

RAGE BEHIND THE SCREEN

RAGE BEHIND THE SCREEN

*The Psychology of Online Hate in the Age of
Social Media*

Farah Zidan



There is a world elsewhere

Copyright © 2026 by Elsewhere Publishing House
All rights reserved.

No part of this publication may be reproduced, distributed, stored in a retrieval system, or transmitted in any form or by any means — including electronic, mechanical, photocopying, recording, or otherwise — without the prior written permission of the publisher, except in the case of brief quotations used in reviews or scholarly references.

Published by Elsewhere Publishing House

First Edition: 2026

Author: Farah Zidan



There is a world elsewhere

" For those who hurt — on both sides of the screen."

Table of Contents

Introduction: Why Are We So Ready to Attack?	9
PART ONE The Birth of Digital Rage.....	12
Chapter 1: How the Internet Changed What It Means to Be Human	13
Chapter 2: The Algorithm of Anger — How Outrage Gets Engineered.....	18
Chapter 3: The Joy of the Attack — Collective Aggression Online	24
PART TWO The Psychology of the Attacker.....	28
Chapter 4: Projection — Seeing Our Shadows in Others.....	29
Chapter 5: Shame, Envy, and the Invisible Wounds	33
Chapter 6: Powerlessness, Tribalism, and the Need to Belong... ..	37
PART THREE Life Under the Gaze of the Crowd	41
Chapter 7: The Influencer's Burden — Visibility as Vulnerability... ..	42
Chapter 8: The Psychological Toll — What Digital Hate Does to Human Beings.....	48
Chapter 9: The Cycle of Violence — When Targets Become Attackers.....	53
PART FOUR Survival in the Age of Rage	56
Chapter 10: Building the Psychological Architecture of Resilience	57
Chapter 11: The Practice of Boundaries — Protecting Your Presence Online	63
Chapter 12: Rebuilding Humanity in Digital Spaces	67
Conclusion: The Long Work Ahead	71

Introduction: Why Are We So Ready to Attack?

Every morning, before the first cup of coffee, millions of people scroll through their phones and enter a battlefield. Not one they were consciously recruited into — but one they joined by tapping an app icon. Social media, in all its luminous promise of connection, has quietly become one of the most reliably hostile environments human beings have ever created for themselves. And yet we keep returning.

This book is not a moral indictment of the internet, nor a nostalgia trip for a pre-digital past that was never as peaceful as we romanticize. It is an honest, scientifically grounded, and deeply human attempt to understand a phenomenon that touches virtually everyone who inhabits the modern world: online hate, digital rage, and the peculiar cruelty that emerges when human beings interact through screens.

The question I want to begin with — and the one that animates every chapter that follows — is not "why are people so bad online?" That framing is too simple and, ultimately, too convenient. It lets us off the hook. The harder, more honest question is: why are we so ready to attack?

Because here is what the research consistently reveals: the people who leave hateful comments, who pile onto strangers, who coordinate harassment campaigns, who cancel public figures with righteous fury — these are not a separate subspecies of humanity. They are, by and large, ordinary people. People who may volunteer at their local food bank.

People who love their children. People who, in person, might hold the door open for you without thinking twice.

Something happens between the physical world and the digital one. Something loosens. Some constraint that ordinarily governs our worst impulses gets suspended, and what rushes in to fill that space is often shocking in its intensity and breadth.

This book is structured around four interconnected inquiries. First, we examine what online hate actually is — its forms, its mechanisms, and how the architecture of social media platforms amplifies it in ways that are not accidental but engineered. Second, we go deep into the psychology of those who attack: the projections, shame spirals, envy, and tribalism that power digital cruelty. Third, we explore what it feels like to live under the gaze of the crowd — particularly for content creators, influencers, and anyone who has chosen visibility in a culture that simultaneously rewards and punishes it. And fourth, we turn toward survival and healing: not the shallow "just don't read the comments" variety, but genuine, neurobiologically informed strategies for building resilience without losing your humanity.

I want to be clear about something from the start. This book holds compassion as a non-negotiable commitment — compassion for those who are attacked, yes, but also, more uncomfortably, compassion for those who do the attacking. Not because their behavior is acceptable, but because understanding is the only tool sharp enough to cut through the problem at its root. Judgment, however satisfying, changes nothing.

The central thesis of this book is simple and, I believe, radical in its implications: digital rage is most often a mask worn over unhealed wounds. When someone goes online and destroys a stranger, they are rarely doing something to that stranger. They are doing something with their own pain — projecting it, displacing it, temporarily relieving the pressure of carrying it in private. The screen makes this process easy, consequence-free, and even socially rewarded. That is the horror at the heart of this story.

But there is also a different story available. One where we learn to see the scream inside the cruelty. One where we stop mistaking public punishment for moral virtue. One where the people most likely to be targeted — those brave or reckless or ambitious enough to be visible — learn to receive hostility without being dismantled by it.

That story is possible. This book is an attempt to tell it.

PART ONE

The Birth of Digital Rage

Chapter 1: How the Internet Changed What It Means to Be Human

The Promise and the Paradox

The internet was, in its earliest public imagination, a utopian technology. Its architects and evangelists spoke of it as the great equalizer: a space where geography collapsed, where the powerless could find voice, where ideas could travel freely across barriers of class and nation. And in many ways, this vision came true. The Arab Spring was organized on social media. LGBTQ+ teenagers in isolated rural communities found lifelines in online forums. Medical patients connected across the world to share information that kept them alive. The internet genuinely expanded what was possible for human connection.

But somewhere in the expansion, something else expanded too. Something that the utopian vision had not accounted for — or perhaps had deliberately looked away from. Because human beings carry their psychology wherever they go, including into new technological spaces. And the human psyche, it turns out, is not a smooth instrument. It is ancient, tribal, status-obsessed, and exquisitely sensitive to threat. It did not take long for these older instincts to colonize the new digital territory.

The Disinhibition Effect: Science of Saying the Unsayable

In 2004, psychologist John Suler published what became one of the foundational papers in cyberpsychology: "The Online Disinhibition Effect." Suler identified a consistent phenomenon

he observed across early internet communities: people said things online that they would never say in person. They were more aggressive, more sexually explicit, more emotionally raw, and — critically — more willing to be cruel.

Suler identified several factors that produced this disinhibition. Anonymity was the most obvious: when you cannot be identified, you cannot be held accountable, and with accountability removed, the social penalties that ordinarily constrain behavior disappear. But anonymity alone is insufficient to explain the full phenomenon. After all, people continue to behave aggressively online even when using their real names.

More fundamental, Suler argued, is the phenomenon of dissociative imagination — the unconscious ability to treat one's online persona as a fictional character, a separate entity who exists in a different world than one's physical self. "It's not really me saying this," the psyche implicitly tells itself. "It's my avatar. My username. A version of me that exists only in this other world." This dissociation from one's actions is not a conscious strategy; it operates below awareness, and it is remarkably effective at suspending the empathy and self-regulation systems that ordinarily moderate how we treat each other.

Neuroscience offers a complementary lens. When we interact with another person face-to-face, our mirror neuron systems are fully engaged. We see their facial expressions, hear the quiver in their voice, feel viscerally the impact of our words on their physiology. This somatic loop — the way we literally feel each other — is one of the primary brakes on human cruelty. It

is why it is much harder to be genuinely vicious to someone who is weeping in front of you.

Online, this loop is severed. Text on a screen carries no facial expression, no trembling hands, no flushed skin. The person on the receiving end of our words is not absent — they exist, they suffer — but we cannot perceive that suffering in real time, and our nervous systems, calibrated over millions of years for face-to-face interaction, respond to their absence by treating the interaction as consequence-free. In a very real neurological sense, hurting someone online does not feel like hurting them.

“The person on the receiving end of our words is not absent — they exist, they suffer — but we cannot perceive that suffering in real time.”

From Village to Stadium: The Scale Problem

Human beings evolved in small groups. Anthropologist Robin Dunbar famously proposed that the human neocortex is calibrated for maintaining stable social relationships with approximately 150 people — a number that corresponds to the typical size of hunter-gatherer communities. Our cognitive and emotional systems for navigating social life were built for village-scale intimacy, where everyone knew everyone, where reputations were slow to build and slow to collapse, where conflict had to be managed because you would see this person again tomorrow.

Social media put us in a stadium. Suddenly, a single post could be seen by thousands, then millions. A single misstep could be

broadcast to an audience so large it defied comprehension. Our village-scale psychology was dropped, without preparation or warning, into a stadium-scale environment. The results have been, in many ways, catastrophic.

At stadium scale, the village mechanisms of conflict resolution — gradual conversation, mutual understanding, the soft pressure of ongoing relationship — become impossible. Instead, what emerges is something closer to mob dynamics. The crowd becomes a character in itself, with its own emotional contagion, its own momentum, its own logic. And in that crowd, individual inhibitions further dissolve. Social psychologists call this deindividuation: the reduction of self-awareness and personal responsibility that occurs in group contexts. What feels impossible to do alone feels permitted — even righteous — in a crowd.

The Architecture of Outrage: How Platforms Are Designed to Inflammate

It would be a serious error to treat online hate purely as a psychological phenomenon — something that lives in individuals — without attending to the structural conditions that produce it. Social media platforms are not neutral pipes through which human nature simply flows. They are engineered environments, designed by teams of brilliant people with specific behavioral objectives, and the behavior they are optimized to produce is engagement.

Engagement, in the language of Silicon Valley, means time spent on platform, interactions generated, content created and consumed. Engagement is what sells advertising. Engagement

is how these companies survive and grow. And the sobering research finding — confirmed now across multiple independent studies — is that the content that generates the most engagement is not the content that makes people feel good. It is the content that makes people feel angry.

A 2021 study published in *Science* examined Facebook's algorithmic amplification systems and found that content designed to provoke moral outrage was consistently boosted over content that was merely informative, beautiful, or kind. Twitter (now X) researchers found similar patterns: posts containing so-called "moral-emotional language" spread roughly 20% further with each word of outrage added. Anger is the most viral emotion. Platforms discovered this early, and their recommendation algorithms — largely opaque to outside scrutiny — were tuned accordingly.

This is not a conspiracy. It is simply the logic of optimization. When you build a system that rewards what generates engagement, and anger generates more engagement than joy, you build a system that produces anger. The humans who built these systems may not have intended to make the world more hostile. But the systems they built have done so regardless.

The result is what researchers have begun calling the outrage economy: an attention marketplace in which the currency is moral indignation, the producers are users who have learned (often without being fully aware of it) that outrage generates attention and attention generates rewards, and the infrastructure is owned by corporations whose financial incentives are perfectly aligned with keeping the outrage flowing.

Chapter 2: The Algorithm of Anger — How Outrage Gets Engineered

Dopamine, Feedback Loops, and the Slot Machine

To understand why online outrage is so difficult to resist — both for those who generate it and those who consume it — you need to understand something about how the brain's reward circuitry works, and specifically how it was co-opted by the designers of social media platforms.

Dopamine is, in popular imagination, the "pleasure molecule" — the neurotransmitter released when we experience something pleasurable. But this is a significant oversimplification. Dopamine is more precisely understood as a signal related to anticipated reward and uncertainty. The dopamine system fires not when you receive a reward, but when you are uncertain whether you will. This is why slot machines are so powerfully addictive: they provide unpredictable, variable rewards, which is the most potent possible trigger for dopamine release. It is also why social media notifications are designed as they are.

When you post something online, you do not know whether people will respond. You check — and sometimes nothing has happened, and sometimes you have thirty notifications, and sometimes your post is being shared by thousands of strangers. This unpredictability is not an accident. Former Google design ethicist Tristan Harris, who spent years working inside Silicon Valley before becoming its most prominent internal critic, has documented in detail how the variable reward structure of social media platforms was deliberately designed on the model

of the slot machine. The pull-to-refresh gesture, Harris observes, is functionally identical to pulling the lever on a slot machine.

Now add the specific rewards associated with expressing outrage. When you post something angry — a righteous denunciation, a call-out, an expression of moral disgust — and it goes viral, you receive an enormous dopamine hit. Thousands of people validate your anger. They repost it, add their own outrage, award you the currency of attention and moral authority. For many people, particularly those who otherwise feel powerless or invisible, this experience is intoxicating. The research on this is consistent: aggressive and outrage-generating behavior online is reinforced not just by platform algorithms but by the social rewards that come from one's community responding approvingly.

Over time, this conditioning shapes behavior in ways that feel, to the person experiencing them, like genuine moral conviction. You come to experience yourself as righteous and principled when you are, neurologically, largely responding to the reward circuitry of your own brain.

Outrage as Community: The Tribal Logic of Online Hate

Human beings have always defined themselves in relation to groups. The tribal psychology that evolved over millions of years as a survival strategy — the deep need to belong to an in-group and differentiate it from an out-group — does not disappear in the age of social media. It finds new expression in new arenas.

Psychologist Jonathan Haidt, in his influential work on moral psychology, describes how moral emotions like disgust, contempt, and righteous anger evolved not primarily to guide individual behavior but to coordinate group behavior. When a community collectively expresses outrage at a transgressor, it is doing something socially important: it is reaffirming the group's values, consolidating group identity, and communicating to potential defectors that violations will be punished.

Online, this dynamic plays out at enormous scale and with terrifying efficiency. When a target is identified — whether they genuinely violated a community norm or simply made a poor-taste joke or held an unpopular opinion — the pile-on that follows is not primarily about them. It is about us: our collective identity, our shared values, our pleasure in unity. The target is, in a meaningful sense, a prop in a performance of community identity.

This is why pile-ons can feel, to participants, deeply righteous and even joyful. They are social bonding experiences. They create exactly the sense of belonging, purpose, and shared meaning that modern life — with its atomized individuals staring at screens in separate apartments — so often fails to provide. The tragedy is that this belonging is built on someone else's destruction.

The Pleasure of the Fall: Schadenfreude and Social Leveling

There is another dimension to online hate that is rarely discussed directly because it is embarrassing to acknowledge: