

### Wie gefährlich ist der Hass im Netz? Wirkungen von Hasskommentaren gegen Geflüchtete auf das prosoziale Verhalten von Rezipierenden

Köhler, Christina; Ziegele, Marc; Weber, Mathias

Erstveröffentlichung / Primary Publication

Sammelwerksbeitrag / collection article

#### Empfohlene Zitierung / Suggested Citation:

Köhler, C., Ziegele, M., & Weber, M. (2019). Wie gefährlich ist der Hass im Netz? Wirkungen von Hasskommentaren gegen Geflüchtete auf das prosoziale Verhalten von Rezipierenden. In I. Engemann, M. Legrand, & H. Marzinkowski (Hrsg.), *Politische Partizipation im Medienwandel* (S. 299-319). Berlin <https://doi.org/10.17174/dcr.v6.12>

#### Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:

<https://creativecommons.org/licenses/by/4.0/deed.de>

#### Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:

<https://creativecommons.org/licenses/by/4.0>

**Empfohlene Zitierung:** Köhler, C., Ziegele, M., & Weber, M. (2019). Wie gefährlich ist der Hass im Netz? Wirkungen von Hasskommentaren gegen Geflüchtete auf das prosoziale Verhalten von Rezipierenden. In I. Engelmann, M. Legrand, & H. Marzinkowski (Hrsg.), *Politische Partizipation im Medienwandel* (S. 299-319). doi: 10.17174/dcr.v6.12

**Zusammenfassung:** Inzivilität in Nutzerkommentaren kann sich negativ auf die Einstellungen von Rezipierenden auswirken. Inwieweit solche Kommentare auch das reale Verhalten der Lesenden beeinflussen, ist bislang weitgehend ungeklärt. Die vorliegende Studie untersucht deshalb die Auswirkungen von Hasskommentaren unter journalistischen Artikeln über Geflüchtete auf das prosoziale Verhalten von Rezipierenden. Basierend auf der Theory of Planned Behavior und auf Forschung zu sozialen Gruppenidentitäten wird mit einem Online-Experiment gezeigt, dass Hasskommentare gegen Geflüchtete – vermittelt über die Einstellungen der Rezipierenden – den tatsächlich gespendeten Betrag für eine Flüchtlingshilfsorganisation reduzieren. Gleichzeitig erhöhen Hasskommentare die Spende an eine Hilfsorganisation für Obdachlose. Für negative, aber zivil formulierte Kommentare finden sich keine vergleichbaren Effekte. Diese Befunde implizieren eine identitätsbasierte Polarisierung des prosozialen Verhaltens durch Hasskommentare.

**Lizenz:** Creative Commons Attribution 4.0 (CC-BY 4.0)

*Christina Köhler, Marc Ziegele & Mathias Weber*

# Wie gefährlich ist der Hass im Netz?

Wirkungen von Hasskommentaren gegen Geflüchtete auf das prosoziale Verhalten von Rezipierenden

## 1 Einleitung

Digitale Medientechnologien erweitern nicht nur die verfügbaren Informationsquellen und Kommunikationskanäle (Walther, 2011), sie ermöglichen es Nutzenden auch, ihre Positionen und Gedanken öffentlich kundzutun und miteinander zu diskutieren. Eine der verbreitetsten Formen öffentlich-interpersonaler Anschlusskommunikation sind Nutzerkommentare unter journalistischen Artikeln. In westlichen Gesellschaften schreiben zwischen 10 und 28 Prozent der Nutzenden mindestens einmal wöchentlich selbst Kommentare (Newman, Fletcher, Levy, & Nielsen, 2016). Eine weitaus größere Zahl rezipiert die Inhalte anderer Nutzender, um sich ein Bild der öffentlichen Meinung sowie der Positionen anderer zu den Themen der Nachrichtenberichterstattung zu machen (Lee & Jang, 2010; Springer, Engelmann, & Pfaffinger, 2015). Möglich ist dies nicht zuletzt dadurch, dass immer mehr Medienorganisationen ihre Inhalte über soziale Medien verbreiten und somit nicht nur für die eigenen Beiträge, sondern auch für die korrespondierenden Nutzerbeiträge große Reichweiten generieren (Rowe, 2015).

Verschiedene Studien argumentieren, dass Nutzerkommentare eine neue Form des sozialen Einflusses darstellen (Winter, Brückner, & Krämer, 2015;

Yun & Park, 2011): Tatsächlich dienen die Positionen, die Nutzende in ihren Beiträgen kundtun, anderen Nutzenden als Orientierungspunkte für ihre eigenen Wahrnehmungen und Einstellungen (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014; Hsueh, Yogeewaran, & Malinen, 2015; Prochazka, Weber, & Schweiger, 2016; Walther, DeAndrea, Kim, & Anthony, 2010). Fraglich ist allerdings, ob diese Einflüsse im Sinne einer demokratischen Gemeinschaft sind: Da ein Großteil der Nutzerkommentare einen deutlich negativen Ton bezüglich des Themas oder des Artikels aufweist und sich oftmals durch Inzivilität oder sogar Hass gegenüber den dargestellten Personen und Gruppen auszeichnet (Coe, Kenski, & Rains, 2014), werden insbesondere inzivile und hasserfüllte Nutzerbeiträge unter journalistischen Inhalten zunehmend als Gefahr für demokratische Gesellschaften diskutiert (Anderson et al., 2014). Vor allem im Kontext der Flüchtlingskrise im Jahr 2015 wurde der Hass von Nutzenden in den digitalen Kommentarspalten zum Gegenstand intensiver gesellschaftlicher Debatten. Befürchtet wurde insbesondere, dass sich das integrative Potenzial digitaler Medientechnologien in eine destruktive Kraft wandelt, die die Hilfsbereitschaft und Solidarität gesellschaftlicher Gruppen untereinander bedroht.

Ziel der vorliegenden Studie ist es, auf Basis von Ansätzen zur Informationsverarbeitung, der Theory of Planned Behavior (TPB) sowie Forschung zu sozialen Gruppenbeziehungen zu untersuchen, inwieweit inzivile und Hasskommentare gegen hilfsbedürftige Gruppen nicht nur die Wahrnehmungen und Einstellungen, sondern auch das prosoziale Verhalten der Rezipierenden gegenüber ebenjenen Gruppen beeinflussen. Darüber hinaus soll analysiert werden, inwiefern bereits zivil artikulierte Ablehnung und ein sachlicher, jedoch negativer Ton in Nutzerkommentaren ähnliche Effekte nach sich ziehen können, da diese Charakteristika in Online-Diskussionen noch weitaus verbreiteter sind als Hass und Inzivilität (Ziegele & Quiring, 2017). Um schließlich einen empirischen Beitrag zur Frage nach den Folgen von Hass und Inzivilität online für den gesellschaftlichen Zusammenhalt zu liefern, wird in einem dritten Schritt analysiert, inwiefern soziale Gruppenidentitäten (Ingroup vs. Outgroup) differenzierte Effekte auf prosoziales Verhalten bedingen können. Konkret wird geklärt, inwiefern Hasskommentare gegen Geflüchtete auch Effekte auf das Spendenverhalten zugunsten anderer hilfsbedürftiger Gruppen in der Gesellschaft – insbesondere aus der eigenen Ingroup (z. B. Obdachlose) – haben.

## 2 **Negativ, inzivil und hasserfüllt: Theoretische Grundlagen zur Wirkung von Nutzerkommentaren**

Die Frage nach den Effekten von negativen und inzivilen Nutzerkommentaren zieht zunehmend die Aufmerksamkeit zahlreicher Forschender auf sich (Anderson et al., 2014; Rösner, Winter, & Krämer, 2016). Inzivilität ist eine Kommunikationsform, mit der einer gegnerischen Position oder Sichtweise auf respektlose Art die Rechtmäßigkeit abgesprochen wird (Coe et al., 2014; Hwang, Kim, & Kim, 2016). Beispiele sind herabwürdigende Sprache, Vorurteile gegenüber sozialen Gruppen sowie verbale Aggressionen und Bedrohungen (Blom, Carpenter, Bowe, & Lange, 2014; Coe et al., 2014; Gervais, 2014; Hsueh et al., 2015). Ausgehend von diesem grundlegenden Verständnis werden unterschiedliche Schweregrade von Inzivilität differenziert: Während manche Formen wie Unhöflichkeit und Schimpfworte als relativ „harmlos“ erachtet werden (Coe et al., 2014), gelten andere wie etwa Hassrede, die den öffentlichen Diskurs vor allem im Kontext der Flüchtlingskrise in der jüngeren Zeit geprägt hat (Gagliardone, Gal, Alves, & Martinez, 2015), als weitaus schädlicher (Papacharissi, 2004). Hasskommentare in Online-Diskussionen zeichnen sich dadurch aus, dass sie andere aufgrund ihrer Zugehörigkeit zu einer sozialen Gruppe (z. B. basierend auf Religion, Rasse oder Geschlecht) dehumanisieren, reduzieren und bedrohen (Gagliardone et al., 2015).

Ein Blick auf die Prävalenz von Inzivilität in Online-Nutzerdiskussionen offenbart, dass es sich nicht um ein Randphänomen handelt: Eine Inhaltsanalyse auf einer US-Nachrichtenplattform kommt zu dem Ergebnis, dass 22 Prozent der Nutzerbeiträge mindestens eine Form von Inzivilität enthielten (Coe et al., 2014). Mit Blick auf Hasskommentare gaben zwei Drittel der deutschen Internetnutzenden an, solchen Kommentaren bereits online begegnet zu sein (Landesanstalt für Medien NRW [LfM], 2017). Dabei zeigen experimentelle Studien, dass die Folgen der Rezeption von inzivilen Kommentaren gravierend sein können: Neben aggressiven Kognitionen und negativen Emotionen (Gervais, 2014; Rösner et al., 2016) fördern inzivile Kommentare auch stereotypes Denken über soziale Gruppen (Hsueh et al., 2015) und können zur Meinungspolarisierung beitragen (Anderson et al., 2014). Schließlich führt das Lesen von inzivilen Nutzerkommentaren auch dazu, dass sich Nutzende in ihren eigenen Beiträgen in einem weniger zivilen Ton äußern (Hsueh et al., 2015; Ziegele, Weber, Quiring, & Breiner, 2017).

Erklärt werden diese Effekte unter anderem mit Ansätzen zur Informationsverarbeitung sowie zu Priming (Prochazka et al., 2016; Rösner et al., 2016). Folgt man den Überlegungen des Elaboration-Likelihood-Modells (Petty & Cacioppo, 1986), so verarbeiten Menschen Informationen auf einer peripheren Route bzw. in einem heuristischen Modus, wenn es ihnen in einer Situation an Motivation, Fähigkeit oder Involvement fehlt. Dies führt dazu, dass sie in ihrer Verarbeitung insbesondere für periphere Botschaftsmerkmale wie Emotionen oder die Tonalität empfänglich sind. Da gerade inzivile Nutzerkommentare eine starke affektive Komponente und einen prägnanten negativen Ton aufweisen, könnten sie bei Menschen in einem Stadium geringen Involvements besonders persuasiv wirken. Über ihre Wirkung auf die Einstellungen der Rezipierenden könnten schließlich auch Effekte auf das Verhalten resultieren (Petty & Cacioppo, 1990) – beispielsweise auf den Ton des eigenen Kommentars, den sie anschließend verfassen (Hsueh et al., 2015) oder mit Blick auf die Hilfsbereitschaft, die sie anderen entgegenbringen.

Ein ähnlicher Wirkmechanismus lässt sich aus der Forschung zu Priming ableiten: Die Rezeption einer bestimmten Information aktiviert korrespondierende Kognitionen, die – meist unterbewusst (Murphy & Zajonc, 1993) – einen Einfluss darauf haben, wie sich Rezipierende in Bezug auf einen nachfolgenden Stimulus verhalten (Gilliam & Iyengar, 2000; Roskos-Ewoldsen, Klinger, & Roskos-Ewoldsen, 2007). So zeigt die Forschung zur Wirkung von Gewalt in den Medien zum Beispiel, dass Gewaltdarstellungen feindselige und aggressionsbezogene kognitive Konzepte aktivieren, die im Anschluss wiederum aggressives Verhalten begünstigen (Roskos-Ewoldsen et al., 2007). Folglich kann vermutet werden, dass auch Hasskommentare gegen Flüchtlinge feindselige kognitive Muster aktivieren und darüber antisoziales Verhalten auslösen oder zumindest die Bereitschaft der Rezipierenden zu prosozialem Verhalten reduzieren.

In der Gesamtschau implizieren sowohl die Ansätze zur Informationsverarbeitung als auch die Forschung zu Priming durch Medieninhalte, dass inzivile und Hasskommentare das prosoziale Verhalten der Nutzenden beeinträchtigen können, indem sie die Einstellungen der Rezipierenden negativ beeinflussen. Wie jedoch der Zusammenhang zwischen den Einstellungen und dem tatsächlichen Verhalten der Rezipierenden konkret begriffen werden kann, vermag die Theory of Planned Behavior (TPB) zu erklären. Demnach ist die Intention, eine bestimmte Handlung auszuführen, zentral dafür, ob die entsprechende Handlung auch tatsächlich ausgeführt wird (Ajzen, 1985). Diese Intention zu

agieren wird wiederum durch Einstellungen (d. h. allgemeine positive oder negative Bewertung eines Verhaltens), subjektive Normen (d. h. wahrgenommener sozialer Druck zu einem Verhalten) sowie die wahrgenommene eigene Verhaltenskontrolle (d. h. wahrgenommene Selbstwirksamkeit mit Blick auf ein bestimmtes Verhalten) beeinflusst (Armitage & Conner, 2001). Die Annahmen der TPB wurden empirisch umfangreich bestätigt (Armitage & Conner, 2001) und eignen sich auch zur Erklärung von verschiedenen Formen des prosozialen Verhaltens (France, France, & Himawan, 2007; Smith & McSweeney, 2007). Allerdings wurde bislang nur selten der kausale Einfluss von Botschaften mit spezifischen Merkmalen – etwa Hasskommentaren – auf anschließendes Verhalten im Rahmen der TPB experimentell untersucht. Daher ist das Ziel der vorliegenden Untersuchung, die empirische Evidenz hinsichtlich der TPB um diesen Aspekt zu erweitern.

Zusammenfassend führen die zuvor dargestellten theoretischen Überlegungen zu folgender Argumentation: Inzivile und Hasskommentare entwickelten sich gerade im Kontext der gesellschaftlichen und politischen Entwicklungen in der Flüchtlingskrise zu einem problematischen Phänomen (Gagliardone et al., 2015). Es kann angenommen werden, dass das Lesen von inzivilen Kommentaren gegen soziale Minderheiten dazu führt, dass negative, stereotype Einstellungen gegenüber diesen sozialen Gruppen aktiviert werden, die wiederum das prosoziale Verhalten der Rezipierenden unterbinden. Eine potenzielle Form des prosozialen Verhaltens gegenüber den diskreditierten Gruppen stellen Spenden für entsprechende Hilfsorganisationen dar. Die TPB liefert ein fruchtbares theoretisches Konzept, um die Effekte von Hasskommentaren gegen Flüchtlinge auf das Spendenverhalten der Rezipierenden zugunsten von Flüchtlingshilfsorganisationen zu untersuchen. Entsprechend nehmen wird an, dass inzivile Nutzerkommentare gegen Geflüchtete einen negativen Effekt auf die Einstellungen der Rezipierenden bezüglich einer Spende für eine Flüchtlingshilfsorganisation haben. Diese Einstellungen hängen wiederum mit der Spendenintention zusammen, die letztendlich die Höhe des tatsächlich gespendeten Geldbetrags determiniert. Insgesamt wird somit die Hypothese aufgestellt, dass Hasskommentare – vermittelt über die Einstellungen der Rezipierenden und ihre Spendenintention – den gespendeten Betrag zugunsten einer Flüchtlingshilfsorganisation reduzieren (H1). Darüber hinaus soll die Wirkung von inzivilen Nutzerkommentaren auf die subjektiven Normen sowie auf die wahrgenommene Verhaltenskontrolle bezüglich

Spenden untersucht werden – Aspekte, die in der bisherigen Forschung bislang nicht betrachtet wurden (FF1).

Abschließend bleibt hinsichtlich der Wirkung inziviler Nutzerkommentare auf das prosoziale Verhalten der Rezipierenden eine grundlegende Frage zu klären: Da inzivile und auch Hasskommentare – als Steigerung ersterer – von Natur aus negativ im Ton sind, könnten potenzielle Effekte bereits der Negativität und nicht ausschließlich der Inzivilität zuzurechnen sein. Dies scheint insofern relevant, als zivile Negativität in Nutzerdiskussionen weitaus verbreiteter ist als ihr inziviles Pendant. Zudem zeigen Studien, dass Negativität und Ablehnung in einem Nutzerkommentar auch dann einen Einfluss auf die Meinungen der Rezipierenden haben können, wenn sie in einem zivilen Ton geäußert werden (Lee & Jang, 2010; Winter & Krämer, 2016). Ungeklärt ist jedoch, ob zivil-negative und inzivile bzw. hasserfüllte Kommentare einen vergleichbaren Einfluss auf die Einstellungen und das prosoziale Verhalten von Nutzenden haben, weshalb wir diesbezüglich eine Forschungsfrage formulieren (FF2).

### **3 Soziale Gruppenidentitäten als Kanalisierungsinstrument für die Wirkung von Nutzerkommentaren**

Im Zuge der europäischen Flüchtlingskrise im Jahr 2015 wurden die ankommenden Migrantinnen und Migranten, die vor allem aus Ländern mit einer muslimischen Bevölkerungsmehrheit stammten, in einigen Debatten als Gefahr für die westliche Kultur, ihre Werte und auch die sozialen Sicherungssysteme dargestellt (Topinka, 2017). Gleichzeitig versuchte insbesondere die politische Rechte einen sozialen Vergleich zu aktivieren, der die „fremden“ Geflüchteten „einheimischen“ Hilfsbedürftigen (z. B. Obdachlose) gegenüberstellte. Die Etablierung solcher sozialer Vergleichsmechanismen zwischen der eigenen Ingroup und einer Outgroup beschränkt sich nicht nur auf die Flüchtlingskrise, sondern scheint ein übliches Narrativ in gesellschaftspolitischen Debatten zu sein; kürzlich konnte auch in einer US-Debatte die Gegenüberstellung zwischen „fremden“ Geflüchteten und einheimischen Veteraninnen und Veteranen vernommen werden, wobei letztere zuerst das Mitgefühl der Gesellschaft verdient hätten (Darda, 2017). Im Kontext von Nutzerdiskussionen wirft die Etablierung solcher Vergleichsmechanismen die Frage auf, inwieweit inzivile Kommentare gegen soziale Minderheiten ausschließlich das

prosoziale Verhalten gegenüber der diskreditierten sozialen Gruppe oder gleichzeitig auch das Verhalten gegenüber anderen sozialen Gruppen in der Bevölkerung beeinflussen.

Die Self-Concept Maintenance Theorie bietet zur Erklärung möglicher Effekte einen fruchtbaren Ausgangspunkt. Prosoziales Verhalten zugunsten einer hilfsbedürftigen Gruppe kann als eine spezifische Form von moralischem Verhalten begriffen werden. Folgt man der Theorie, so stellen Menschen ein Kosten-Nutzen-Kalkül an, um sich für oder gegen ein moralisch wünschenswertes Verhalten zu entscheiden: Individuen streben danach, die Vorzüge aus einer Situation zu maximieren und dabei gleichzeitig ein moralisches Selbstbild zu erhalten (Mazar, Amir, & Ariely, 2008; Welsh & Ordonez, 2014). Spenden zugunsten einer hilfsbedürftigen Gruppe stellt ein Verhalten dar, in dem diese beiden Ziele in Widerspruch geraten können: Zwar führt die Spende zugunsten von Geflüchteten zu einem monetären Verlust. Gleichzeitig bietet sie die Chance, das eigene moralische Selbstkonzept zu stärken. Wenn nun jedoch über Hasskommentare gegen Geflüchtete negative Einstellungen gegenüber dem Spendenobjekt aktiviert werden, so ist zu vermuten, dass sich Individuen eher dafür entscheiden, ihre eigenen monetären Vorzüge zu sichern, da die moralische Verhaltensoption (Spende für eine hilfsbedürftige Gruppe) in Misskredit geraten ist. Darüber hinaus könnte jedoch auch das Verhalten gegenüber anderen sozialen Gruppen in der Gesellschaft beeinflusst werden: Forschung zu sozialen Gruppenidentitäten impliziert, dass die Ablehnung einer vermeintlichen Outgroup sowie die positive Bindung an die eigene Ingroup unter bestimmten Bedingungen negativ reziprok zusammenhängen: Nämlich dann, wenn sich Intergruppenbeziehungen dadurch auszeichnen, dass eine Outgroup als unmittelbare Bedrohung für die Interessen der eigenen Ingroup wahrgenommen wird, scheint die Identifikation der oder des Einzelnen mit der eigenen Ingroup und die gleichzeitige Feindseligkeit gegenüber der Outgroup besonders ausgeprägt (Brewer, 1999).

Zusammenfassend führen diese Überlegungen zu folgenden Annahmen: Inzivile und Hasskommentare aktivieren negative Einstellungen gegenüber Geflüchteten, was wiederum zu einer geringeren Spende zugunsten einer Flüchtlingshilfsorganisation führt (vgl. H1). Die Forschung zu sozialen Gruppenidentitäten impliziert, dass diese negativen Einstellungen gegenüber den Geflüchteten als vermeintliche Outgroup mit einer ausgeprägten Identifikation mit hilfsbedürftigen sozialen Gruppen der eigenen Ingroup einhergehen (z. B. einheimische

Obdachlose). Basierend auf den Thesen der Self-Concept Maintenance Theorie sollten demnach Spenden zugunsten von Hilfsorganisationen, die sich dem Wohl von Obdachlosen annehmen, üppiger ausfallen, wenn Rezipierende zuvor mit einem Hasskommentar gegen Geflüchtete konfrontiert wurden (H2).

#### **4 Methodisches Vorgehen: Online-Experiment zur Wirkung von Hasskommentaren**

Die Hypothesen und Forschungsfragen wurden mit einem Online-Experiment im 3 x 2-Between-Subjects-Design (Valenz des Nutzerkommentars x Empfänger der Spende) untersucht. Den Teilnehmenden des Experiments wurden drei aufeinanderfolgenden Screenshots gezeigt, die Nachrichtenbeiträge auf der Facebook-Präsenz eines fiktionalen lokalen Nachrichtenmediums darstellten. Zwei der drei Beiträge thematisierten die Flüchtlingskrise. Dazwischen diente eine Nachricht ohne jeglichen Bezug zur Flüchtlingsthematik als Puffer. Unter jedem Nachrichtenbeitrag standen zudem mehrere Nutzerkommentare.

##### *4.1 Stichprobe*

Die für deutsche Online-Nutzende ab 18 Jahren repräsentativ quitierte Stichprobe des Online-Experiments umfasste 497 Teilnehmende und wurde über ein kommerzielles Online-Accesspanel rekrutiert. 51 Prozent waren weiblich, das mittlere Alter der Probandinnen und Probanden betrug 44 Jahre und 63 Prozent hatten maximal den Realschulabschluss.

##### *4.2 Design*

Auf der Startseite des Onlinefragebogens wurden die Teilnehmenden zunächst informiert, dass die Studie gewalthaltiges Material enthalten könnte. Zudem wurde angekündigt, dass die Teilnehmenden einen monetären Anreiz von 5 Euro erhalten würden, wenn sie den Fragebogen vollständig ausfüllten. Nach demographischen Angaben und verschiedenen Fragen zu ihrer persönlichen Me-

diennutzung wurden die Teilnehmenden mit dem Stimulusmaterial konfrontiert: Sie sahen jeweils ein zufällig zugewiesenes Set aus drei aufeinanderfolgenden Nachrichtenteasern und Nutzerkommentaren. Die Nachrichtenbeiträge unterschieden sich dabei nicht zwischen den Teilnehmenden. Im ersten Beitrag wurden die Befunde einer aktuellen Studie berichtet, wonach viele Geflüchtete in Europa bleiben und arbeiten wollen. Der Pufferbeitrag an zweiter Stelle hatte einen Zugunfall zum Gegenstand. Im dritten Nachrichtenteaser ging es erneut um die Flüchtlingsthematik: Thema war ein rechtsextremer Anschlag auf Geflüchtete und die zunehmende Gewalt gegen diese soziale Gruppe. Um den Einfluss von Hasskommentaren und zivil geäußelter Kritik auf das prosoziale Verhalten zu untersuchen, wurde der erste von drei Nutzerkommentaren unter dem ersten und dritten Nachrichtenbeitrag manipuliert. Ein Drittel der Teilnehmenden sah Kommentare, die sich in einem zivilen Ton neutral über Flüchtlinge äußerten (z. B. „Es hängt von uns ab, wie wir mit Flüchtlingen umgehen und welches Leben wir ihnen hier ermöglichen wollen.“). Die zweite Experimentalgruppe sah Kommentare, die negativ im Ton, aber dennoch zivil formuliert waren (z. B. „Schon wieder Flüchtlinge. Leider sind einige Asylbewerber wirklich gefährlich – selbst sehr aggressiv und oft daran interessiert, sich an uns zu bereichern.“). Die dritte Gruppe sah Hasskommentare gegen Geflüchtete (z. B. „Schon wieder diese DRECKSFLUECHTLINGE?!? Das versiffte Gesindel ist ne Gefahr für jeden anständigen Bürger! Selbst aggressiv wie die Pest, saudumm und wollen uns nur aussaugen!!“). Die zweiten und dritten Kommentare unter den entsprechenden Nachrichtenbeiträgen unterschieden sich nicht zwischen den Experimentalgruppen und enthielten keine Elemente von Hass oder Negativität. Insgesamt handelt es sich bei den Kommentaren zwar um fiktionale Beiträge, die jedoch an realen Kommentaren auf verschiedenen Facebook-Seiten angelehnt waren. Die Ergebnisse des Treatmentchecks bestätigen, dass die Manipulation erfolgreich war: Die Versuchspersonen wurden auf einer siebenstufigen Skala (1 = ‘überhaupt nicht’ 7 = ‘sehr stark’) gefragt, wie inzivil, negativ, hasserfüllt und realistisch sie die zuvor gesehenen Nutzerkommentare empfanden. Die Teilnehmenden schätzten die Kommentare mit Blick auf die Inzivilität ( $F(2,495) = 132,07, p < ,001, \eta_p = ,35$ ), den Grad an Hass ( $F(2,495) = 160,13, p < ,001, \eta_p = ,39$ ) sowie die empfundene Negativität ( $F(2,250) = 73,99, p < ,001, \eta_p = ,23$ ) signifikant unterschiedlich ein. Darüber hinaus wurden die Kommentare im Durchschnitt als recht realistisch eingeschätzt ( $M = 5,25, SD = 1,52$ ), wobei sich diese Einschätzung nicht signifikant zwischen den verschiedenen Experimentalgruppen unterschied

( $F(2,495) = 2,10, p = ,13$ ). Bonferroni-Post-Hoc-Tests zeigten, dass die Hasskommentare ( $M_{unziv} = 5,65, SD_{unziv} = 1,54, M_{hass} = 5,48, SD_{hass} = 1,60$ ) signifikant inziviler und hassereffüllter wahrgenommen wurden als ihre negativ-zivilen Pendanten ( $M_{unziv} = 3,52, SD_{unziv} = 1,73, p < ,001, M_{hass} = 3,86, SD_{hass} = 1,75, p < ,001$ ). Jedoch wurden hassereffüllte Kommentare ( $M = 5,11, SD = 1,57$ ) im Vergleich zu den negativ-zivilen Kommentaren nicht als negativer wahrgenommen ( $M = 5,12, SD = 1,40, p = ,40$ ). Schließlich zeigte sich, dass die neutralen Nutzerbeiträge als weniger inzivil ( $M = 2,25, SD = 1,48$ ), weniger hassereffüllt ( $M = 2,29, SD = 2,29$ ) und weniger negativ ( $M = 3,55, SD = 1,46$ ) wahrgenommen wurden als die negativ-zivilen und Hasskommentare.

Nachdem die Teilnehmenden die Nachrichtenbeiträge sowie die korrespondierenden Nutzerbeiträge gesehen hatten, wurden sie gebeten, einige Pufferfragen in Bezug auf den Nachrichtenbeitrag zum Zugangsglück zu beantworten und die Items zum Treatmentcheck wurden erfasst. Anschließend wurde den Versuchspersonen eröffnet, dass sie den Anreiz von fünf Euro für die erfolgreiche Teilnahme an der Studie entweder für sich behalten oder ganz bzw. in Teilen für eine Hilfsorganisation spenden können. Dabei wurden die Teilnehmenden zufällig einer von zwei Spendenkonditionen zugeteilt: Die eine Hälfte konnte für die Flüchtlingshilfe einer etablierten Hilfsorganisation spenden, die andere für die Obdachlosenhilfe derselben Hilfsorganisation. Bevor die Probandinnen und Probanden allerdings über die tatsächliche Höhe der Spende entscheiden konnten, wurden die relevanten Items zur Operationalisierung der TPB erfasst. Abschließend wurden die Teilnehmenden über den fiktiven Charakter der Nachrichtenbeiträge sowie der Nutzerkommentare aufgeklärt. Zudem wurde den Versuchspersonen ihr Anreizbetrag ausgezahlt (organisiert über einen kommerziellen Online-Access-Anbieter) und der zur Spende vorgesehene Betrag an die entsprechenden Hilfsorganisationen überwiesen.

#### 4.3 Operationalisierung der Konstrukte

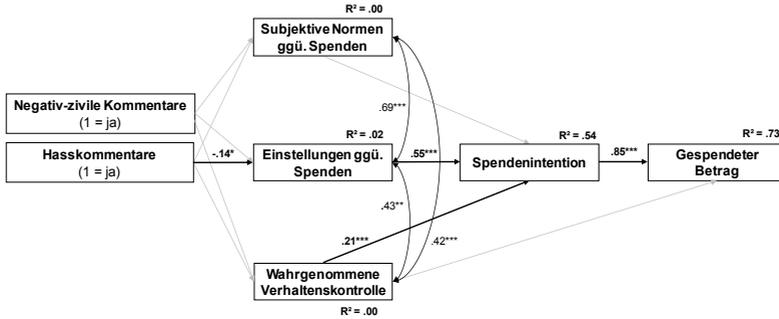
Die Items der TPB wurden in Anlehnung an die Skalen von Fishbein und Ajzen (2010) operationalisiert. Die Handlung, die bei allen Fragen zu den Einstellungen der Teilnehmenden, ihren subjektiven Normen sowie ihrer wahrgenommenen Verhaltenskontrolle als Referenz diente, war das Spenden ihres Incentives zugunsten der Flüchtlings- bzw. Obdachlosenhilfe. Alle Items wurden auf einer Skala von 1 = 'stimme überhaupt nicht zu' bis 7 = 'stimme voll und

ganz zu' erfasst. Die Einstellungen wurden über fünf siebenstufige semantische Differentiale abgefragt (z. B. „Mein Incentive ganz oder teilweise für die Hilfsorganisation zu spenden finde ich sehr schlecht/sehr gut“,  $\alpha = ,96$ ). Zur Messung der subjektiven Normen dienten vier Items (z. B. „Die meisten Leute, die mir wichtig sind, finden, dass ich mein Incentive ganz oder teilweise für die Hilfsorganisation spenden sollte“,  $\alpha = ,88$ ) und die wahrgenommene Verhaltenskontrolle wurde über drei Items erfasst (z. B., „Wenn ich wollte, wäre es einfach für mich, mein Incentive ganz oder teilweise an die Hilfsorganisation zu spenden“,  $\alpha = ,85$ ). Die Intention zur Spende ihres Anreizbetrages (oder Teilen davon) wurde über zwei Items abgefragt (z. B. „Ich habe vor, mein Incentive ganz oder teilweise an die Hilfsorganisation zu spenden.“,  $\alpha = ,92$ ). Über einen Schieberegler konnten die Teilnehmenden abschließend angeben, wie viel ihres Anreizbetrages sie zugunsten der Flüchtlings- bzw. Obdachlosenhilfe spenden wollen.

## 5 Ergebnisse: Einfluss von negativ-zivilen und Hasskommentaren auf prosoziales Verhalten

Randomisierungschecks zeigten keine signifikanten Unterschiede zwischen den Experimentalgruppen hinsichtlich Alter ( $F(2,495) = 1,01, p = ,37$ ), Geschlecht ( $\chi^2(2) = 2,81, p = ,25$ ) und Bildungsniveau der Probandinnen und Probanden ( $\chi^2(6) = 4,21, p = ,65$ ). Um H1 zu prüfen, wurde ein Strukturgleichungsmodell mit AMOS 23 geschätzt, wobei nur diejenigen Teilnehmenden einbezogen wurden, die zugunsten von Geflüchteten spenden konnten ( $n = 253$ ). Das Modell umfasste die unabhängige Variable (recodiert in die beiden dichotomen Variablen „negativ-ziviler Kommentar“ und „Hasskommentar“, wobei die neutrale Kommentarversion jeweils als Referenzkategorie diente), die subjektiven Normen, Einstellungen, wahrgenommene Verhaltenskontrolle und Spendenintention als latente Mediatorvariablen und schließlich die abhängige Variable („gespendeter Betrag“). Das Modell weist eine gute Anpassungsgüte auf (Hu & Bentler, 1999):  $\chi^2/df = 1,42$  ( $df = 74$ ),  $p = ,01$ , CFI = ,99, RMSEA = ,04, SRMR = ,03, siehe Abbildung 1). Verglichen mit neutralen Kommentaren verschlechterten Hasskommentare die Einstellungen der Teilnehmenden bezüglich einer Spende für die Flüchtlingshilfsorganisation ( $\beta = -,14, p = ,02$ ). Die Einstellungen selbst hatten einen positiven Effekt auf die Spendenintention der Teilnehmenden ( $\beta = ,55, p < ,001$ ), die

Abbildung 1: Strukturgleichungsmodell der Wirkungen von negativ-zivilen und Hasskommentaren auf das Spendenverhalten der Probandinnen und Probanden



Anmerkung: Maximum Likelihood-Schätzung,  $n = 253$ ,  $\chi^2/df = 1,42$  ( $df = 74$ ),  $p = ,01$ , CFI = ,99, RMSEA = ,04, SRMR = ,03, standardisierte Koeffizienten. Indirekter Effekt von Hasskommentaren über Einstellungen und Spendenintention ( $b = -0,12$ , LLCI = -0,28, ULCI = -0,01,  $p = ,03$ ) wurde über Bootstrapping ermittelt ( $m = 2.000$ , 95% Konfidenzintervalle). Graue Linien zeigen nicht-signifikante Pfade. Die (nicht-signifikanten) direkten Effekte von negativ-zivilen und Hasskommentaren auf die Spendenintention und den gespendeten Betrag wurden im Modell kontrolliert.

wiederum stark mit der tatsächlichen Höhe der Spende korrelierte ( $\beta = ,85$ ,  $p < ,001$ ). Der über Bootstrapping ermittelte indirekte Effekt von Hasskommentaren über die Einstellungen und die Intention der Teilnehmenden auf ihr tatsächliches Spendenverhalten war negativ und ebenfalls signifikant ( $\beta = -,12$ ,  $p = ,03$ , LLCI = -0,28, ULCI = -0,01). Diese Befunde bestätigen H1. Mit Blick auf FF1 zeigte sich, dass die wahrgenommene Verhaltenskontrolle sowie die Intention zur Spende zwar in einem positiven Zusammenhang standen ( $\beta = ,21$ ,  $p < ,001$ ), erstere jedoch nicht signifikant durch die Rezeption der Hasskommentare beeinflusst wurde ( $\beta = -,01$ ,  $p = ,91$ ). Die subjektiven Normen der Teilnehmenden bezüglich einer Spende für Geflüchtete waren weder von der Rezeption von Hasskommentaren beeinflusst ( $\beta = -,05$ ,  $p = ,49$ ) noch hatten sie einen Effekt auf die Spendenintention ( $\beta = ,08$ ,  $p = ,25$ ). FF2 zielte auf die differenzierten Effekte von negativ-zivilen im Gegensatz zu inzivilen Kommentaren: Hierbei zeigte sich, dass die Rezeption von negativen, jedoch sachlichen Nutzerkommentaren zur Flüchtlingsthematik weder signifikante Einflüsse auf die relevanten Konstrukte der TPB noch auf das Spendenverhalten der Teilnehmenden hatten.

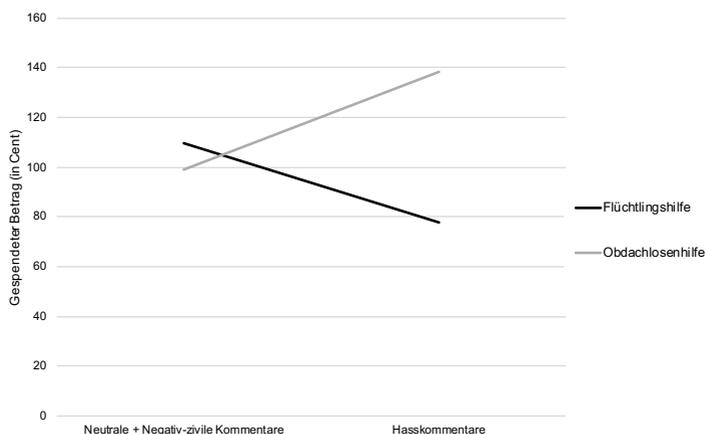
Um H2 zu prüfen, wurde eine mehrfaktorielle Varianzanalyse mit allen 497 Teilnehmenden gerechnet (vgl. Abbildung 2). Ziel war es, zu untersuchen, inwieweit das Lesen von Hasskommentaren gegen Geflüchtete unterschiedliche Effekte mit Blick auf den gespendeten Betrag zugunsten einer Flüchtlingshilfsorganisation im Vergleich zu einer Obdachlosenhilfsorganisation nach sich zieht. Die Rezeption von zivil-negativen Kommentaren wurde hierbei als Kontrollvariable im Modell berücksichtigt. Tatsächlich zeigten die Ergebnisse einen Interaktionseffekt zwischen dem Vorhandensein von Hass in den Nutzerkommentaren und der die Spende empfangenden Organisation: Das Lesen von Hasskommentaren führte einerseits zu einer geringer ausgefallenen Geldspende für Geflüchtete. Andererseits führten ebendiese Kommentare dazu, dass die Teilnehmenden einen signifikant größeren Anteil ihres Anreizbetrages zugunsten der Obdachlosenhilfe spendeten ( $F(3,494) = 5,326, p < ,05, \eta_p = ,01$ ). Damit bestätigen die Ergebnisse H2.

## 6 Diskussion und Fazit

Ziel der vorliegenden Untersuchung war es, den Einfluss von negativ-zivilen und Hasskommentaren auf das prosoziale Verhalten von Rezipierenden zu untersuchen. In Einklang mit früheren Erkenntnissen der Forschung zu Inzivilität in Online-Diskussionen (Anderson et al., 2014; Hsueh et al., 2015) zeigte sich, dass Hasskommentare die Kognitionen und Einstellungen anderer Nutzender negativ beeinflussten. Darüber hinaus hatten Hasskommentare einen negativen indirekten Effekt auf das Spendenverhalten der Rezipierenden zugunsten einer Flüchtlingshilfsorganisation. Gleichzeitig führte die Rezeption von Hasskommentaren zu einer umfangreicheren Spende, wenn das Geld einheimischen Obdachlosen zugutekam. Überraschenderweise hatten Kommentare, die ähnliche Kritik und Ablehnung äußerten, dabei aber zivil im Ton blieben, keine vergleichbaren Folgen.

Die Ergebnisse zeigen zunächst, dass Online-Nutzerkommentare tatsächlich sozialen Einfluss ausüben können (Hsueh et al., 2015; Winter et al., 2015) und anderen Nutzenden als Basis dienen, um ihre eigenen Einstellungen zu bilden und ihr Verhalten abzuleiten. Während frühere Studien insbesondere auf die Konsequenzen von inzivilen Kommentaren auf das Onlinediskussionsverhalten der Nutzenden fokussierten (Borah, 2014; Hsueh et al., 2015; Rösner et al., 2016), liefert die vorliegende Untersuchung empirische Evidenz dafür, dass Hasskommentare auch

Abbildung 2: Interaktionseffekt zwischen Tonalität der Kommentare und Spende empfangender Organisation auf den gespendeten Betrag



Anmerkung:  $F(3,494) = 5,326, p < ,05, \eta p = ,01$ .

mit Blick auf eine Form von prosozialem Verhalten – nämlich Spenden zugunsten von hilfsbedürftigen Minderheiten – einen Einfluss haben. Dies ist nicht zuletzt deshalb problematisch, weil nur sehr wenige Nutzende Hasskommentare verfassen (Blom et al., 2014) und sich somit offenbar eine große Anzahl an ‚Unbeteiligten‘ von einer kleinen Minderheit beeinflussen lässt.

Gleichzeitig scheint dies nicht für jeden Nutzerkommentar zu gelten: Negativ-zivile Kommentare, die im Kern dieselbe Kritik enthielten wie die Hasskommentare, hatten in der vorliegenden Untersuchung keine sichtbaren Wirkungen auf Einstellungen und Verhalten der Rezipierenden. Dass sachlich geäußerte Kritik keine vergleichbaren Effekte hervorbrachte, impliziert, dass Rezipierende weniger die Argumente der Kommentarverfasserinnen und -verfasser berücksichtigen, sondern sich vielmehr von einem scharfen Ton sowie affektiven Komponenten beeindrucken lassen. Im Lichte des Elaboration-Likelihood-Modells (Petty & Cacioppo, 1986) spricht dieses Ergebnis dafür, dass Nutzerkommentare eher oberflächlich und in einem heuristischen Modus verarbeitet werden.

Die hier vorgestellten Resultate fügen sich auch in die Annahmen der Priming-Forschung ein: Basierend auf den Erkenntnissen früherer Forschung (Hsueh et al., 2015; Rösner et al., 2016) erscheint die Erklärung plausibel, dass die Hasskommentare in unserer Studie feindselige Kognitionen und/oder Emotionen der Teilnehmenden aktiviert haben, die wiederum als Grundlage zur Bewertung dienen, ob und wie viel zugunsten einer Flüchtlingshilfsorganisation gespendet werden soll. Der Befund, dass negativ-zivile Kommentare keine Effekte auf die Einstellungen zum Spenden sowie das tatsächliche Spendenverhalten hatten, kann damit erklärt werden, dass die Stärke des Primes (Roskos-Ewoldsen et al., 2007) im Falle von negativ-zivilen Kommentaren zu gering war, um starke negative Kognitionen und Emotionen zu aktivieren.

Nichtsdestotrotz vermag die Priming-Perspektive keine vollständige Erklärung dafür zu liefern, warum das Lesen von Hasskommentaren gegen Geflüchtete die Spende zugunsten von einheimischen Obdachlosen positiv beeinflusste. Hier scheinen die Self-Concept Maintenance Theorie sowie Forschung zu sozialen Gruppenidentitäten fruchtbarer: Die Flüchtlingsthematik wurde in Europa sehr kontrovers diskutiert und in manchen Bevölkerungsteilen scheint sich die Wahrnehmung etabliert zu haben, dass Migrantinnen und Migranten eine vermeintliche Bedrohung für deren eigene Interessen darstellen. In einem solchen Meinungsklima vermag die Diskreditierung der Outgroup durch einen Hasskommentar zu einer stärkeren Bindung an die vermeintliche Ingroup zu führen (Brewer, 1999), die wiederum positive Einstellungen und wohlwollendes Verhalten gegenüber dieser bedingt. Dieser vermeintlich positive Effekt auf das prosoziale Verhalten gegenüber einer hilfsbedürftigen Gruppe (einheimische Obdachlose) könnte jedoch gleichzeitig eine problematische Ingroup-Outgroup-Polarisierung auf der gesamtgesellschaftlichen Ebene andeuten und den gesellschaftlichen Zusammenhalt insgesamt gefährden. Ähnliche Polarisierungseffekte durch inzivile Kommentare konnten auch für die Debatte um Risikotechnologien in der Studie von Anderson et al. (2014) aufgezeigt werden.

In Anbetracht dieser in Teilen besorgniserregenden Effekte scheint die Frage angebracht, ob und wie ihnen entgegengewirkt werden kann. Während viele Social-Media-Redakteurinnen und -Redakteure sowie Community-Managerinnen und -Manager inzivile und Hasskommentare löschen (Singer et al., 2011), werden solche Interventionen mit Blick auf eine potenzielle Einschränkung der freien Meinungsäußerung auch kritisch betrachtet (Ziegele & Jost, 2016). Ein alternatives,

transparentes Instrument könnte die interaktive Moderation von Nutzerdiskussionen sein (Stroud, Scacco, Muddiman, & Curry, 2015): Dabei antworten Journalistinnen und Journalisten auf inzivile Kommentare, engagieren sich aktiv in Gegengrede zu diesen und zeigen damit auf zivile Art alternative Perspektiven auf (Ziegele & Jost, 2016). Insbesondere auf Plattformen wie Facebook werden solche moderierenden Beiträge dann auch direkt unter dem inzivilen Kommentar angezeigt. Dies kann wiederum auch auf die rezipierenden Nutzenden wirken: Die Forschung zu Moralurteilen und moralisch wünschenswertem Verhalten zeigt, dass die Aktivierung von Moralstandards in den Kognitionen der Rezipierenden dazu führt, dass diese dann eher bereit sind, Kosten auf sich zu nehmen, um sich moralisch wünschenswert zu verhalten und damit ihr moralisches Selbstkonzept zu verbessern (Welsh & Ordonez, 2014). Im Kontext von Online-Nutzerdiskussionen könnten moderierende Beiträge von Moderatorinnen und Moderatoren – aber auch anderen Nutzenden – ebensolche Moralstandards aktivieren.

Die vorliegenden Ergebnisse müssen im Lichte einiger Limitationen betrachtet werden. Zunächst vermag das verwendete Experimentaldesign lediglich kurzfristige Effekte aufzuzeigen. Die Priming-Forschung legt zudem nahe, dass solche Effekte nur von kurzer Dauer sind (Roskos-Ewoldsen et al., 2007). Folglich könnte vermutet werden, dass die Effekte der Hasskommentare auf die Einstellungen und das prosoziale Verhalten der Nutzenden nur kurzlebige Ausreißer sind, die sich nach der Konfrontation mit dem Stimulus schnell wieder normalisieren. Bedenkt man jedoch, dass zahlreiche Onlinediskussionen zumindest einen kleinen Anteil an inzivilen Kommentaren enthalten (Coe et al., 2014; Rowe, 2014), so kann vermutet werden, dass Nutzende in regelmäßigen – wenn auch kleinen – Dosen mit respektloser Ablehnung oder gar Hass konfrontiert werden. Daher sollten sich künftige Studien insbesondere der Frage nach den langfristigen und kumulativen Effekten von inzivilen Nutzerkommentaren widmen. Zweitens muss bei der Ergebnisinterpretation berücksichtigt werden, dass in der vorliegenden Untersuchung radikale Beispiele von Hasskommentaren verwendet wurden. Viele Onlinediskussionsplattformen folgen jedoch der Praxis, solche Beiträge zu löschen (Singer et al., 2011), sodass viele Nutzende diese womöglich gar nicht erst sehen. Dem entgegen steht jedoch, dass 75 Prozent der Studienteilnehmenden angegeben haben, ähnlichen Beiträgen bereits in Nutzerdiskussionen begegnet zu sein. Dies spricht für eine gewisse externe Validität der Ergebnisse. Schließlich muss auf die geringe Effektstärke der hier dargestellten Befunde hingewiesen werden. Eine Erklärung hierfür könnte in der

heterogenen Stichprobe der Studie liegen. Diese könnte ein zusätzliches Rauschen in den Daten verursacht haben. Tatsächlich sind die Effektstärken der vorliegenden Studie vergleichbar mit anderen experimentellen Untersuchungen zu Online-Inzivilität, die mit ähnlich heterogenen Stichproben gearbeitet haben (Anderson et al., 2014). Abseits der methodologischen Ursachen könnten die kleinen Effektstärken aber auch einen inhaltlichen Befund indizieren – nämlich, dass Nutzende nicht blind die Positionen aus einem Hasskommentar übernehmen, sondern vielmehr interne Schutzmechanismen anwenden, um dem Hass zu widerstehen.

Trotz der Limitationen liefert die vorliegende Studie differenzierte Einblicke in die Wirkung von negativen und inzivilen Hasskommentaren gegen soziale Minderheiten auf das reale prosoziale Verhalten von Leserinnen und Lesern. Insbesondere die Theory of Planned Behavior sowie die Self-Concept Maintenance Theorie haben sich dabei als solide Rahmen erwiesen, um differenzierte Verhaltenskonsequenzen aus der Rezeption von Hasskommentaren abzuleiten. Der Befund, dass zivile Kritik keine negativen Effekte auf die Einstellungen und das prosoziale Verhalten der Nutzenden hatte, spricht dafür, dass kritische öffentliche Debatten und abweichende Meinungen keine Gefahr für das deliberative Ideal darstellen müssen. Dass Hasskommentare die Einstellungen und das prosoziale Verhalten der Nutzenden beeinträchtigen und gleichzeitig zur Polarisierung von Intergruppenbeziehungen beitragen können, stellt die Politik und den Journalismus jedoch vor die Herausforderung, diesen schädlichen Effekten zu begegnen. Insbesondere Gesellschaften, in denen Werte wie Diversität, Toleranz und gegenseitige Solidarität kultiviert werden, sollten sich hasserfüllten Kommentaren mit alternativen, weniger aggressiven und weniger stereotypisierten Positionen entgegenstellen.

*Christina Köhler, M.A.* ist wissenschaftliche Mitarbeiterin am am Institut für Publizistik der Johannes Gutenberg-Universität Mainz

*Dr. Marc Ziegele* ist Juniorprofessor für politische Online-Kommunikation am Institut für Sozialwissenschaften der Heinrich-Heine-Universität Düsseldorf.

*Dr. Mathias Weber* ist wissenschaftlicher Mitarbeiter am Institut für Publizistik der Johannes Gutenberg-Universität Mainz

## Quellenverzeichnis

- Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl (Hrsg.), *Springer series in social psychology. Action control. From cognition to behavior* (S. 11–39). Berlin: Springer.
- Anderson, A. A., Brossard, D., Scheufele, D. A., Xenos, M. A., & Ladwig, P. (2014). The „nasty effect“: Online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3), 373–387. doi:10.1111/jcc4.12009
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behavior: A meta-analytic reviews. *British Journal of Social Psychology*, 40, 471–499. doi:10.1348/014466601164939
- Blom, R., Carpenter, S., Bowe, B. J., & Lange, R. (2014). Frequent contributors within U.S. newspaper comment forums: An examination of their civility and information value. *American Behavioral Scientist*, 58(10), 1314–1328. doi:10.1177/0002764214527094
- Borah, P. (2014). Does it matter where you read the news story? Interaction of incivility and news frames in the political blogosphere. *Communication Research*, 41(6), 809–827. doi:10.1177/0093650212449353
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3), 429–444. doi:10.1111/0022-4537.00126
- Coe, K., Kenski, K., & Rains, S. A. (2014). Online and uncivil? Patterns and determinants of incivility in newspaper website comments. *Journal of Communication*, 64(4), 658–679. doi:10.1111/jcom.12104
- Darda, J. (2017, 10. Juli). Trump taps a long U.S. cultural history of pitting immigrants against veterans. *Dallas News*. Abgerufen von <https://www.dallasnews.com>
- Fishbein, M., & Ajzen, I. (2010). *Predicting and changing behavior: The reasoned action approach*. New York: Psychology Press.
- France, J. L., France, C. R., & Himawan, L. K. (2007). A path analysis of intention to redonate among experienced blood donors: An extension of the theory of planned behavior. *Transfusion*, 47(6), 1006–1013. doi:10.1111/j.1537-2995.2007.01236.x
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering Online Hate Speech*. Paris: United Nations Educational, Scientific and Cultural Organisation (UNESCO). Abgerufen von <http://unesdoc.unesco.org/images/0023/002332/233231e.pdf>

- Gervais, B. T. (2014). Incivility online: Affective and behavioral reactions to uncivil political posts in a web-based experiment. *Journal of Information Technology & Politics*, 12(2), 167–185. doi:10.1080/19331681.2014.997416
- Gilliam, F. D., & Iyengar, S. (2000). Prime suspects: The influence of local television news on the viewing public. *American Journal of Political Science*, 44(3), 560. doi:10.2307/2669264
- Hsueh, M., Yogeewaran, 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. doi:10.1111/hcre.12059
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6, 1–55. doi: 10.1080/10705519909540118
- Hwang, H., Kim, Y., & Kim, Y. (2016). Influence of discussion incivility on deliberation: An examination of the mediating role of moral indignation. *Communication Research*. doi:10.1080/08838151.2014.966365
- Lee, E.-J., & Jang, Y. J. (2010). What do others' reactions to news on internet portal sites tell us? Effects of presentation format and readers' need for cognition on reality perception. *Communication Research*, 37, 825–846. doi:10.1177/0093650210376189
- Landesanstalt für Medien NRW (2017). *Hate speech*. Abgerufen von <https://goo.gl/78unIs>
- Mazar, N., Amir, O., & Ariely, D. (2008). The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, 45(6), 633–644. doi:10.1509/jmkr.45.6.633
- Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology*, 64(5), 723–739.
- Newman, N., Fletcher, R., Levy, D. A. L., & Nielsen, R. K. (2016). *Reuters Institute Digital News Report 2016*. Abgerufen von <https://reutersinstitute.politics.ox.ac.uk/sites/default/files/Digital-News-Report-2016.pdf>
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(6), 259–283. doi:10.1177/1461444804041444

- Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Berlin, New York: Springer.
- Petty, R. E., & Cacioppo, J. T. (1990). Involvement and persuasion: Tradition versus integration. *Psychological Bulletin*, *107*, 367–374.  
doi:10.1037//0033-2909.107.3.367
- Prochazka, F., Weber, P., & Schweiger, W. (2016). Effects of civility and reasoning in user comments on perceived journalistic quality. *Journalism Studies*, 1–17. doi:10.1080/1461670X.2016.1161497
- Roskos-Ewoldsen, D. R., Klinger, M. R., & Roskos-Ewoldsen, B. (2007). Media priming: A meta-analysis. In R. W. Preiss (Hrsg.), *LEA's communication series. Mass media effects research. Advances through meta-analysis* (S. 53–80). New York: Lawrence Erlbaum.
- Rösner, L., Winter, S., & Krämer, N. C. (2016). Dangerous minds? Effects of uncivil online comments on aggressive cognitions, emotions, and behavior. *Computers in Human Behavior*, *58*, 461–470.  
doi:10.1177/2056305116664220
- Rowe, I. (2014). Civility 2.0: A comparative analysis of incivility in online political discussion. *Information, Communication & Society*, *18*(2), 121–138.  
doi:10.1080/1369118X.2014.940365
- Rowe, I. (2015). Deliberation 2.0: Comparing the deliberative quality of online news user comments across platforms. *Journal of Broadcasting & Electronic Media*, *59*(4), 539–555. doi:10.1080/08838151.2015.1093482
- Singer, J. B., Hermida, A., Domingo, D., Heinonen, A., Paulussen, S., Quandt, T., Reich, Z., & Vujnovic, M. (Hrsg.). (2011). *Participatory journalism: Guarding open gates at online newspapers*. Malden, MA: Wiley-Blackwell.
- Smith, J. R., & McSweeney, A. (2007). Charitable giving: The effectiveness of a revised theory of planned behaviour model in predicting donating intentions and behaviour. *Journal of Community & Applied Social Psychology*, *17*(5), 363–386. doi:10.1002/casp.906
- Springer, N., Engelmann, I., & Pfaffinger, C. (2015). User Comments: Motives and Inhibitors to Write and Read. *Information, Communication & Society*, *18*(7), 798–815. doi:10.1080/1369118X.2014.997268
- Stroud, N. J., Scacco, J. M., Muddiman, A., & Curry, A. L. (2015). Changing deliberative norms on news organizations' Facebook sites. *Journal of Computer-Mediated Communication*, *20*(2), 188–203. doi:10.1111/jcc4.12104

- Topinka, R. J. (2017). Politically incorrect participatory media: Racist nationalism on r/ImGoingToHellForThis. *New Media & Society*. Online first. doi:10.1177/1461444817712516
- Walther, J. B. (2011). Theories of computer-mediated communication and interpersonal relations. In M. L. Knapp & J. A. Daly (Hrsg.), *The Sage handbook of interpersonal communication* (4. Auflage, S. 443–479). Los Angeles: Sage.
- Walther, J. B., DeAndrea, D., Kim, J., & Anthony, J. C. (2010). The influence of online comments on perceptions of Antimarijuana Public Service Announcements on YouTube. *Human Communication Research*, 36(4), 469–492. doi:10.1111/j.1468-2958.2010.01384.x
- Welsh, D. T., & Ordonez, L. D. (2014). Conscience without cognition: The effects of subconscious priming on ethical behavior. *Academy of Management Journal*, 57(3), 723–742. doi:10.5465/amj.2011.1009
- Winter, S., Brückner, C., & Krämer, N. C. (2015). They Came, They Liked, They Commented: Social Influence on Facebook News Channels. *Cyberpsychology, Behavior And Social Networking*, 18(8), 431–436. doi:10.1089/cyber.2015.0005
- Winter, S., & Krämer, N. C. (2016). Who's right: The author or the audience? Effects of user comments and ratings on the perception of online science articles. *Communications*, 41(3), 339–360. doi:10.1515/commun-2016-0008
- Yun, G. W., & Park, S.-Y. (2011). Selective posting: Willingness to post a message online. *Journal of Computer-Mediated Communication*, 16, 201–227. doi:10.1111/j.1083-6101.2010.01533.x
- Ziegele, M., & Jost, P. B. (2016). Not funny? The effects of factual versus sarcastic journalistic responses to uncivil user comments. *Communication Research*, 1–30. doi:10.1177/0093650216671854
- Ziegele, M., & Quiring, O. (2017, Mai). *The discussion value of online news: How news story characteristics affect the deliberative quality of user discussions in SNS comment sections*. Vortrag auf der Jahrestagung der International Communication Association (ICA), San Diego, USA.
- Ziegele, M., Weber, M., Quiring, O. & Breiner, T. (2017). The dynamics of online news discussions: Effects of news articles and reader comments on users' involvement, willingness to participate, and the civility of their contributions. *Information, Communication & Society* 7, 1–17. doi:10.1080/1369118X.2017.1324505