

Sebastian Enghofer

Nudge and Crime – Verhaltenspsychologische und kriminologische Aspekte von Desinformation in sozialen Netzwerken am Beispiel von COVID-19

Verschiedene Gewaltdelikte der letzten Jahre – sei es in Christchurch, Hanau oder der Pizzagate-Vorfall in Washington, D.C. – haben gezeigt, wie fließend der Übergang von virtueller zu realer Gewalt ist: In all diesen Fällen wurden die Täter durch Verschwörungsmethoden und Desinformation insbesondere auf Sozialen Medien zu ihren Taten motiviert. Was als Desinformation und Verschwörungsmethoden seinen Anfang im Cyber- und Informationsraum nimmt, endet im physischen Raum mit der Ausübung von körperlicher Gewalt. Dieser Übergang zeigt sich auch in ‚Infodemie‘ der Coronakrise. Der Kampf gegen Desinformation und Falschnachrichten ist damit von zentraler gesellschaftlicher und kriminologischer Bedeutung. Hier werden bisher erprobte Strategien aus verhaltenspsychologischer, demokratietheoretischer und kriminologischer Sicht evaluiert werden, um daraus Vorschläge für neue Ansätze zur Eindämmung der Informationskrise zu entwickeln.

Schlagerwörter: Corona; Desinformation; Nudging; Situational Crime Prevention; Virtuelle Gewalt

Disinformation on Social Media Based on the Example of COVID-19: Insights From Behavioural Psychology and Criminology

A range of violent offences in recent years – whether in Christchurch, Hanau or the Pizzagate incident in Washington, D.C. – have shed light on the transition from virtual to physical violence. In all these cases, the perpetrators were motivated by conspiracy theories and disinformation that circulated on social media. Conspiracy theories and disinformation, which took off in the virtual reality of social media, led to physical violence. This transition can also be clearly seen in the current coronavirus ‘infodemic’. The fight against disinformation and fake news is thus of vital interest from both a societal and a criminological perspective. In this essay, current strategies against this phenomenon shall be evaluated from a behavioural psychology, political theory and criminology perspective. Finally, new solutions to contain the ‘infodemic’ of our time will be proposed.

Keywords: coronavirus, disinformation, nudging, situational crime prevention, virtual violence

1. Desinformation als kriminologisches und gesellschaftliches Problem

Am 4. Dezember 2016 stürmte der 28-jährige Amerikaner Edgar Maddison Welch das Pizzeria *Comet Ping Pong* in Washington, D.C. und feuerte mit seinem AR-15-Sturmgeschütz drei Schüsse ab. Bei dem Vorfall wurde glücklicherweise niemand verletzt. Als Motiv gab Welch an, er habe einen Kinderpornografie-Ring ausheben wollen, welcher angeblich von Poli-

tikern der demokratischen Partei im Keller des Restaurants betrieben werde. Dabei bezog Welch seine Informationen aus einer als #Pizzagate bekannt gewordenen Verschwörungserzählung (Lopez, 2016), welche einen zentralen Baustein des gegenwärtig allseits diskutierten QAnon-Verschwörungsmythos darstellt.¹

In ähnlicher Weise inspiriert tötete Tobias Rathjen in Hanau am 19. Februar 2020 neun Menschen. Bereits vor seiner Tat hatte der Täter aufgrund einer vermeintlichen Verschwörung gegen ihn Strafanzeige gegen Unbekannt gestellt (Giebel et al., 2020). Diese und andere, an QAnon erinnernde Versatzstücke von Verschwörungsmythos dienten dem Täter mutmaßlich als Motiv für seine Taten (Laufer, 2020).

Diese beiden Beispiele illustrieren den fließenden Übergang von virtueller² und realer Gewalt: Was als Desinformation und Verschwörungsmythos seinen Anfang im Cyber- und Informationsraum nimmt, endet im physischen Raum mit der Ausübung von körperlicher Gewalt. Diese Verbindung wird in der Coronapandemie umso deutlicher. So sah WHO-Generaldirektor Tedros Adhanom Ghebreyesus in seiner Rede auf der Münchener Sicherheitskonferenz 2020 die zentrale Herausforderung durch COVID-19 nicht nur in der Erkrankung selbst, sondern auch in der damit verbundenen Informationskrise: „We’re not just fighting an epidemic, we’re fighting an infodemic. Fake news spreads faster and more easily than this virus, and is just as dangerous“ (Ghebreyesus, 2020).

Dabei erscheinen jedoch nicht nur die mittelbar medizinischen Folgen problematisch. Beispielsweise können falsche Informationen darüber, wie das Virus übertragen wird, welche Maßnahmen gegen eine Infektion schützen, welche Behandlungsmethoden es gibt oder wie gefährlich das Virus für welche Bevölkerungsgruppen ist, zu falschem Leichtsinns bei individuellem und kollektivem Verhalten, bspw. durch Verzicht auf Social Distancing oder dem Vertrauen auf unwirksame oder gar schädliche Selbsttherapien, verleiten (Bridgman et al., 2020). Darüber hinaus wirken ‚Fake News‘ im Rahmen der Coronapandemie auch unmittelbar kriminogen: Die Identifikation von Asiaten als vermeintliche Sündenböcke oder absurde Unterstellungen, der Mobilfunk-Standard 5G verursache COVID-19, haben reale Straftaten wie Körperverletzung und Sachbeschädigung nach sich gezogen (Enghofer & Müller, 2020). Verschwörungstheorien, wonach die Pandemie nur dazu diene, die Bevölkerung einer vermeintlichen neuen Weltordnung zu unterwerfen, stellen mutmaßlich auch das Motiv von Brandanschlägen auf das Robert-Koch-Institut (RKI) am 25. Oktober 2020 und einem Sprengsatz, welcher in Berlin am selben Morgen explodierte, dar. In einem anbei aufgefundenen Bekenner schreiben wurde das sofortige Ende der Corona-Beschränkungen gefordert (Flade, 2020).

Auch in der Coronakrise kann Desinformation somit reale, physische Gewalt nach sich ziehen und – sehr überspitzt formuliert – der Verbreiter oder Rezipient der Desinformation von heute der Terrorist von morgen werden (Enghofer et al., 2020). Damit erscheint sowohl aus kriminologischer als auch aus gesellschaftlicher Perspektive die Eindämmung von Fake News nicht nur als zentrale Herausforderung bei der Bekämpfung von COVID-19, sondern auch insgesamt

¹ Zum Unterschied zwischen konkreter Verschwörungserzählung und allgemeinem Verschwörungsmythos siehe Lamberty (2020).

² Der Begriff ‚virtuelle Gewalt‘ soll hier darauf verweisen, dass Sprechakte im Cyber- und Informationsraum als virtuellem Raum reale psychische und physische Gewalt nach sich ziehen können, selbst aber noch nicht unmittelbar als gewalttätiger Akt gelten müssen. Damit ist jedoch ausdrücklich nicht gemeint, dass Desinformation virtuell im Sinne von ‚nicht existent‘ sei.

als wesentliches Feld zur Prävention von Straftaten, welche ihre Motive dem Cyber- und Informationsraum entnehmen, aber in der realen Welt wirken.³

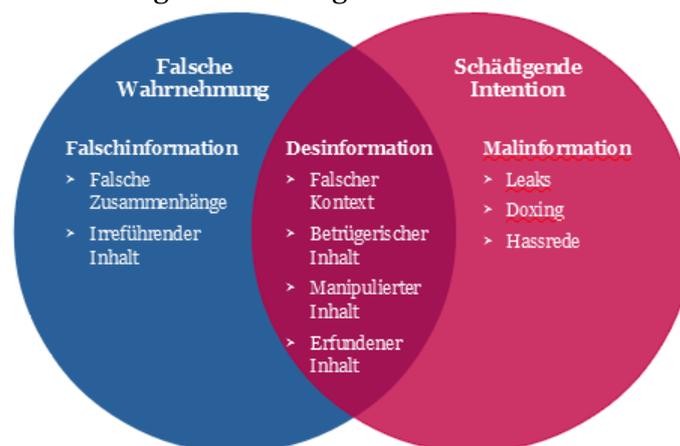
Im Folgenden soll daher vor allem theoretisch und mit kurzen Ausblicken auf praktische Anwendungen skizziert werden, wie verschiedene Perspektiven der Verhaltenspsychologie, Kriminologie und Politikwissenschaft zusammengesetzt werden können, um den potenziell kriminogenen Einfluss von Desinformation erklären und mögliche Präventionsstrategien darstellen zu können.

2. Desinformation – eine typologische Einordnung

In der öffentlichen Diskussion wird im Kontext dieser Herausforderung meist allgemein von Fake News gesprochen. Unter diesem Sammelbegriff werden dabei jedoch drei unterschiedliche Konzepte von Falschnachrichten subsumiert, die es zu unterscheiden gilt (Wardle & Derakhshan, 2017, S. 20):

1. Mis-information (Falschinformation): Nachrichten, die falsch sind, aber von Sender und Empfänger fälschlicherweise für wahr gehalten werden, jedoch ohne boshafte Intention
2. Malinformation: Nachrichten, die wahr sind, aber mit der bewussten Intention (teils illegalerweise) veröffentlicht werden, Anderen zu schaden
3. Desinformation: Nachrichten, die bewusst erstellt werden, um anderen Personen, sozialen Gruppen oder Ländern zu schaden, dabei jedoch einen wahren Kern besitzen können. Im Gegensatz zur Malinformation wird dabei bewusst z.B. der Kontext ignoriert, um eine an sich wahre Nachricht möglichst schädigend wirken zu lassen

Abbildung 1. Eigene Darstellung auf Grundlage von Wardle & Derakhshan (2017, S. 20)



³ Damit soll freilich nicht der Auffassung das Wort geredet werden, Desinformation sei ein alleiniger oder zwingend ausschlaggebender Faktor bei der Erklärung von Gewalttaten. Analog zur Debatte um die Rolle von „Killerspielen“ (Gollwitzer et al., 2014) dürfen andere Faktoren wie bspw. Persönlichkeit, unmittelbar Tatgelegenheiten oder Sozialisation nicht vernachlässigt werden. Im vorliegenden Aufsatz soll vielmehr die Perspektive eröffnet werden, inwiefern die Informationsumwelt einen möglichen Erklärungs- und Präventionsansatz liefern kann.

Dabei ist allerdings zu beachten, dass im Zuge der Verbreitung die ursprüngliche Desinformation häufig zur Falschinformation wird. Während die boshafte Intention beim ursprünglichen Erstellen der Nachricht noch vorhanden war, ist diese Intention für viele Rezipienten, welche die Nachricht indirekt weiterleiten, nicht mehr erkennbar.

Die Unterscheidung zwischen den Produzenten von Desinformation, welche aus monetären oder ideologischen Gründen bewusst Desinformation verbreiten (Allcott & Gentzkow, 2017, S. 217) und denjenigen, welche unbewusst zur Weiterverbreitung beitragen, indem sie Artikel teilen und weiterleiten, ist insofern relevant, als sie unterschiedliche Ansätze der Bekämpfung erfordert. Während Akteure, welche bewusst und intentional Desinformation verbreiten, strafrechtlich belangt werden könnten (Klein & Wueller, 2017; Rückert, 2018) und ihnen zumindest Verletzungen der journalistischen Sorgfaltspflicht vorgeworfen werden können, ist dies bei der Rezeption und unbewussten Weiterverbreitung durch Nutzer deutlich schwieriger.

Im vorliegenden Aufsatz soll daher – als Gegenpol zur strafrechtlichen Perspektive der bewussten Verbreitung – die Frage gestellt werden, welche Faktoren zur Weiterverbreitung von Desinformation beitragen und wie jenseits rein strafrechtlicher Erwägungen präventiv dagegen vorgegangen werden kann.

3. Desinformation im Spannungsfeld zwischen Wahrheit und Demokratie

Die Prävention von Falschnachrichten ist jedoch gerade in westlichen Demokratien nicht trivial und ohne weiteres möglich. Dies liegt nicht nur an pragmatischen oder juristischen Hemmschwellen, sondern verweist vielmehr auf ein grundsätzliches philosophisches Spannungsmoment von Demokratien.

Als „offene Gesellschaften“ (Popper, 2003), welche für sich die Schaffung eines freien, diskursiven Klimas, in welchem verschiedene Sichtweisen, Argumente und Meinungen miteinander konkurrieren und niemand für sich eine ‚alternativlose Wahrheit‘ beanspruchen kann, gehört für moderne Demokratien der Meinungspluralismus zu ihren zentralen Pfeilern. Deswegen stehen alleinige Wahrheitsansprüche über ‚richtige‘ und ‚falsche‘ Meinungen für viele Demokratietheoretiker im Widerspruch zur Demokratie (Hidalgo, 2014). Prägnant fasst dies der österreichische Staatsrechtler Hans Kelsen in seiner berühmten „Verteidigung der Demokratie“ zusammen:

„Gewiß [sic], wenn die Frage nach dem, was sozial richtig, was das Gute, das Beste ist, in einer absoluten, objektiv gültigen, für alle unmittelbar einleuchtenden Weise beantwortet werden könnte: dann wäre die Demokratie schlechthin unmöglich. Denn was könnte es für einen Sinn haben, über eine Maßnahme, deren Richtigkeit über allem Zweifel erhaben feststeht, abstimmen und die Mehrheit entscheiden zu lassen?“ (Kelsen, 2006, S. 236)

Daraus kann allerdings nicht geschlossen werden, dass in einer Demokratie keine legitimen Einschränkungen der Meinungsfreiheit geben kann. Die Demokratie braucht zwar einen Pluralismus an verschiedenen Meinungen, zugleich aber einen kleinsten gemeinsamen Nenner, auf den sich alle einigen können und der die Basis aller Debatten darstellt.

Dieser kleinste gemeinsame Nenner wurde von Karl Popper im Prinzip des kritischen Rationalismus zusammengefasst als Lebenseinstellung, „die zugibt, dass ich mich irren kann, dass du recht haben kannst und dass wir zusammen vielleicht der Wahrheit auf die Spur kommen

werden“ (Popper, 2003, S. 281). In ähnlicher Art argumentiert die deliberative Demokratietheorie, dass in demokratischen Diskursen der „zwanglose Zwang des besseren Argumentes“ (Habermas, 2009, S. 144) alleiniger Maßstab sein müsse. Hannah Arendt schließlich sieht in der Verabschiedung von Fakten den ersten Schritt hin zum totalitären Denken, da sich „Propaganda durch eine bemerkenswerte Verachtung für Tatsachen überhaupt“ (Arendt, 2015, S. 742) auszeichne. Zusammengefasst stellt also „eine auf sachlichen Informationen beruhende Deliberation [...] eine notwendige Bedingung für demokratische Legitimität dar“ (Vestergaard & Hendricks, 2017).

Desinformationen, welche einer „auf sachlichen Informationen beruhende[n] Deliberation“ (Vestergaard & Hendricks, 2017) fundamental zuwiderlaufen und einen gegenseitigen Austausch im Sinne Karl Poppers verunmöglichen, tragen zur Vertiefung gesellschaftlicher Konfliktlinien bei und sind insofern – jenseits ihres unmittelbar kriminogenen Potentials – hochproblematisch. Allerdings ist der Spagat zwischen dem Problem, als Staat einerseits nicht pauschal vorschreiben zu können, welche Meinung als legitim und wahr gilt, und andererseits bei einer Kollision der Meinungsfreiheit mit anderen Grundrechten Einschränkungen der Meinungsfreiheit rechtfertigen zu können, dabei weder juristisch (Löber & Roßnagel, 2020; Rückert, 2018) noch demokratietheoretisch endgültig geklärt.

Im Versuch, diese Herausforderung zu meistern und eine angemessene und geeignete Lösung im Umgang mit Des- und Falschinformationen zu finden, haben liberale Demokratien bisher vor allem mit zwei (einander nicht ausschließenden) idealtypisch zu unterscheidenden Prinzipien gearbeitet: einer regulativen sowie einer edukativen Herangehensweise.⁴

3.1 Regulation: Löschung und Reichweiteneinschränkung

Die regulative Reaktion besteht darin, falsche Nachrichten soweit möglich zu löschen, zu blockieren oder anderweitig ihre Verbreitung zu verhindern oder zumindest zu minimieren. Dies wird in westlichen Demokratien, welche anders als China kein umfassendes Zensurnetzwerk besitzen, in Kooperation mit Plattformbetreibern wie Facebook, Twitter, Youtube usw. vorgenommen. Beispielsweise werden im Zuge des Netzwerkdurchsetzungsgesetz (NetzDG) Plattformbetreiber dazu verpflichtet, eine Infrastruktur zum raschen Melden und Löschen offensichtlich rechtswidriger Inhalte zu schaffen. Schutzzweck des Gesetzes ist dabei, strafbare Inhalte nach § 1 Abs. 3 NetzDG (darunter fallen unter anderem §§ 129, 129b, 130, 131 StGB) zu verhindern. Das Gesetz dient explizit nicht der in der Gesetzesbegründung angegebenen allgemeinen „Bekämpfung von strafbaren Falschnachrichten („Fake News“) in sozialen Netzwerken“. Dies widerspräche den Prinzipien der fdGO (Hoven & Gersdorf, 2019).

Angesichts der hohen Strafandrohung gab es seit Verabschiedung des NetzDG die Befürchtung, soziale Netzwerke könnten deswegen im Zweifel lieber zu viel, als zu wenig blocken (overblocking) (Kurz, 2018). Diese Befürchtung hat sich allerdings laut dem im September 2020 erschienen Evaluierungsbericht des Bundesjustizministeriums nicht bewahrheitet (Evaluierungsbericht NetzDG, 2020). Soziale Medien löschen jedoch auch eigenverantwortlich jenseits der Anforderungen des NetzDG Inhalte, wenn diese den eigenen Nutzungsbedingungen widersprechen. Dabei arbeiten (nicht nur in der Coronakrise) soziale Medien wie Facebook, Youtube oder Twitter mit einer Vielzahl an privaten ‚Faktencheckern‘ sowie offiziellen Stellen wie bspw.

⁴ Einen kurzen Überblick über diese beiden Idealtypen der Bekämpfung von Desinformation bieten auch Högden et al. (2020, S. 83-88).

der WHO zusammen, um Falschnachrichten zu identifizieren und dann zu löschen oder in ihrer Reichweite einzuschränken (EU Disinfo Lab, 2020; Hacıyakupoglu et al., 2018).

Sowohl die Löschung im Rahmen der vom NetzDG vorgegebenen Rahmenbedingungen als auch die Löschung aufgrund eines Verstoßes gegen plattformeigene Nutzungsbedingungen erscheint dabei jedoch problematisch.

Zum einen kann eine Löschung erst erfolgen, nachdem der Inhalt bereits veröffentlicht und dann geprüft wurde – selbst kurze Zeiträume reichen aber aus, um virulente Nachrichten tausendfach zu kopieren und weiter zu verbreiten. Aufgrund der verhaltenspsychologischen Aspekte bei der Rezeption von Falsch- und Desinformation (siehe Abschnitt 3.2) kommt eine Löschung somit meist zu spät.

Zum anderen erschwert die enorme Zahl an Veröffentlichungen in sozialen Medien die sorgfältige Bewertung einzelner Posts – dies ist weder technisch mithilfe von Algorithmen (Marx, 2020) noch durch menschliche Redakteure immer zuverlässig zu leisten. Dieses Problem wird darüber hinaus dadurch verschärft, dass nicht nur auf praktischer Ebene diese Erkennung nicht immer eindeutig erfolgen kann, sondern auch dass auch die Maßstäbe, wann etwas Satire oder Ironie ist, nicht nur von Land zu Land verschieden sind, sondern auch prinzipiell juristisch und demokratietheoretisch nicht ohne Weiteres allgemeingültig beantwortet werden können. Wie schwer die juristische Bekämpfung von Desinformationen ist, sofern diese nicht eindeutig unter andere Straftatbestände wie bspw. Volksverhetzung, Beleidigung oder Verleumdung subsumiert werden können, belegt § 276 des österreichischen Strafgesetzbuches, der Anfang 2016 außer Kraft gesetzt wurde, da es nie zu einer Verurteilung auf seiner Grundlage kam (Schmid, 2016).

Im Fall der eigenmächtigen Löschung von Inhalten aufgrund von Verstößen gegen die eigenen Nutzungsbedingungen entscheiden zudem soziale Netzwerke aufgrund nicht immer intersubjektiv nachvollziehbarer Regeln darüber, welche Inhalte gelöscht werden und welche nicht. Damit handeln sich soziale Medien allerdings den Vorwurf ein, nicht neutral zu sein. Wiederholt wurde, speziell von konservativer Seite, der Vorwurf einer angeblichen ‚links-liberalen Voreingenommenheit‘ sozialer Medien erhoben (Satariano et al., 2018).⁵ Die damit einhergehende Verantwortung sozialer Medien, aufgrund ihrer ohnehin intransparenten Algorithmen zwangsläufig mit darüber zu entscheiden, welche Meinungen zugelassen und sichtbar bleiben, wurde jedoch bspw. von Mark Zuckerberg zurückgewiesen. Mit dem Argument, man wolle eben kein „arbiter of truth“ (Levy, 2020) sein, zögerte Facebook lange damit, insbesondere Falschnachrichten von Politikern zu löschen, da die Blockade von Hassrede zu einer „schleichenden Zensur“ (Citron, 2018, S. 1039) führen könne.

Dies führt zurück zur demokratietheoretischen Problematik des Spannungsverhältnisses von Wahrheit und Demokratie. In der regulativen Vorgehensweise spiegelt sich nämlich das Problem der „streitbaren Demokratie“ (Löwenstein, 1937a, 1937b): Ein Staat, der den möglichst freien Zugang zu Informationen und zu Diskursen zu seinem Fundament erklärt, kann nur schwer zugleich eine Einschränkung genau dieser Freiheit fordern, möchte er sich nicht dem Verdacht aussetzen, seine eigenen Prinzipien nicht ernst zu nehmen (Rijpkema & Asbury, 2019). Das Löschen von Nachrichten in sozialen Medien – ob aufgrund gesetzlicher Veranlassung oder eigenmächtig aufgrund der Nutzungsbedingungen – kann somit schnell als Zensur wahrgenommen werden und im Sinne der extremistischen „Aktions-Repressions-Spirale“ (Paris, 1989) eine weitere Radikalisierung fördern.

⁵ Zur kritischen Bewertung dieser These siehe Scott (2020), wonach gemäß einer Datenauswertung des Magazins *Politico* konservative Positionen in sozialen Medien mitnichten marginalisiert seien.

Bereits die derzeit praktizierte Lösung des moderaten Löschens und der Einschränkung von Sichtbarkeit durch Anpassung der algorithmenbasierten Empfehlungen führt dazu, dass viele Akteure aus radikalen und extremistischen politischen Milieus, ebenso wie Verschwörungstheoretiker, Impfgegner und andere Akteure ‚alternativer Wirklichkeiten‘⁶, einfach zu anderen sozialen Medien wechseln, um so der vermeintlichen Zensur von Facebook, Youtube oder Twitter zu entgehen. Eine verstärkte Löschung von Falschnachrichten und Desinformation würde somit tendenziell nur zu einer noch weiteren Binnenradikalisierung der alternativen Realitätswahrnehmung führen und einen sinnvollen, rationalen Diskurs, wie er für Demokratien zwingend notwendig ist, weiter erschweren. Dies bestätigt sich in der Coronakrise durch die vermehrte Abwanderung extremistischer Akteure und Kanäle auf weitgehend unmoderierte Plattformen wie Telegram, wodurch deren Radikalisierung eher beschleunigt als gebremst wurde (Rogers, 2020; Urman & Katz, 2020).

3.2 Edukation: Richtigstellungen und Faktenprüfung

Deswegen setzt die edukative Variante weiterhin auf die Vernunft der Nutzer und vertraut dem „zwanglosen Zwang des besseren Arguments“ (Habermas, 2009, S. 144), indem versucht wird, durch Faktenchecker Falschmeldungen und emotionalisierende Diskurselemente zu widerlegen, ohne diese einfach zu entfernen. Dabei haben sich verschiedenste Organisationen der „Gegenaufklärung“ (debunking, Appel, 2020, S. 206), also dem Richtigstellen von falschen Informationen, verschrieben. Sowohl spezialisierte Abteilungen der öffentlich-rechtlichen sowie privaten Medienanstalten als auch eine Vielzahl an privaten Initiativen versuchen, den Wahrheitsgehalt von Nachrichten und die Glaubwürdigkeit der Absender zu überprüfen und dadurch Falschnachrichten und Desinformation aufzudecken (Appel, 2020, S. 206).

Alle diese Initiativen kranken jedoch daran, dass die menschliche Natur gegen sie arbeitet. Zwar wird häufig auf die Bedeutung technischer Aspekte wie bspw. Algorithmen bei der Verbreitung von Fake News verwiesen; diese entfalten ihre Wirkung und schnelle Verbreitung jedoch vor allem aufgrund psychologischer Charakteristika: Es sind eben zumeist reale Menschen, und nicht (nur) Social Bots⁷, welche freiwillig Falschnachrichten und Desinformation teilen (Appel & Doser, 2020, S. 16). Dabei agieren Menschen selten rational, sondern unterliegen einer Reihe von kognitiven Verzerrungen (cognitive biases) (Tversky & Kahneman, 1974), welche ihre Rationalität einschränken (bounded rationality).

Menschliches Denken lässt sich in zwei Entscheidungssysteme unterscheiden: System 1 handelt schnell, intuitiv und assoziativ (Kahneman, 2012) mittels Heuristiken (Gigerenzer et al., 2011), während System 2 langsam, nach bewusster Überlegung und logisch arbeitet.

⁶ Für einen Ausschnitt der ‚alternativen Realitäten‘ in der Coronakrise siehe den Beitrag von Enghofer und Müller (2020).

⁷ Unter Social Bots versteht man Accounts in sozialen Medien, hinter welchen kein Mensch, sondern ein Algorithmus steht, der automatisiert bestimmte Inhalte weiterleitet und verbreitet (siehe dazu Neis & Mara, 2020).

Table 1. Gegenüberstellung von System 1 und System 2, eigene Darstellung auf Basis von (Kahneman, 2012)

	System 1	System 2
Geschwindigkeit	Schnell	Langsam
Aktiv bei	Emotionaler Erregung, unter Zeitdruck	Zeit zum Nachdenken
Modus	Heuristisch	Logisch
Ebene	Unterbewusst / automatisch	Bewusst / willentlich
Menschenbild	Homo sapiens („humans“)	Homo oeconomicus („econs“)

Ausgehend von diesem Ideal des homo oeconomicus werden Ergebnisse von System 1 als „verzerrt“ bezeichnet.⁸

Dass Menschen nicht immer rational und logisch Entscheidungen erst nach Verarbeitung aller relevanten Informationen fällen, sondern sich dabei mehr von ihrem „Bauchgefühl“ (Gigerenzer, 2009) leiten lassen, ist jedoch kein prinzipieller Mangel, sondern eine notwendige Anpassung an eine dynamische und von Informationsmangel gekennzeichnete Umwelt (Weber & Knorr, 2020, S. 113). In der wissenschaftlichen Diskussion um die Bewertung von Heuristiken und kognitiven Verzerrungen ist dementsprechend auch umstritten, wie ‚negativ‘ oder ‚positiv‘ die jeweiligen Verarbeitungsmuster zu sehen sind. So argumentieren Befürworter von ‚Nudging‘ (s.u.) eher, dass kognitive Verzerrungen Menschen prinzipiell an der Umsetzung ihrer ‚wahren‘ Interessen hindern und daher überwunden werden sollten (Thaler & Sunstein, 2008); in ähnlicher Weise findet sich diese Perspektive auch bei Daniel Kahneman (Kahneman, 2012; Kahneman et al., 2021).

Dagegen betont z.B. Gerd Gigerenzer die adaptive Rolle von Heuristiken, wonach ‚kognitive Verzerrungen‘ eine Anpassung an bestimmte Umwelten darstellten und innerhalb dieser Umwelten auch gute Entscheidungen ermöglichten (Gigerenzer, 2009). Gemäß dieser Perspektive der ‚ecological rationality‘ (Todd & Gigerenzer, 2012) ist die ‚Vernünftigkeit‘ von System 1 oder System 2 nicht unabhängig von der Umwelt zu betrachten, sondern kann nur vor dem Hintergrund der Umgebung bewertet werden: während heuristische Verarbeitungsmuster nach System 1 in manchen Situation zu ausreichend guten (oder sogar besseren) Ergebnissen als das analytische System 2 führen, können sie in anderen Situationen zu deutlich schlechteren Resultaten bei der Entscheidungsfindung führen. Gleiches gilt analog für System 2.

Bei der Frage, inwiefern Desinformation und Informationsverarbeitung eine kriminogene Wirkung entfalten können und ob Richtigstellungen dem entgegenwirken können, ist somit primär danach zu fragen, inwiefern innerhalb des Cyber- und Informationsraums eher die positive Rolle von Heuristiken überwiegt, oder sich kognitive Verzerrungen als Ursache von kriminologisch relevantem Verhalten identifizieren lassen. Dabei soll im Folgenden der Fokus auf Aspekte der Informationsverarbeitung gelegt werden, welche die Anfälligkeit für Desinformation erhöhen und gerade im Kontext des Cyber- und Informationsraums besonders relevant sind, insofern sie erklären können, warum das Richtigstellen von Informationen als Maßnahme gegen Desinformation nicht ausreicht.

Allgemein erfolgt menschliche Informationsverarbeitung (unbewusst) oft nicht mit dem Ziel, ergebnisoffen und unvoreingenommen die ‚Wahrheit‘ herauszufinden, sondern dient eher

⁸ Diese Gegenüberstellung wird in der Forschung auch kritisiert. So weist Thomas Schramme darauf hin, dass im Vergleich mit dem angenommenen *homo oeconomicus* alle Menschen als kognitiv *beeinträchtigt* erscheinen – einen Vergleich, den man in Bezug auf körperliche Unterschiede, gemessen am „Mollathlet“, so kaum ziehen würde (siehe Schramme, 2016, S. 539).

dazu, Vorurteile zu bestätigen und zu rationalisieren.⁹ Gemäß der Theorie der „Motivated Recognition“ (Kunda, 1990) lassen sich Menschen bei der Informationsverarbeitung primär vom erwünschten Ergebnis leiten. Informationen werden beispielsweise danach selektiert, inwiefern sie eine bereits vorgefasste Meinung bestätigen – hierbei spricht man vom so genannten Bestätigungsfehler (confirmation bias). Dieser Effekt wird durch die Wiederholung von (Des-)information verstärkt – je häufiger etwas wahrgenommen wird, desto eher wird es für wahr gehalten (Illusory Truth Effekt). Außerdem fällt es Menschen leichter, zuerst wahrgenommene Informationen im Kopf zu behalten (Primacy Effekt). Dies kann dazu führen, dass Richtigstellungen unfreiwillig die Glaubwürdigkeit der ursprünglichen Desinformation erhöhen, wenn sie diese vor der Richtigstellung erneut sichtbar machen (Pennycook et al., 2018). Dadurch wird jedoch nicht nur die ursprüngliche Falschinformationen wieder auf die Agenda gehoben, sondern gemäß der „Dividende des Lügners“ (Chesney & Citron, 2018, S. 1785) insgesamt die Glaubwürdigkeit von Medien zur Disposition gestellt.

Außerdem zeigen verschiedene Studien (Bail et al., 2018; Levy, 2017; Nyhan & Reifler, 2010), dass die Konfrontation mit abweichenden Sichtweisen zu einer weiteren Selbstbestätigung führen kann. Der erwünschte Effekt, das Gegenüber durch den zwanglosen Zwang des besseren Arguments zu überzeugen, ist also häufig psychologisch nicht haltbar, ja kann sich sogar in sein Gegenteil verkehren (backfire-Effekt).¹⁰ Selbst wenn dieser Effekt nicht auftritt und sich auf kognitiver Ebene die Meinung ändert, kann die (eigentlich bereits überwundene) Falschinformationen noch Auswirkungen auf das Handeln von Individuen haben (Ecker et al., 2011; Johnson & Seifert, 1994; Levy, 2017, S. 496; Nyhan et al., 2020; Wood & Porter, 2019).

Die systematischen Verzerrungen sind dabei jedoch nicht nur der zeitlichen und kognitiven Ressourcenknappheit geschuldet, sondern bilden einen wesentlichen Bestandteil der individuellen und kollektiven Identität (Raab et al., 2013; Schulz et al., 2020). Wie der Ingroup- und Outgroup-Bias zeigen, orientieren sich Menschen bei ihren Wertvorstellungen und Ansichten an der sozialen Gruppe, der sie sich zugehörig fühlen, wohingegen Werte der Outgroup eher abgelehnt werden (Mullen et al., 1992). Dementsprechend werden Informationen, welche von der ‚eigenen‘ Gruppe kommen, als glaubwürdiger empfunden, als ‚fremde‘ Informationen (Clark & Winegard, 2020), wobei Menschen als „naive Realisten“ (Weber & Knorr, 2020, S. 108) fälschlicherweise davon ausgehen, sie selbst seien rational und objektiv und unterlägen keinen kognitiven Verzerrungen, wohingegen alle anderen subjektiv und voreingenommen seien.

Die Motivation im Umgang mit neuen Informationen ist somit meist nicht, sie unvoreingenommen zu prüfen, um der Wahrheit auf die Spur zu kommen. Vielmehr versuchen Menschen, kognitive Dissonanz zu vermeiden, indem Informationen so interpretiert werden, dass sie ins Weltbild des Individuums und seiner sozialen Bezugsgruppe passen.¹¹ Wenn aber Menschen gar nicht von Wahrheitssuche motiviert sind, sondern gemäß des *motivated reasoning* Infor-

⁹ Der Psychologie Jonathan Haidt umschreibt dies mit seiner Metapher des „emotionalen Hundes und seines rationalen Schwanzes“ (Haidt, 2001).

¹⁰ Zur kritischen Beurteilung des „backfire-Effekts“ siehe Ecker et al. (2019); Guess & Coppock (2016); Wood & Porter (2019).

¹¹ Wie stark die soziale Erwünschtheit kognitive Prozesse beeinflusst, ist allerdings umstritten. So gehen etwa Pennycook und Rand (2019) davon aus, dass im Kontext von Fake News die Bereitschaft zu reflexivem Denken Unterschiede in der Anfälligkeit von Individuen besser erklärt als ideologische Voreingenommenheit.

mationen selektiv nach Ähnlichkeit, Verfügbarkeit oder Kompatibilität mit der eigenen (Gruppen-)Meinung auswerten, und dabei sich von der empfundenen ideologischen Distanz zur Quelle beeinflussen lassen, erscheinen Faktenchecker als wenig vielversprechend.

Schließlich kommen – gerade bei stark polarisierenden Themen wie der Coronakrise – die Richtigstellungen zumeist von ‚etablierten‘ Medien, denen die Verbreiter von Desinformation ihrerseits das Etikett der Fake News anheften.¹² Richtigstellungen wirken somit umso schlechter, je mehr deren Quelle beim Rezipienten bereits als unglaubwürdig, ja gar ihrerseits als Fake News, wahrgenommen wird (Nekmat, 2020). Dieser Effekt ist umso stärker, je mehr bereits vorher ein Grundmisstrauen gegen die ‚korrigierende‘ Instanz vorherrscht. Gemäß der Repräsentativheuristik bewerten Menschen Informationen anhand ihrer Ähnlichkeit zu mental bereits vorhandenen Prototypen (Weber & Knorr, 2020, S. 111). Entspricht nun beispielsweise eine Richtigstellung durch öffentlich-rechtliche Sender beim Empfänger der mentalen Schablone ‚Die Lügenpresse will mich beeinflussen‘, ist kaum damit zu rechnen, dass ein Umdenken stattfindet. Darüber hinaus verweisen die kognitiven Verzerrungen auf fundamentale Timing- und Verfügbarkeitsprobleme: bis eine Fehlinformation korrigiert ist, wurde sie meistens schon tausendfach geteilt und erreicht eine größere Reichweite als die Richtigstellung. Die Fehlinformation ist dabei der Korrektur jedoch nicht nur zeitlich voraus, wodurch gemäß dem Primacy Effekt die Korrektur weniger nachhaltig im Gedächtnis verankert wird. Sie muss auch überhaupt erst einmal ihren Weg zum Rezipienten finden, was wenig wahrscheinlich ist, wenn sie auf spezialisierte Webseiten, die erst aktiv aufgerufen werden müssen, beschränkt bleibt. Gemäß der Verfügbarkeitsheuristik beschränken sich Menschen auf unmittelbar verfügbare Informationen und suchen oft nicht aktiv nach alternativen Erklärungsansätzen, welche noch dazu der vorgefassten Meinung widersprechen (Sängerlaub, 2018; Thornhill et al., 2019, S. 2).

3.3 Zwischenfazit

Zusammengefasst können also weder regulative noch edukative Maßnahmen gegen Desinformationen als ausreichend angesehen werden – während die regulatorische Methode an demokratiethoretischen Problemen sowie pragmatischen Hindernissen krankt, erscheinen die edukativen Maßnahmen aus psychologischer Perspektive als unzureichend durchdacht.

Angesichts von Befunden, wonach die Anfälligkeit für Falschnachrichten und Desinformation mit der generellen Fähigkeit zum reflektierten Denken nach System 2 korreliert (Pennycook & Rand, 2020) könnte der Schluss naheliegend erscheinen, dies könne vor allem auf problematische individuelle Denkmuster zurückgeführt und letztlich sogar lediglich als Ausdruck von Paranoia (Hofstadter, 1964) pathologisiert werden.

Dabei handelt es sich jedoch um den fundamentalen Attributionsfehler (Gilbert & Malone, 1995), welcher die Bedeutung individueller Faktoren über- und Umweltfaktoren unterschätzt. Gerade bei Desinformation zeigt sich nämlich, dass die Umgebung, innerhalb derer Menschen sich entscheiden müssen, welche Nachricht sie für vertrauenswürdig erachten, ob sie eine Nachricht weiterverbreiten und sich dazu zustimmend oder ablehnend positionieren, einen

¹² Diese Auseinandersetzung um den Begriff ‚Fake News‘ zeigte sich prägnant 2017 in Donald Trumps Vorwurf gegen CNN beim Streit um die Besucherzahl bei der Amtseinführung (Times, 2017). Im Kontext der Coronakrise wird das Etikettieren von etablierten Medien als vermeintliche Fake-News-Produzenten etwa an den vielfältigen Vorwürfen des Senders RT Deutsch gegen die öffentlich-rechtlichen Rundfunksender in Deutschland (RT Deutsch, 2020) erneut sichtbar.

wesentlichen Einfluss darauf ausübt, wie leicht Menschen Falschnachrichten glauben und weiterverbreiten. Soziale Medien, welche Nachrichten vor allem nach Ähnlichkeit und Salienz selektieren, begünstigen das Auftreten kognitiver Verzerrungen bei der Informationsverarbeitung, und tragen damit indirekt zur Verbreitung von Desinformation bei (Humborg & Nguyen, 2018, S. 29-31).

Kognitive Verzerrungen sind somit nicht an und für sich das Problem, sondern – im Sinne der ecological rationality (Todd & Gigerenzer, 2012) – vor allem die Tatsache, dass innerhalb des Cyber- und Informationsraums, und hier speziell innerhalb sozialer Netzwerke, heuristische Verarbeitungsmuster die Anfälligkeit für Desinformation erhöhen, anstatt sie zu verringern. Deswegen besteht ein zentraler Ansatzpunkt zur Prävention in der Frage, wie die Umwelt als Entscheidungsarchitektur so gestaltet werden kann, dass heuristische Informationsverarbeitung nicht so leicht zur Weiterverbreitung von Desinformation beiträgt.

4. Entscheidungsarchitekturen als Hebel zur Prävention von Desinformation

4.1 Raumbezogene Kriminalitätsprävention

In der Kriminologie wird die Frage, inwieweit die (geografische) Umgebung als Entscheidungsraum das Auftreten von deviantem und delinquenten Verhalten beeinflusst, bereits seit den 1970ern unter den Schlagworten „Crime Prevention through Environmental Design“ (Jeffery, 1971) und „Defensible Space“ (Newman, 1972) behandelt.¹³ Der Architekt Oscar Newman hatte anhand der Beobachtung von verfallenden Nachbarschaften wie der Sozialsiedlung Pruitt-Igoe in St. Louis, Missouri, festgestellt, dass Sicherheit und Ordnung in öffentlichen Räumen nur aufrecht erhalten können, wenn Bewohner sich dort geborgen fühlen, eine natürliche soziale Kontrolle der Bewohner untereinander besteht und der Raum als ‚Defensible Space‘ ein positives Image besitzt, also nicht in Verruf gerät (Newman, 1972). In ähnlicher Weise geht der spätere Broken-Windows-Ansatz davon aus, dass eine allgemeine Verwahrlosung der Umgebung weitere Normbrüche nach sich zieht, da dann der Regelverstoß als allgemein sozial akzeptiertes Verhalten gilt (Wilson & Kelling, 1982). Mit Ronald Clarkes Situational Crime Prevention (SCP) fokussierte sich raumbezogene Kriminalitätsprävention schließlich primär auf unmittelbare Tatgelegenheiten sowie konkreten Maßnahmen zu deren Minimierung (Clarke, 2018).

All diesen verschiedenen Ansätzen liegt die gemeinsame Annahme zugrunde, dass die Umgebung einen wesentlichen Einfluss darauf ausübt, ob sich Menschen deviant oder regelkonform verhalten. Dabei ist jedoch nicht zwingend eine geographisch-materielle Umgebung gemeint: der Cyber- und Informationsraum kann ebenfalls als ‚Raum‘ gelten, innerhalb dessen das Auftreten devianten Verhaltens auf das Fehlen eines „Defensible Digital Space“ (Ehlert & Rüdiger, 2020) zurückgeführt werden kann.

Neuere Vertreter dieser Theoriefamilie gehen dabei nicht mehr von einem ‚perfekt rationalen‘ Täter aus, sondern integrieren die psychologischen Einsichten in die beschränkte Rationalität (Wortley & Tilley, 2018). Dies lässt sich gut an den 25 Techniken der Situational Crime Prevention illustrieren.

¹³ Für eine detaillierte Chronologie der *Environmental Criminology* siehe Wortley & Townsley (2017b).

Table 2. Übersicht der Situational Crime Prevention, eigene Darstellung auf Basis von (Freilich & Newman, 2017, S. 8-10)

Aufwand erhöhen	Risiko erhöhen	Belohnung minimieren	Provokationen minimieren	Ausreden entfernen
1. Ziele härten	6. Soziale Zuständigkeit und Überwachung	11. Ziele verschleiern	16. Frustration und Stress reduzieren	21. Regeln setzen
2. Eingangs- und Zugangskontrollen	7. Natürliche Überwachung	12. Ziele entfernen	17. Streit verhindern	22. klare Anweisungen
3. Ausgangskontrollen	8. Anonymität reduzieren	13. Eigentum identifizieren	18. Emotionale Erregung verhindern	23. Bewusstmachung von Regeln
4. Straftäter abhalten	9. lokal zuständige Personen einbinden	14. Schwarzmarkt kontrollieren	19. Gruppendruck reduzieren	24. Regelbefolgung vereinfachen
5. Zugang zu Werkzeugen und Waffen kontrollieren	10. technische Überwachung ausbauen	15. Belohnung vorenthalten	20. Nachahmung verhindern	25. Drogen und Alkohol kontrollieren

Hier fallen unmittelbar Anknüpfungspunkte für die oben besprochenen kognitiven Verzerrungen ins Auge. So können unter anderem die Punkte „Ziele verschleiern“, „Ziele härten“ und „Streit verhindern“ als Aspekt des *availability bias* gelesen werden. Wenn Menschen ihr Verhalten nach unmittelbar verfügbaren Optionen ausrichten, kann die Vermeidung unmittelbar zugänglicher Tatgelegenheiten die Wahrscheinlichkeit krimineller Handlungen reduzieren. Der Punkt 18 „Emotionale Erregung verringern“ kann als Ausdruck des Versuchs, System 2-Denken zu erleichtern, gelesen werden, und die Punkte 19 und 20 können als Versuch gesehen werden, den Ingroup- und Outgroup-Bias zu nutzen, um kriminelles Verhalten zu minimieren. Die Punkte 22, 23 und 24 zielen schließlich darauf ab, dass selbst bei System 1-Denken regelkonformes Verhalten wahrscheinlicher wird als regelwidriges, indem nicht lange darüber nachgedacht werden muss, was die richtige Verhaltensweise ist.

4.2 Nudging

Außerhalb der Kriminalprävention hat sich für den Versuch, Verhalten unter Ausnutzung kognitiver Verzerrungen durch die Veränderung der Entscheidungsarchitektur zu beeinflussen, vor allem der Begriff „Nudging“ eingebürgert. Der Begriff (englisch für „anstupsen“) geht auf die Monographie *Nudge* von Richard Thaler und Cass Sunstein (2008) zurück. Ausgehend von der Erkenntnis, dass Menschen meist nicht rational als homo oeconomicus entscheiden und kognitiven Verzerrungen unterliegen, plädieren die Autoren dafür, diese Erkenntnisse aktiv zu nutzen, um das Verhalten von Menschen beeinflussen zu können:

„A nudge, as we will use the term, is any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. [...] a nudge is any factor that significantly alters the behavior of Humans, even though it would be ignored by Econs. Econs respond primarily to incentives. If the government taxes candy, they will buy less candy, but they are not influenced by such “irrelevant” factors as the order in which options are displayed.“ (Thaler & Sunstein, 2008, S. 5, 8)

Durch eine bewusste Veränderung der Entscheidungsarchitektur, also der Art und Weise, wie sich Handlungsoptionen dem Entscheider darstellen, kann somit das Verhalten von Individuen in eine bestimmte Richtung ‚gestupst‘ werden, ohne dass Wahlmöglichkeiten ausgeschlossen werden oder finanzielle Anreize gesetzt werden.

Damit bietet sich der Nudge-Ansatz in doppelter Hinsicht als mögliche Lösung für die angesprochenen Probleme zur Bekämpfung von Fake News an:

Einerseits betonen Thaler und Sunstein den freiheitsbewahrenden Aspekt des Nudging (Sunstein spricht dabei auch von „libertärem Paternalismus“ (2015, S. 58)). Es geht beim Nudging gerade nicht darum, Optionen wie beispielsweise eine Falschnachricht einfach zu löschen und dadurch die Meinungsfreiheit zu beschränken, womit ein zentraler Kritikpunkt an regulativen Maßnahmen gegen Desinformation umgangen wird.

Andererseits greift die Theorie explizit die psychologische Einsicht der kognitiven Verzerrungen und der beschränkten Rationalität auf, was ihr eine realistischere anthropologische Basis verschafft als edukative Ansätze, welche auf die Vernunft und den „zwanglosen Zwang des besseren Arguments“ vertrauen.

Dabei ist allerdings festzuhalten, dass beim Nudging-Ansatz der rationale Entscheider als Idealvorstellung bestehen bleibt. Während die Theorie zwar davon ausgeht, dass der homo oeconomicus keine adäquate Beschreibung menschlicher Entscheidungsfindung darstellt, sollen Menschen in die Richtung ‚gestupst‘ werden, so zu handeln ‚als ob‘ sie dem Ideal entsprächen (Sunstein, 2015, S. 89). Nudges, welche nicht zum rationalen Denken anregen¹⁴, sondern lediglich bestehende kognitive Defizite ausnutzen, werden in der Forschung überwiegend abgelehnt. Einerseits sei ihre Wirkung weniger nachhaltig, da das Individuum nicht selbst lerne, sondern von einer wohlwollenden Entscheidungsarchitektur abhängig bleibt (Grüne-Yanoff & Hertwig, 2016, S. 164), andererseits widersprechen Intervention, welche kognitive Verzerrungen ausnutzen, anstatt Menschen darüber aufzuklären und sie selbst in die Lage zu versetzen, überlegter zu handeln, zentralen demokratischen Prinzipien von Menschenwürde und Autonomie (Button, 2018; Hertwig & Grüne-Yanoff, 2017; Lodge & Wegrich, 2016). Diese Sichtweise deckt sich auch mit der Perspektive Robert Mertons auf das emanzipatorische Potential von Verhaltensbeeinflussung: „Mass persuasion is not manipulative when it provides access to the pertinent facts; it is manipulative when the appeal to sentiment is used to the exclusion of pertinent information.“ (Merton et al., 2004, S. 186)

Im Kontext von Desinformation bedeutet dies: Nudging-Maßnahmen müssen dazu führen, dass Menschen bei der Rezeption von Nachrichten reflektierter darüber nachdenken, wie plausibel die Quelle ist, welche eigenen Vorurteile sie dabei beeinflussen und welche Konsequenzen eine Weiterverbreitung von Nachrichten haben könnte.

4.3 Nudging und raumbezogene Kriminalitätsprävention

Angesichts ähnlicher Ausgangspunkte und Zielsetzungen lassen sich zwischen Nudging und raumbezogenen Kriminalitätspräventionsmechanismen wie SCP vielfältige Verbindungslinien

¹⁴ Zur Unterscheidung der verschiedenen Typen von Nudges siehe exemplarisch Hansen & Jespersen (2013); Hertwig & Grüne-Yanoff (2017).

ziehen. Die möglichen Verknüpfungen¹⁵ liegen beispielsweise darin, regelkonformes Verhalten durch soziale Vorbilder sichtbarer zu machen, Konsequenzen der eigenen Handlung aufzuzeigen, die natürliche Kontrolle durch andere zu verstärken und dadurch den ‚Mitläufer-Effekt‘ auszunutzen und ein Gefühl der Verantwortung und Geborgenheit zu schaffen, um Anreize für deviantes Verhalten zu minimieren.

Übertragen auf die Eindämmung von Desinformation bedeutet dies: Soziale Medien müssen Anreize dafür schaffen, Nachrichten nicht einfach ungeprüft weiterzuleiten, ohne über deren Glaubwürdigkeit nachzudenken, das Verbreiten und Hochladen von Desinformation und Falschnachrichten zu sanktionieren, indem andere Nutzer dies kritisieren (soziale Kontrolle) und die Verfügbarkeit von Desinformationen einzuschränken, ohne sich aber dem Verdacht der Zensur auszusetzen.

5. Konkrete Umsetzbarkeit der bisherigen Überlegungen

Soziale Medien haben jedoch bisher vor allem mit edukativen und regulativen Methoden gearbeitet¹⁶: So werden in Suchergebnissen und bei automatischen Empfehlungen Informationen offizieller Stellen priorisiert, es wird enger mit externen Faktencheckern kooperiert und Methoden zur automatischen Erkennung und Löschung von Falschnachrichten wurden ausgebaut. Allerdings kann keine dieser Maßnahmen als „nudge to reason“ gelten, da sie zwar beispielsweise den *availability bias* ausnutzen, um den Zugang zu gesicherten Informationen erleichtern, aber den einzelnen Nutzer dadurch eben nicht zu einem reflektierteren Umgang mit Informationen anregen.

Solche Interventionen könnten in der Praxis vor allem an drei Punkten ansetzen:

1. Erstens muss das leichte und gedankenlose Weiterleiten von Nachrichten ohne Reflexion über den Inhalt erschwert werden.
2. Zweitens muss die Belohnung für emotionale, moralisierende und polarisierende Nachrichten in Form von ‚Likes‘ und ‚Shares‘ unterbunden werden.
3. Drittens muss die soziale Kontrolle erhöht und sichtbarer werden: bleiben Desinformationen unwidersprochen stehen, erhöht dies – ganz im Sinne der Theorien ‚Defensible Space‘ und ‚Broken Windows‘ – die empfundene Legitimität solcher Aussagen.

5.1 Reflexionsfähigkeit der Nutzer erhöhen

Twitter führte 2009 mit dem ‚Retweet-Button‘ eine einfache Möglichkeit zum Teilen und Weiterleiten von Nachrichten anderer Accounts ein. Während vorher Nutzer einen Tweet, um ihn zu teilen, manuell kopieren und als neuen eigenen Tweet veröffentlichen mussten, ermöglicht der Retweet-Button eine Automatisierung dieses Prozesses. Dies verkürzt jedoch erheblich die

¹⁵ Eine Übersicht der möglichen, aber noch nicht systematisch erforschten Anknüpfungspunkte bieten unter anderem: Bouhana (2013); Clarke (2014, S. 34); Pease & Graham (2014, S. 82-85); Pease & Laycock (2012); Wortley & Townsley (2017a, S. 53).

¹⁶ Für eine Übersicht im Kontext der COVID-19 Pandemie siehe bspw. das EU Disinfo Lab (2020).

Zeit, welche zwischen Lesen und Weiterleitung eines Tweets erfolgt. Besonders emotionalisierende und moralisierende Tweets werden somit schnell und unmittelbar unter Einfluss des kognitiven Systems 1 weitergeleitet, ohne dass System 2 intervenieren und eine Reflexion über Glaubwürdigkeit, Wirkung und Absicht der ursprünglichen Nachricht anstoßen kann, weswegen Retweets eine zentrale Rolle bei der Verbreitung von Falschnachrichten zukommt (Vossoughi et al., 2018).

In Experimenten (Wang et al., 2014, 2013) konnte am Beispiel von Verzögerungen beim Posten auf Facebook gezeigt werden, dass Nudges, welche das Verbreiten von Nachrichten verzögern, zu reflektiertem Denken anregen und damit das unbewusste Weiterleiten von Fake News erschweren könnten. In Bezug auf Twitter konnte der Effekt des *throttling mindless activity* ebenfalls experimentell belegt werden (Caraban et al., 2019; Konstantinou et al., 2019). Bei den genannten Studien wurde jeweils eine Verzögerung von einigen Sekunden zwischen dem ersten Klicken auf ‚Absenden‘ und dem tatsächlichen Posten der Nachricht eingeblendet. Nutzer mussten also mit leicht erhöhtem (Zeit-)Aufwand rechnen, um eine Nachricht auf Twitter oder Facebook zu versenden – Zeit, die ausreichte, um System 2 zu aktivieren und eine Reflexion über den Inhalt anzustoßen.

Erste Schritte in diese Richtung werden bei sozialen Medien nur zaghaft unternommen: Beispielsweise schränkte WhatsApp die Teilbarkeit von Nachrichten ein und kennzeichnet weitergeleitete Nachrichten (WhatsApp FAQ - Weiterleitungsbeschränkungen, o. J.), während Facebook bei Falschnachrichten eine Warnmeldung einblendet, bevor diese weitergeleitet werden können (Smith, 2017). In ähnlicher Weise überblendet Twitter Falschnachrichten mit einem Banner, welches als erstes – noch bevor der eigentliche Tweet sichtbar ist – darauf hinweist, dass er folgende Inhalt irreführend oder falsch sein könnte. Diese Maßnahme greift sowohl den Primacy Effect als auch den Verzögerungseffekt auf: Wenn Nutzer zuerst den Warnhinweis sehen und dann bewusst klicken müssen, um einen Tweet anschauen zu können, erhöht dies die Wahrscheinlichkeit, dass Desinformationen als solche erkannt und behandelt werden.

Allerdings bleibt bei den Ansätzen von Facebook und Twitter das Grundproblem bestehen, dass zuverlässige Maßstäbe der Bewertung, was Falschnachrichten sind und was nicht, Voraussetzung der Maßnahmen bleiben.

5.2 Verringerung von Emotionalisierung in sozialen Medien

Auch der zweite Aspekt – die Einschränkung der emotionalen Belohnung für polarisierende Nachrichten in Form von Likes, welche das Streben nach sozialer Bestätigung innerhalb der empfundenen Bezugsgruppe befriedigt, und die damit verbundene leichte Verbreitung durch Empfehlungen – wird von den sozialen Medien nur unzureichend angegangen, da dies ihr Geschäftsmodell unmittelbar berühren würde. Weder Facebook noch Twitter sind gegenwärtig als Plattformen für deliberative, rationale Argumentation angelegt, sondern dienen vor allem dem schnellen, oberflächlichen Austausch, der emotionale und polarisierende Inhalte bevorzugt (Boehme-Neßler, 2018, S. 55). Aufgrund der ideologisch verzerrten Wissensaneignung (Clark & Winegard, 2020) und der Tendenz zur Selbstbestätigung generieren vor allem polarisierende Inhalte eine hohe Aufmerksamkeit, erscheinen im Newsfeed weit oben, werden entsprechend häufig geteilt und geliked und damit noch weiter nach oben befördert (Brady et al., 2017). Damit verbunden ist jedoch eine oberflächlichere, automatische Auseinandersetzung

mit den Inhalten: diese werden entweder gemäß dem Bestätigungsfehler unkritisch angenommen oder aber als gegnerische Meinung der Outgroup einfach abgelehnt. Eine deliberative, um Faktentreue und Wahrheitsfindung bemühte Diskussionskultur kann in einem solchen Umfeld kaum entstehen.

Würden Diskursteilnehmer hingegen aufgefordert, Erklärungen und Belege für ihre Ansichten vorzubringen, anstatt nur ihre politischen Präferenzen kundzutun, schaltet sich System 2 ein, wodurch die Anfälligkeit für Fake News sinkt (Bago et al., 2020). Dies führt auch zu einer Verringerung von Polarisierung und einer Mäßigung der Haltung (Fernbach et al., 2013), da die Einflüsse des tribalistischen Denkens und der damit verbundenen Verzerrungen minimiert werden.

5.3 Soziale Kontrolle und Gegenrede

Von besonderer Bedeutung schließlich ist der dritte Punkt, die sichtbare soziale Kontrolle: so wie im Broken-Window-Ansatz¹⁷ kaputte Scheiben als Sinnbild fehlender Normenkontrolle möglichst schnell repariert werden sollen, um nicht andere zur Nachahmung anzuregen, müssen im digitalen Raum Falschnachrichten möglichst schnell sichtbar korrigiert und richtiggestellt werden.

Automatische Hinweise auf Korrekturen können erfolgreich sein, sofern sie schnell und einfach genug wahrgenommen werden (Bode & Vraga, 2018). Facebooks neu eingeführte Praxis, Nutzer aktiv auf verlässliche Informationen ‚hin zu stupsen‘, sofern diese vorher Falschnachrichten geteilt oder gelesen haben, kann als erster Schritt in diese Richtung gelten (Ghaffary, 2020). In ähnlicher Weise funktionieren Vorschläge, zu jedem Artikel verschiedene Sichtweisen einzublenden, um Falschnachrichten nicht unwidersprochen stehen zu lassen (Thornhill et al., 2019). Auch Twitter versieht mittlerweile irreführende Nachrichten mit einem Faktencheck, welcher unmittelbar sichtbar ist und Tweets mit Hintergrundinformationen versieht (Fung, 2020). Allerdings bleibt hier das Problem des ‚Senders‘ der Korrekturen bestehen – wird Facebook oder die verlinkten Faktenchecker ohnehin vom Rezipienten bereits als unglaubwürdig angesehen, werden Nutzer einer solchen Korrektur eher keinen Glauben schenken.

Da – wie oben gezeigt wurde – Informationsverarbeitung vor allem im Hinblick auf soziale Kohäsion erfolgt, sind Korrekturen von Mitgliedern der ‚Ingroup‘ – insbesondere dem nahen sozialen Umfeld – wesentlich glaubwürdiger als externe und durch das soziale Medium lancierte Richtigstellungen (Margolin et al., 2018). Eine wesentliche Rolle fällt dabei Familien und Freunden zu: Widersprechen diese offensichtlichen Falschnachrichten, sinkt die Wahrscheinlichkeit signifikant, dass diese weiterverbreitet und für wahr befunden werden (Bode & Vraga, 2018). Auch kritische Nutzerkommentare, welche nicht aus dem direkten sozialen Umfeld stammen, können eine Weiterleitung falscher Nachrichten minimieren (Kluck et al., 2019). Prinzipiell sind der Gegenrede allerdings Grenzen gesetzt. Da ein wesentlicher Antrieb in der Informationsverarbeitung darin liegt, kognitive Dissonanz zwischen dem eigenen (Gruppen-) Weltbild und der Realität zu vermeiden, ist die Bedeutung der Inhalte beschränkt. Vielmehr

¹⁷ Für die prinzipielle Anwendbarkeit der Broken-Window-Metapher auf den digitalen Raum siehe Hoheisel-Gruler (2020).

geht es Menschen meist darum, überhaupt ein resonantes Weltbild¹⁸ konstruieren zu können – welches, ist letztlich egal (Koehler, 2020). Zudem darf der Widerspruch nicht den Anschein erwecken, die ursprüngliche Desinformation sei prinzipiell diskussionswürdig und damit potentiell glaubwürdig – dies würde bedeuten, in die Falle der „Dividende des Lügners“ (Chesney & Citron, 2018, S. 1785) zu tappen. Deswegen ist von besonderer Bedeutung, dass Richtigstellungen nicht nur die Falschmeldung negieren, sondern ein eigenständiges Gegenarrativ erzeugen, welches ohne ausführlichen Rekurs auf die ursprüngliche Falschmeldung auskommt (Cook & Lewandowsky, 2012) und die emotionalen Bedürfnisse nach einem resonanten Weltbild berücksichtigt.

6. Zusammenfassung und Ausblick

In der Coronakrise zeigt sich erneut, wie schwierig der Kampf gegen Desinformation insbesondere in den sozialen Medien ist. Edukative und regulative Maßnahmen, welche auch leichter direkt von staatlicher Seite vorgenommen werden können, erscheinen als wenig vielversprechend, das Problem zu lösen. Hingegen zeigen Ansätze, welche die psychologischen Erkenntnisse zur Informationsverarbeitung nutzen, um durch Beeinflussung der Entscheidungsarchitektur die Verbreitung von Fake News zu minimieren, eine größere Erfolgswahrscheinlichkeit. Allerdings sind viele der möglichen Nudge-Interventionen nur in Laborstudien erprobt. Außerdem setzen Interventionen, welche eine kritischere Reflexion von Nachrichten anstoßen sollen, vor allem bei der Rezeption von Mis-informationen an, also dem Weiterverbreiten von falschen Informationen, ohne dass man sich der Fehlerhaftigkeit der Information bewusst ist. Eine stärkere Reflexion kann dann helfen, die Weiterverbreitung zu minimieren.

Das willentliche Verbreiten von Falschnachrichten mit dem Ziel, das Gegenüber zu manipulieren (Desinformation), erfolgt hingegen meist rational und bewusst zur Erreichung politischer oder ökonomischer Ziele (Allcott & Gentzkow, 2017, S. 217). „Nudges to reason“ (Levy, 2017) sind dabei kaum hilfreich, da die Ursache der Verbreitung von Desinformation nicht in mangelnder Reflexion, sondern eben in manipulativer Intention besteht (Stahl, 2006, S. 91).

Eine Veränderung der Entscheidungsarchitektur sozialer Medien durch Nudges kann somit zwar als „Impfung“ (Linden, 2018) die Weiterverbreitung von *Falschinformation* eindämmen, indem potentielle Empfänger sensibilisiert werden. Das bewusste Verbreiten von Desinformation wider besseren Wissens hingegen kann durch solche Maßnahmen nicht verhindert werden. In Analogie zum biologischen Impfen bleiben solche Intervention dennoch äußerst relevant, da sie zumindest die Weiterverbreitung der ‚Infodemic‘ verlangsamen können.

Eine derart groß angelegte, koordinierte Anstrengung des „designing against Fake News“¹⁹ durch alle großen Anbieter sozialer Medien ist bisher allerdings nicht in Sicht. Gerade in der Coronapandemie hat sich gezeigt, dass soziale Netzwerke zwar weitergehende Maßnahmen ergriffen haben, dabei jedoch die psychologischen Ursachen der Verbreitung von Desinformation nur unzureichend aufgreifen.

Solche grundlegenden Veränderungen in der Entscheidungsarchitektur liefern letztlich auf eine weniger reibungslose Nutzung der Plattform hinaus: Idealerweise könnten dann Inhalte nicht

¹⁸ Zum Zusammenhang zwischen der sozialwissenschaftlichen Resonanztheorie Hartmut Rosas (2016) und Radikalisierung siehe exemplarisch die Arbeit von Frank Buchheit (2019).

¹⁹ Dies ist als Analogie zur raumbezogenen Kriminalprävention unter dem Schlagwort „Design against Crime“ zu lesen (siehe Davey & Wootton, 2017).

mehr so schnell und einfach geteilt werden, Postings müssten mit Begründungen und Quellenangaben versehen werden, und inhaltlich gehaltvoller Widerspruch anstatt Bestätigung würde zur Grundlage der Nutzungskultur. Diese Idealvorstellung deliberativer und um Wahrheitsfindung und Faktentreue bemühten Nutzung sozialer Medien widerspricht jedoch genau jenen kognitiven Eigenarten, weswegen soziale Medien so häufig genutzt werden. Der Erfolg sozialer Medien liegt nämlich genau darin, leicht und ohne kognitive Anstrengung mit Gleichgesinnten in Kontakt zu kommen (Boehme-Neßler, 2018, S. 55). Daher ist kaum zu erwarten, dass soziale Medien so tiefgreifende Veränderungen, wie sie oben skizziert wurden, vollumfänglich implementieren werden.

Außerdem erscheint es fraglich, ob die Nutzer und Nutzerinnen selbst dies überhaupt wollen. Angesichts der zunehmenden Abwanderung von Nutzern zu ‚alternativen‘ Plattformen wie Telegram (Eckert & Hornung, 2020) oder Parler (Roth, 2020), welche explizit damit werben, keinerlei Beeinflussung oder Moderation der Inhalte vorzunehmen, scheinen nicht wenige Menschen Meinungsfreiheit als ‚anything goes‘ zu verstehen. Dies zeigt sich in der Coronakrise auch daran, dass auf Querdenken-Demonstrationen gemäßigte Demonstranten häufig kein Problem darin sehen, Seite an Seite mit rechtsextremen Verfassungsfeinden, Reichsbürgern und Hooligans zu demonstrieren, getreu dem Motto ‚Der Feind meines Feindes ist mein Freund‘ (Lauer, 2020; Soldt, 2020).

Damit zeigt sich in der Diskussion um Desinformation (nicht nur in der Coronakrise) ein grundlegendes Problem: der Verlust einer demokratischen Geisteshaltung (Enghofer, 2020) im Sinne Kelsens oder Poppers, „die zugibt, dass ich mich irren kann, dass du recht haben kannst und dass wir zusammen vielleicht der Wahrheit auf die Spur kommen werden“ (Popper, 2003, S. 281). So scheint beispielsweise das Grundrecht auf Meinungsfreiheit zunehmend mit einem Grundrecht auf unwidersprochene Meinung verwechselt zu werden, welches allen anderen Grundrechten übergeordnet zu sein habe. Solche vereinfachten Vorstellungen von Politik und Rechtsstaatlichkeit korrelieren nicht nur mit extremistischen Denkmustern (Gambetta & Hertog, 2016, S. 146-150), sondern verweisen auch auf die enorme Bedeutung von Prävention durch Bildung und Medienkompetenz (Walton & Hepworth, 2011). Das Herausbilden eines demokratischen Charakters nämlich kann weder durch das Strafrecht noch durch die Struktur sozialer Medien allein geleistet werden.

Literaturverzeichnis

- Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives*, 31(2), 211-236. doi.org/10.1257/jep.31.2.211
- Appel, M. (2020). Wie lässt sich das Postfaktische eindämmen? In M. Appel (Hrsg.), *Die Psychologie des Postfaktischen: Über Fake News, „Lügenpresse“, Clickbait & Co.* (S. 205-210). Springer Berlin Heidelberg. doi.org/10.1007/978-3-662-58695-2_18
- Appel, M., & Doser, N. (2020). Fake News. In M. Appel (Hrsg.), *Die Psychologie des Postfaktischen: Über Fake News, „Lügenpresse“, Clickbait & Co.* (S. 9-20). Springer Berlin Heidelberg. doi.org/10.1007/978-3-662-58695-2_2
- Arendt, H. (2015). *Elemente und Ursprünge totaler Herrschaft: Antisemitismus, Imperialismus, totale Herrschaft* (Ungekürzte Taschenbuchausgabe, 18. Auflage). Piper.
- Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General*, 149(8), 1608-1613. doi.org/10.1037/xge0000729

- Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, *115*(37), 9216-9221. doi.org/10.1073/pnas.1804840115
- Bericht der Bundesregierung zur Evaluierung des Gesetzes zur Verbesserung der Rechtsdurchsetzung in sozialen Netzwerken (Netzwerkdurchsetzungsgesetz—NetzDG). (2020). Bundesregierung.
- Bode, L., & Vraga, E. K. (2018). See Something, Say Something: Correction of Global Health Misinformation on Social Media. *Health Communication*, *33*(9), 1131-1140. doi.org/10.1080/10410236.2017.1331312
- Boehme-Neßler, V. (2018). *Das Ende der Demokratie? Effekte der Digitalisierung aus rechtlicher, politologischer und psychologischer Sicht* (1. Auflage). Springer.
- Bouhana, N. (2013). The reasoning criminal vs. Homer Simpson: Conceptual challenges for crime science. *Frontiers in Human Neuroscience*, *7*. doi.org/10.3389/fnhum.2013.00682
- Brady, W. J., Wills, J. A., Jost, J. T., Tucker, J. A., & Bavel, J. J. V. (2017). Emotion shapes the diffusion of moralized content in social networks. *Proceedings of the National Academy of Sciences*, *114*(28), 7313-7318. doi.org/10.1073/pnas.1618923114
- Bridgman, A., Merkley, E., Loewen, P. J., Owen, T., Ruths, D., Teichmann, L., & Zhilin, O. (2020). The causes and consequences of COVID-19 misperceptions: Understanding the role of news and social media. *Harvard Kennedy School Misinformation Review*, *1*(3). doi.org/10.37016/mr-2020-028
- Buchheit, F. (2019). Resonanzachsen und ideologische Deradikalisierung. In E. Marks & H. Fünfsinn (Hrsg.), *Gewalt und Radikalität Ausgewählte Beiträge des 23. Deutschen Präventionstages (11. und 12. Juni 2018 in Dresden)*. (S. 199-216). Forum Verlag Godesberg GmbH.
- Button, M. E. (2018). Bounded Rationality without Bounded Democracy: Nudges, Democratic Citizenship, and Pathways for Building Civic Capacity. *Perspectives on Politics*, *16*(4), 1034-1052. doi.org/10.1017/S1537592718002086
- Caraban, A., Karapanos, E., Gonçalves, D., & Campos, P. (2019). 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, 1-15. doi.org/10.1145/3290605.3300733
- Chesney, R., & Citron, D. K. (2018). *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security* (SSRN Scholarly Paper ID 3213954). Social Science Research Network. doi.org/10.2139/ssrn.3213954
- Citron, D. K. (2018). Extremist Speech, Compelled Conformity, and Censorship Creep. *Notre Dame Law Review*, *93*(9), 1035-1072.
- Clark, C. J., & Winegard, B. M. (2020). Tribalism in War and Peace: The Nature and Evolution of Ideological Epistemology and Its Significance for Modern Social Science. *Psychological Inquiry*, *31*(1), 1-22. doi.org/10.1080/1047840X.2020.1721233
- Clarke, R. V. (2014). Affect and the reasoning criminal—Past and future. In J.-L. Van Gelder, H. Elffers, & D. Nagin (Hrsg.), *Affect and cognition in criminal decision making* (S. 20-41). Routledge/Taylor & Francis Group.
- Clarke, R. V. (2018). The Theory and Practice of Situational Crime Prevention. *Oxford Research Encyclopedia of Criminology and Criminal Justice*. doi.org/10.1093/acrefore/9780190264079.013.327
- Cook, J., & Lewandowsky, S. (2012). *The debunking handbook*. http://www.skepticalscience.com/docs/Debunking_Handbook.pdf (2021, Juni 25).
- Davey, C. L., & Wootton, A. B. (2017). *Design against crime: A human-centred approach to designing for safety and security* (1 Edition). Routledge, Taylor & Francis Group.
- Ecker, U. K. H., Lewandowsky, S., Jayawardana, K., & Mladenovic, A. (2019). Refutations of Equivocal Claims: No Evidence for an Ironic Effect of Counterargument Number. *Journal of Applied Research in Memory and Cognition*, *8*(1), 98-107. doi.org/10.1016/j.jarmac.2018.07.005

- Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin & Review*, 18(3), 570-578. doi.org/10.3758/s13423-011-0065-1
- Eckert, S., & Hornung, P. (2020, August 4). Analyse: Wie sich Verschwörungsmythen auf Telegram verbreiten. *tagesschau.de*. <https://www.tagesschau.de/investigativ/ndr-wdr/corona-telegram-101.html> (2021, Juni 25).
- Ehlert, C., & Rüdiger, T.-G. (2020). Defensible Digital Space: Die Übertragbarkeit der Defensible Space Theory auf den digitalen Raum. In T.-G. Rüdiger & P. S. Bayerl (Hrsg.), *Cyberkriminologie* (S. 151-171). Springer Fachmedien Wiesbaden. doi.org/10.1007/978-3-658-28507-4_6
- Enghofer, S. (2020). Demokratie und Weltanschauung. Zur Krise des demokratischen Charakters. In K. Y. Albrecht, L. Kirchmair, & V. Schwarzer (Hrsg.), *Die Krise des demokratischen Rechtsstaats im 21. Jahrhundert oder Wie sich die Geschichte gegen ihr Ende wehrt. Tagung des Jungen Forums Rechtsphilosophie (JFR) in der Internationalen Vereinigung für Rechts- und Sozialphilosophie (IVR) im April 2019 in Salzburg* (S. 151-164). Franz Steiner Verlag.
- Enghofer, S., & Müller, D. (2020). Fakt vs. Fake in der Corona-Krise—Kriminologische und gesellschaftliche Implikationen. *Die Polizei*, 111(10), 400-408.
- Enghofer, S., Müller, D., & Parrino, A. (2020). Wie Cyberterrorismus stattfindet – und warum wir ihn nicht sehen. Wirklichkeitsmanipulation als virtuelle Gewalt. In T.-G. Rüdiger & P. S. Bayerl (Hrsg.), *Cyberkriminologie: Kriminologie für das digitale Zeitalter* (S. 217-239). Springer Fachmedien Wiesbaden. doi.org/10.1007/978-3-658-28507-4
- EU Disinfo Lab. (2020). *Platforms' responses to COVID-19 mis- and disinformation*. EU Disinfo Lab. <https://www.disinfo.eu/resources/covid-19/platforms-responses-to-covid-19-mis-and-disinformation/> (2021, Juni 25).
- Fernbach, P. M., Rogers, T., Fox, C. R., & Sloman, S. A. (2013). Political Extremism Is Supported by an Illusion of Understanding. *Psychological Science*, 24(6), 939-946. <https://doi.org/10.1177/0956797612464058>
- Flade, F. (2020, November 2). Corona-Demos: Vom Protest zur Gewalt? *tagesschau.de*. <https://www.tagesschau.de/investigativ/wdr/corona-protest-radikalisierung-101.html> (2021, Juni 25).
- Freilich, J. D., & Newman, G. R. (2017). Situational Crime Prevention. In J. D. Freilich & G. R. Newman, *Oxford Research Encyclopedia of Criminology and Criminal Justice*. Oxford University Press. doi.org/10.1093/acrefore/9780190264079.013.3
- Fung, B. (2020, Mai 27). *Twitter labeled Trump tweets with a fact check for the first time*. CNN. <https://www.cnn.com/2020/05/26/tech/twitter-trump-fact-check/index.html> (2021, Juni 25).
- Gambetta, D., & Hertog, S. (2016). *Engineers of jihad: The curious connection between violent extremism and education*. Princeton University Press.
- Ghaffary, S. (2020, April 16). *Facebook will start nudging users who have "liked" coronavirus hoaxes*. Vox. <https://www.vox.com/recode/2020/4/16/21223972/facebook-coronavirus-hoaxes-warning-misinformation-avaaz>
- Ghebreyesus, T. A. (2020, Februar 15). *Rede vor der Münchner Sicherheitskonferenz 2020*. Münchner Sicherheitskonferenz 2020. <https://www.who.int/dg/speeches/detail/munich-security-conference> (2021, Juni 25).
- Giebel, M., Naumann, F., Greil, A., Lanzinger, J., & Kaindl, A. (2020, März 4). Hanau: Benefizkonzert für Anschlagopfer wird verschoben - Coronavirus erreicht Stadt. *Merkur.de*. <https://www.merkur.de/politik/hanau-benefiz-konzert-anschlag-attentat-bundesinnenministerium-news-massnahme-coronavirus-zr-13549464.html> (2021, Juni 25).
- Gigerenzer, G. (2009). *Bauchentscheidungen: Die Intelligenz des Unbewussten und die Macht der Intuition* (H. Kober, Übers.; 10. Aufl). Goldmann.
- Gigerenzer, G., Hertwig, R., & Pachur, T. (Hrsg.). (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.

- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, *117*(1), 21-38. <https://doi.org/10.1037/0033-2909.117.1.21>
- Gollwitzer, M., Rothmund, T., Klimmt, C., Nauroth, P., & Bender, J. (2014). Gründe und Konsequenzen einer verzerrten Darstellung und Wahrnehmung sozialwissenschaftlicher Forschungsbefunde: Das Beispiel der „Killerspiele-Debatte“. *Zeitschrift für Erziehungswissenschaft*, *17*(S4), 101-117. <https://doi.org/10.1007/s11618-014-0511-8>
- Grüne-Yanoff, T., & Hertwig, R. (2016). Nudge Versus Boost: How Coherent are Policy and Theory? *Minds and Machines*, *26*(1-2), 149-183. <https://doi.org/10.1007/s11023-015-9367-9>
- Guess, A., & Coppock, A. (2016). *The Exception, Not the Rule? The Rarely Polarizing Effect of Challenging Information*. APSA. <https://pdfs.semanticscholar.org/d53a/59ce58efed9bb67b3cbd23b5ae8a3b2b70ba.pdf> (2021, Juni 25).
- Habermas, J. (2009). *Philosophische Texte. Bd. 1: Sprachtheoretische Grundlegung der Soziologie*. Suhrkamp.
- Hacıyakupoglu, G., Hui, J. Y., Suguna, V. S., & Leong, D. (2018). *Countering Fake News. A survey of recent global initiatives* (S. 24). S. Rajaratnam School of International Studies. https://www.think-asia.org/bitstream/handle/11540/8063/PR180307_Countering-Fake-News.pdf (2021, Juni 25).
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814-834. <https://doi.org/10.1037/0033-295X.108.4.814>
- Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the Manipulation of Choice: A Framework for the responsible Use of the Nudge Approach to Behavior Change in Public Policy. *European Journal of Risk Regulation*, *4*(1), 3-28.
- Hertwig, R., & Grüne-Yanoff, T. (2017). Nudging and Boosting: Steering or Empowering Good Decisions. *Perspectives on Psychological Science*, *12*(6), 973-986. <https://doi.org/10.1177/1745691617702496>
- Hidalgo, O. (2014). *Die Antinomien der Demokratie*. Campus-Verl.
- Hofstadter, R. (1964). The Paranoid Style in American Politics. *Harper's Magazine*. <https://harpers.org/archive/1964/11/the-paranoid-style-in-american-politics/> (2021, Juni 25).
- Högdén, B., Krämer, N., Meinert, J., & Schaewitz, L. (2020). Kapitel 3: Desinformation aus medienpsychologischer Sicht. In M. Steinebach, K. Bader, L. Rinsdorf, N. Krämer, & A. Roßnagel (Hrsg.), *Desinformation aufdecken und bekämpfen* (S. 77-100). Nomos Verlagsgesellschaft mbH & Co. KG. doi.org/10.5771/9783748904816-77
- Hoheisel-Gruler, R. (2020). Der digitale Raum ist kein (grund-)rechtsfreier Raum. In T.-G. Rüdiger & P. S. Bayerl (Hrsg.), *Cyberkriminologie* (S. 71-108). Springer Fachmedien Wiesbaden. doi.org/10.1007/978-3-658-28507-4_3
- Hoven, & Gersdorf. (2019). NetzDG § 1 Rn. 4—Schutzziel des NetzDG. In Gersdorf & Paal (Hrsg.), *Beck'scher Online Kommentar Informations- und Medienrecht* (29.).
- Humborg, C., & Nguyen, T. A. (2018). Fake News – Fact Checking. In C. Humborg & T. A. Nguyen, *Die publizistische Gesellschaft* (S. 29-31). Springer Fachmedien Wiesbaden. doi.org/10.1007/978-3-658-20959-9_9
- Jeffery, R. C. (1971). *Crime Prevention Through Environmental Design*. SAGE.
- Johnson, H. M., & Seifert, C. M. (1994). Sources of the continued influence effect: When misinformation in memory affects later inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(6), 1420-1436. doi.org/10.1037/0278-7393.20.6.1420
- Kahneman, D. (2012). *Thinking, fast and slow*. Penguin Books.
- Kahneman, D., Sibony, O., & Sunstein, C. R. (2021). *Noise: Was unsere Entscheidungen verzerrt – und wie wir sie verbessern können* (T. Schmidt, Übers.). Siedler.
- Kelsen, H. (2006). *Verteidigung der Demokratie: Abhandlungen zur Demokratietheorie* (M. Jestaedt & O. Lepsius, Hrsg.). Mohr Siebeck.
- Klein, D., & Wueller, J. (2017). *Fake News: A Legal Perspective* (SSRN Scholarly Paper ID 2958790). Social Science Research Network. <https://papers.ssrn.com/abstract=2958790> (2021, Juni 25).

- Kluck, J. P., Schaewitz, L., & Krämer, N. C. (2019). Doubters are more convincing than advocates. The impact of user comments and ratings on credibility perceptions of false news stories on social media. *Studies in Communication and Media*, 8(4), 446-470. doi.org/10.5771/2192-4007-2019-4-446
- Koehler, D. (2020). Switching Sides: Exploring Violent Extremist Intergroup Migration Across Hostile Ideologies. *Political Psychology*, 41(3), 499-515. doi.org/10.1111/pops.12633
- Konstantinou, L., Caraban, A., & Karapanos, E. (2019). Combating Misinformation Through Nudging. In D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, & P. Zaphiris (Hrsg.), *Human-Computer Interaction – INTERACT 2019* (Bd. 11749, S. 630-634). Springer International Publishing. doi.org/10.1007/978-3-030-29390-1_51
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108(3), 480-498. doi.org/10.1037/0033-2909.108.3.480
- Kurz, C. (2018, Februar 18). Die Meinungsfreiheit und das NetzDG: Schwerwiegender Verstoß gegen Grundrecht. *netzpolitik.org*. <https://netzpolitik.org/2018/die-meinungsfreiheit-und-das-netzdg-schwerwiegender-verstoss-gegen-grundrecht/>
- Lamberty, P. (2020). Verschwörungserzählungen. *Informationen zur politischen Bildung aktuell*, 2020(35), 1-16.
- Lauer, S. (2020, August 31). „Querdenken“-Demo in Berlin: Warum es keine gute Idee ist, mit Rechtsextremen zu demonstrieren. *Belltower.News*. <https://www.belltower.news/querdenken-demo-in-berlin-warum-es-keine-gute-idee-ist-mit-rechtsextremen-zu-demonstrieren-103445/>
- Laufer, D. (2020, Februar 21). Attentat in Hanau—Der Kampf um die Deutungshoheit. *netzpolitik.org*. <https://netzpolitik.org/2020/der-kampf-um-die-deutungshoheit/>
- Levy, N. (2017). Nudges in a post-truth world. *Journal of Medical Ethics*, 43(8), 495-500. doi.org/10.1136/medethics-2017-104153
- Levy, S. (2020, Mai 6). Mark Zuckerberg Is an Arbiter of Truth—Whether He Likes It or Not. *Wired*. <https://www.wired.com/story/mark-zuckerberg-is-an-arbiter-of-truth-whether-he-likes-it-or-not/> (2021, Juni 25).
- Linden, S. V. D. (2018). The future of behavioral insights: On the importance of socially situated nudges. *Behavioural Public Policy*, 2(2), 207-217. doi.org/10.1017/bpp.2018.22
- Löber, L. I., & Roßnagel, A. (2020). Kapitel 5: Desinformation aus der Perspektive des Rechts. In M. Steinebach, K. Bader, L. Rinsdorf, N. Krämer, & A. Roßnagel (Hrsg.), *Desinformation aufdecken und bekämpfen* (S. 149-194). Nomos Verlagsgesellschaft mbH & Co. KG. doi.org/10.5771/9783748904816-149
- Lodge, M., & Wegrich, K. (2016). The Rationality Paradox of Nudge: Rational Tools of Government in a World of Bounded Rationality: The Rationality Paradox of Nudge. *Law & Policy*, 38(3), 250-267. doi.org/10.1111/lapo.12056
- Lopez, G. (2016, Dezember 5). *Pizzagate, the totally false conspiracy theory that led a gunman to a DC pizzeria, explained*. *Vox*. <https://www.vox.com/policy-and-politics/2016/12/5/13842258/pizzagate-comet-ping-pong-fake-news> (2021, Juni 25).
- Löwenstein, K. (1937a). Militant Democracy and Fundamental Rights, I. *American Political Science Review*, 31(3), 417-432. doi.org/10.2307/1948164
- Löwenstein, K. (1937b). Militant Democracy and Fundamental Rights, II. *American Political Science Review*, 31(4), 638-658. doi.org/10.2307/1948103
- Margolin, D. B., Hannak, A., & Weber, I. (2018). Political Fact-Checking on Twitter: When Do Corrections Have an Effect? *Political Communication*, 35(2), 196-219. doi.org/10.1080/10584609.2017.1334018
- Marx, K. (2020). Warum automatische Verfahren bei der Detektion von Hate Speech nur die halbe Miete sind. In T.-G. Rüdiger & P. S. Bayerl (Hrsg.), *Cyberkriminalologie* (S. 707-725). Springer Fachmedien Wiesbaden. doi.org/10.1007/978-3-658-28507-4_27
- Merton, R. K., Lowenthal, M. F., & Curtis, A. (2004). *Mass persuasion: The social psychology of a war bond drive* (1st Howard Fertig pbk. ed). H. Fertig.

- Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, 22(2), 103-122. doi.org/10.1002/ejsp.2420220202
- Neis, M., & Mara, M. (2020). Social Bots – Meinungsroboter im Netz. In M. Appel (Hrsg.), *Die Psychologie des Postfaktischen: Über Fake News, „Lügenpresse“, Clickbait & Co.* (S. 189-203). Springer Berlin Heidelberg. doi.org/10.1007/978-3-662-58695-2_17
- Nekmat, E. (2020). Nudge Effect of Fact-Check Alerts: Source Influence and Media Skepticism on Sharing of News Misinformation in Social Media. *Social Media + Society*, 6(1). doi.org/10.1177/2056305119897322
- Newman, O. (1972). *Defensible Space: Crime prevention through urban design*. Macmillan Press LTD.
- Nyhan, B., Porter, E., Reifler, J., & Wood, T. J. (2020). Taking Fact-Checks Literally But Not Seriously? The Effects of Journalistic Fact-Checking on Factual Beliefs and Candidate Favorability. *Political Behavior*, 42(3), 939-960. doi.org/10.1007/s11109-019-09528-x
- Nyhan, B., & Reifler, J. (2010). When Corrections Fail: The Persistence of Political Misperceptions. *Political Behavior*, 32(2), 303-330. doi.org/10.1007/s11109-010-9112-2
- Paris, R. (1989). Der kurze Atem der Provokation. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 41(1989).
- Pease, K., & Graham, F. (2014). What Have Criminologists Done for Us Lately? In M. Gill (Hrsg.), *The Handbook of Security* (S. 65-88). Palgrave Macmillan.
- Pease, K., & Laycock, G. (2012). Ron and the Schiphol Fly. In R. V. Clarke, N. Tilley, & G. Farrell (Hrsg.), *The reasoning criminologist: Essays in honour of Ronald V. Clarke* (S. 172-183). Routledge.
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *Journal of Experimental Psychology: General*, 147(12), 1865-1880. doi.org/10.1037/xge0000465
- Pennycook, G., & Rand, D. G. (2019). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39-50. doi.org/10.1016/j.cognition.2018.06.011
- Pennycook, G., & Rand, D. G. (2020). Who falls for fake news? The roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of Personality*, 88(2), 185-200. doi.org/10.1111/jopy.12476
- Popper, K. R. (2003). *Gesammelte Werke: Band 6: Die offene Gesellschaft und ihre Feinde. Band II: Falsche Propheten: Hegel, Marx und die Folgen* (H. Kieseewetter, Hrsg.; 8., A.). Mohr Siebeck.
- Raab, M. H., Ortlieb, S. A., Auer, N., Guthmann, K., & Carbon, C.-C. (2013). Thirty shades of truth: Conspiracy theories as stories of individuation, not of pathological delusion. *Frontiers in Psychology*, 4. doi.org/10.3389/fpsyg.2013.00406
- Rijkema, B., & Asbury, A. (2019). *Militant Democracy: The limits of democratic tolerance*. Routledge.
- Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. *European Journal of Communication*, 35(3), 213-229. doi.org/10.1177/0267323120922066
- Rosa, H. (2016). *Resonanz: Eine Soziologie der Weltbeziehung* (Erste Auflage). Suhrkamp.
- Roth, J. (2020, November 12). Parler: In der Echokammer der Trump-Anhänger. *Die Zeit*. <https://www.zeit.de/politik/ausland/2020-11/parler-donald-trump-unterstuetzer-plattform-facebook-twitter-fake-news> (2021, Juni 25).
- RT Deutsch. (2020). *Die Könige der Desinformation: Wie die ARD in vier Minuten vierzehnmal manipuliert*. RT Deutsch. <https://deutsch.rt.com/meinung/101646-koenige-der-desinformation-wie-die-ard-vier-minuten-vierzehnmal-manipuliert/> (2021, Juni 25).
- Rückert, C. (2018). Fake News und Social Bots – Demokratieschutz durch Strafrecht? In A. H. Albrecht, J. Geneuss, A. Giraud, & E. Pohlreich (Hrsg.), *Strafrecht und Politik* (S. 167-186). Nomos Verlagsgesellschaft mbH & Co. KG. doi.org/10.5771/9783845283760-167

- Sängerlaub, A. (2018). *Feuerwehr ohne Wasser? Möglichkeiten und Grenzen des Fact-Checkings als Mittel gegen Desinformation*. Stiftung Neue Verantwortung. https://www.stiftung-nv.de/sites/default/files/grenzen_und_moeglichkeiten_fact_checking.pdf (2021, Juni 25).
- Satariano, A., Wakabayashi, D., & Kang, C. (2018, August 28). Trump Accuses Google of Burying Conservative News in Search Results. *The New York Times*. <https://www.nytimes.com/2018/08/28/business/media/google-trump-news-results.html> (2021, Juni 25).
- Schmid, F. (2016, Dezember 19). Abgeschafft: Österreich hatte bis 2016 Gesetz gegen Fake-News. *DER STANDARD*. <https://www.derstandard.at/story/2000049437140/abgeschafft-oesterreich-hatte-bis-2016-gesetz-gegen-fake-news> (2021, Juni 25).
- Schramme, T. (2016). Die politische Quacksalberei des libertären Paternalismus. *Zeitschrift für Praktische Philosophie*, 3(1), 531-558. doi.org/10.22613/zfpp/3.1.15
- Schulz, A., Wirth, W., & Müller, P. (2020). We Are the People and You Are Fake News: A Social Identity Approach to Populist Citizens' False Consensus and Hostile Media Perceptions. *Communication Research*, 47(2), 201-226. doi.org/10.1177/0093650218794854
- Scott, M. (2020, Oktober 27). Despite cries of censorship, conservatives dominate social media. *POLITICO*. <https://www.politico.com/news/2020/10/26/censorship-conservatives-social-media-432643> (2021, Juni 25).
- Smith, J. (2017, Dezember 20). *Designing Against Misinformation*. Medium. <https://medium.com/facebook-design/designing-against-misinformation-e5846b3aa1e2> (2021, Juni 25).
- Soldt, R. (2020, Oktober 6). Studie zu Corona-Protesten: Wer sind die „Querdenker“? *FAZ.NET*. <https://www.faz.net/1.6989153> (2021, Juni 25).
- Stahl, B. C. (2006). On the Difference or Equality of Information, Misinformation, and Disinformation: A Critical Research Perspective. *Informing Science: The International Journal of an Emerging Transdiscipline*, 9, 83-96. doi.org/10.28945/473
- Sunstein, C. R. (2015). *Why nudge? The politics of libertarian paternalism* (paperback). Yale University Press.
- Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Thornhill, C., Meeus, Q., Peperkamp, J., & Berendt, B. (2019). A Digital Nudge to Counter Confirmation Bias. *Frontiers in Big Data*, 2. doi.org/10.3389/fdata.2019.00011
- Times, T. N. Y. (2017, Januar 11). Trump Calls CNN 'Fake News'. *The New York Times*. <https://www.nytimes.com/video/us/politics/10000004865825/trump-calls-cnn-fake-news.html> (2021, Juni 25).
- Todd, P. M., & Gigerenzer, G. (2012). What is ecological rationality? In *Ecological rationality: Intelligence in the world* (S. 3-30). Oxford University Press. doi.org/10.1093/acprof:oso/9780195315448.003.0011
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, New Series*, 185(4157), 1124-1131.
- Urman, A., & Katz, S. (2020). What they do in the shadows: Examining the far-right networks on Telegram. *Information, Communication & Society*, 1-20. doi.org/10.1080/1369118X.2020.1803946
- Vestergaard, M., & Hendricks, V. F. (2017). Verlorene Wirklichkeit? An der Schwelle zur postfaktischen Demokratie. *APuZ*, 67(13). <https://www.bpb.de/apuz/245212/verlorene-wirklichkeit-an-der-schwelle-zur-postfaktischen-demokratie> (2021, Juni 25).
- Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. doi.org/10.1126/science.aap9559
- Walton, G., & Hepworth, M. (2011). A longitudinal study of changes in learners' cognitive states during and following an information literacy teaching intervention. *Journal of Documentation*, 67(3), 449-479. doi.org/10.1108/0022041111124541

- Wang, Y., Leon, P. G., Acquisti, A., Cranor, L. F., Forget, A., & Sadeh, N. (2014). A field trial of privacy nudges for facebook. *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems - CHI '14*, 2367-2376. doi.org/10.1145/2556288.2557413
- Wang, Y., Leon, P. G., Scott, K., Chen, X., Acquisti, A., & Cranor, L. F. (2013). Privacy nudges for social media: An exploratory Facebook study. *Proceedings of the 22nd International Conference on World Wide Web - WWW '13 Companion*, 763-770. doi.org/10.1145/2487788.2488038
- Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policy making*. Europäischer Rat. <https://rm.coe.int/information-disorder-toward-an-interdisciplinary-framework-for-research/168076277c> (2021, Juni 25).
- Weber, S., & Knorr, E. (2020). Kognitive Verzerrungen und die Irrationalität des Denkens. In M. Appel (Hrsg.), *Die Psychologie des Postfaktischen: Über Fake News, „Lügenpresse“, Clickbait & Co.* (S. 103-115). Springer Berlin Heidelberg. doi.org/10.1007/978-3-662-58695-2_10
- WhatsApp FAQ - Weiterleitungsbeschränkungen. (o. J.). WhatsApp.com. <https://faq.whatsapp.com/general/coronavirus-product-changes/about-forwarding-limits/> (2020, Juni 24).
- Wilson, J. Q., & Kelling, G. L. (1982). Broken Windows. *The Atlantic*. <https://www.theatlantic.com/magazine/archive/1982/03/broken-windows/304465/> (2021, Juni 25).
- Wood, T., & Porter, E. (2019). The Elusive Backfire Effect: Mass Attitudes' Steadfast Factual Adherence. *Political Behavior*, 41(1), 135-163. doi.org/10.1007/s11109-018-9443-y
- Wortley, R., & Tilley, N. (2018). Does situational crime prevention require a rational offender? In D. M. Reynald & B. Leclerc (Hrsg.), *The future of Rational Choice for Crime Prevention* (S. 8-27). Routledge. doi.org/10.4324/9781315441207-2
- Wortley, R., & Townsley, M. (Hrsg.). (2017a). *Environmental criminology and crime analysis* (Second Edition). Routledge, Taylor & Francis Group.
- Wortley, R., & Townsley, M. (2017b). Environmental criminology and crime analysis. Situating the theory, analytic approach and application. In R. Wortley & M. Townsley (Hrsg.), *Environmental criminology and crime analysis* (Second Edition, S. 1-25). Routledge, Taylor & Francis Group.

Kontakt | Contact

Sebastian Enghofer | Universität Regensburg | Lehrstuhl für Vergleichende Politikwissenschaft | sebastian.enghofer@politik.uni-regensburg.de