

"Weg mit dem Ungeziefer!": Hasskommentare aus Sicht der Nutzenden; Bewertung und Meldeverhalten

Wilhelm, Claudia; Ziegler, Isabell; Jöckel, Sven

Erstveröffentlichung / Primary Publication

Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Wilhelm, C., Ziegler, I., & Jöckel, S. (2019). "Weg mit dem Ungeziefer!": Hasskommentare aus Sicht der Nutzenden; Bewertung und Meldeverhalten. In I. Engelmann, M. Legrand, & H. Marzinkowski (Hrsg.), *Politische Partizipation im Medienwandel* (S. 277-297). Berlin <https://doi.org/10.17174/dcr.v6.11>

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: Wilhelm, C., Ziegler, I., & Jöckel, S. (2019). „Weg mit dem Ungeziefer!“ Hasskommentare aus Sicht der Nutzenden: Bewertung und Meldeverhalten. In I. Engelmann, M. Legrand, & H. Marzinkowski (Hrsg.), *Politische Partizipation im Medienwandel* (S. 277-297). doi: 10.17174/dcr.v6.11

Zusammenfassung: Hasskommentare in sozialen Medien und in den Kommentarbereichen von Nachrichtenwebseiten haben zu einer Debatte über inzivile Formen politischer Online-Partizipation geführt. Im vorliegenden Beitrag untersuchen wir, wie Devianzmerkmale und Neutralisierungstechniken in potenziellen Hasskommentaren die Bewertung und das Melden solcher Kommentare durch die Nutzenden beeinflussen. Dabei unterscheiden wir Kommentare nach Art der Normverletzung, Subtilität, Opfertyp und dem Vorhandensein von Rechtfertigungen oder Entschuldigungen. Ergebnisse einer Online-Befragung (n = 457), die ein wahlbasiertes Conjoint-Design beinhaltete, legen nahe, dass Gewaltaufrufe mit höherer Wahrscheinlichkeit gemeldet werden als Verschwörungstheorien, Gerüchte, Beleidigungen und Agitation. Direkte Normverstöße wurden häufiger gemeldet als indirekte Normverletzungen. Hasskommentare wurden eher akzeptiert, wenn sie eine Rechtfertigung oder Entschuldigung enthielten. Der Opfertyp hatte keinen Einfluss auf die Bewertung und das Melden von Hasskommentaren.

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

Claudia Wilhelm, Isabell Ziegler & Sven Jöckel

„Weg mit dem Ungeziefer!“

Hasskommentare aus Sicht der Nutzenden:
Bewertung und Meldeverhalten

1 Einleitung

„Hoffentlich werden diese Antifa-Schlampen, von ihren muslimischen Flüchtlings-Proleten vergewaltigt!“ (sic). Wer regelmäßig in Kommentarbereichen und Sozialen Netzwerken unterwegs ist, kommt nicht an solchen (mutmaßlich) norm-verletzenden Aussagen vorbei. In akademischer Forschung und journalistischer Praxis werden solche Aussagen als „Hasskommentare“ bezeichnet, die sich meist gegen eine bestimmte Gruppe aber auch Einzelpersonen richten. Laut einer Forsa-Studie (Forsa, 2016) haben zwei Drittel der befragten Nutzenden schon einmal einen Hasskommentar im Internet gelesen. Spätestens im Zuge der 2015 einsetzenden Debatte um Flucht und Asyl ist auch ins Bewusstsein der breiten Öffentlichkeit gedrungen, wie virulent menschenfeindliche Äußerungen im Internet sind. Neben Diffamierungen enthalten solche Kommentare auch explizite Aufrufe zur Gewalt vor allem gegenüber Geflüchteten, Menschen aus dem Ausland oder islamischen Glaubens (Hanzelka & Schmidt, 2017; Santana, 2015). Eine Analyse der Nutzerkommentare auf der Webseite der britischen Zeitung *The Guardian* zeigt, dass Menschen mit Immigrationshintergrund eine

der am häufigsten angefeindeten Personengruppe stellen (Gardiner, Mansfield, Anderson, Holder, Louter, & Ulmanu, 2016).

Die vermehrte Sichtbarkeit inziviler und devianter Beiträge hat den Ruf nach Regulierung befördert. Die Politik hat mittlerweile in Deutschland mit einer Gesetzesverschärfung zu Hasskommentaren in Internet reagiert (Reinbold, 2017) und treibt somit zur Eindämmung des Phänomens die Strafverfolgung voran bzw. nimmt dazu auch die Plattformanbieter stärker in die Pflicht. Diese setzen auf die Bereitschaft ihrer Nutzenden, inzivile Inhalte zu melden, jedoch fallen die Reaktionen der Nutzenden unterschiedlich aus: Neben der Möglichkeit den Kommentar zu melden, reagieren andere Nutzende auch mit Gegenrede und ein geringer Teil sogar mit Unterstützung des Hasskommentars (Forsa, 2016). Das Melden von Hasskommentaren setzt voraus, dass die Nutzenden über ein gewisses Verständnis für den Grenzverlauf zwischen Meinungsäußerung und Hassrede verfügen (Leets, 2001; Samaratinge & Hattotuwa, 2014).

Höflichkeit und Zivilität gelten als maßgebliche Kriterien der Qualität politischer (Online-)Partizipation, sei es wie aktuell in den Sozialen Medien (Anderson, Brossard, Scheufele, Xenos, & Ladwig, 2014; Coe, Kenski, & Rains, 2014; Sobieraj & Berry, 2011) oder wie – schon länger – im Internet allgemein (Papacharissi, 2004): Als unhöflich charakterisierte Kommentare beinhalten Beschimpfungen, Verleumdungen, Übertreibungen, abwertende Formulierungen, vulgäre Sprache oder Sarkasmus (Papacharissi, 2004) sowie mangelhafte Umgangs- und Ausdrucksformen (Brown & Levinson, 1987). Inzivilität bezieht sich auf den Verstoß gegen Standards des demokratischen Diskurses. Gemeint sind Verhaltensweisen, die persönliche Freiheit und Demokratie bedrohen und zur Stereotypisierung sozialer Gruppen beitragen (Papacharissi, 2004). Während inzivile Ausdrucksformen teilweise von Journalistinnen und Journalisten oder Nutzenden geduldet werden, wenn sie mit Nachrichtenwerten und journalistischen Normen in Einklang stehen oder der Kommentar die Meinung anderer Nutzenden wiedergibt (Muddiman & Stroud, 2017), wird die Wirkung von Hasskommentaren als gewaltfördernd und diskriminierend beschrieben (Gagliardone, Gal, Alves, & Martinez, 2015). Normverletzungen in Hasskommentaren gehen demzufolge über Inzivilität hinaus. Der Begriff der Devianz greift sowohl bei schlicht unhöflichen, inzivilen Kommentaren als auch bei den darüber hinaus gehenden Hasskommentaren. Devianz meint jeglichen Verstoß gegen soziale Normen des jeweiligen Bezugssystems (Dyade, Gruppe, Kollektiv, Gesellschaft), der potentiell Schaden anrichtet

und negative Sanktionen nach sich ziehen kann (Cohen, 1968; Opp, 1974). Soziale Devianz wird definiert als Differenz bzw. Abweichung von der Norm, die von der Gesellschaft als moralisch verwerflich und bedrohlich bewertet wird (Henry, 2009). Hasskommentare sind demnach Akte devianter Kommunikation, die soziale oder gesellschaftliche Regeln und Normen – in der alltäglichen zwischenmenschlichen Kommunikation wie auch in Foren im Internet – verletzen (Henry, 2009) und dabei Vorurteile, Diskriminierung und Gewalt gegenüber sozialen Gruppen und deren Mitgliedern begünstigen (Gagliardone et al., 2015).

Doch nicht immer sind Kommentare wie im eingangs zitierten Beispiel so offensichtlich normverletzend formuliert, sondern verschleiert und indirekt. Wann wird also ein Kommentar als deviant bewertet und wann nicht? Der vorliegende Beitrag beschäftigt sich zum einen mit der Frage, wie Merkmale sogenannter Hasskommentare deren Bewertung durch die Nutzenden beeinflussen. Des Weiteren interessiert uns die Wirkung von Strategien der Verschleierung in devianten Kommentaren. Dabei greifen wir auf die Neutralisierungstheorie von Sykes und Matza (1957) zurück, die davon ausgeht, dass deviant handelnde Personen mit Hilfe von Rechtfertigungen und Entschuldigungen versuchen ihre Regelverletzungen zu relativieren und soziale bzw. gesellschaftliche Akzeptanz zu erlangen. Im Fokus dieses Beitrags steht die Frage, inwiefern Rechtfertigungen und Entschuldigungen Normverstöße relativieren können und so die Akzeptanz devianter Kommentare unterstützen. In methodischer Hinsicht untersucht der vorliegende Beitrag mittels wahlbasierter Conjoint-Analyse, welche Formen von Hasskommentaren Nutzende als deviant und meldewürdig wahrnehmen und welche Merkmale der Kommentare ihre Beurteilung beeinflussen.

2 Melden von devianten Kommentaren

Für die nutzerseitige Regulierung von Inhalten stehen zur sozialen Sanktionierung, neben bspw. Counter Speech oder dem Weiterleiten von Kommentaren an sogenannte „Pranger-Seiten“, auf den meisten Social Media-Plattformen technische Optionen zum Melden von normverletzenden Äußerungen zur Verfügung. Mittels Melde-Option können Nutzende die Plattformanbieter über Beiträge informieren, in denen sie einen Normverstoß wahrnehmen. Manche Plattformen (z. B. YouTube, Twitter) werden erst aktiv, wenn sie von Nutzenden auf kritische Inhalte

aufmerksam gemacht werden, andere nutzen die Melde-Option in Kombination mit der eigenen proaktiven Überprüfung von Inhalten (z. B. Facebook; Crawford & Gillespie, 2014). Darüber hinaus soll interaktive Moderation in Kommentarbereichen von Nachrichtenseiten, z. B. indem sich Journalistinnen und Journalisten in die Diskussion einbringen und sich für einen respektvollen Umgang einsetzen, die Qualität von Online-Diskussionen verbessern (Springer, Engelmann, & Pfaffinger, 2015; Stroud, Scacco, Muddiman, & Curry, 2015; Ziegele & Jost, 2016). Dabei sind unterschiedliche Auffassungen zwischen (multinationalen) Plattformanbietern und einem diversen Publikum wie auch zwischen unterschiedlichen Nutzergruppen darüber unvermeidbar, welche Inhalte akzeptable Meinungsäußerung darstellen und welche Anstoß erregen (Crawford & Gillespie, 2014). Die Forschung hat bisher vor allem auf die Wahrnehmung der Online-Debattenkultur durch Journalistinnen und Journalisten (z. B. Bergström & Wadbring, 2015; Vobič, 2014) oder Politikerinnen und Politiker fokussiert – wenig bekannt ist, wie die Nutzenden Hasskommentare wahrnehmen (Kenski, Coe, & Rains, 2017).

2.1 *Formen und Dimensionen von Devianz in Nutzerkommentaren*

Allgemein betrachtet verstoßen Hasskommentare als Form devianter Kommunikation gegen grundlegende Werte und Normen des sozialen Miteinanders. Häufig werden in ihnen Personen aufgrund ihrer Zugehörigkeit zu einer bestimmten sozialen Gruppe in mehr oder weniger extremer Form abgewertet. Monitoring-Berichte zu Nutzerkommentaren auf Nachrichtenseiten und in sozialen Medien zeigen, dass sich die Kommentare hinsichtlich der Art der Normverletzung unterscheiden (Amadeu Antonio Stiftung, 2016a). In Hasskommentaren finden sich Herabsetzungen (z. B. „Sozialschmarotzer“) und Aufrufe zu Taten (z. B. „Moscheen gehören abgefackelt“) oder es wird gegen Personen oder Gruppen agitiert (z. B. „Ausländer müssen draußen bleiben“) (Amadeu Antonio Stiftung, 2015; 2016a). Hasskommentare gegen Menschen mit Immigrationshintergrund oder islamischen Glaubens und Geflüchtete enthalten beleidigende, inzivile Aussagen, aber auch explizite Gewaltaufrufe (Hanzelka & Schmidt, 2017). Herabsetzende Kommentare beinhalten verbale Beschimpfungen und Beleidigungen von Personen oder sozialen Gruppen (Amadeu Antonio Stiftung, 2016b; Coe et al., 2014). Aufrufe zu Taten beinhalten Gewaltandrohungen und

Anstiftung zu Gewalt (Hanzelka & Schmidt, 2017). Des Weiteren können sich Formen politischer Agitation und Hetze gegen soziale Gruppen wie Geflüchtete richten und stellen somit eine weitere Art der Normverletzung dar, da sie eine feindliche und aggressive Stimmung schaffen (Meibauer, 2013). Studien zur Verbreitung von Hasskommentaren zeigen, dass Nutzende sich auch Gerüchten und Verschwörungstheorien bedienen, um eine feindselige Atmosphäre zu schüren (Antonio Amadeu Stiftung, 2016a; Mocanu, Rossi, Zhang, Karsai, & Quattrociocchi, 2015; Stefanowitsch & Flach, 2016). Das Verbreiten nicht bestätigter oder unwahrer Gerüchte wird als grobes politisches Fehlverhalten bewertet (Muddiman, 2013). Im Vergleich zu Gewaltaufrufen oder Agitation wirkt die in Gerüchten beschworene Gefahr konkreter, hat meist einen lokalen Bezug, legt somit eine unmittelbare Bedrohung nahe und ist emotional zugänglicher (Amadeu Antonio Stiftung, 2016a). Verschwörungstheorien erklären bedeutsame soziale oder politische Ereignisse wie z. B. die Flüchtlingskrise oder generelle politische Praktiken als von machtvollen Eliten inszeniert, um damit ihre Führungsrolle zu festigen und die von ihnen angestrebten Ziele zu erreichen (Sunstein & Vermeule, 2009). Gerüchte und Verschwörungstheorien sind in sozialen Medien weit verbreitet und Nutzende „alternativer“ Medien neigen im besonderem Maße dazu diese weiterzuverbreiten (Mocanu et al., 2015). Insgesamt lassen sich also fünf Arten von Normverletzungen unterscheiden (siehe Tabelle 1).

Tabelle 1: Arten der Normverletzung in Hasskommentaren

<i>Art der Normverletzung</i>	<i>Definition</i>	<i>Beispiel</i>
Herabsetzung	Aussagen, die Beschimpfungen und Verunglimpfungen enthalten (Antonio Amadeu Stiftung, 2016b) oder abwertend formuliert sind (Coe et al., 2014)	„Asylanten sind undankbare Bastarde.“
Aufruf zu Taten	Aussagen, die Gewalt androhen oder dazu aufrufen (Hanzelka & Schmidt, 2017)	„Heime sollen brennen!“

Agitation	Provozierende Aussagen zu politischen oder sozialen Problemen zur Steuerung des öffentlichen Interesses, die dabei häufig eine diskriminierende und vorurteilsbehaftete Perspektive einnehmen (“The American heritage dictionary of the English language“, 2000)	„Ausländer müssen draußen bleiben.“
Gerücht	Nicht bestätigte Behauptung, die sich gegen soziale, kulturelle oder ethnische Gruppen bzw. deren Mitglieder richtet (Bhavnani, Findley, & Kuklinski, 2009)	„Flüchtlinge räumen REWE leer und sagen zur Kassiererin: ‚Das bezahlt Deutschland!‘.“
Verschwörungstheorie	Aussagen, die die Vorstellung befördern, Eliten seien böse, z. B. Regierungen und große Firmen verdächtigen (Campion-Vincent, 2009; Sunstein & Vermeule, 2009)	„Flüchtlinge werden von europäischen Regierungen gezielt hierhergeholt, um die Islamisierung des Abendlandes zu erreichen, bezahlt von islamischen Mächten im Hintergrund.“

Anmerkung: Kommentare sind entnommen aus Amadeu Antonio Stiftung (2016a; 2016b), Meibauer (2013) sowie Stefanowitsch und Flach (2016; auch Wilhelm, Ziegler, & Jöckel, im Druck).

Wie werden diese von den Nutzenden wahrgenommen? Auf Basis der vorgenommenen Systematisierung wird untersucht, ob und in welchem Ausmaß die Art des Normverstoßes die Beurteilung der Devianz des Kommentars beeinflusst.

FF1: Wie beeinflusst die Art des Normverstoßes (a) die Beurteilung der Devianz des Kommentars und (b) das Meldeverhalten?

Nicht immer sind deviante Aussagen so gestaltet, dass der Normverstoß offensichtlich ist und von allen Beteiligten entdeckt werden kann (Meibauer, 2013). Kommentare unterscheiden sich also in der Subtilität der in ihnen getätigten Aussagen. Direkte Verstöße werden zwar eher als deviant erkannt als indirekte, jedoch können auch indirekte, versteckte Gewaltaufrufe deviantes Verhalten begünstigen (Leets, 2001). Indirekte Diffamierungen werden in Hasskommentaren strategisch genutzt, um ein Melden oder Löschen des Kommentars zu verhindern (Meibauer, 2013). Je expliziter Hassbotschaften formuliert sind, desto eher werden sie abgelehnt (Lee & Leets, 2002). Die Subtilität der Normverletzung ist dementsprechend ein weiteres relevantes Merkmal der Devianz von Hasskommentaren, welche die Devianzbewertung und das Meldeverhalten beeinflusst.

H1: Direkte Normverletzungen werden als (a) devianter bewertet und (b) rufen eine höhere Meldebereitschaft hervor als indirekte Normverletzungen.

Deviante Kommentare richten sich außerdem oft gegen abstrakte Gruppen oder Institutionen (Reich 2011), im Zuge der Fluchtdebatte z. B. gegen „die Flüchtlinge“ bzw. „den Islam“ oder „das Gebilde EU“. Im Gegenzug kommt es jedoch auch zu Normverstößen, wenn es um konkrete Personen geht (Bär, 2008; zitiert nach Springer, 2014), also um andere Kommentierende, um die Autoren und Autorinnen von journalistischen Beiträgen oder um Politikerinnen und Politiker. Hasskommentare, die sich gegen konkrete Personen richten, werden als gefährlicher eingeschätzt (Citron & Norton, 2011). Abstraktheit reduziert in diesem Zusammenhang die wahrgenommene Nähe und Bedeutung des Opfers (Oliver, Dillard, Bae, & Tamul, 2012) und erleichtert eine moralische Distanzierung (Bandura, 1999).

H2: Kommentare, die sich gegen ein konkretes Opfer richten, werden als (a) devianter bewertet und rufen (b) eine höhere Meldebereitschaft hervor als Kommentare gegen abstrakte Opfer.

Deviante Kommentare lassen sich somit nach folgenden Kriterien systematisieren (siehe Tabelle 2): nach Art des Normverstoßes (Hanzelka & Schmidt, 2017),

nach Subtilität der Botschaft (Meibauer, 2013) und ob eine konkrete Person Ziel des Angriffs ist oder eine abstrakte Gruppe oder Institution (Opfertyp; Naab, Kalch, & Meitz, 2017). Wobei sich – wie in Tabelle 2 aufgezeigt – auch meist Kombinationen dieser Merkmale finden, also z. B. direkte Kommentare mit abstraktem Opfertyp und einem Aufruf zu Taten. Gerade solche Merkmalskombinationen machen es schwierig zu ermitteln, welchen Beitrag jedes einzelne Merkmal letztlich für die wahrgenommene Devianz leistet.

Tabelle 2: Merkmale devianter Kommentare

<i>Art des Normverstoßes</i>	<i>Opfertyp: abstrakt</i>	<i>Subtilität:</i>		
		<i>individualisiert</i>	<i>direkt</i>	<i>indirekt</i>
Herabsetzung	X	X	X	X
Aufruf zu Taten	X	X	X	X
Gefährliche Agitation	X	X	X	X
Verschwörungstheorie	X	X	–	X
Gerüchte	X	X	–	X

2.2 Neutralisierung von Devianz in Online-Kommentaren

Die Neutralisierungstheorie (Sykes & Matza, 1957) geht davon aus, dass deviantes Verhalten begünstigt wird, wenn die Delinquentin oder der Delinquent in der Lage ist, Rechtfertigungen oder Entschuldigungen zu verbalisieren. Diese Relativierungen des devianten Verhaltens werden auch als Neutralisierungstechniken bezeichnet (Henry, 2009; Maruna & Copes, 2005). Beispielsweise gesteht die Person sich zwar ein, einen Normverstoß begangen zu haben, streitet jedoch ab, in negativer Absicht gehandelt zu haben („war doch nur Spaß“). Weitere Formen der Neutralisierung bestehen u. a. darin, schlimmere Vergehen zur Relativierung der Devianz des eigenen Verhaltens anzuführen oder den Opferstatus der Betroffenen in Frage zu stellen. Der Erfolg solcher Neutralisierungstechniken im Sinne einer Etablierung oder Aufwertung des oder der deviant Handelnden, hängt auch von der Akzeptanz der Relativierungen durch das jeweilige Umfeld

und relevante Institutionen (Schulen, Gerichte) ab. Die Forschung zeigt hier verschiedene Formen der Neutralisierung auf, die sich auch bei der Rechtfertigung unterschiedlicher Formen devianter Mediennutzung wie beispielsweise bei digitaler Piraterie finden (z. B. Hinduja, 2007; Malin & Fowers, 2009; Morris & Higgins, 2009). So wird die Devianz des Verhaltens zwar eingestanden, die negative Absicht jedoch abgestritten („just having fun“, „just a joke“) (Denial of Negative Intent), es wird argumentiert, dass das Opfer eigentlich als Täter anzusehen ist und es verdient hat (Henry, 2009). Die Tat wird als ausgleichende Gerechtigkeit angesehen (Denial of Victim). Vergleichbar werden vermeintlich schlechte Taten durch ebensolche guten Taten relativiert (Metaphor of the Ledger; z. B. Harris & Dumas, 2009) oder die Devianz des Verhaltens generell wird in Frage gestellt („Everyone’s doing it!“), indem das Verhalten normalisiert wird (Claim of Normality; Harris & Dumas, 2009; Henry, 2009; Maruna & Copes, 2005). Übertragen auf das Verfassen von Hasskommentaren interessiert nicht, welche dieser Neutralisierungstechniken am erfolgreichsten erscheint. Dies scheint vom konkreten Kontext abhängig zu sein (Maruna & Copes, 2005). Was jedoch im Fokus des Interesses steht, ist die Frage, inwieweit die Verwendung solcher Techniken per se zu Veränderungen in der Wahrnehmung des Devianzgrades führt. Wir können dabei folgende Hypothesen ableiten:

H3: *Die Verwendung von Neutralisierungstechniken führt dazu, dass vermeintliche Hasskommentare a) als weniger deviant wahrgenommen werden und b) weniger häufig gemeldet werden als Kommentare ohne Neutralisierungstechnik.*

3 Methode

Wir überprüfen unsere Hypothesen im Rahmen einer Online-Befragung, die über verschiedene Facebook-Gruppen im Juli/August 2016 verteilt wurden. Ein Großteil der Teilnehmenden rekrutierte sich dabei aus einem Forenbeitrag des NDR Magazins Panorama auf Facebook über Hasskommentare im Netz. Wir gehen also von einer am Thema interessierten Gruppe aus und nicht von einer generalisierbaren Online-Population. Es nahmen insgesamt 457 Personen an der Studie teil ($M_{\text{Alter}} = 37,40$, $SD = 12,67$; 52% weiblich). Auf Basis der Rekrutierung ist die Stichprobe verzerrt hin zu Personen mit einem höheren Bildungsabschluss

(50%, $n = 223$) mit Universitätsabschluss gegenüber Personen mit Haupt-/Realschulabschluss (21%, $n = 94$). Wie bei einem solchen Thema zu erwarten – dem Melden von Hasskommentaren gegenüber Geflüchteten und einer Asyl-positiven Politik – zeichnet sich die Stichprobe ferner durch eine eher linksgerichtete politische Einstellung aus ($M = 1,94$, $SD = 0,71$ auf einer 7er-Skala von links (1) zu rechts (7); Inglehart & Klingemann, 1976).

Die zentrale methodische Besonderheit der empirischen Studie stellt der relative Vergleich verschiedener Hasskommentarformen dar. Dazu wurden im Vorfeld 48 Kommentare konzipiert, die sich im Hinblick auf die Art des Normverstoßes, des Opfertyps und der Subtilität sowie in Bezug auf den Einsatz von Neutralisierungstechniken unterschieden. Die Kommentare wurden aus vorhandenen Forenbeiträgen mit den entsprechenden sprachlichen Eigenheiten (Grammatik, Tippfehler) übernommen, jedoch für den eigenen Forschungszweck angepasst. Die vermeintlichen Normverletzungen beinhalteten Äußerungen zu Geflüchteten, Flüchtlingspolitik und Migration. Tabelle 3 gibt einen Überblick verschiedener Kommentare und ihrer jeweiligen Merkmalskombination. Die 48 Kommentare wurden im Rahmen eines Pretests mit $n = 24$ Studierenden getestet und jeweils von zwei Kommentaren derjenige für die eigentliche Studie ausgewählt, der deutlicher die jeweils intendierten Merkmale aufwies. Es wurden insgesamt 24 Kommentare in der eigentlichen Studie verwendet. Diese wurden dann im Stil eines wahlbasierten Conjoint-Designs im Fragebogen präsentiert: Den Studienteilnehmenden wurden randomisiert jeweils sechs von insgesamt 12 Kommentarpaaren als sog. Choice-Sets präsentiert.

Die Beurteilung der Devianz der Kommentare wurde anhand zweier Indikatoren ermittelt: In jedem Choice-Set wurden die Teilnehmenden gebeten, anzugeben, welchen der beiden Kommentare sie als „schlimmer“ bewerten und welchen sie melden würden („Kommentar 1 melden“, „Kommentar 2 melden“, „Beide Kommentare melden“ und „keinen der beiden Kommentare melden“). Der Vorteil eines solchen Designs liegt darin, dass die Kommentare sowohl absolut als auch vor allem relativ bewertet wurden.

Wichtig ist an dieser Stelle festzuhalten, dass das Melden eines Kommentars hier ein relativ abstraktes Unterfangen ist und keinerlei realweltliche Konsequenzen nach sich zieht. Es ist demnach nicht Ziel der Studie, zu ermitteln ob ein Kommentar tatsächlich gemeldet wird oder nicht. Um dies zu erklären müssten neben den Eigenschaften des Kommentars auch vor allem Eigenschaften der

Tabelle 3: Beispiel-Stimuli und Merkmalskombinationen

<i>Kommentar</i>	<i>Art der Norm- verletzung</i>	<i>Subtilität</i>	<i>Opfertyp</i>	<i>Neutra- lisierung</i>
Du hirnamputierter Drecks- Gutmensch	Abwertung	direkt	indivi- duell	nein
wer die adresse von herrn mustafa e. kennt, weiß was passieren muss ... Keine sorge, das mein ich nicht ernst...	Aufruf zu Taten	indirekt	indivi- duell	ja
Frau Merkel will mit ihrer Flüchtlingspolitik einen Bür- gerkrieg heraufbeschwören, um den Volkstod herbeizu- führen.	Verschwö- rungstheorie	indirekt	indivi- duell	nein
Deutschland den Deut- schen!! Grünenpolitiker cem özdemir raus aus Deutsch- land!	Agitation	direkt	indivi- duell	nein
Bislang hab ich ja echt nichts gegen Flüchtlinge, hab denen auch Wintersachen gespendet, aber eben hab ich gehört: Flüchtlinge räumen REWE leer und sagen zur Kassiererin: „Das bezahlt Deutschland!“.	Gerücht	indirekt	abstrakt	ja

Person als auch der Situation, des jeweiligen Forums und der dort herrschenden Netiquette mitberücksichtigt werden. Das Melden des Kommentars stellt in dieser Studie eine Art Proxy-Indikator für die wahrgenommene Devianz dar.

4 Ergebnisse

Die Conjointdaten wurden unter Anwendung einer geschichteten Cox-Regression (Backhaus, Erichson, & Weiber, 2013) analysiert, welche typischerweise bei Überlebensanalysen verwendet wird. Es handelt sich um eine Regression für Daten mit zensierten Beobachtungen. Das bedeutet, die unzensierten Fälle sind die im Choice-Set gewählten Alternativen. Jede Schicht repräsentiert eine Auswahl-situation. Die Regressionskoeffizienten B werden als Teilnutzenwerte interpretiert und zeigen demnach die Präferenz für eine Merkmalsausprägung in Relation zu allen anderen Merkmalsausprägungen an. $\text{Exp}(B)$ gibt das Chancenverhältnis (Odds), in der Überlebensanalyse Hazard Ratio genannt, also das Verhältnis Wahrscheinlichkeit/Gegenwahrscheinlichkeit, dass eine Alternative mit einer bestimmten Merkmalsausprägung gewählt wird, an.

4.1 Bewertung der Devianz des Kommentars

Zunächst prüfen wir, ob die Art des Normverstoßes einen signifikanten Einfluss auf die Bewertung der Devianz eines Kommentars hatte und welche Art der Normverletzung – relativ betrachtet – von den Befragten als devianter wahrgenommen wurde (FF1a). Tabelle 4 zeigt, dass es signifikante Unterschiede in der Devianzbewertung bezüglich der Art des Normverstoßes gibt. Aufruf zu Taten wurde in Relation zu den anderen vier Normverstößen als schlimmster Normverstoß bewertet. Enthielt ein Kommentar einen Aufruf zu Taten, war die Wahrscheinlichkeit, dass er als devianter wahrgenommen wurde etwa achtmal höher ($\text{Exp}(B) = 7,96$) als bei Kommentaren, die Agitation enthalten. Danach folgen Gerüchte und Verschwörungstheorien, die ebenfalls als deutlich stärker normverletzend bewertet werden als Herabsetzung und gefährliche Agitation.

In H1a nahmen wir an, dass direkte Normverstöße als devianter beurteilt werden als indirekte. Dies lässt sich anhand der Daten bestätigen: Bei direkten Normverstößen verdoppelte sich die Chance, dass sie als problematisch bewertet werden ($\text{Exp}(B) = 2,08$). Entgegen unserer Annahme in H2a hatte der Opfertyp keinen signifikanten Einfluss darauf, ob ein Kommentar als schlimmer bewertet wurde. Entsprechend kann H2a nicht bestätigt werden.

Des Weiteren wurde angenommen, dass ein Kommentar als weniger deviant beurteilt wird, wenn er neben Normverstößen auch Neutralisierungstechniken enthält (H3a). Dies konnte bestätigt werden: Enthielten Kommentare Rechtfertigungen oder Entschuldigungen, wurden sie als weniger problematisch bewertet als Kommentare ohne Neutralisierungstechniken.

Tabelle 4: Bewertung der Devianz des Kommentars

Art des Normverstoßes	B	SE	Exp(B)
Herabsetzung	0,03	0,09	1,03
Aufruf zu Taten	2,08***	0,11	7,96
Verschwörungstheorie	0,77***	0,12	2,17
Gerüchte	0,88***	0,13	2,40
gefährliche Agitation	als Referenzkategorie auf Null gesetzt		
<i>Subtilität</i>			
direkt	0,73***	0,09	2,08
indirekt	als Referenzkategorie auf Null gesetzt		
<i>Opfertyp</i>			
individualisiert	0,07	0,07	1,07
abstrakt	als Referenzkategorie auf Null gesetzt		
<i>Neutralisierung</i>			
ja	-0,33***	0,07	0,72
nein	als Referenzkategorie auf Null gesetzt		

Anmerkung: AV: „Welcher Kommentar ist schlimmer?“ LLR = 810,18; Cox&Snell = ,14; Nagelkerke = ,14; *** $p < ,001$.

4.2 Meldeverhalten

Neben der relationalen Devianzbewertung („Welcher Kommentar ist schlimmer?“) wurde auch untersucht, ob und welche Normverletzungen eine höhere Meldebereitschaft hervorrufen (FF1b). Auch hier zeigten sich signifikante Unterschiede nach Art der Normverletzung (siehe Tabelle 5). Enthielt ein Kommentar einen Aufruf zu Taten, stieg die Meldebereitschaft um mehr als das 5-fache und war damit am größten ($\text{Exp}(B) = 5,64$). Danach folgten Verschwörungstheorie, Gerüchte und Herabsetzung. Bei Kommentaren, die gefährliche Agitation enthielten, war die Meldebereitschaft in Relation zu den anderen Normverstößen am geringsten ausgeprägt.

Bestätigen ließ sich die Annahme zur Subtilität (H1b). Handelte es sich um einen direkten Normverstoß, so stieg die Meldebereitschaft um das 1,5-fache. Der Opfertyp hatte kaum Einfluss auf das Meldeverhalten. Entsprechend muss H2b verworfen werden. Wie vermutet, hatte die Neutralisierung auch bezüglich der Meldebereitschaft einen negativen Effekt (H3b): Enthielt der Kommentar neben dem Normverstoß eine neutralisierende Aussage, so sank die Meldebereitschaft bzw. halbierte sich fast ($\text{Exp}(B) = 0,54$).

5 Diskussion

Kommentare von Nutzenden zu politischen Themen wie Asylpolitik sind oftmals von Merkmalen gekennzeichnet, die sie für eine Vielzahl an Nutzenden als deviant kennzeichnen. Sie brechen mit etablierten Normen des Umgangs, weil sie andere Menschen, sowohl einzelne Personen als auch Gruppen diffamieren, zu Gewalt aufrufen oder mehr oder weniger gezielt Verschwörungstheorien verbreiten. Nicht all diese Merkmale werden, so die Grundannahme dieses Beitrags, als gleichermaßen normverletzend wahrgenommen. Ziel dieser Arbeit war es demnach, diejenigen Merkmale eines potentiellen Hasskommentars zu identifizieren, die als besonders kritisch wahrgenommen werden. Damit verbunden ist aber auch ein weiteres Erkenntnisziel: Nicht jeder potentiell normverletzende Kommentar wird gleich als solcher identifiziert und es lassen sich Techniken beobachten, wie die soziale Kontrolle gegenüber solchen Kommentaren gezielt unterlaufen werden kann. Hierbei spielen neben subtileren und indirekten Formen

Tabelle 5: Ergebnisse zur Meldebereitschaft devianter Kommentare

Art des Normverstoßes	B	SE	Exp(B)
Herabsetzung	0,21*	0,09	1,23
Aufruf zu Taten	1,73***	0,09	5,64
Verschwörungstheorie	0,61**	0,12	1,84
Gerüchte	0,47**	0,12	1,60
gefährliche Agitation	als Referenzkategorie auf Null gesetzt		
<i>Subtilität</i>			
direkt	0,45***	0,08	1,56
indirekt	als Referenzkategorie auf Null gesetzt		
<i>Opfertyp</i>			
individualisiert	0,11	0,06	1,12
abstrakt	als Referenzkategorie auf Null gesetzt		
<i>Neutralisierung</i>			
ja	-0,62***	0,06	0,54
nein	als Referenzkategorie auf Null gesetzt		

Anmerkung: AV: Kommentar melden. LLR = 891,59; Cox&Snell = ,10; Nagelkerke = ,10; ** $p < ,01$, *** $p < ,001$.

der Normabweichung vor allem Neutralisierungstechniken eine Rolle, mit deren Hilfe, die Devianz in einem Kommentar gezielt kaschiert werden kann.

In theoretischer Hinsicht baute dieser Beitrag somit auf dem vor allem in den letzten Jahren wachsenden Kanon zur Analyse der Wirkung von Kommentaren auf (Kenski et al., 2017; Muddiman & Stroud, 2017; Naab, et al., 2017; Ziegele & Jost, 2016;) legte den Fokus aber auf Formen inziviler (Coe et al., 2014; Kenski et al., 2017; Papacharisi, 2004), devianter Kommentare. Es zeigt sich, dass wie erwartet, gerade indirekte, weniger offensichtliche Normverletzungen wie gefährliche Agitation in Kommentaren als weniger deviant wahrgenommen und weniger häufig gemeldet werden. Auch Herabsetzungen werden offenbar als weniger problematisch wahrgenommen. Besonders interessant ist hierbei, dass unsere Hypothesen in Bezug auf

den Einsatz von Neutralisierungstechniken ebenfalls Bestätigung fanden: Neutralisierungen tragen dazu bei, dass Normverletzungen in Nutzerkommentaren als moralisch akzeptabel angesehen werden und somit deviantes Verhalten begünstigen. Folgt man Henry (2009), so können Neutralisierungen zu einer gesellschaftlichen Demoralisierung beitragen, da sie die Akzeptanz von Normverletzungen begünstigen. Die vorliegende Studie bietet dazu – zumindest bezogen auf die politische Partizipation in sozialen Medien – erste Hinweise. Was im Rahmen dieser Studie nicht untersucht werden konnte, ist die Wirkung solcher indirekten, subtilen und vor allem mit Hilfe von Neutralisierungstechniken abgeschwächten Kommentare. Da es solche Formen der Kommunikation leichter haben, in einem Forum stehen zu bleiben, also nicht gelöscht zu werden, steigt somit die Wahrscheinlichkeit, dass Nutzende genau mit solchen Kommentaren konfrontiert werden, wenn sie sich an Debatten zu einem kontroversen Thema beteiligen wollen. Welche persuasiven Wirkungen können nun solche abgemilderten, dennoch aber zum gewissen Grad deviante Kommentare bei Nutzenden entfalten? Es wäre möglich, dass eben genau jene Kommentare, durch ihre Unterschwelligkeit im Devianzgrad es leichter haben, wahrgenommen zu werden und gegebenenfalls somit auch als Gegenargumente so in einen Diskurs eingebracht werden können. Dies scheint gerade der Mechanismus von beispielsweise Verschwörungstheorien zu sein. Eine Studie von Mocanu und Kollegen (2015) konnte zeigen, dass gerade Nutzende von Nachrichtenangeboten, die politisch weniger ausgewogene, extreme Positionen vertreten, offen für derartige Inhalte sind. Die Ergebnisse dieses Beitrags öffnen somit das Feld für eine kritische Untersuchung gerade solcher nicht offensichtlich devianten Kommentare im Internet.

Was diese Studie jedoch nicht beantworten konnte, ist die Frage, was tatsächlich dazu führt, dass ein Kommentar auf einer bestimmten Plattform gemeldet wird. Zum einen wurden persönliche Disposition aber auch situative Faktoren im Rahmen unserer Studie nicht berücksichtigt. Zum anderen fand die Untersuchung mit Bezug zu einem spezifischen politischen Thema statt – dem der Debatte um Flüchtlings- und Asylpolitik – und fokussierte auf eine ausgewählte Gruppe – eher höher gebildete, eher tendenziell links eingestellte Personen mit einer, demnach zu vermutenden positiven Einstellung gegenüber Geflüchteten und asylfreundlichen Politikerinnen und Politikern. Auch wenn wir die untersuchten Devianzmerkmale als allgemeingültig betrachten, so gilt die spezifische Manipulation wie wir sie vorgenommen haben, nur für dieses besondere Sample. Andere Gruppen mögen eine anders geartete Ausgestaltung der Merkmale als besonders

deviant betrachten. Eine eher rechtsgerichtete, Geflüchteten gegenüber negativ eingestellt Gruppe wird sich vermutlich stärker von Kommentaren angegriffen fühlen, die die Eigengruppe („die Deutschen“) verletzt. Eine Ausweitung des Samples, aber auch der Themenstellung könnte hier kritisch hinterfragen, inwieweit die Erkenntnisse dieser Studie auch für andere Kontexte von Relevanz sind.

Dr. Claudia Wilhelm ist akademische Rätin (a. Z.) am Seminar für Medien- und Kommunikationswissenschaft der Universität Erfurt.

Dipl.-Medienwirtin Isabell Ziegler ist Promovendin im Nachwuchskolleg „Communication and Digital Media“ der Universität Erfurt.

Dr. Sven Jöckel ist Professor für Kommunikationswissenschaft mit dem Schwerpunkt Kinder- und Jugendmedien am Seminar für Medien- und Kommunikationswissenschaft der Universität Erfurt.

Quellenverzeichnis

- Amadeu Antonio Stiftung (2015). „Geht sterben!“ Umgang mit Hate Speech und Kommentaren im Internet. Abgerufen von <https://www.amadeu-antonio-stiftung.de/w/files/pdfs/hatespeech.pdf>
- Amadeu Antonio Stiftung (2016a). *Monitoringbericht 2015/16. Rechtsextreme und menschenverachtende Phänomene im Social Web*. Abgerufen von <https://www.amadeu-antonio-stiftung.de/w/files/pdfs/monitoringbericht-2015.pdf>
- Amadeu Antonio Stiftung (2016b). *Hetze gegen Flüchtlinge in Sozialen Medien. Handlungsempfehlungen*. Abgerufen von <https://www.amadeu-antonio-stiftung.de/w/files/pdfs/hetze-gegen-fluechtlinge.pdf>
- 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
- Backhaus, K., Erichson, B., & Weiber, R. (2013). *Multivariate Anaylsemethoden. Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3, 193–209. doi:10.1207/s15327957pspr0303_3

- Bär, J. (2008). *Qualität von Öffentlichkeit im Netz. Eine quantitative Inhaltsanalyse von Leserkommentaren in Onlinezeitungen*. Unveröffentlichte Masterarbeit, Ludwig-Maximilians-Universität, München.
- Bergström, A., & Wadbring, I. (2015). Beneficial yet crappy: Journalists and audiences on obstacles and opportunities in reader comments. *European Journal of Communication*, 30(2), 137-151. doi:10.1177/0267323114559378
- Bhavnani, R., Findley, M.G., & Kuklinski, J.H. (2009). Rumor dynamics in ethnic violence. *Journal of Politics*, 71, 876-892. doi:10.1017/S002238160909077X
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage. Studies in interactional sociolinguistics: Vol. 4*. Cambridge: Univ. Press.
- Campion-Vincent, V. (2009). From evil others to evil elites. In G. A. Fine & C. Heath (Hrsg.), *Rumor mills: The social impact of rumor and legend* (S.103-122). New Brunswick, NJ: Transaction Publishers.
- Citron, D.K., & Norton, H.L. (2011). Intermediaries and hate speech: Fostering digital citizenship for our information age. *Boston University Law Review*, 91, 1435-1484. Abgerufen von <http://scholar.law.colorado.edu/articles/178>
- 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
- Cohen, A. K. (1968). *Abweichung und Kontrolle*. München: Juventa-Verlag.
- Crawford, K. & Gillespie, T. (2014). What is a flag for?: Social media reporting tools and the vocabulary of complaint. *New Media & Society*, 18(3), 410-428. doi:10.1177/1461444814543163
- Forsa (2016). *Ethik im Netz. Hate Speech*. Abgerufen von http://www.lfm-nrw.de/fileadmin/user_upload/lfm-nrw/Service/Veranstaltungen_und_Preise/Medienversammlung/2016/EthikimNetz_Hate_Speech-PP.pdf
- Gagliardone, I., Gal, D., Alvez, T., & Martinez, G. (2015). *Countering online hate speech*. Abgerufen von <http://unesdoc.unesco.org/images/0023/002332/233231e.pdf>
- Gardiner, B., Mansfield, M., Anderson, I., Holder, J., Louter, D., & Ulmanu, M. (12. April 2016). The dark side of Guardian comments. *The Guardian*. Abgerufen von <https://www.theguardian.com/technology/2016/apr/12/the-dark-side-of-guardian-comments>

- Hanzelka, J., & Schmidt, I. (2017). Dynamics of cyber hate in social media: A comparative analysis of anti-Muslim movements in the Czech Republic and Germany. *International Journal of Cyber Criminology*, 11, 143-160. doi:10.5281/zenodo.495778
- Henry, S. (2009). *Social deviance*. Cambridge: Polity Press.
- Hinduja, S. (2007). Neutralization theory and online software piracy: An empirical analysis. *Ethics and Information Technology*, 9, 187-204. doi:10.1007/s10676-007-9143-5
- Inglehart, R. F., & Klingemann, H.D. (1976). Party identification, ideological preference, and the left-right dimension among western mass publics. In I. Budge (Hrsg.), *Party identification and beyond: representations of voting and party competition* (S. 143-204). Chester: Wiley.
- Kenski, K., Coe, K., & Rains, S. A. (2017). Perceptions of uncivil discourse online: An examination of types and predictors. *Communication Research*, 26. doi:10.1177/0093650217699933
- Lee, E., & Leets, L. (2002). Persuasive storytelling by hate groups online. *American Behavioral Scientist*, 45(6), 927-957. doi:10.1177/0002764202045006003
- Leets, L. (2001). Responses to internet hate sites: Is speech too free in cyberspace? *Communication Law and Policy*, 6(2), 287-317. doi:10.1207/S15326926CLP0602_2
- Malin, J., & Fowers, B.J. (2009). Adolescent self-control and music and movie piracy. *Computers in Human Behavior*, 25, 718-722. doi:10.1016/j.chb.2008.12.029
- Maruna, S., & Copes, H. (2005). What have we learned from five decades of neutralization research? *Crime and Justice*, 32, 221-320.
- Meibauer, J. (2013). Hassrede – von der Sprache zur Politik. In J. Meibauer (Hrsg.), *Hassrede / Hate Speech. Interdisziplinäre Beiträge zu einer aktuellen Diskussion*, 2., korrigierte Fassung (S. 1-16). Abgerufen von http://geb.uni-giessen.de/geb/volltexte/2013/10121/pdf/HassredeMeibauer_2013_korr.pdf
- Mocanu, D., Rossi, L., Zhang, Q., Karsai, M., & Quattrociocchi, W. (2015). Collective attention in the age of (mis) information. *Computers in Human Behavior*, 51, 1198-1204. doi:10.1016/j.chb.2015.01.024
- Morris, R. G., & Higgins, G. E. (2009). Neutralizing potential and self-reported digital piracy. A multitheoretical exploration among college graduates. *Criminal Justice Review*, 34, 173-195. doi:10.1177/0734016808325034
- Muddiman, A., & Stroud, N. J. (2017). News values, cognitive biases, and partisan incivility in comment sections. *Journal of Communication*. doi:10.1111/jcom.12312

- Muddiman, A. (2013). *The instability of incivility: How news frames and citizen perceptions shape conflict in American politics*. University of Texas at Austin.
- Naab, T. K., Kalch, A., & Meitz, T. G. K. (2017). Flagging uncivil user comments: Effects of intervention information, type of victim, and response comments on bystander behavior. *New Media & Society*, 70(1). doi:10.1177/1461444816670923
- Oliver, M. B., Dillard, J. P., Bae, K., & Tamul, D. J. (2012). The effect of narrative news format on empathy for stigmatized groups. *Journalism & Mass Communication Quarterly*, 89(2), 205–224. doi:10.1177/1077699012439020
- Opp, K. D. (1974). *Abweichendes Verhalten und Gesellschaftsstruktur*. Darmstadt: Luchterhand.
- Papacharissi, Z. (2004). Democracy online: Civility, politeness, and the democratic potential of online political discussion groups. *New Media & Society*, 6(2), 259–283. doi:10.1177/1461444804041444
- Reich, Z. (2011). User Comments. The transformation of participatory space. In J. B. Singer, A. Hermida, D. Domingo, A. Heinonen, S. Paulussen & T. Quandt (Hrsg.), *Participatory Journalism: Guarding Open Gates at Online Newspapers* (S. 96–117). Malden, MA: Wiley-Blackwell.
- Reinbold, F. (5. April 2017). Gesetz gegen Hassrede. So will Maas Facebook und Co. büßen lassen. *Spiegel Online*. Abgerufen von <http://www.spiegel.de/netzwelt/netzpolitik/heiko-maas-und-facebook-bussgelder-in-hassrede-gesetz-geaendert-a-1141853.html>
- Samaratunge, S., & Hattotuwa, S. (2014). *Liking violence. A study of hate speech on Facebook in Sri Lanka*. Abgerufen von <http://www.cpalanka.org/wp-content/uploads/2014/09/Hate-Speech-Final.pdf>
- Santana, A. D. (2015). Incivility dominates online comments on immigration. *Newspaper Research Journal*, 36, 92–107. doi:10.1177/0739532915580317
- Sobieraj, S., & Berry, J. (2011). From incivility to outrage: Political discourse in blogs, talk radio, and cable news. *Political Communication*, 28, 19–41. doi:10.1080/10584609.2010.542360
- Springer, N. (2014). *Beschmutzte Öffentlichkeit? Warum Menschen die Kommentarfunktion auf Online-Nachrichtenseiten als öffentliche Toilettenwand benutzen, warum Besucher ihre Hinterlassenschaften trotzdem lesen, und wie die Wände im Anschluss aussehen*. Berlin: LIT.

- Springer, N., Engelmann, I., & Pfaffinger, C. (2015). User comments: Motives and inhibitors to write and read. *Information, Communication & Society*, 18, 798-815. doi:10.1080/1369118X.2014.997268
- Stefanowitsch, A., & Flach, S. (2016). *Auswertung von Userkommentaren auf der offiziellen Facebook-Seite von PEGIDA, Januar bis Dezember 2015 im Auftrag der Sueddeutschen Zeitung*. Abgerufen von https://drive.google.com/file/d/0B9mLol0BxlQ_Z053SXZ6S2NVR3M/view
- 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
- Sunstein, C. R., & Vermeule, A. (2009). Conspiracy theories: Causes and cures. *Journal of Political Philosophy*, 17(2), 202-227. doi:10.1111/j.1467-9760.2008.00325.x
- Sykes, G., & Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, 22, 664-670.
- The American heritage dictionary of the English language*. (2000). Boston: Houghton Mifflin.
- Vobič, I. (2014). Historicising the journalist-audience relationships in the internet era: A case study of the Slovenian newspaper Delo. *Communication Management Quarterly*, 30, 131-156. doi:10.5937/comman1430141V
- Wilhelm, C., Jöckel, S., & Ziegler, I. (im Druck). Reporting hate comments – investigating the effects of deviance characteristics, neutralization strategies and users' moral orientation. *Communication Research*.
- Ziegele, M., & Jost, P. B. (2016). *Not funny?: The effects of factual versus sarcastic journalistic responses to uncivil user comments*. *Communication Research*. doi:10.1177/0093650216671854