

How easy it is to hate on the net: who the haters are and why they do it

Quanto è facile odiare in rete: chi sono gli odiatori e perché lo fanno

Gabriele Rocca | Alessio Lechiara | Giovanni Battista Gnecco |
Caterina Magni | Giulia Gibelli | Alfredo Verde

OPEN ACCESS

Double blind peer review

How to cite this article: Rocca G. et al. (2024). How easy it is to hate on the net: who the haters are and why they do it. *Rassegna Italiana di Criminologia*, XVIII, 4, 307-316 <https://doi.org/10.7347/RIC-042024-p307>

Corresponding Author: Gabriele Rocca email: gabriele.rocca@unige.it

Copyright: © 2024 Author(s). This is an open access, peer-reviewed article published by Pensa Multimedia and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. *Rassegna Italiana di Criminologia* is the official journal of Italian Society of Criminology.

© The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

Received: 23.12.2023

Accepted: 29.06.2024

Published: 31.12.2024

Pensa MultiMedia

ISSN 1121-1717 (print)

ISSN 2240-8053 (on line)

[doi10.7347/RIC-042024-p307](https://doi.org/10.7347/RIC-042024-p307)

Abstract

The phenomenon of cyberhate is one of the most obvious side effects of the advent of social network. According to an analysis conducted by Amnesty International in 2022, the phenomenon is widespread and steadily increasing.

This paper aims to focus on the personality of online haters and their motives for doing so.

Together with the intrinsic characteristics of the web, such as anonymity, physical absence, asynchrony and the consequent difficulty in containing the phenomenon, the personal characteristics of the individual and the desires for approval and social identification contribute to the spread of cyberhate.

At the end of this work we will try to propose new strategies to intervene on the phenomenon, intercepting haters and supporting victims. However, the most effective solution remains prevention in childhood: particularly if it focuses on strengthening victims' coping strategies and on the digital and emotional literacy of potential haters.

Keywords: social network, cyberhate, haters, prevention, treatment

Riassunto

Il fenomeno del *cyberhate* è uno dei più evidenti effetti collaterali dell'avvento dei social network. Secondo un'analisi condotta da Amnesty International nel 2022, il fenomeno è molto diffuso e in costante aumento. Questo contributo intende focalizzarsi sulla personalità degli odiatori online e sui motivi che li spingono ad agire in questo modo. Insieme alle caratteristiche intrinseche del web, come l'anonimato, l'assenza fisica, l'asincronia e la conseguente difficoltà di contenimento del fenomeno, concorrono alla diffusione del *cyberhate* le caratteristiche personali del singolo e i desideri di approvazione e identificazione sociale.

Alla fine del presente lavoro cercheremo di proporre nuove strategie per intervenire sul fenomeno, intercettando gli haters e sostenendo le vittime. La soluzione più efficace rimane tuttavia la prevenzione in età evolutiva: in particolare se essa si concentra sul potenziamento delle strategie di coping delle vittime e sull'alfabetizzazione digitale ed emotiva dei potenziali odiatori.

Parole chiave: social network, odio on line, odiatori, prevenzione, trattamento

Gabriele Rocca, Health Sciences Department, University of Genoa | Alessio Lechiara, Neurosciences Department, University of Genoa | Giovanni Battista Gnecco, Neurosciences Department, University of Genoa | Caterina Magni, Neurosciences Department, University of Genoa | Giulia Gibelli, Health Sciences Department, University of Genoa | Alfredo Verde, Health Sciences Department, University of Genoa

How easy it is to hate on the net: who the haters are and why they do it

Introduzione

Negli ultimi anni lo sviluppo tecnologico e il crescente utilizzo del web, in particolare dei social network, hanno cambiato le abitudini delle persone. Oggi, esistono “app” che permettono di svolgere molti compiti, tra cui vendere oggetti, trovare un lavoro o un partner e gestire le proprie risorse economiche.

I social network sono quindi divenuti i luoghi virtuali di maggiore impatto sulla nostra vita: Whatsapp, Facebook, Instagram, Twitter, Snapchat, TikTok, Telegram, Youtube e tanti altri offrono la possibilità di connettersi con persone lontane in modo veloce ed efficace, di accedere a informazioni su qualsiasi argomento e di disporre di uno spazio di espressione personale tramite la pubblicazione di foto e più o meno brevi narrative relative a sé stessi, alle proprie relazioni personali e lavorative, alle proprie convinzioni ideologiche e politiche, alle proprie preferenze: uno spazio virtuale potenzialmente infinito e disponibile a tutti.

In questo spazio di presenza senza lacune, il bene si converte in male, l'amore in odio, molto facilmente. Essendo nati per creare relazioni tra le persone, i social media possono infatti veicolare anche aspetti negativi e permettere alle persone di stigmatizzare quanto non amano, o addirittura disprezzano e odiano.

L'odio online costituisce quindi il risvolto della relazionalità resa sempre più maneggevole, scorrevole e consumabile, tanto che mai come oggi la rete viene utilizzata anche dalle frange più radicali di odiatori: razzisti, terrapiattisti, creazionisti, complottisti, estremisti politici, incels, non esitano a utilizzare le comunicazioni online per esprimere disprezzo verso i loro avversari. La rete, quindi, diventa il veicolo di potenti difese proiettive e, parallelamente, anche il modo per mondarsi la coscienza e proclamare la propria verità al mondo.

Secondo Zhang e Luo (2019), l'odio online può essere definito come la distorsione dei contesti comunicativi disponibili in rete al fine di diffondere messaggi, commenti e discorsi saturi di disprezzo, odio e violenza nei confronti di persone o gruppi sulla base delle loro caratteristiche di razza, colore, etnicità, genere, orientamento sessuale, nazionalità, religione od orientamento politico.

Il Centro per le pari opportunità e l'opposizione al razzismo di Bruxelles propone una definizione più ampia definendo tali fenomeni come “*propagazione dell'incitamento all'odio su Internet*”.

Esso può presentarsi sotto varie forme: fenomeni di bullismo (il c.d. cyberbullismo), insulti, discriminazioni contro individui o gruppi di persone sulla base del colore

della pelle, della presunta razza, dell'origine etnica, del sesso, dell'orientamento sessuale o delle convinzioni politiche o religiose (Rieger et al., 2021).

L'odio online include anche molestie deliberate nei confronti di individui specifici in nuove forme di espressione, che vanno dal *trolling*, il persistente tormento nei confronti di qualcuno attraverso provocazioni irritanti e continui spostamenti del discorso (Rieger et al., 2017), al *doxxing*, la documentazione pubblica delle informazioni personali offline di qualcuno (Eckert et al., 2020), passando per *meme* degradanti o offensivi, per le *shitstorm* (“tempeste” di insulti di massa contro uno specifico individuo, gruppo o organizzazione che si risolvono in poche ore o giorni) fino ad arrivare alla condivisione non consensuale di materiale intimo.

La ricerca ha inoltre distinto dagli attacchi contro i singoli individui appena citati, quelli contro l'identità collettiva o sociale detti *hate speech* (Gagliardone et al., 2016).

Altre forme ancora di odio online sono il *cyber-stalking*, ossia le condotte persecutorie online, e il *flaming*, definito come attacco deliberatamente ostile e opprimente nei confronti di uno specifico individuo.

I numeri evidenziano l'ampia diffusione del fenomeno: in Unione Europea l'80% delle persone ha incontrato fenomeni di incitamento all'odio online e il 40% si è sentito attaccato o minacciato tramite l'utilizzo dei social network (Gagliardone et al., 2015).

Nel 2020 Amnesty International ha condotto un'analisi quantitativa e qualitativa del fenomeno del *cyberhate* basandosi su oltre 22 milioni di contenuti tra Facebook e Twitter. I risultati parlano chiaro: il fenomeno è in aumento. Nei dibattiti sui social, un commento su dieci è discriminatorio e offensivo: diffusi sono i contenuti sessisti, islamofobi e razzisti. Infine, viene rilevato come i discorsi d'odio siano aumentati del 40% rispetto agli anni precedenti.

Nel 2021, VoxDiritti – Osservatorio Italiano sui diritti (un'associazione che ha lo scopo di vigilare sul rispetto dei diritti delle fasce deboli e promuovere l'eguaglianza sostanziale) ha svolto un'indagine per monitorare l'uso dell'odio online in Italia attraverso l'analisi dei *tweet*: è emerso che il 69% dei messaggi inviati era discriminatorio.

Una ricerca svolta nel 2016 sempre su Twitter, in un periodo di 10 mesi, ha rilevato la diffusione delle espressioni violente, di denigrazione del genere o della disabilità, antisemite, omofobe e razziali (Musto et al., 2016). Le più diffuse sembrano essere quelle violente, seguite subito dopo dal tema della disabilità e dai commenti razzisti.

Tutto questo non risparmia alcun Paese: una ricerca svolta negli Stati Uniti negli anni 2019-2020 per moni-

torare l'odio espresso contro le persone di origine asiatica a causa del Covid-19 testimonia l'enorme diffusione dell'odio online, e rileva quanto i social abbiano contribuito alla diffusione dell'ondata paranoide proiettiva, volta alla ricerca dei presunti colpevoli della pandemia da virus che ha colpito tutto il mondo (Hohl, 2022).

In generale, è comunque molto difficile reperire dati certi sulla diffusione del fenomeno perché esso è caratterizzato da un altissimo numero oscuro, aspetto paradossale considerando l'estrema visibilità di questi comportamenti.

Comprendere il fenomeno del *cyberhate* richiede una categorizzazione dei vari comportamenti violenti sul web, avendo cura di preservare la libertà di espressione degli utenti e garantendo allo stesso tempo la protezione da qualsiasi forma di discriminazione.

Incoraggiando il diritto fondamentale degli utenti che è, come appena sottolineato, la libertà di espressione, i social network implicano il rischio di dare la possibilità di odiare una persona o un gruppo esprimendo opinioni discriminatorie o violente. L'identificazione dei potenziali bersagli del discorso di odio è la chiave per distinguere l'odio online da argomentazioni che rappresentano punti di vista politici protetti dai diritti di libertà di espressione (Meza, 2018).

La difficoltà di fare ordine in questa materia deriva principalmente dalla mancanza di definizioni chiare dei comportamenti violenti online (Peterson & Densley, 2017). Ad esempio, è proprio dalla complessità nel definire il concetto di cyberbullismo che Grigg (2010) giunge alla conclusione che sia necessario passare ad un concetto più ampio di *cyberhate* che includa tutti i comportamenti online caratterizzati da un alto livello di aggressività, introducendo così il concetto di cyber-aggressione.

Un'altra distinzione importante da tenere a mente è quella tra odio online e crimine d'odio: ci sono casi in cui il *cyberhate* può trasformarsi in un vero e proprio atto illecito.

In Italia non è stata formulata dal legislatore una definizione specifica di crimine d'odio, in quanto i comportamenti online vengono ricondotti alle comuni fattispecie dei delitti contro la persona. Ai fini della ricerca, si può utilizzare la definizione elaborata dall'Ufficio per le Istituzioni Democratiche e i Diritti Umani (Odihr) dell'Organizzazione per la Sicurezza e la Cooperazione in Europa (Osce) in base alla quale il crimine d'odio è un reato, commesso contro un individuo e/o beni ad esso associati, motivato da un pregiudizio che l'autore nutre nei confronti della vittima, in ragione di una "caratteristica protetta" di quest'ultima. Il crimine d'odio, quindi, si caratterizza per la presenza di due elementi: un fatto previsto dalla legge penale come reato (cosiddetto reato base) e la motivazione di pregiudizio rispetto a una caratteristica che fa sentire l'autore ingiustamente trascurato per la protezione fornita dal diritto a un soggetto più debole (Chirico et al., 2020).

Chi sono gli odiatori? Perché odiano?

Il presente lavoro intende analizzare i principali contributi sulla personalità degli haters, per poi proporre soluzioni in termini di intervento e prevenzione. Cercheremo di capire chi sono ma soprattutto perché lo fanno e cosa contraddistingua la loro controparte, le vittime.

In uno studio condotto da Shachaf e Hara nel 2010, gli autori hanno identificato come giustificazione dei comportamenti aggressivi esperienze quali la noia, o ancora la ricerca di attenzione, la vendetta, il piacere e il desiderio di fare un danno alla comunità rispetto alla quale gli *haters* si percepiscono come *outsider* o addirittura come oppositori.

Sembra che gli individui caratterizzati da impulsività abbiano maggiori probabilità di offendere o minacciare gli altri online, mostrando meno capacità di riflettere sui propri vissuti, fattore che potrebbero portarli a fallire più spesso di altri nell'inibire il proprio comportamento (Van Royen et al., 2017). L'impulsività è altresì correlata ai disturbi di personalità, ed è possibile che l'associazione con i reati d'odio sia in parte spiegata dai tratti della personalità antisociale (Few et al., 2015). Bassi livelli di autocontrollo, connessi a una disponibilità al rischio e un'avversione a ritardare la gratificazione, sono positivamente associati alla produzione di odio *online*, suggerendo come l'impulsività giochi un ruolo importante nella decisione di produrre *cyberhate* (Bernatzky et al. 2022).

L'odio online è inoltre associato a sintomi internalizzanti, che possono causare intensi stati affettivi negativi e una regolazione emotiva e comportamentale disfunzionale (Selby et al., 2008), che possono a loro volta portare ad avere un comportamento aggressivo (Krakowsk et al., 2013), specialmente sui social media dove, per via dell'anonimato e della mancanza di contatto fisico con le vittime, l'autore del comportamento si sente più protetto (Peterson et al., 2017).

Altre caratteristiche attribuibili agli *haters* sono un senso di identità scarso, bassa consapevolezza di sé (Chao & Tao, 2012), mancanza di autostima (Bishop, 2013), tratti psicopatici (Sorokowski et al, 2020) e alessitimia (Pizzi, 2022).

Un fattore importante dei social network è l'identificazione sociale con i gruppi online (Kenski et al., 2017). La comunicazione sul web facilita l'interazione in cui le persone vedono sé stesse e gli altri in termini di gruppo piuttosto che di identità individuale (Tang & Fox, 2016), cosicché il contenuto offensivo o minaccioso non viene inteso come espressione contro un altro in quanto tale, semmai come membro di un gruppo contrapposto. L'appartenenza a gruppi con idee omogenee induce ostilità verso gruppi che si discostano da tali contenuti (Densley & Peterson, 2018).

Attenzione particolare deve essere posta nei confronti degli adolescenti, popolazione nata e cresciuta con l'avvento dei social network. Le analisi sulle differenze di genere hanno mostrato come i ragazzi siano più propensi a usare la violenza verbale online e offline rispetto alle

ragazze (Pace et al., 2021). Anche la presenza di legami affettivi familiari è correlata a una ridotta probabilità di produrre comportamenti di odio *online*, così come avviene rispetto ai comportamenti devianti e criminali adulti, sia offline che online (Bernatzky et al., 2022); mentre i legami amicali e l'appartenenza a comunità religiose non costituiscono una variabile rilevante. Gli adolescenti online spesso perdono il senso del limite e delle proprie responsabilità senza pensare agli effetti che il loro comportamento può avere sulle vittime e le ricadute penali dello stesso (Chetty & Alathur, 2018). La letteratura ha anche evidenziato che gli individui che commettono atti di odio sono caratterizzati da un'elevata presenza di distorsioni cognitive che possono essere alla base del mantenimento di quegli stessi atteggiamenti (Grieve & Panebianco, 2013).

Un altro fattore che può influire sullo sviluppo di odio online è l'esperienza di vittimizzazione subita, quando le vittime iniziano a vendicarsi della persona che perpetra il *cyberhate*; i sentimenti di 'impotenza percepita elevano poi il rischio di sviluppare schemi disadattivi e di cronicizzare lo stato di vittimizzazione (Calvete et al., 2018), creando circoli viziosi e aumentando così il rischio di coinvolgimento. Questa ipotesi è ulteriormente supportata da un'altra ricerca che mostra che molti adolescenti perpetrano discorsi di odio *online* come reazione all'essere stati vittimizzati, per vendicarsi e combattere i sentimenti di impotenza. Come si può notare, questi dati confermano anche per l'odio online quanto la vittimizzazione favorisce la perpetrazione di comportamenti antisociali, (c.d. "ipotesi del ciclo della violenza": Widom, 1989). In effetti, la ricerca criminologica e psicologica ha stabilito che l'essere una vittima e l'essere un carnefice sono fortemente associati (Forsman & Langstrom, 2012).

Altre ricerche hanno peraltro individuato quanto siano le caratteristiche specifiche della rete a facilitare il propagarsi di odio e violenza sul *web*.

Un primo elemento è legato all'assenza fisica dell'altro, presente solo su alcuni canali sensoriali, che può portare addirittura alla cosiddetta *disattivazione dei cinque sensi*. Spesso, nemmeno l'udito e la vista sono disponibili (Tripathi, 2017), e l'altro si manifesta solo a livello verbale, con comunicazioni istantanee che possono portare all'attivazione massiccia delle difese proiettive. Secondariamente, la difficoltà o la mancanza di un processo sanzionatorio immediato può portare ad una valutazione di minore gravità dei propri atti violenti (Lumsden & Morgan, 2017). La velocità con cui si può esprimere l'emotività online è un altro fattore: gli autori della violenza difficilmente vengono contenuti da reazioni ai loro comportamenti, o feedback negativi da parte dell'altro o del sistema (Tripathi, 2017). Anche l'iper-accessibilità, la non simultaneità, la pervasività nella vita di ognuno e la mancanza di un codice etico di regole precise (Verbena, 2020). Sul *web* è poi spesso possibile non essere riconosciuti grazie all'anonimato tramite l'utilizzo di siti in cui si può celare totalmente la propria identità o creare un profilo falso (Patton, 2014). Questo sembrerebbe favorire la de-

individuazione, ossia la perdita di consapevolezza e autocontrollo che si sperimenta nei gruppi e in alcune situazioni sociali, che a sua volta aumenterebbe la probabilità di mettere in atto di comportamenti violenti online (Peterson & Densley, 2017): in altre parole, la deafferentazione sensoriale (che riduce l'altro a un entità spoglia della maggior parte delle caratteristiche umane), può facilitare il collasso dell'empatia e azzerare la possibilità di riconoscere l'altro in quanto tale (caduta della teoria della mente: Franco et al., 2017), come risulta anche da alcune meta-analisi (cfr. Jolliffe et al., 2004). Altre ricerche (Peterson et al., 2017, Steffgen et al., 2011) confermano che la difficoltà ad attivare un processo empatico verso l'altro è legata all'impossibilità di vedersi in faccia: d'altronde, le basi della capacità di comunicare e relazionarsi risiedono nelle inferenze che facciamo ogni momento sugli stati d'animo propri e altrui, cogliendo ed interpretando la prosodia e la mimica facciale dell'altro (Franco et al., 2017). Secondo Pasta (2018), addirittura alcuni "nativi digitali", per la deafferentazione provocata dalla minore quantità di interazione fisica con il mondo sociale dovuta al troppo tempo trascorso in rete, potrebbero aver sviluppato capacità empatiche ridotte: questa condizione sembrerebbe promuovere l'hate speech online. Dalle analisi è emerso che, per contrastare gli hate speech, l'attività che raggiunge migliori risultati a lungo termine interessa la cultura in termini di educazione, informazione e comunicazione, soprattutto se avviate in età evolutiva, in contesti come la scuola e la famiglia, in una parola la prevenzione.

Per converso, anche la presenza di scarse capacità empatiche *offline* sembrerebbe riflettersi sul *web*, e per questo pare necessario educare all'utilizzo dei *social network* e alla creazione di contenuti in maniera corretta e consapevole – la cosiddetta alfabetizzazione digitale – prevedendo anche una alfabetizzazione emotiva, relativa cioè al riconoscere le emozioni altrui tramite le espressioni e la gestualità (Pizzi, 2022). In base al rapporto di EU Kids Online, uno studio europeo che analizza le condotte dei giovani sul *web* in 19 paesi per studiare i rischi, le opportunità e il livello di sicurezza della rete, sembrerebbe inoltre che i nativi digitali manchino delle competenze critiche necessarie per controllare la veridicità delle informazioni rinvenute *online*, con il rischio di diffonderle a loro volta; e rischiando anche di diventare maggiormente tolleranti rispetto al linguaggio violento (Mascheroni & Ólafsson, 2018).

In sintesi, quindi, la letteratura più recente sostiene che quando perpetratore e vittima sono online i comuni meccanismi che portano a riconoscere l'altro come un altro Sé con pensieri, emozioni e sentimenti propri sono bloccati dalla natura stessa del media.

Alle spiegazioni psicologiche, da un lato, e situazionali, dall'altro, si sono recentemente aggiunte ipotesi ancora più complesse, che aggiungono ai fattori sovradescritti anche i fattori relazionali, connettendo le caratteristiche di personalità con le risposte da parte degli altri significativi. Al momento il modello più gettonato è quello bio-

psico-sociale, tratto da un recente contributo di Walther (2023), secondo cui i tratti psicologici e personologici appena analizzati si interseccherebbero con le teorie dell'approvazione sociale e dell'identificazione sociale: le motivazioni e le gratificazioni di coloro che pubblicano messaggi di odio non deriverebbero principalmente dal piacere legato all'inimicarsi le presunte vittime; piuttosto le persone genererebbero messaggi di odio principalmente per ottenere ammirazione e lodi da altri utenti (Walther, 2032). Tutto ciò non tanto per propagandare le proprie credenze e convincere ipotetici seguaci, ma in cerca di riconoscimento, status e prestigio (Douglas et al., 2005).

Riassumendo le teorie proposte finora, si può asserire che un individuo con certe caratteristiche psicologiche cercherebbe di sentirsi approvato dai consociati e di ricevere input positivi; questo accadrebbe soprattutto a persone caratterizzate da narcisismo immaturo e connotate da un forte desiderio di sentirsi riconosciuti e approvati. Infine, per colmare il proprio scarso senso di identità personale, l'individuo utilizzerebbe l'odio come veicolo per entrare a far parte di un gruppo consolidato, e rafforzare la propria autostima.

Passando ora a descrivere le caratteristiche della vittima, è possibile identificare le persone più a rischio di subire atti di *cyberhate*?

Considerando che il fenomeno è relativamente nuovo, la metodologia della ricerca nel campo è ancora poco standardizzata e questo potrebbe rendere conto dell'ampia variabilità dei valori di prevalenza (Hamm et al., 2015).

Anche gli studi sulle differenze di genere nelle vittime non sono completamente concordi: molti concludono che le ragazze sarebbero più frequentemente vittimizzate rispetto ai ragazzi (cfr. ad esempio Barranco *et al.*, 2022), ma altri autori sostengono che non ci siano differenze statisticamente significative fra i sessi. Ad ogni modo, in una ricerca condotta da Sourander sul cyberbullismo (Sourander et al, 2010), si è concluso che i ragazzi sono più propensi delle ragazze ad essere parte attiva, mentre le ragazze hanno maggiori probabilità di essere vittimizzate. I risultati di ricerche canadesi rilevano infatti che le ragazze hanno una probabilità doppia rispetto ai ragazzi di essere vittime di cyberbullismo.

Per quanto concerne le modalità con cui la violenza online si estrinseca: le ragazze ricevono attacchi che criticano il loro aspetto fisico, le relazioni sentimentali o la loro popolarità, e a ciò seguono comportamenti di esclusione sociale; i ragazzi spesso ricevono messaggi omofobici o transfobici, o vengono scherniti sulla base dell'aspetto fisico (Hoff & Mitchell, 2009).

Un'altra variabile è quella del tempo trascorso online: e maggiore è il tempo che i giovani trascorrono sui media elettronici, maggiori sono le probabilità di essere vittime di odio online (Sampasa-Kanyinga, 2014).

Quanto agli effetti dell'odio online, come è noto, essere vittima di bullismo in età adolescenziale è considerato uno dei tanti possibili fattori responsabili della genesi di disturbi psicologici, autolesionismo e suicidio (Lereya et al., 2015). Le situazioni più gravi e impattanti sul piano

psicologico riguardano l'invio online di immagini o video in cui le vittime sono messe in imbarazzo, svalutate, dileggiate o umiliate, e la diffusione non consensuale di foto e video intimi (Smith et al., 2008, Miller & Shirley, 2009; Ybarra et al., 2006). Le manifestazioni di odio attraverso il web hanno un grave impatto sulla salute mentale delle vittime e determinano conseguenze paragonabili a quelle dal bullismo nella realtà, oltre che a condividere alcune caratteristiche, come gli insulti verbali e l'esclusione sociale. Alcuni autori sostengono che il cyberbullismo sia addirittura in grado di generare maggiore distress psicologico rispetto al bullismo tradizionale per la completa invasione della privacy dell'individuo, per la rapidità di trasmissione di eventuali contenuti indesiderati, e per la difficoltà nel controllare e perseguire tali comportamenti a causa dell'anonimato dietro cui spesso si nasconde chi diffonde e incita all'odio online; tutto ciò determinerebbe una maggior probabilità di sviluppare disturbi mentali, autolesionismo e comportamenti suicidari (Bonanno et al., 2013).

Come prevenire e arginare il fenomeno?

Un fenomeno ampio e complesso come l'odio online va affrontato con norme giuridiche adeguate, ma richiede anche e soprattutto un lavoro di tipo culturale. In questo senso, le migliori iniziative sono riscontrabili a livello educativo e di comunicazione (Pizzi, 2022).

Negli ultimi anni numerosi sono stati gli interventi di educazione e formazione e progetti di ricerca dedicati alla prevenzione e al contrasto all' *hate speech*, con lo scopo di implementare le capacità di critica di chi naviga e insegnargli a comunicare nella maniera corretta ed efficace, prevenendo la produzione di contenuti violenti.

Possiamo identificare due strategie di prevenzione e lotta all'odio online: rafforzare il quadro normativo al fine di sviluppare programmi di identificazione automatica dei messaggi di odio per intervenire efficacemente e prontamente, ed educando all'utilizzo corretto delle tecnologie insegnando al contempo una corretta contro-comunicazione.

Dal punto di vista normativo, va detto che la consapevolezza della necessità di combattere l'intolleranza e la crescita dell'odio fra le componenti sociali si è sviluppata da tempo, anche da prima dello sviluppo della rete: infatti, già nel 1993 era stata creata presso il Consiglio d'Europa la Commissione Europea contro il Razzismo e l'Intolleranza (ECRI), al fine di fornire suggerimenti ai paesi membri allo scopo di promuovere l'eguaglianza tra tutti i cittadini e la protezione da qualsiasi forma di discriminazione.

Per arginare i fenomeni di cyberhate è stato necessario definire il ruolo e la possibile responsabilità delle multinazionali che gestiscono i grandi social media su internet, che contribuiscono alla diffusione e alla permanenza dei contenuti digitali in rete, ma che hanno anche la possibilità di rimuovere eventuali messaggi lesivi dei diritti di altri. A tale scopo, la Commissione Europea nel 2018 ha stipulato

assieme alle più importanti piattaforme web (come Facebook, Google, Microsoft, etc.), un “Codice di Condotta per contrastare l’illecito incitamento all’odio online” (come Facebook, Google, Microsoft)¹ che fa riferimento anche a un altro documento della Commissione Europea del 2008 riguardante la lotta contro le forme di razzismo e xenofobia; il testo prevede che queste ultime adeguino le proprie procedure interne garantendo una risposta rapida e incisiva in caso di *hate speech* sui rispettivi siti. L’Unione Europea compie annualmente un monitoraggio sul suddetto Codice, che dura sei settimane, allo scopo di verificare i risultati dell’azione di contenimento: i dati dimostrano che si sta ottenendo qualche successo in questo senso.

Non si è ritenuto sufficiente, però, il vincolo proposto con gli atti sopra citati, e si è reso necessario un atto direttamente valido in tutti i Paesi dell’Unione, di valore regolamentare². Per questo è stato stipulato il Digital Service Act (Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio dell’Unione Europea) del 19 ottobre 2022, proposto dalla Commissione Europea nel 2020, entrato in vigore nello scorso agosto solo per i più grandi social media, e che riguarderà anche tutti gli altri dal prossimo gennaio 2024, riguarderà anche tutti gli altri. L’atto predispone una disciplina dell’intero contesto web, tutela i diritti alla riservatezza dei dati degli utenti, e li protegge dagli abusi perpetrati dai gestori delle piattaforme nonché dai comportamenti attribuibili ad altri utenti; accentua la tempestività della protezione imponendo una sorveglianza continua e la messa in atti di interventi immediati ai più noti social media, mirando a creare uno spazio digitale più sicuro in cui i diritti fondamentali degli utenti siano protetti, inclusa la libertà di espressione, e prevede sanzioni pecuniarie di rilevante entità in caso di violazione da parte delle *corporations* che gestiscono le piattaforme online.

In Europa, attualmente, l’unico Paese con una specifica legislazione riguardante l’odio online è la Germania, mentre gli altri Paesi, compreso il nostro stanno studiando il problema e/o predisponendo o esaminando proposte di regolamentazione del campo.

Nel nostro Paese, la Camera dei Deputati ha costituito nel maggio 2016 la “Commissione Jo Cox sull’intolleranza, la xenofobia, il razzismo e i fenomeni di odio” e composta da parlamentari, esperti e rappresentanti di associazioni impegnati nella ricerca, che ha prodotto e ap-

provato una relazione finale dal titolo “La piramide dell’odio” (Camera dei Deputati, 2017). Il riferimento alla piramide propone una rappresentazione geometrica dell’odio online che parte dal linguaggio d’odio, passa per la discriminazione e gli stereotipi e le false rappresentazioni, per giungere infine agli *hate crimes*: si teorizza che sia necessario partire dalla base della piramide per prevenire escalation verso comportamenti concretamente dannosi. La relazione si concentra quindi sui metodi di prevenzione e lotta all’odio, proponendo lo sviluppo di misure di educazione del cittadino al rispetto e al dialogo interculturale e interreligioso, e la diffusione di informazione, tramite la promozione di narrative e campagne contro gli hate speech. Basandosi sull’esperienza degli altri Paesi, e riconoscendo a libertà di informazione e di parola sul web, la relazione suggerisce infine la necessità di imporre alle piattaforme online un’autoregolazione imponendo loro di prevenire o rimuovere i discorsi d’odio, e stabilendo la responsabilità dei provider e delle piattaforme nel caso di mancata rimozione tempestiva dei contenuti considerati e segnalati come lesivi dagli utenti.

Sulla base della relazione Jo Cox, quindi, nel marzo 2021, è stato depositato il Disegno di Legge n.2936/21, intitolato “Misure per la prevenzione e il contrasto della diffusione di manifestazioni d’odio mediante la rete internet” a prima firma Boldrini, che prevede passi importanti per arginare il fenomeno e, soprattutto, sollecita strategie di gestione del problema. Questo disegno di legge introduce procedure di monitoraggio per verificare l’eventuale illegittimità dei contenuti sulle piattaforme online tramite un resoconto ogni 24 ore, e un conseguente intervento tempestivo di rimozione degli *hate speech*, pur garantendo ragionevolmente libertà di espressione a tutti gli utenti; se le piattaforme non dovessero dotarsi di questi sistemi, incorrerebbero in sanzioni, inoltre sarebbero tenute a produrre relazioni annuali circa l’attività descritta qui sopra, nonché a renderle pubbliche. Il DDL introduce anche il concetto di “educazione civica digitale”, ovvero l’educazione all’uso della rete internet in modo consapevole e rispettoso, facendo specifico riferimento alla legislazione del 2019 sul Codice Rosso, che introduce l’articolo 612-ter del Codice Penale, il quale prevede anche in Italia il reato di diffusione illecita di immagini o video sessualmente espliciti, in cui rientra la condivisione non consensuale di materiale intimo.

Quanto alla prevenzione attiva dello sviluppo dei fenomeni di *hate speech*, sono state realizzate dal Consiglio e dalla Commissione Europea varie iniziative come la campagna “No Hate Speech Movement” per promuovere l’educazione ai diritti umani, e il progetto “CO.N.T.R.O.” (COunter Narratives Against Racism Online): quest’ultimo ha lo scopo di selezionare e analizzare, a livello delle singole nazioni e poi europeo, strumenti e pratiche utili a prevenire e combattere efficacemente il razzismo, la xenofobia e altre forme di intolleranza diffuse tramite l’odio online. Emerge, come una delle migliori metodologie, quella della promozione delle “counter-narratives” o “counter-speech” (letteralmente repli-

- 1 Il codice di condotta per contrastare l’illecito incitamento all’odio online è stato firmato il 31 maggio 2016 dalla Commissione e da Google (YouTube), Facebook, Twitter e i servizi per i consumatori ospitati da Microsoft (es. i servizi di giochi Xbox o LinkedIn). Nel 2018 e nel 2019 hanno aderito Instagram, Google+, Dailymotion, Snap e Jeuxvideo.com. Ciò significa che il codice copre ora il 96% della quota di mercato UE delle piattaforme online che possono essere toccate da contenuti di incitamento all’odio.
- 2 Come è noto, i Regolamenti dell’Unione europea sono norme applicabili ai soggetti giuridici di tutti gli Stati dell’Unione, di livello superiore rispetto alle leggi di ogni singolo paese.

care, controbattere), termine che sta a identificare l'opposizione agli *hate speech* online tramite contenuti prodotti in risposta a messaggi di odio.

Una metodologia, progredita con lo sviluppo di sistemi informatici in grado di analizzare il lessico online, è quella della c.d. "Mappa dell'Intolleranza", realizzata da Vox - Osservatorio Italiano sui Diritti ormai in successive: questa mappatura comprende la selezione di sei gruppi specifici (donne, persone omosessuali, migranti, persone con disabilità, ebrei e musulmani) e, tramite l'estrazione e la localizzazione di parole e messaggi online considerati sensibili, identifica i social network dove l'intolleranza è maggiormente diffusa. Il livello di aggressività delle comunicazioni online viene analizzato con riferimento alle categorie della scala MOAS (Modified Overt Aggression Scale) (Margari et al., 2005). In base a tale valutazione, Twitter si è rivelato essere il social in cui questi fenomeni accadono di più.

Prendendo l'esempio della piattaforma Facebook, quest'ultima ha provveduto ad adeguarsi al contesto normativo in evoluzione, considerando il *counter-speech* uno strumento fondamentale per contrastare i contenuti d'odio, riconosciuti dalla piattaforma come "contenuti che attaccano direttamente una persona o un gruppo di persone in virtù di razza, etnia, nazionalità di origine, affiliazione religiosa, orientamento sessuale, sesso, genere o identità di genere, disabilità o malattia". Tali contenuti vengono proibiti e rimossi sulla base degli Standard della Comunità richiamati da Facebook stesso e adottati come propria politica interna. Questo metodo ha sicuramente dei vantaggi essendo immediato, flessibile ed efficiente, anche se non sembra completamente adeguato nel tutelare l'importanza del principio di libera espressione *online*, essendo alcune volte il dibattito, la discussione e il disaccordo che si generano preferibili alla censura e alla rimozione dei contenuti. Inoltre, il *counter-speech* può addirittura diventare controproducente se questo è definito non costruttivo, ad esempio la risposta a un messaggio d'odio da parte di un altro utente tramite attacchi personali e offese.

Nel 2020 il governo italiano ha siglato un decreto ministeriale per la nascita di un gruppo di lavoro sull'Odio Online formato da sedici esperti, al fine per contrastare i fenomeni dell'*hate speech*. Secondo questo gruppo, seguendo i suggerimenti di Ziccardi (2016), è necessario creare un programma di consapevolezza digitale, non solo indirizzato ai più giovani, ma alla società tutta. In Italia dal 2019 qualcosa si è mosso: è stata introdotta in alcune scuole, nell'ambito dell'insegnamento di educazione civica, l'educazione all'uso consapevole delle tecnologie digitali.

Come fa notare Blaya (2019), non sono ancora stati sviluppati metodi di intervento standardizzati nei confronti degli *haters* e la maggior parte dei progetti di ricerca è focalizzata sulle vittime del *cyberhate*, attraverso la strutturazione di specifiche strategie di *coping*, il rafforzamento della confidenza in sé stessi e l'insegnamento di capacità di comunicazione e di reazione in maniera adeguata: tutto ciò per impedire alla vittima di sentirsi legittimata a uti-

lizzare l'*hate speech* come risposta, evitando la propagazione di un circolo vizioso (Wachs et al., 2022): in particolare, la strategia di coping tecnico (bloccare l'autore) e la ricerca di *distal advice* (richiedere consiglio a soggetti informati su come gestire la situazione) sembrano le più adatte per ridurre le probabilità di vittimizzazione, mentre l'impegnarsi nella relazione rischia, come abbiamo visto, di produrre la reazione negativa e trasformare la vittima in autore.

Alla luce di questi dati, potrebbe essere utile anche sperimentare programmi di prevenzione direttamente su potenziali autori di odio online, già in età evolutiva. Secondo alcuni studiosi (Perasso et al., 2020; Serrais, 2013; Nikiforou et al., 2013), i campi più adeguati dovrebbero essere quelli della promozione dell'alfabetizzazione digitale, del potenziamento delle competenze critiche, della gestione dell'emotività, del livello di comunicazione interna alla famiglia e della qualità del supporto dei genitori: sarebbe utile che i genitori promuovessero un corretto uso di internet, non limitato all'operatività e alla sicurezza ma inclusivo anche del modo di stare in relazione con gli altri in rete, cercando di trasmettere i valori del rispetto reciproco, dell'amicizia, dell'empatia (Machimbarrena et al., 2019). È ormai chiaro in letteratura che la famiglia è il luogo dove il bambino sviluppa, sentendosi oggetto delle cure dei genitori, l'accettazione di sé. Da qui poi, a cascata, si svilupperebbero tutta una serie di capacità, tra cui quella di comunicare i propri stati d'animo senza arrecare danno a nessuno (Suarez & Velez, 2018). Quest'ultima costituisce una competenza fondamentale nelle relazioni, anche in quelle online.

Ancora, un'altra caratteristica che andrebbe potenziata per la prevenzione è quella della prudenza, la quale, se presente, permette di riflettere sulle modalità di espressione dei propri bisogni e sulle conseguenze delle proprie azioni (Ahedo, 2021). Una buona dose di prudenza rende molto poco probabile l'utilizzo della stessa in tutti i contesti relazionali, compresi quelli *online* (Rumayor, 2008). Negli ultimi anni ha infine suscitato attenzione una precisa modalità di prevenzione: quella che si serve dei giochi di ruolo, utilizzando gli ambienti virtuali per rendere i giovani più consapevoli della necessità di considerare il contesto relazionale sviluppando competenze empatiche e di riflessione (Del Moral & Villalustre, 2018).

Ma si potrebbe anche andare oltre, immaginando, ad esempio, che i siti in cui tali comportamenti sono maggiormente presenti possano essere anche presidiati da operatori professionali, che svolgano una sorta di "educativa di strada" in rete, cercando di prevenire il fenomeno dall'interno dei *social network*. In tal modo, si potrebbero prevedere strategie per contattare le vittime e gli autori dell'odio online e offrire loro supporto, sia per tutelarli che per stimolare la comprensione del disvalore delle azioni offensive eventualmente commesse, allo scopo di renderli consapevoli delle conseguenze delle proprie azioni e di stimolare in loro una riflessione critica.

La prevenzione, quindi, sembra fondamentale allo scopo di ridurre lo sviluppo dell'odio online e aiutare le

potenziali vittime ad affrontare in modo adeguato tali situazioni, al fine di diminuire la possibilità di trasformarsi a loro volta in autori.

Riferimenti bibliografici

- Ahedo, J. (2021). Educar en la prudencia y la justicia segun la propuesta de Leonardo Polo. *Studia Poliana*, 23, 43-64.
- Amnesty International (2021) *Barometro dell'odio: intolleranza pandemica*. Amnesty International, Sezione Italiana. Amnesty-barometro-odio-2021 PDF (d21zrvtkxt6ae.cloudfront.net)
- Barranco, R., Gatti, U., Verde, A., & Rocca, G. (2023). Psychosocial factors of risk and protection associated with juvenile cyberbullying victimization: results from an international multi-city study (International Self-Report Delinquency Study 3, ISRD3). *Psychiatry, Psychology and Law*, 30(3), 397-417.
- Bartlett, J., & Krasodomski-Jones A. (2015). *Counter-Speech: Examining Content that Challenges Extremism Online*. London: Demos.
- Bernatzky, C., Costello, M., & Hawdon, J. (2021). Who produces online hate?: An examination of the effects of self-control, social structure, & social learning. *American Journal of Criminal Justice*, 47(1), 1-20.
- Bishop J. (2013) The effect of de-individuation of the Internet Troller on criminal procedure implementation: AAn interview with a hater. *International Journal of Cyber Criminology*, 7(1), 28-48.
- Blaya, C. (2019). Cyberhate: A review and content analysis of intervention strategies. *Aggression and Violent Behavior*, 45, 163-172.
- Bonanno, R. A., & Hymel, S. (2013). Cyber bullying and internalizing difficulties: Above and beyond the impact of traditional forms of bullying. *Journal of Youth and Adolescence*, 42, 685-697.
- Calvete, E., Fernández-González, L., González-Cabrera, J. M., & Gámez-Guadix, M. (2018). Continued bullying victimization in adolescents: Maladaptive schemas as a mediational mechanism. *Journal of Youth and Adolescence*, 47, 650-660.
- Camera dei Deputati (2021). Proposta di legge n. 2936, Misure per la prevenzione e il contrasto della diffusione di manifestazioni d'odio mediante la rete internet. Retrieved from <https://documenti.camera.it/leg18/pdl/pdf/leg.18.pdl.camera.2936.18PDL0133410.pdf>
- Camera dei Deputati, Commissione Jo Cox su fenomeni di odio, intolleranza, xenofobia e razzismo (2017). *La piramide dell'odio in Italia*. Relazione finale.
- Chao, C. H., & Tao, Y. H. (2012). Human flesh search: A supplemental review. *Cyberpsychology, Behavior, and Social Networking*, 15(7), 350-356.
- Chetty, N., & Alathur, S. (2018). Hate speech review in the context of online social networks. *Aggression and Violent Behavior*, 40, 108-118.
- Chirico, S., Gori, L., & Esposito, I. (2020). Quando l'odio diventa reato. Caratteristiche e normativa di contrasto degli hate crimes. Insetto di *Polizia Moderna*.
- Del Moral, M., & Villalustre, L. (2018). Analisis de serious games anti-bullying: recursos ludicos para promover habilidades prosociales en escolares. *Revista Complutense de Educacion*, 29(4), 1345-1364.
- Densley, J., & Peterson, J. (2018). Group aggression. *Current Opinion in Psychology*, 19, 43-48.
- Douglas, K. M., McGarty, C., Bliuc, A. M., & Lala, G. (2005). Understanding cyberhate: Social competition and social creativity in online white supremacist groups. *Social Science Computer Review*, 23(1), 68-76.
- Eckert, S., & Metzger-Riftkin, J. (2020). Doxxing, privacy and gendered harassment. The shock and normalization of veillance cultures. *M&K Medien & Kommunikationswissenschaft*, 68(3), 273-287.
- European Commission (2016). *The EU code of conduct on countering illegal hate speech online*. Retrieved from https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en
- European Union (2008). *Council Framework Decision 2008/913/JHA of 28 November 2008 on Combating Certain Forms and Expressions of Racism and Xenophobia by Means of Criminal Law*. Retrieved from https://eur-lex.europa.eu/eli/dec_framw/2008/913/oj
- European Union (2022). *Regolamento (UE) 2022/2065 del Parlamento Europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (regolamento sui servizi digitali)*.
- Ferrini, C., & Paris, O. (2019). *I discorsi dell'odio. Razzismo e retoriche xenofobe sui social*. Roma: Carocci Editore.
- Few, L. R., Lynam, D. R., & Miller, J. D. (2015). Impulsivity-related traits and their relation to DSM-5 section II and III personality disorders. *Personality Disorders: Theory, Research, and Treatment*, 6(3), 261-266.
- Forsman, M., & Langstrom, N. (2012). Child maltreatment and adult violent offending: population-based twin study addressing the "cycle of violence" hypothesis. *Psychological Medicine*, 42(9), 1977-1983.
- Franco, G., Vianzone, S., & Chiapasco, E. (2017). Teoria della mente e neuroni specchio per capire il cyberbullismo. *Educare.it*, 7(12), 104-108.
- Fulantelli, G., Taibi, D., Scifo, L., Schwarze, V., & Eimler, S. C. (2022). Cyberbullying and cyberhate as two interlinked instances of cyber-aggression in adolescence: a systematic review. *Frontiers in Psychology*, 13, <https://doi.org/10.3389/fpsyg.2022.909299>.
- Gagliardone, I., Gal, D., Alves, T., & Martinez, G. (2015). *Countering Online Hate Speech*. Paris: Unesco.
- Gagliardone, I., Pohjonen, M., Beyene, Z., & Zerai, A. (2016). *Mechachal: Online Debates and Elections in Ethiopia. From Hate Speech to Engagement in Social Media*. University of Oxford, The Programme in Comparative Media Law and Policy.
- Ghisleri, E., & Samada, C. (2022). Prevention of cyberhate and cyberbullying in adolescents: family strategies for personal growth. *Teoría de la Educación*, 34(1), 105-124.
- Grieve, R., & Panebianco, L. (2013). Assessing the role of aggression, empathy, and self-serving cognitive distortions in trait emotional manipulation. *Australian Journal of Psychology*, 65, 79-88.
- Grigg, D. W. (2010). Cyber-aggression: Definition and concept of cyberbullying. *Journal of Psychologists and Counsellors in Schools*, 20(2), 143-156.
- Hamm, M. P., Newton, A. S., Chisholm, A., Shulhan, J., Milne, A., Sundar, P., ... & Hartling, L. (2015). Prevalence and effect of cyberbullying on children and young people: A scoping review of social media studies. *JAMA Pediatrics*, 169(8), 770-777.
- Hoff, D. L., & Mitchell, S. N. (2009). Cyberbullying: Causes, effects, and remedies. *Journal of Educational Administration*, 47(5), 652-665.

- Hohl, A., Choi, M., Yellow Horse, A.J., Medina, R.M., Wan, N., & Wen M. (2022). Spatial Distribution of Hateful Tweets Against Asians and Asian Americans During the COVID-19 Pandemic, November 2019 to May 2020. *American Journal of Public Health, 112*(4), 646-649.
- IRS - Istituto per la Ricerca Sociale (2019). *C.O.N.T.R.O "COunter Narratives Against Racism Online". Mappatura delle principali metodologie italiane ed europee di contronarrativa per il contrasto degli "hate speech online"*. Roma: Presidenza del Consiglio dei Ministri, Dipartimento per le pari opportunità.
- Krakowski, M. I., & Czobor, P. (2014). Depression and impulsivity as pathways to violence: implications for anti-aggressive treatment. *Schizophrenia Bulletin, 40*(4), 886-894.
- Lereya S.T. et al. (2015) Adult mental health consequences of peer bullying and maltreatment in childhood: two cohorts in two countries. *The Lancet Psychiatry, 2*(6), 524-531.
- Lumsden, K. & Morgan, H. (2017). Media framing of trolling and online abuse: silencing strategies, symbolic violence, and victim blaming. *Feminist Media Studies, 17*(6), 926-940.
- Machimbarrena, J. M., Gonzalez-Cabrera, J., & Garaigordobil, M. (2019). Variables familiares relacionadas con el bullying y el cyberbullying: una revision sistematica. *Pensamiento Psicologico, 17*(2), 37-56.
- Margari, F., Matarazzo, R., Casacchia, M., Roncone, R., Dieci, M., Safran, S., Fiori, G., Simoni, L., and The Epica Study Group (2005). Italian validation of MOAS and NOSIE: a useful package for psychiatric assessment and monitoring of aggressive behaviours. *International Journal of Methods in Psychiatric Research, 14* (2), 109-118.
- Mascheroni, G., & Ólafsson, K. (2018) *Accesso, usi, rischi e opportunità di internet per i ragazzi italiani. I primi risultati di EU Kids Online 2017*. EU Kids Online e OssCom.
- Meza, R., Vincze, H. O., & Mogoş, A. (2018). Targets of online hate speech in context. A comparative digital social science analysis of comments on public Facebook pages from Romania and Hungary. *East European Journal of Society and Politics, 4*(4), 26-50.
- Miller, J., & Shirley, M. (2009). "Cyberbullying knows no borders". *Australian Teacher Education Association*.
- Ministero per l'Innovazione Tecnologica e la Digitalizzazione, Ministero della Giustizia, Presidenza del Consiglio dei Ministri (2021). *Gruppo di lavoro Odio Online Rapporto finale*. Roma: Presidenza del Consiglio dei Ministri, <https://innovazione.gov.it/notizie/articoli/odio-online-il-report-finale-del-gruppo-di-lavoro/>.
- Musto, C., Smeraro, G., de Gemmis, M., & Lops, P. (2016). "Modeling community behavior through semantic analysis of social data: The Italian hate map experience". In Vassileva, J., J. Blustein, L. Aroyo, S.K. D'Mello (eds.), *UMAP 2016 - Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*. New York: Association for Computing Machinery Press.
- Pace, U., D'Urso, G. & Zappulla, C. (2021). Hating among adolescents: Common contributions of cognitive distortions and maladaptive personality traits. *Current Psychology, 40*(7), 3326-3331.
- Pasta, S. (2018). *Razzismi 2.0. Analisi educativa del razzismo online*. Brescia: Morcelliana.
- Patton, D.U., et al. (2014) Social media as a vector for youth violence: A review of the literature. *Computers in Human Behavior 35*, 548-553.
- Peterson, J., & Densley, J. (2017). Cyber violence: What do we know and where do we go from here? *Aggression and Violent Behaviour, 34*, 193-200.
- Pizzi, F. (2022). Online hate speech. Education as prevention. *Journal of Inclusive Methodology and Technology in Learning and Teaching, 2*(1), 1-8.
- Rains, S. A., Kenski, K., Coe, K., & Harwood, J. (2017). Incivility and political identity on the Internet: Intergroup factors as predictors of incivility in discussions of news online. *Journal of Computer-Mediated Communication, 22*(4), 163-178.
- Reynolds, P. (2021). *6th evaluation of the Code of Conduct on Countering Illegal Hate Speech*, European Commission. https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/eu-code-conduct-countering-illegal-hate-speech-online_en.
- Rieger, D., Kumpel, A.S., Wich, M., Kiening, T., & Groh, G. (2021). Assessing the extent and types of hate speech in fringe communities: a case study of alt-right communities on 8chan, 4chan, and Reddit. *Social Media Society, 7*(4), 1-14.
- Rumayor, M. (2008). *Ciudadania y Democracia en la Educacion*. Pamplona: Universidad de Navarra.
- Sampasa-Kanyinga, H., Roumeliotis, P., & Xu, H. (2014). Associations between cyberbullying and school bullying victimization and suicidal ideation, plans and attempts among Canadian schoolchildren. *PLoS One, 9*(7), 102-145.
- Selby, E.A., Anestis, M.D., & Joiner, T.E. (2008). Understanding the relationship between emotional and behavioral dysregulation: Emotional cascades. *Behaviour Research and Therapy, 46*(5), 593-611.
- Shachaf, P., & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science, 36*(3), 357-370.
- Smith, P.K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: its nature and impact in secondary school pupils. *Journal of Child Psychology and Psychiatry, 49*(4), 376-385.
- Sorokowski, P., Kowal, M., Zdybek, P., & Oleszkiewicz, A. (2020). Are Online Haters Psychopaths? Psychological Predictors of Online Hating Behavior. *Frontiers in Psychology, 11*, 5. <https://doi.org/10.3389/fpsyg.2020.00553>
- Sourander, A., Brunstein Klomek, A., Ikonen, M., Lindroos, J., Luntamo, T., Koskelainen, M., Ristkari, T., & Helenius, H. (2010). Psychosocial risk factors associated with cyberbullying among adolescents: a population-based study. *Archives of General Psychiatry, 67*(7), 720-728.
- Steffgen, G., König, A., Pfetsch, J., & Melzer, A. (2011). Are Cyberbullies Less Empathic? Adolescents' Cyberbullying Behavior and Empathic Responsiveness. *Cyberpsychology, Behavior, and Social Networking, 14*(11), 643-648.
- Suarez, P., & Velez, M. (2018). El papel de la familia en el desarrollo social del niño: una mirada desde la afectividad, la comunicación familiar y estilos de educación parental. *Revista Psicoespacios, 12*(20), 173-198.
- Tang W.Y. & Fox J. (2016). Men's harassment behavior in online video games: Personality traits and game factors. *Aggressive Behaviour, 42*, 513-521.
- The London School of Economics and Political Sciences (2020). *EU Kids Online. New European Study on children and the internet in 19 countries*. Retrieved from <https://www.lse.ac.uk/media-and-communications/research/research-projects/eu-kids-online/eu-kids-online-2020>
- Tripathi, V. (2017). Youth Violence and Social Media. *Journal of Social Sciences, 52* (1-3), 1-7.
- Van Royen, K., Poels, K., Vandebosch, H., & Adam, P. (2017).

- “Thinking before posting?” Reducing cyber harassment on social networking sites through a reflective message. *Computers in Human Behaviour* 66, 345-352.
- Verbena, S. (2021). ODIO 2.0: una prospettiva psicosociale sui processi di radicalizzazione online. In Cristante, S., S. Pinnelli, A. Rinella (eds.), *La comunicazione: media, processi produttivi e narrazioni. Atti del Workshop 2020*, Lecce: Università del salento Press, 22-33.
- Vitullo, A. (2021). L'intersezione tra media education e hate speech. I futuri cittadini digitali crescono online. In Giorgi, A., F. Rispoli (eds.), *HATE SPEECH. Riflessioni, pratiche e proposte contro l'odio in rete* (59-70). I Quaderni di Benvenuti in Italia.
- Vox – Osservatorio Italiano sui Diritti (2021). *La mappa dell'intolleranza*. 7° anno.
- Wachs, S., Machimbarrena, J. M., Wright, M. F., Gámez-Guadix, M., Yang, S., Sittichai, R., Singhu, R. Biswal, R., Flora, K., Daskalou, V., Maziridou, E., Sung Hong, J. & Krause, N. (2022). Associations between coping strategies and cyberhate involvement: Evidence from adolescents across three world regions. *International Journal of Environmental Research and Public Health*, 19(11), 6749.
- Walther, J.B. (2023). Online hate: a prosocial explanation of antisocial behavior and affordances of social media. In R. Nabi & J. Myrick (eds.) *Our Online Emotional Selves: The Link between Digital Media and Emotional Experience*. Oxford: Oxford University Press.
- Widom, C.S. (1989). Does violence beget violence? A critical examination of the literature. *Psychological Bulletin*, 106(1), 3–38.
- Ybarra, M.L., Mitchell, K.J., Wolak, J. & Finkelhor, D. (2006). Examining characteristics and associated distress related to internet harassment: findings from the Second Youth Internet Safety Survey. *Pediatrics*, 118(4), 1169-1177.
- Zhang, Z., Luo, L. (2019). Hate speech detection: A solved problem? The challenging case of long tail on Twitter. *Semantic Web*, 10(5), 925-945.
- Ziccardi, G. (2016). *L'Odio online. Violenza verbale e ossessioni in rete*. Bologna: Il Mulino.