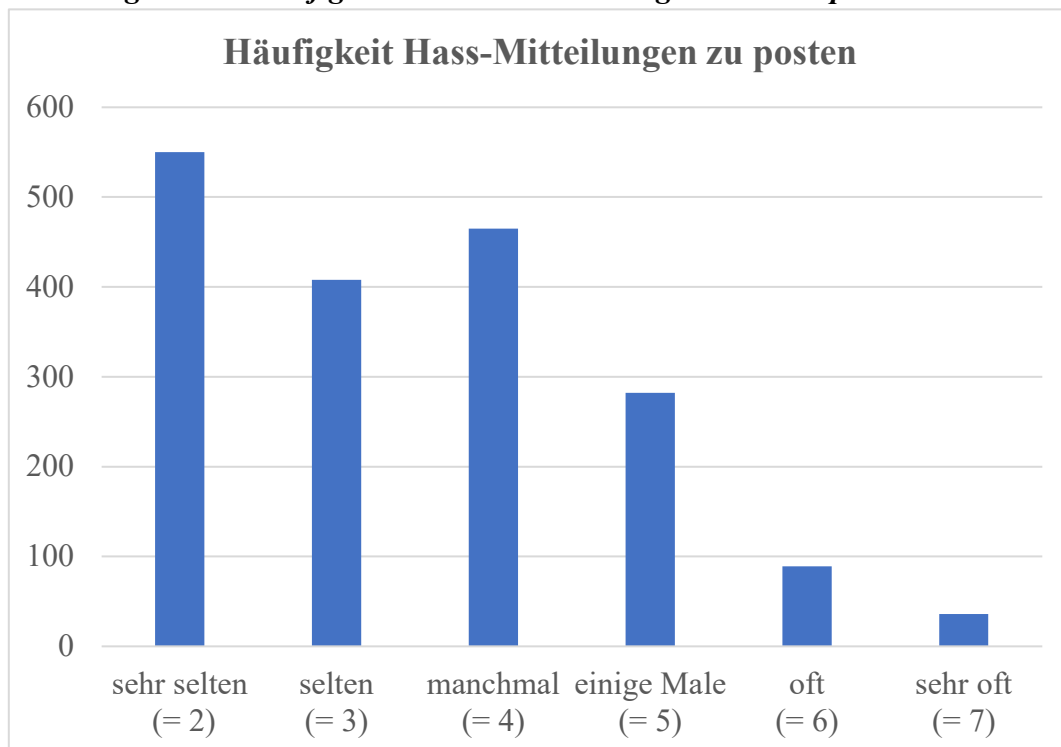
Tabelle 3 zeigt die Nutzungshäufigkeit verschiedener Social Media Plattformen.

Tabelle 3: Social Media Nutzung der Stichprobe von Studie 2

| | Social Media Nutzung | | | | | | | |
|-------------------|----------------------|-----------|-----------|-----------|-----------|--|--|--|
| | YouTube | Facebook | Instagram | Twitter | TikTok | Kommentarbereiche von Online-Zeitungen | Gaming Streaming Plattformen (z.B. Twitch) | Online games (z.B. Minecraft, World of Warcraft, Counter-Strike) |
| M | 4,19 | 5,03 | 3,49 | 2,43 | 2,21 | 2,78 | 1,81 | 1,97 |
| SD | 1,93 | 1,94 | 2,28 | 1,99 | 1,89 | 1,80 | 1,53 | 1,68 |
| | Gültige % | Gültige % | Gültige % | Gültige % | Gültige % | Gültige % | Gültige % | Gültige % |
| nie (= 1) | 13,40 | 10,10 | 35,30 | 57,50 | 63,90 | 36,70 | 72,10 | 68,30 |
| sehr selten (= 2) | 9,70 | 5,00 | 8,50 | 8,30 | 7,30 | 15,80 | 7,30 | 6,80 |
| selten (= 3) | 11,90 | 5,90 | 6,70 | 6,30 | 4,90 | 12,60 | 4,80 | 5,50 |
| manchmal (= 4) | 18,50 | 10,70 | 11,40 | 8,30 | 7,00 | 16,10 | 6,80 | 7,70 |
| einige Male (= 5) | 15,00 | 14,80 | 10,20 | 6,90 | 6,40 | 9,00 | 4,10 | 5,00 |
| oft (= 6) | 19,00 | 26,40 | 15,20 | 7,10 | 6,00 | 6,10 | 3,00 | 3,80 |
| sehr oft (= 7) | 12,60 | 27,00 | 12,70 | 5,60 | 4,50 | 3,70 | 2,00 | 3,00 |
| Gesamt | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |

Abbildung 2 verdeutlicht wie oft Hass-Mitteilungen gepostet wurden.

Abbildung 2: Post-Häufigkeit von Hass-Mitteilungen der Stichprobe von Studie 2

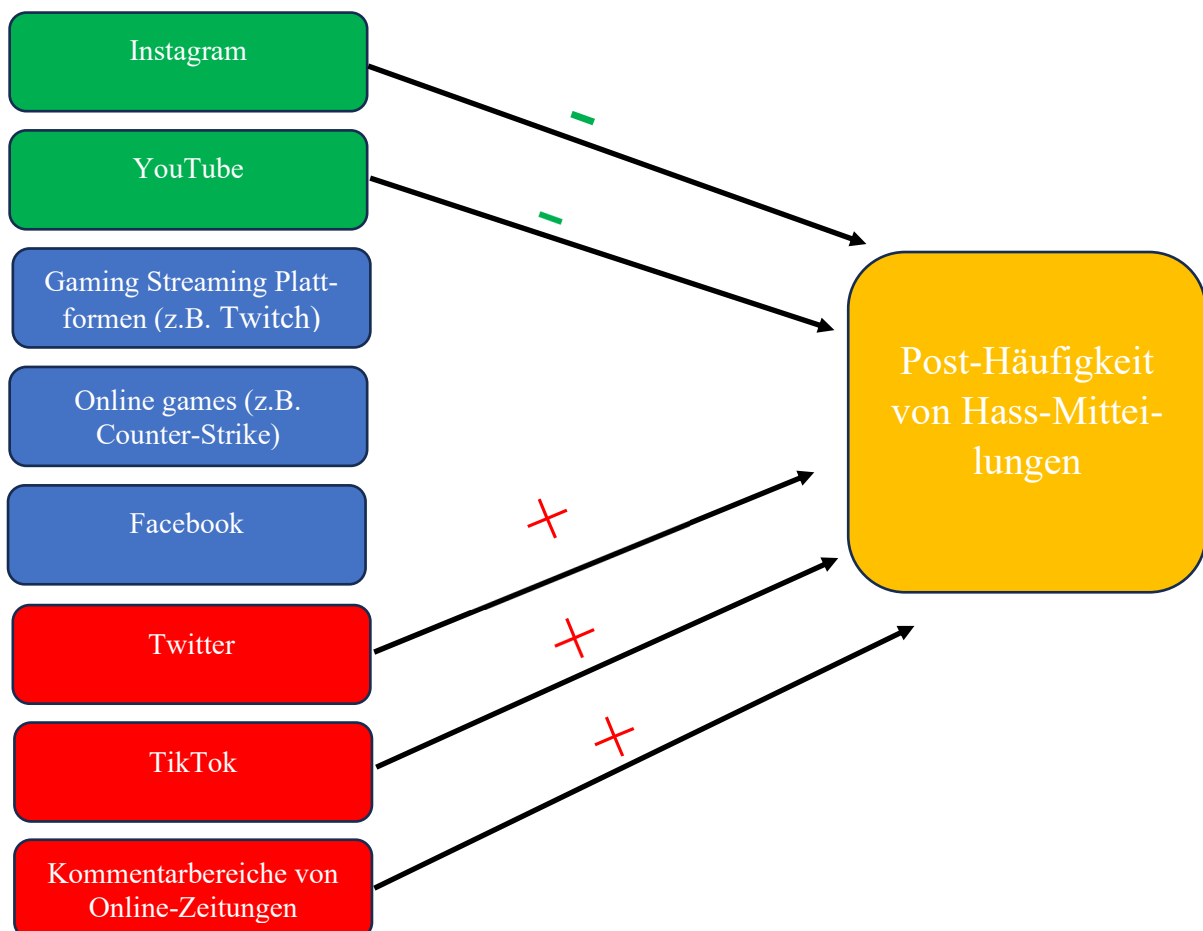


Ergebnisse

Auswirkung verschiedener social media Plattformen auf das Posten von Hass-Mitteilungen

Hat die Nutzungshäufigkeit unterschiedlicher social media Plattformen einen Effekt darauf wie oft online-hater:innen Hass-Mitteilungen posten? → Ja. Je häufiger die online-hater:innen Twitter ($\beta = 0,15$; $p < 0,001$), TikTok ($\beta = 0,12$; $p < 0,001$) und Kommentar-bereiche von Online-Zeitungen ($\beta = 0,09$; $p < 0,001$) nutzen, desto häufiger posten sie Hass-Mitteilungen. Je häufiger die online-hater:innen Instagram ($\beta = -0,09$; $p < 0,001$) und YouTube ($\beta = -0,06$; $p < 0,05$) nutzen, desto seltener posten sie Hass-Mitteilungen. Die Nutzungshäufigkeit von Gaming Streaming Plattformen (z.B. Twitch), Online games (z.B. Counter-Strike) und Facebook hat keine Auswirkung auf die Häufigkeit des Postens von Hass-Mitteilungen. Abbildung 3 fasst die Ergebnisse gruppenweise zusammen.

Abbildung 3: Auswirkung verschiedener social media Plattformen auf das Posten von Hass-Mitteilungen (Studie 2)



Motive für das Posten von Hass-Mitteilungen

Was sind die stärksten Motive für das Posten von Hass-Mitteilungen? → Das stärkste Motiv ist Ideologie (M = 5,17; SD = 1,78), also weil es der Überzeugung der Online-Hater:innen entspricht. Das zweitstärkste Motiv ist Rache (M = 3,69; SD = 1,90; $t = 31,62$; $p < 0,001$), also weil es die Person(en), die im Posting erwähnt wurden, die Online-Hater:innen wütend gemacht hat/haben. Das drittstärkste Motiv ist Erheiterung (M = 2,26; SD = 1,49, $t = 28,57$; $p < 0,001$), also weil die Online-Hater:innen es lustig fanden. Das viertstärkste Motiv ist Gruppenkonformität (M = 2,09; SD = 1,43 $t = 5,36$; $p < 0,001$), also weil dieses Verhalten in der jeweiligen community der Online-Hater:innen erwartet wird. Das fünftstärkste Motiv ist Statusverbesserung (M = 1,89; SD = 1,38; $t = 8,49$; $p < 0,001$), also um respektiert zu werden und zu den angesehenen Mitgliedern der jeweiligen community zu gehören. Die stärksten Motive werden in Abbildung 4 dargestellt:

Abbildung 4: Die fünf stärksten Motive für das Posten von Hass-Mitteilungen in Studie 2 (n = 1830) (Quelle: eigene Darstellung mittels <https://chatgpt.com/>)



Kontext-Faktoren für einen Anstieg an Hasskommentaren

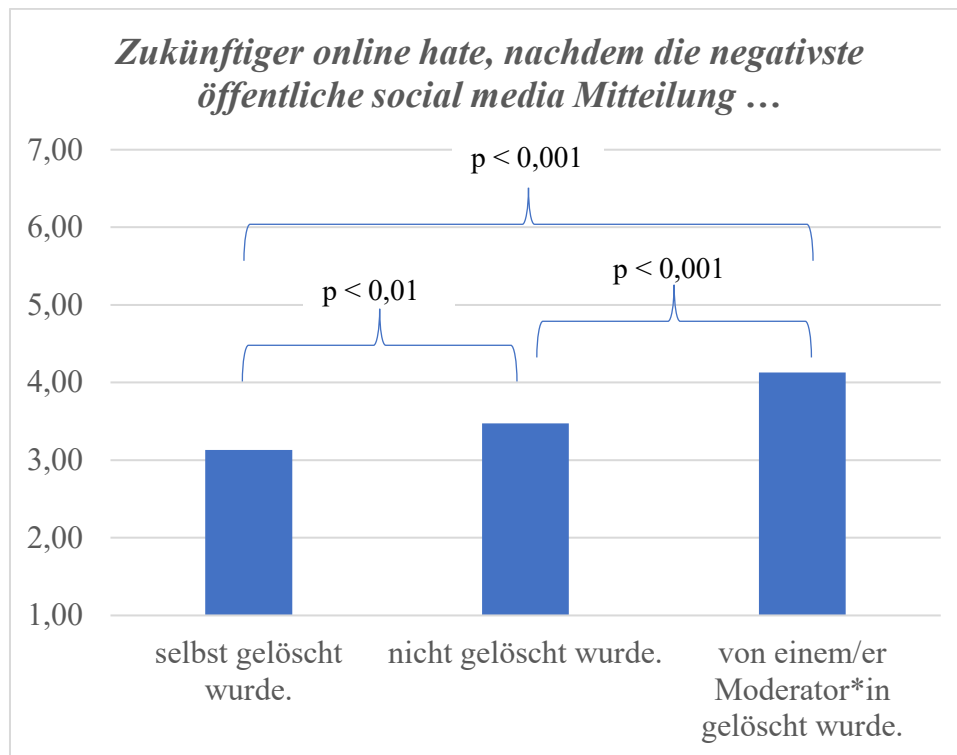
Die Online-Hater:innen wurden zusätzlich gebeten, an Ihre negativste, in den letzten 12 Monaten öffentlich auf social media gepostete Mitteilung, die eine negative Einstellung über andere Personen ausgedrückt hat, zu denken.

Welche Kontext-Faktoren stimulieren Personen, vermehrt Hasskommentare zu posten? → Likes, Shares, Positive Comments. Je mehr Likes ("gefällt mir" (Daumen nach oben, Herz etc.) durch andere Nutzer:innen) Online-Hater:innen auf ihre negativste social media Mitteilung bekommen, desto höher ist ihre Tendenz in Zukunft wieder vermehrt Hasskommentare zu posten ($r = 0.20$; $p < 0.001$). Je mehr Shares (Teilen der Mitteilung durch andere Nutzer:innen) Online-Hater:innen auf ihre negativste social media Mitteilung bekommen, desto höher ist ihre Tendenz in Zukunft wieder vermehrt Hasskommentare zu posten ($r = 0.22$; $p < 0.001$). Je mehr Supportive Comments (unterstützende Kommentare mit ähnlicher Meinung) Online-Hater:innen auf ihre negativste social media Mitteilung bekommen, desto höher ist ihre Tendenz in Zukunft wieder vermehrt Hasskommentare zu posten ($r = 0.19$; $p < 0.001$).

Reaktion von online-hater:innen, wenn ihre Hasskommentare gelöscht werden

Wie reagieren Personen, die Hasskommentare posten, wenn ihre Kommentare gelöscht werden? → Er kommt darauf an. Das Löschverhalten der negativsten öffentlichen social media Mitteilung wirkt sich unterschiedlich auf zukünftigen online hate aus ($F(2, 1827) = 16,92$; $p < 0.001$). Wird die negativste öffentliche social media Mitteilung selbst gelöscht ($M = 3,13$; $SD = 1,68$) verbreiten die online-hater:innen zukünftig weniger Hass, als wenn die Mitteilung nicht gelöscht wird ($M = 3,47$; $SD = 1,69$; $p < 0,01$). Wird jedoch die negativste öffentliche social media Mitteilung von einem/er Moderator:in gelöscht ($M = 4,13$; $SD = 1,86$) verbreiten die online-hater:innen zukünftig mehr Hass, als wenn die Mitteilung nicht gelöscht wird ($M = 3,47$; $SD = 1,69$; $p < 0,001$). Abbildung 5 fasst diese Ergebnisse zusammen.

Abbildung 5: Auswirkung des Löschverhaltens auf zukünftigen online hate (Studie 2)



Studie 3 „Auswirkungen auf Online-Hate-Beobachter:innen auf österreichischen Online-Nachrichtenportalen“ (n = 968)

Die Studie 3 hat das Ziel einen Beitrag zur Beantwortung der Forschungsfrage 5 „Welche Verhaltensabsichten werden bei Beobachter:innen von Hasskommentaren ausgelöst?“ zu leisten.

Sample und deskriptive Ergebnisse

Für Studie 3 wurde im Juni 2023 eine Stichprobe von n = 968 Menschen erhoben⁵, die in Österreich leben, zwischen 16 und 75 Jahren alt sind und Social Media nutzen. 54.4% der befragten Personen waren Frauen, 45.6% Männer (M = 46,98 Jahre; SD = 14,39). Studie 3 ist ein online Experiment und es wurden etablierte wissenschaftliche Messinstrumente verwendet.⁶ Alle Teilnehmer:innen lasen den Text in Abbildung 6, welcher sehr nahe an echten online hate postings dran ist.

⁵ Die Erhebung wurde vom Online-Marktforschungsdienstleister Bilendi <https://www.bilendi.de/> durchgeführt.

⁶ Bei Fragen zur Befragung und zu den verwendeten Messinstrumenten, wenden Sie sich bitte an clemens.hutzinger@uni-seeburg.at.

Abbildung 6: Stimulustext des online Experiments in Studie 3

Auf einer österreichischen Online-Zeitung lesen Sie einen Artikel, der von ukrainischen Flüchtlingen handelt, die vor dem Ukraine-Krieg nach Österreich geflüchtet sind. Unter dem Artikel lesen Sie folgenden Kommentar eines social media Nutzers:

Hans W. 1975 vor 2 Stunden

Jetzt geht das schon wieder los. Wir haben scho genug Leute aus der Ukraine bei uns in Österreich aufgenommen. Es soll ja nur ein kleiner Teil (so 10%) des Landes von den Russen besetzt sein. Und die wenigsten Leute aus der Ukraine wollen arbeiten. Die werden sich auch nicht an unsere Kultur anpassen. Glaubts mir, die werden wir nicht mehr los. Geh, schleichts eich zurück oder besser bleibts doch einfach dort in eurem gschissenen Land, ich heiße euch NICHT herzlich willkommen!!!

Die Teilnehmerinnen haben diesen Kommentar als äußerst negativ ($M = 1,84$; $SD = 1,35$; gemessen auf einer 7-stufigen Skala von 1 = „sehr negativ“ bis 7 = „sehr positiv“) und äußerst hasserrfüllt ($M = 5,90$; $SD = 1,43$; gemessen auf einer 7-stufigen Skala von 1 = „gar nicht hasserrfüllt“ bis 7 = „sehr hasserrfüllt“) empfunden.

Auf die Aussage „Ich möchte, dass der online-Kommentar von Hans W. von den online-Moderator:innen der online-Zeitung gelöscht wird.“ Antworten 57,3% der Teilnehmer:innen mit „ja“ und 42,7% mit „nein“.

Danach erfolgt die experimentelle Manipulation des Löschens des online-Kommentars. Die Teilnehmer:innen werden zufällig zu einer der beiden Experimentalgruppen zugewiesen, entweder zur „Kommentar nicht gelöscht“-Gruppe oder zur „Kommentar gelöscht“-Gruppe. Die Teilnehmer:innen haben das beschriebene Szenario mit dem online-Kommentar auf der Online-Zeitung als überdurchschnittlich realistisch eingeschätzt ($M = 5,36$; $SD = 1,71$; gemessen auf einer 7-stufigen Skala von 1 = „stimme gar nicht zu“ bis 7 = „stimme völlig zu“).

Ergebnisse

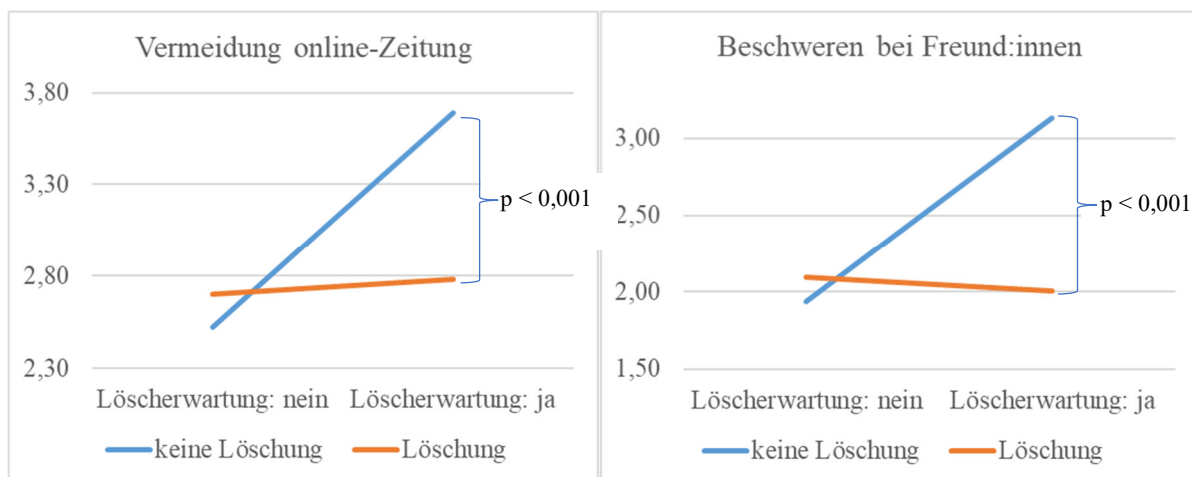
Auswirkung des Löschens des online-Kommentars auf die Verhaltensabsichten der Beobachter:innen

Wie wirkt sich das Löschen des online-Hasskommentars darauf aus, ob die Beobachter:innen von Hasskommentaren hinsichtlich Vermeidung der online-Zeitung und Beschwerden bei Freund:innen aus? → Es kommt darauf an. Die Ergebnisse zeigen einen signifikanten Interaktionseffekt zwischen der Löscherwartung und der tatsächlichen Löschung des Hasskommentars ($F(1, 964) = 19,34$; $p < 0,001$) auf die Vermeidung der online-Zeitung. Haben die Beobachter:innen keine Löscherwartung, hat die tatsächliche Löschung des Hasskommentars keine Auswirkung auf die Vermeidung der online-Zeitung ($Sig. = 0,31$). Haben die Beobachter:innen im

Gegensatz dazu die Erwartung, dass der Hasskommentar gelöscht wird, ist die Vermeidung der online-Zeitung höher, wenn der Hasskommentar tatsächlich nicht gelöscht wird ($M = 3,69$; $SD = 2,09$), als wenn er gelöscht wird ($M = 2,78$; $SD = 1,91$; $t = 5,36$; $p < 0,001$).

Weiters zeigen die Ergebnisse einen signifikanten Interaktionseffekt zwischen der Löscherwartung und der tatsächlichen Löschung des Hasskommentars ($F(1, 964) = 34,31$; $p < 0,001$) auf das Beschweren bei Freund:innen. Haben die Beobachter:innen keine Löscherwartung, hat die tatsächliche Löschung des Hasskommentars keine Auswirkung auf das Beschweren bei Freund:innen (Sig. = 0,28). Haben die Beobachter:innen im Gegensatz dazu die Erwartung, dass der Hasskommentar gelöscht wird, ist das Beschweren bei Freund:innen höher, wenn der Hasskommentar tatsächlich nicht gelöscht wird ($M = 3,13$; $SD = 2,05$), als wenn er gelöscht wird ($M = 2,01$; $SD = 1,54$; $t = 7,34$; $p < 0,001$). Abbildung 7 fasst diese Ergebnisse zusammen.

Abbildung 7: Das Zusammenspiel der Löscherwartung und der tatsächlichen Löschung auf die Verhaltensabsichten der online-hate Beobachter:innen (Studie 3)



Diskussion

Österreichische und deutsche Online-Nachrichtenportale sind zentrale Informations- und Kommunikationsplattformen, dienen jedoch auch der Verbreitung von Hasskommentaren. Online Hate Speech wird aus verschiedenen Perspektiven erforscht (Ouvrein, 2024; Park et al., 2023; Weber et al., 2024). Dieses Projekt untersucht die Häufigkeit öffentlicher Social-Media-Posts mit negativen Einstellungen gegenüber anderen (Malecki et al., 2021; Wachs & Wright, 2018) sowie deren beleidigende, bedrohliche und aggressive Inhalte (Kaakinen et al., 2020; Soral et al., 2018; Pace et al., 2021; Wachs & Wright, 2018). Die Ergebnisse von Studie 1, die national-repräsentativ für Geschlecht, Alter und Region Bundesland ist, zeigt drastisch auf, wie weit

online-hate in Österreich verbreitet ist: 37% der Menschen in Österreich zwischen 16 und 75 Jahren, die in den letzten 12 Monaten öffentlich auf social media Mitteilungen gepostet haben, haben in diesem Zeitraum öffentlich Hass-Mitteilungen gepostet.

Dieses Forschungsprojekt hatte zum Ziel, ein besseres Verständnis von Hate Speech auf Onlineportalen österreichischer und deutscher Nachrichtenmedien und daraus abgeleitet Lösungswege zu erlangen. So gab es bis dato nur wenige wissenschaftliche Studien, welche sich mit dem sozialen Phänomen online Hasskommentare auseinandersetzen. Dies ist verwunderlich, bedenkt man die große gesellschaftliche Relevanz und Aktualität des Themas.

Beiträge für Medienunternehmen. Nachrichtenmedien sollen die Bevölkerung mit glaubwürdigen Informationen versorgen, doch Hasskommentare stören diesen Auftrag. Sie greifen Personen oder Gruppen an und stellen deren Glaubwürdigkeit infrage. Dies kann die wahrgenommene Berichterstattungsqualität beeinträchtigen. Medienunternehmen müssen gegen Hate Speech vorgehen, um Betroffene zu schützen. Dieses Forschungsprojekt liefert wichtige Erkenntnisse darüber, wann und wie Hate Speech auf österreichischen und deutschen Online-Nachrichtenportalen entsteht und wer sich daran beteiligt (siehe Forschungsfragen 1 und 2). Die Ergebnisse von Studien 1 zeigen, dass Geschlecht, politische Orientierung und Bildung sich weder auf die Häufigkeit Hass-Mitteilungen zu posten noch auf die Hasserfülltheit der Postings auswirken. Mit zunehmendem Lebensalter der Verfasser:innen, werden die verfassten Postings immer weniger hasserfüllt. Vor allem die dunkle Seite der Persönlichkeit spielt eine große Rolle. Je mehr die Verfasser:innen beispielsweise gefühllos und zynisch sind (Dimension: Psychopathie) und bewundert und beachtet werden wollen (Dimension: Narzissmus), desto öfter posten sie Hass-Mitteilungen und desto mehr Hass verpacken sie in diesen.

Hinsichtlich der Beweggründe für hate speech, liefert Studie 2 die fünf Top Motive für das Posten von Hass-Mitteilungen in absteigender Stärke: 1) Ideologie (also der Überzeugung der Verfasser:innen entsprechend); 2) Rache (weil die Person(en), die im Posting erwähnt wurden, die Verfasser:innen wütend gemacht haben); 3) Erheiterung (weil die Verfasser:innen es lustig fanden); 4) Gruppenkonformität (weil dieses Verhalten in der jeweiligen community der Verfasser:innen erwartet wird) und 5) Statusverbesserung (also um respektiert zu werden und zu den angesehenen Mitgliedern der jeweiligen community zu gehören). Nun kann beispielsweise das Wissen über diese Hauptbeweggründe und Ziele von Hassposter:innen dazu verwendet

werden, diese zu entschärfen bzw. unreachbar zu machen. Nur durch ausreichende Kenntnis der Sender:innenperspektive ist es möglich, effektive Maßnahmen gegen Hate Speech zu ergreifen.

Weiters erzeugt dieses Forschungsprojekt neue Erkenntnisse zu den Faktoren, die ursächlich dafür sind, dass sich Hasskommentare im Internet viral verbreiten (siehe Forschungsfrage 3). So zeigen die Ergebnisse von Studie 2 ganz klar dass zukünftiges Verbreiten von Hass-Nachrichten deutlich ansteigt, je mehr die Online-Hater:innen auf ihre negativste social media Mitteilung Likes ("gefällt mir" (Daumen nach oben, Herz etc.), supportive comments (unterstützende Kommentare mit ähnlicher Meinung) und shares (Teilen der Mitteilung durch andere Nutzer:innen) bekommen. Medienunternehmen wollen konstruktive Diskussionen ermöglichen, doch destruktive Äußerungen gefährden dies. Erkenntnisse zur Viralität von Hasskommentaren helfen, Gegenmaßnahmen zu ergreifen. Erste Schritte wie das Verbot von Dislikes wurden bereits umgesetzt, doch unsere Ergebnisse zeigen auch negative Folgen positiven Online-Verhaltens.

Dieses Forschungsprojekt untersucht die Auswirkungen von Hate Speech auf verschiedene Stakeholder (siehe Forschungsfrage 5), darunter Medienunternehmen und die Online-Community. Erkenntnisse zur Wahrnehmung und Reaktion der Beobachter:innen helfen, negative Rückkopplungseffekte frühzeitig zu erkennen und entgegenzuwirken. Die Ergebnisse von Studie 3 liefern einen spannenden Einblick in die derzeit unterschiedliche Wahrnehmung des Themas in Österreich aus sich der online-hate observer:innen. So zeigt sich, dass der im online Experiment verwendeter post auf einer österreichischen Online-Zeitung, in dem ukrainische Flüchtlinge negativ beschrieben werden, zwar als äußerst hasserfüllt wahrgenommen wurden, jedoch nur etwas mehr als jede(r) Zweite wollte, dass dieser gelöscht wird. Hinsichtlich der Verhaltensreaktionen spielen die eigenen Erwartungen der Beobachter:innen eine entscheidende Rolle. Wollen die Beobachter:innen nicht, dass der Kommentar gelöscht wird, hat die tatsächliche Löschung des Hasskommentars keine Auswirkung auf die Vermeidung der online-Zeitung oder das Beschweren bei Freund:innen über die online-Zeitung. Wollen die Beobachter:innen jedoch dass der Kommentar gelöscht wird, ist die Vermeidung der online-Zeitung und das Beschweren bei Freund:innen über die online-Zeitung signifikant höher, wenn der Hasskommentar tatsächlich nicht gelöscht wird, als wenn er gelöscht wird.

Beiträge für staatliche und nichtstaatliche Organisationen. Das Löschen von Hate Speech wird in der Forschung kontrovers diskutiert, insbesondere die Rolle politischer Akteur:innen (Kalsnes & Ihlebæk, 2021). Unklar ist, ob Zensur Hass reduziert oder ihn auf andere Plattformen verlagert. Dieses Forschungsprojekt liefert erste Antworten auf diese Frage (siehe Forschungsfrage 4). Die Ergebnisse von Studie 2 zeigen verbreiten online-hater:innen zukünftig weniger Hass wenn Sie von selbst umdenken und den Kommentar löschen, als wenn die Mitteilung nicht gelöscht wird. Im Gegenzug führt das Löschen von Hasskommentaren durch Moderator:innen der online-Zeitung dazu, dass die online-hater:innen zukünftig noch mehr Hass verbreiten. Durch diese ersten Ergebnisse können entsprechende Kommunikationsmaßnahmen getroffen werden. Das Projekt sensibilisiert für Hate Speech und bietet konkrete Handlungsempfehlungen. Es liefert Einblicke in Mechanismen von Hasskommentaren und schafft Grundlagen für gezielte Bekämpfungsstrategien.

Berichterstattung und Dissemination der Ergebnisse und Dialog mit der Gesellschaft

Über die Ergebnisse des Projekts wurde in verschiedenen Medien und bei ausgewählten Veranstaltungen berichtet:

- Diverse LinkedIn-Postings (2023 bis 2024)
- Interview für Artikel in Salzburger Nachrichten „Ein Drittel der österreichischen Poster verteilt Hass“ (März 2024)
- Organisation der Gesprächspartner:innen und Teilnahme an Podiumsdiskussion an den 5. Seeburger Dialogen „The Dark Side of Social Media: Auswirkungen auf Unternehmen und Gesellschaft“ (März 2024)
- Workshop am Gymnasium Seekirchen mit Maturant:innen „Online hate speech: Hass im Netz wissenschaftlich verstehen und dadurch die Gesellschaft und Unternehmen verbessern.“ (März 2024)
- Beitrag zum Artikel in Salzburger Nachrichten „Das wird man wohl noch sagen dürfen!?! Nein, darf man nicht.“ (März 2024)
- Radio-Interview bei Ö1 für die Sendereihe „Moment“ zur Sendung „Cybermobbing, Shitstorm, Hass im Netz: Die schlechten Umgangsformen in der Online-Welt - keine Sache der Netiquette“ (April 2024)
- Vorträge und Diskussionen in der Lange Nacht der Forschung 2024 „Warum entsteht Hass im Internet und was löst dieser bei Beobachter:innen aus?“ (Mai 2024)

- Vortrag beim Bildungsnachmittag HAK/HAS Bad Ischl „Online hate: Eigenschaften und Motive der online haters.“ (Juni 2024)
- Vortrag bei Eröffnung des Seestadt OPEN HUB „Wer verbreitet aus welchen Gründen Hass im Internet und was macht das mit Beobachter:innen?“ (Juni 2024)



Salzburger Nachrichten



Quellen

- Barnidge, M., Kim, B., Sherrill, L. A., Luknar, Ž., & Zhang, J. (2019). Perceived exposure to and avoidance of hate speech in various communication settings. *Telematics and Informatics, 44*, 101263.
- Borges-Tiago, M. T., Tiago, F., & Cosme, C. (2019). Exploring users' motivations to participate in viral communication on social media. *Journal of Business Research, 101*, 574–582.
- Calvert, C. (1997). Hate speech and its harms: A communication theory perspective. *Journal of Communication, 47*(1), 4–19.
- Cheng, Y.-H., & Ho, H.-Y. (2015). Social influence's impact on reader perceptions of online reviews. *Journal of Business Research, 68*(4), 883–887.
- Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior, 40*, 108–118.
- Cheung, C. M. K., & Thadani, D. R. (2012). The impact of electronic word-of-mouth communication: A literature analysis and integrative model. *Decision Support Systems, 54*(1), 461–470.
- Cohen-Almagor, R. (2015). *Confronting the internet's dark side: Moral and social responsibility on the free highway*. New York: Cambridge University Press.
- Cowan, G., & Hodge, C. (1996). Judgments of hate speech: The effects of target group, publicness, and behavioral responses of the target. *Journal of Applied Social Psychology, 26*(4), 355–374.
- Dellagiacoma, L., Geschke, D., & Rothmund, T. (2024). Ideological attitudes predicting online hate speech: The differential effects of right-wing authoritarianism and social dominance orientation. *Frontiers in Social Psychology, 2*:1389437.
- Die Presse (18.02.2025). Beleidigungen und Morddrohungen: Ärzte wehren sich gegen Hass im Netz. <https://www.diepresse.com/19376888/beleidigungen-und-morddrohungen-aerzte-wehren-sich-gegen-hass-im-netz>
- Der Spiegel (28.12.2020). Löschung illegaler Postings - Bayerns Polizei unterstützt Opfer von Hasskommentaren. <https://www.spiegel.de/netzwelt/web/hasskriminalitaet-bayerns-polizei-soll-betroffene-staerker-unterstuetzen-a-91c01e04-0985-4dfd-82cc-7a2814e06dd9>
- Der Standard (29.11.2024). "20.000 Mails, 80 Prozent Hass": Mobbing im Netz hat sich "extrem verschlimmert". <https://www.derstandard.at/story/3000000247022/20000-mails-80-prozent-hass-mobbing-im-netz-hat-sich-extrem-verschlimmert>
- Der Standard (16.02.2020). Gegen Hate-Speech hilft kein Staat. <https://www.derstandard.at/story/2000114633601/gegen-hate-speech-hilft-kein-staat>
- Die Zeit (05.07.2018). User nehmen mehr Hass im Internet wahr. https://www.zeit.de/digital/internet/2018-07/hate-speech-internet-soziale-medien-hasskommentare-studie?utm_referer=https%3A%2F%2Fwww.google.com%2F
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering online hate speech*. Paris: United Nations Educational, Scientific and Cultural Organization (UNESCO).
- Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing, 18*(1), 38–52.
- Herr, P. M., Kardes, F. R., & Kim, J. (1991). Effects of word-of-mouth and product-attribute information on persuasion: An accessibility-diagnostics perspective. *Journal of Consumer Research, 17*(4), 454–462.
- Hutzinger, C., & Weitzl, W. J. (2021). Co-creation of online service recoveries and its effects on complaint bystanders. *Journal of Business Research, 130*, 525–538.
- Kaakinen, M., Sirola, A., Savolainen, I., & Oksanen, A. (2020). Impulsivity, internalizing symptoms, and online group behavior as determinants of online hate. *PloS one, 15*(4), e0231052.

- Kalsnes, B., & Ihlebæk, K. A. (2021). Hiding hate speech: Political moderation on Facebook. *Media, Culture & Society*, 43(2), 326–342.
- Kronen Zeitung (18.11.2020). Hass im Netz: Das neue Gesetzespaket im Detail. <https://www.krone.at/2278904>
- Kumar, J., & Nayak, J. K. (2019). Consumer psychological motivations to customer brand engagement: a case of brand community. *Journal of Consumer Marketing*, 36(1), 168–177.
- Kurier (02.04.2019). Beschimpft und bedroht: Islamfeindlichkeit trifft im Alltag vor allem Frauen. <https://kurier.at/politik/inland/islamfeindlichkeit-trifft-im-alltag-vor-allem-frauen/400454221>
- Kurier (25.02.2021). Hass gegen Fußball-Stars: "Ich hoffe du stirbst in der Hölle". <https://kurier.at/sport/fussball/ich-hoffe-du-stirbst-fussball-stars-kroos-und-co-lesen-hassnachrichten-vor/401200534>
- Lederer, L., & Delgado, R. (1995). *The price we pay: The case against racist speech, hate propaganda, and pornography*. New York: Hill and Wang.
- Leets, L. (2002). Experiencing hate speech: Perceptions and responses to anti-semitism and antigay speech. *Journal of Social Issues*, 58(2), 341–361.
- MacAvaney, S., Yao, H.-R., Yang, E., Russell, K., Goharian, N., & Frieder, O. (2019). Hate speech detection: Challenges and solutions. *PLoS ONE*, 14(8), e0221152.
- Malecki, W. P., Kowal, M., Dobrowolska, M., & Sorokowski, P. (2021). Defining online hating and online haters. *Frontiers in Psychology*, 12:744614.
- Meddaugh, P. M., & Kay, J. (2009). Hate speech or “reasonable racism?” The other in stormfront. *Journal of Mass Media Ethics*, 24(4), 251–268.
- Mossie, Z., & Wang, J.-H. (2020). Vulnerable community identification using hate speech detection on social media. *Information Processing & Management*, 57(3), 102087.
- Nieto, J., Hernández-Maestro, R. M., & Muñoz-Gallego, P. A. (2014). Marketing decisions, customer reviews, and business performance: The use of the Toprural website by Spanish rural lodging establishments. *Tourism Management*, 45, 115–123.
- 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.
- Ouvrein, G. (2024). Followers, fans, friends, or haters? A typology of the online interactions and relationships between social media influencers and their audiences based on a social capital framework. *New Media & Society*, 14614448241253770.
- Pace, U., D’Urso, G., & Zappulla, C. (2021). Hating among adolescents: Common contributions of cognitive distortions and maladaptive personality traits. *Current Psychology*, 40(7), 3326–3331.
- Park, A., Kim, M., & Kim, E.-S. (2023). SEM analysis of agreement with regulating online hate speech: Influences of victimization, social harm assessment, and regulatory effectiveness assessment. *Frontiers in Psychology*, 14:1276568.
- Saleem, M., Prot, S., Anderson, C. A., & Lemieux, A. F. (2017). Exposure to muslims in media and support for public policies harming muslims. *Communication Research*, 44(6), 841–869.
- Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems*, 81, 30–40.
- Schäfer, S., Rebasso, I., Boyer, M. M., & Planitzer, A. M. (2024). Can we counteract hate? Effects of online hate speech and counter speech on the perception of social groups. *Communication Research*, 51(5), 553–579.
- Sen, S., & Lerman, D. (2007). Why are you telling me this? An examination into negative consumer reviews on the web. *Journal of Interactive Marketing*, 21(4), 76–94.
- Soral, W., Bilewicz, M., & Winiewski, M. (2018). Exposure to hate speech increases prejudice through desensitization. *Aggressive Behavior*, 44(2), 136–146.

- Wachs, S., & Wright, M. F. (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.
- Wachs, S., Wright, M. F., & Gámez-Guadix, M. (2024). From hate speech to HateLess. The effectiveness of a prevention program on adolescents' online hate speech involvement. *Computers in Human Behavior*, *157*, 108250.
- Weber, I., Vandebosch, H., Poels, K., & Pabian, S. (2024). The ecology of online hate speech: Mapping expert perspectives on the drivers for online hate perpetration with the Delphi method. *Aggressive Behavior*, *50*(2), e22136.
- Weitzl, W., & Hutzinger, C. (2017). The effects of marketer- and advocate-initiated online service recovery responses on silent bystanders. *Journal of Business Research*, *80*, 164–175.
- Weitzl, W., & Hutzinger, C. (2019). Rise and fall of complainants' desires: The role of pre-failure brand commitment and online service recovery satisfaction. *Computers in Human Behavior*, *97*, 116–129.
- Weitzl, W., & Hutzinger, C., & Einwiller, S. (2018). An empirical study on how webcare mitigates complainants' failure attributions and negative word-of-mouth. *Computers in Human Behavior*, *89*, 316–327.
- Wiener Zeitung (15.01.2021). Ende gut, alles Wut. <https://www.wienerzeitung.at/meinung/gastkommentare/2088750-Ende-gut-alles-Wut.html>
- Williams, M. L., Burnap, P., Javed, A., Liu, H., & Ozalp, S. (2020). Hate in the machine: Anti-black and anti-muslim social media posts as predictors of offline racially and religiously aggravated crime. *The British Journal of Criminology*, *60*(1), 93–117
- Zhang, J. (2019). Sexual disgust sensitivity mediates the sex difference in support of censoring hate speech. *Personality and Individual Differences*, *145*, 89–96.