



THE DISINFORMATION
PROJECT

WHITE PAPER

Opportunism and Polarisation:
Presentations of the violence in
Israel and Palestine by Aotearoa
disinformation networks

OCTOBER - NOVEMBER 2023

AUTHORED BY
**DR SANJANA HATTOTUWA
AND KATE HANNAH**

EDITED BY
NICOLE SKEWS-POOLE



CONTENTS

BACKGROUND	3
Purpose	3
Location of study	3
Context	4
Methodology	4
Perspective from TDP’s Director of Research	5
FINDINGS	6
Overarching Observations	6
Polarisation.....	7
Platform interplay.....	8
Far-right and neo-Nazi appropriation and exploitation	9
‘Narrative flattening’	10
Graphic harm, gore and violence	11
SPECTACLE, IRREALISM, AND COGNITIVE WARFARE: IMPLICATIONS FOR SOCIAL COHESION	13
TDP research framing.....	14
HIGH-LEVEL RECOMMENDATIONS	17
CONCLUDING REMARKS	18

BACKGROUND

Purpose

This White Paper was intended for limited distribution to communities, civil society organisations, and agencies, and these groups received the first iteration of this paper in November 2023.

Our purpose in developing this white paper was to capture a summary of high-level themes related to the impact of the atrocities committed between 7-23 October within Aotearoa New Zealand's disinformation ecologies.

This is a time-bound snapshot of data, including footage from Israel and Palestine, and an analysis in near-real time of its impact within existing disinformation communities that initially formed around COVID-19 denial and minimisation.

Location of study

The Disinformation Project's site of study is exclusively Aotearoa New Zealand disinformation networks.

While global events are important contexts that influence these networks, our work lies in understanding how local disinformation environments use global and local contexts to further their agendas. Critically, we seek to understand and provide advice to mitigate the harm posed by disinformation influencers and followers, especially towards already marginalised individuals and communities in Aotearoa New Zealand.

As readers will discover below, antisemitism is a key theme emerging in local disinformation networks. Islamophobia and the targeting of Palestinian communities are also present, though at a different reach and volume. In no way does this observation seek to undermine the harm caused to the Muslim and Palestinian communities in Aotearoa New Zealand, and readers are asked to keep in mind that themes observed in disinformation networks have their own patterns and reasonings, which are, by design, not grounded in truth or nuance.

Our kaupapa requires us to share information relating to targeted local harms across the spectrum of the Aotearoa New Zealand Muslim, Palestinian and Jewish communities, all of whom have been impacted by disinformation networks that are appropriating this conflict for their own agenda.

We trust that readers will understand our need to focus exclusively on local emerging harms in order to best serve diverse communities during a time of heightened targeting.

Context

Since October 7, 2023, when a series of attacks by Hamas on Israeli targets took place, an escalating, contentious and violent conflict in Palestine and Israel has had significant impacts in Aotearoa New Zealand.

This white paper focuses on the uptake of partisanship within Aotearoa New Zealand's existing disinformation communities, resulting in what we describe as increased affective polarisation.

These existing disinformation communities were shaped and formed a shared set of ideas based on denial and minimisation of COVID-19, anti-vaccination sentiment, and distrust of the state.

Palestinian, Jewish, Israeli, and Muslim communities within Aotearoa New Zealand are impacted by the violent conflict, and by its presentation and representation in Aotearoa New Zealand, from mainstream media reporting the use of this content in disinformation communities.

Media reports reflect a disconnect between what security agencies have publicly stated, and what communities, including those impacted by the March 2019 attacks, have communicated regarding fears, anxieties, and concerns in light of the violent conflict¹. Places of worship and community centres have been targeted with threats, graffiti, and arson.

Methodology

- The Disinformation Project (TDP) has studied social media representations and re-representations of the conflict since 9 October 2023, including web-based content.
- Our continuous primary study sites are **content or commentary that is present, referenced, or linked to Aotearoa New Zealand's disinformation ecologies**.
- Our research is open-source and assesses public posts on mainstream (e.g., Facebook, Instagram, Twitter/X, TikTok, YouTube) and other platforms (e.g., Telegram).
- Through off-platform links provided on these primary sites of study, we also reviewed for the purposes of this short snapshot report websites, and far-right video platforms such as Rumble, Gab, and BitChute.
- Since March 2023, when a significant step change took place due to global linkages around disinformation related to the intersex, and trans community², Aotearoa New Zealand's disinformation communities and, increasingly,

¹ New Zealand's Muslim, Jewish communities 'anxious' as Israel-Hamas conflict escalates, <https://www.newshub.co.nz/home/new-zealand/2023/10/new-zealand-s-muslim-jewish-communities-anxious-as-israel-hamas-conflict-escalates.html>

² Working paper: Transgressive transitions, <https://thedisinfoproject.org/2023/05/05/working-paper-transgressive-transitions/>

mainstream media commentary, and discourse is seamlessly integrated with transnational media, news, and information networks.

- Information flows are happening in real-time, in greater volume, much faster, and in more varied ways than the social media ecologies present in, for example, March 2019.
- In our May 2023 report, we described in detail the multidimensional and wholesale community bridging of a suite of far-right talking points into ordinary disinformation communities in Aotearoa New Zealand.
- In October 2023, this now includes the significant influx of violent captures from Palestine and Israel.
- In data collation and analysis from 9 October to 27 October (and ongoing), our situation reports have mapped in detail content flows, narratives, and impacts, totalling over 100,000 words, and nearly 400 pages of analysis dedicated to the manner in which the war has been understood, presented, and spread within domestic information flows.
- This second-order, country and context-specific analysis involved a combination of computational analysis including custom-trained natural language processing models, and the experimental use of ChatGPT-4's multimodality, alongside first-order human analysis of source material to study aspects including narratives, meaning, speaker, audience, framing, expression, and location.
- This mixed-methods approach, which defines TDP's research in general, includes social media platforms, mainstream media reporting, investigative journalism, reports by social media companies themselves, as well as contemporaneous academic research locating trends, patterns, and developments identified in Aotearoa New Zealand in other countries or more globally.
- Notably, this period is categorised by the sharing of graphic harm, gore and violence (GHGV), spread throughout domestic information networks.
- For the first time since 2020, the daily analysis did not embed content or feature links to sources, due to the level of harm and violence.

Perspective from TDP's Director of Research

The Disinformation Project's Director of Research, Dr Sanjana Hattotuwa, has extensive first-hand experience with war, violence, and conflict, including reporting on and documenting war crimes, and crimes against humanity. He has also studied the social media and online representations of violent conflict since 2005, across multiple countries and continents, starting with his home country, Sri Lanka.

This perspective provides a unique understanding of the representations of the war in Aotearoa New Zealand's disinformation ecologies. This background is critical here because violence does not exist in isolation; patterns and motivations often cross borders, even if the manifestations and motivations differ.

Dr Hattotuwa's lived experience, and his trans-continental research into disinformation ecologies and representations of violence, equips TDP to see connections and provide context around violence in more nuanced ways.

The reporting period saw the worst violence Dr Hattotuwa has studied, including Sri Lanka's civil war and its extremely violent denouement in May 2009.

FINDINGS

Overarching Observations

Our examinations of the threats posed to domestic social cohesion, through the rapid dissemination of unprecedented levels of violent, graphic content, include potential offline impacts like radicalisation and kinetic violence.

Specifically, our studies have noted the following in Aotearoa New Zealand disinformation ecologies:

- Telegram is a major hub for sharing propaganda, graphic content and disinformation from pro-Hamas and pro-Israeli perspectives. Telegram channels and groups form echo chambers which reinforce biased narratives.
- Pro-Hamas accounts are very active on Telegram, operating channels and groups that spread graphic violent content, propaganda, and disinformation which co-opts the victimisation of Palestinians to justify Hamas' actions on October 7. *[Post-publication edit 1 March 2023: This point has been amended for clarity].*
- These accounts helped form echo chambers reinforcing anti-Israel, and extremely violent antisemitic narratives, expressed in written, memetic, visual, and video forms.
- Pro-Israel accounts also used Telegram to share their biased perspectives, propaganda and misleading claims, depicting all Palestinians as terrorists and justifying Israeli military actions.
- Twitter/X facilitates heated, polarising debates on the conflict with rigid ideological positions forming. Highly viral misleading claims and conspiracy theories spread rapidly on Twitter.
- 'Narrative flattening', seen through the deliberate misrepresentations of varied, fluid, and complex opinions on the outbreak of war as being partial to, excusing or promoting terrorism resulted in chilling effects, as well as the heightened, networked, and violent targeting of individuals and institutions.
- Swarms of enraged accounts, especially on Twitter/X, engaged in flame wars³ spanning multiple geographies, media, and in some cases, even languages (featuring Arabic, English, Hindi, and Hebrew in the replies, and quotes).
- On Twitter, pro-Palestinian and pro-Israeli accounts engaged in heated debates and ideological polarization around the conflict. They spread viral misinformation and conspiracy theories demonising the other side.
- Domestic verified/Blue check accounts on Twitter/X amplified some of the misleading narratives, despite presenting an air of credibility. Certain influential

³ Flame wars, friendly comments and fake news: Chief comment moderator's year in review, <https://www.smh.com.au/opinion/flame-wars-friendly-comments-and-fake-news-chief-comment-moderators-year-in-review-20171221-h08xnxw.html>, *Sydney Morning Herald*.

accounts emerged as key nodes for cross-platform sharing of deliberately provocative, misleading, untrue or partially true content.

- Conspiracy websites, alternative, far-right web-based platforms, and related online communities helped propagate anti-establishment narratives on the war, aligned with Russian propaganda frames, by exploiting (and adding to the perceptions of) distrust in liberal democratic institutions, and mainstream media.
- Strong in-group, and out-group divisions were evident across platforms, and media, with each side depicting the other as evil, through narratives which lacked any nuance.
- Extensive cross-platform linking, and sharing of provocative content between Telegram, Twitter, YouTube, conspiratorial websites etc., enabling rapid dissemination within, and across online ecosystems.

Overall, the social media ecologies fostered polarisation, confirmation bias and the spread of hate and disinformation. They enabled transnational networks to coordinate widespread propagation of disinformation on a complex geopolitical issue that is evolving at pace.

Polarisation

Content and its presentation and representation on domestic social media have been characterised by:

- The articulation of polarised hate,
- The spread of false, misleading or unconfirmed information; and
- The rapid dissemination of graphic content, including extremist content.

We posit that this informational multidimensional crisis (a polycrisis) is negatively impacting social cohesion, public safety, constructive democratic discourse, trust in public institutions, and respect for human rights and human dignity in Aotearoa New Zealand.

Key features of the information landscape include:

- The liar's dividend⁴,
- Second-order disinformation⁵,
- Fourth generation warfare⁶; and
- Irrealism⁷.

Since 9 October, our research has provided comprehensive, near real-time, multi-faced disinformation research which analyses how the war is being represented in

⁴ Wolf, Wolf! Alarm Over Disinformation and The Liar's Dividend, <https://lab.cccb.org/en/wolf-wolf-alarm-over-disinformation-and-the-liars-dividend/>, Centre de Cultura Contemporània de Barcelona

⁵ More Tech Will Not Save Us from Disinformation, <https://source.opennews.org/articles/more-tech-will-not-save-us-disinformation/>, Source

⁶ Misinformation Is Warfare, <https://time.com/6323387/misinformation-israel-hamas-war-essay/>, by Prof. Joan Donovan, *Time*

⁷ The Worst of Times: Our Year in Irreality, <https://lithub.com/the-worst-of-times-our-year-in-irreality/>, by John Freeman, *Literary Hub*

Aotearoa New Zealand, specifically within communities organised around disinformation.

We have examined how graphic, violent content and disinformation related to the conflict are rapidly spread across varied social media landscapes. This includes the migration of this content between diverse communities on the same product or platform surface (e.g., between different Telegram chat groups).

Overall, we studied sophisticated, cross-platform efforts to distort narratives and sow division. The rapid spread of graphic violence is intended to polarise and radicalise audiences.

In what can be seen as farming anger and outrage at scale, this content is intended to invoke rage in audiences, in the guise of presenting impartial captures of war.

The presentation and representation of content, and the affective polarisation driven by these representations are designed to increase hate, hate crimes, and violence towards communities, and are focused on Jewish, Muslim, and Palestinian communities in Aotearoa New Zealand.

Our research also stresses how social media has collapsed geographic barriers, allowing global conflicts to directly, and enduringly impact local information environments.

Platform interplay

Our research highlighted how social media environments and hyper-partisan ecosystems propagate reductive, misleading perspectives, antisemitism, and Islamophobia including through memetic content.

Hyper-partisan domestic social media accounts played a major role in importing and spreading polarised perspectives, propaganda, and disinformation around the war.

For example, posts studied in just the 24-hour period of 25-26 October demonstrated a complex social media ecology where fringe, conspiratorial voices, and disinformation-driven controversial perspectives were amplified algorithmically and organically.

Ideology-driven narratives from content shared first or foremost on Twitter, TikTok, Instagram, and alternative news sites were re-featured on domestic Telegram ecologies, with some of the videos re-uploaded so that even if taken down in the original location, they would be enduringly accessible.

Mainstream views were largely rejected in favour of highly polarised, incendiary, and accusatory framing of events in Israel-Palestine, Ukraine, climate change, and local political issues. Dangerous speech was the normative discourse.

The interplay between open platforms and closed groups enables the curation of selective content for channel subscribers who likely already agree with such viewpoints. Links were shared from sources across different geographic locations, i.e., foreign accounts, channels, and chat groups, through to domestic Telegram

channel-based posts. This was a rapid, fluid, transnational, multi-modal interlinking of divisive narratives. For the first time, we studied the inclusion of Arabic content, even in domestic social media ecologies, through a range of media including memes, with provenance or links resolving to foreign accounts, and channels.

Day-on-day, Telegram channels provided case studies of how social media discourse and norms enabled the consolidation of white supremacist violent extremist content and voices into mainstream information spaces. This promoted propaganda, disinformation and violent content at scale. Often, this choreography was centred around one bridging account, which cross-pollinated content between diverse social media spaces.

The provenance of material studied on domestic social media ecologies, given the unprecedented volume, and speed of presentation was hard, if not impossible to determine in real time. There are clear indications internationally of state and state-adjacent actors involved in promulgating information and disinformation.⁸ The impact of the amount and tone of content, across every platform, every medium, all the time was to overwhelm users with a constant display of outrage, trauma, violence, grief, and anger.

This was unprecedented at the time of writing (November 2023). The net effect of this content was to exacerbate existing division and accelerate affective polarisation in online and offline discourse. It is important to note that this coincided with the start of the Christchurch Masjidain Attack Coronial Inquiry⁹.

Far-right and neo-Nazi appropriation and exploitation

Our research has increasingly uncovered connections. between and within, domestic social media clusters and accounts characterised by the presence of far-right neo-Nazi ideologies. Specifically, this involves content and/or individuals that promote highly antisemitic content.

We have mapped how the Aotearoa New Zealand far-right and neo-Nazi social media ecologies have responded to the conflict in ways that reflect international trends. This is seen through groups, online spaces and individuals with pre-existing far-right and neo-Nazi ideologies exploiting volatility, polarisation and conflict to spread antisemitic conspiracy theories, as well as Islamophobic and racist worldviews.

Examples from Aotearoa New Zealand included neo-Nazi and white supremacist online communities appropriating pro-Palestinian stances, not out of genuine concern, but to vilify the Jewish community. They aim to radicalise people towards extremist white nationalist ideologies, through performative allyship and outrage.

⁸ <https://www.nytimes.com/2023/11/03/technology/israel-hamas-information-war.html>

⁹ Christchurch terror attack: Crucial Coronial inquiry about 'accountability for actions - or inactions', <https://www.nzherald.co.nz/nz/christchurch-terror-attack-crucial-coronial-inquiry-about-accountability-for-actions-or-inactions/ZEABJUQGUV7LG5XY7NGOSZ7EQ/>, *New Zealand Herald*

This performativity is observed in a number of domestic far-right Telegram channels, and social media accounts, where the crisis is skewed through an antisemitic lens, promoting conspiracy theories around Jewish control, and maliciously portraying Jews as adversaries of Muslims.

Internationally, the outbreak of war has also been used as a pretext to target communities and institutions – Jewish, Muslim, Palestinian, Israeli.

The domestic exploitation extends to promoting historical figures known for their antisemitism, such as Mufti Amin al-Husseini, to substantiate false claims about Jewish conspiracies and to argue for an inherently violent and antisemitic nature within Islam.

Additionally, domestic entities such as Action Zealanda, alongside global far-right networks, have utilised this crisis to glorify Nazi ideologies and incite hatred.

Tactics employed to stoke hatred and radicalise individuals include the blending of archival footage with provocative visuals with videos hosted on far-right platforms like Rumble or BitChute, as well as the promotion of extremist figures and antisemitic conspiracy theories.

Overall, our analysis examines how these extremist networks leverage crises to amplify violent extremist and hateful narratives, mapping domestically, what was contemporaneously studied internationally. It highlights the disturbing nature of ideological alignments between white supremacists and anti-Israeli viewpoints, overwhelmingly founded on antisemitism and appropriation of the Palestinian cause.

Our research surfaced concerning connections between far-right extremism and the representations of the war. We have repeatedly noted that this narrative nexus poses substantial security threats and exacerbates the already violent, volatile discourse surrounding the conflict.

This has ripple effects that extend globally and will impact communities in Aotearoa New Zealand through their resonance in domestic disinformation ecologies, potentially influencing offline behaviours. Mirroring concerns in other countries, our research raises risks to communities in Aotearoa New Zealand as the war progresses. We also note the impact of this content will be felt over the medium, to long term.

‘Narrative flattening’

TDP’s research studied a decontextualisation we have described as "narrative flattening", which we argue is strategic and intentional. This effect is being used by narrators within Aotearoa New Zealand’s disinformation communities to reduce complex identities, histories, and motivations into simplistic binary presentations.

We argue narrative flattening on social media platforms exacerbates social divides, undermines reasoned discourse, and potentially incites harm against targeted individuals or communities.

Examples include:

- Stereotyping all Palestinians – and by extension, all Muslims - as terrorists, despite the religious diversity of the Palestinian people.
- Similarly, in a familiar antisemitic manner, all Jewish people are associated with Israel, and Israeli and Jewish diversity is ignored: all Jewish people and all Israelis are settler-colonists, complicit in or actively promoting what are described as war crimes.
- In the domestic context, some voices promoted flattening narratives that conflated support for Palestine with support for terrorists by accusing academics and politicians of aligning with Hamas.

Antisemitic and Islamophobic rhetoric also manifested through flattening entire religious groups into threats. By erasing nuance, flattened narratives contribute to polarised conflict narratives, the denigration of religions and the dehumanisation of groups. Critical, and nuanced reflection is not just absent – it is actively rejected.

Our research suggests strategic motivations, and divisive intent, linked to well-known disinformation networks, accounts and actors. These accounts often underlined the narrative flattening, with partisan goals and divisive rhetoric taking precedence over considered, civil representations of a complex conflict.

Overall, the weaponisation of reductive, simplistic narratives was identified across social media and studied as a common tactic to spread politically motivated propaganda about the multidimensional Israel-Palestine context, and more immediately, the current war.

Graphic harm, gore and violence

The October 7 attack and subsequent acts of atrocity have signalled the single greatest increase in the production, promotion, and propagation of graphic harms, gore, and violence (GHGV) TDP has studied since the commencement of our daily landscape reporting in mid-2021. This includes our studies of the invasion of Ukraine by Russia in February 2022. It also reflects the algorithmic presentation of GHGV on platforms in addition to the frictionless, borderless discoverability of material through searches and subscriptions.

It is important to state that TDP recognises the legitimate need to record, and strategically share, footage that captures crimes against humanity. However, the nature, volume and availability of the footage we have studied, and its use by disinformation groups locally and internationally, shows these recordings are being weaponised in order to shock, harm and radicalise individuals.

We note that our analysis of graphic harms, gore and violence is driven from and by our site of study, which we reiterate remains domestic social media communities that initially organised around COVID-19 disinformation.

We focus here therefore on the use of GHGV by disinformation networks, and their leveraging of GHGV for information warfare. This is distinct from the necessary and urgent publication and dissemination of content for human rights purposes.

We studied how Hamas content on Telegram in both Arabic and English was shared by domestic accounts when it supported their antisemitic, anti-Israeli stance or, conversely (and as frequently), Islamophobic presentations.

Our research highlighted the disturbing trajectory of GHGV and disinformation proliferating on social media platforms, particularly Twitter/X and Telegram. The research flagged that in addition to the extent of violence, there is increasing public discoverability/accessibility, an enduring presence, and even algorithmic recommendation on Twitter/X which reached an unprecedented level, even in comparison to the study of the invasion of Ukraine in February 2022, which TDP conducted with as much detail.

The availability of GHGV across the spectrum of social media platforms, and more specifically, Telegram, has established a 'total immersion' scenario where users are surrounded by the war's representations, and can also easily discover, download, engage with, and share violent content. The sharing of and appreciation of such content across Twitter/X, and Telegram is highlighted in our research, creating a networked ecosystem of violence-centric content.

The nature, volume, varied vectors, velocity of production, and networked diffusion of GHGV have bypassed all known harm mitigation, and classification protocols, regulations, laws, and frameworks in Aotearoa New Zealand. These include, for example, entities with harm mitigation mandates such as the Classification Office, and The Department of Internal Affairs.

We theorised three intertwined phenomena, unfolding in real-time:

1. Distributed networked harassment
2. GHGV-centric cross-platform networking; and
3. Seamless cross-pollination of GHGV content.

An example of distributed networked harassment involves Green Party MP Golriz Ghahraman, and illustrates how content moved, and was amplified across platforms, contributing to a toxic, abusive environment. This networking allowed harmful content to transcend platform trust, and safety boundaries, rendering policies to mitigate harms entirely ineffective in controlling the spread, and impact of such content.

Concerningly, our studies show a complete absence of empathy and dignity in the presentation of violent content, especially towards women. This was particularly prevalent in the sharing of graphic sexual violence – often equated with pornography - related to women taken as hostages by Hamas. We suggest that we are now seeing the emergence of technology-facilitated gender-based violence as a propaganda tool, and an extension of the use of rape and sexual violence as a war crime.

SPECTACLE, IRRREALISM, AND COGNITIVE WARFARE: IMPLICATIONS FOR SOCIAL COHESION

TDP used the theoretical approach of 'spectacle' in our analysis of the first open-air protests against Level 4 public health measures in late 2021.

French critical theorist Guy Debord's thesis of 'spectacle', expanded on by media scholar Douglas Kellner, as 'media spectacles' are also valid frames to help understand these representations in social media.

In 1967's *La société du spectacle* Debord wrote that:

"The spectacle is not a collection of images, rather, it is a social relation among people, mediated by images."

Social media's production, presentation, engagement and sharing brings into sharp focus how content online can be best understood as a socio-technological phenomenon.

The production of the content is one part of an equation that relies on engagement, (re)presentation and a flurry of commentary, including performative anger and offence-taking, in order for the content to spread amongst communities, including those who are offline and consumers of terrestrial or print media.

Seen this way, online (and social media) content seeks to influence how we see ourselves and, *inter alia*, engage with divergence, dissonance and difference. Based on this foundation, American critical media theorist Douglas Kellner called 'media spectacle' a form of media culture that puts contemporary dreams, nightmares, fantasies, and values on display.

In sum, our application of Debord, and Kellner to disinformation on social media finds:

- Social media content represents a social relation between people, mediated by images/videos/texts. It is not just neutral information but part of a socio-technological system.
- The production, presentation, engagement with, and sharing of content is crucial to its impact.
- Content spreads through performative reactions, outrage, commentary etc. beyond just creators and consumers to influence communities.

These points are further augmented by American writer, critic, and former Editor of *Granta*, John Freeman presentation of 'irreality' or 'irrealism'¹⁰.

Irreality is a state of cultural psychosis characterised by the endless looping of unreality upon unreality at a pace that exceeds our ability to determine truth. In this

¹⁰ The Worst of Times: Our Year in Irreality, <https://lithub.com/the-worst-of-times-our-year-in-irreality/>

mediated environment where shared reality feels impossible, consolidated semantic authority dictates narratives while degrading critical language, and rewarding subjective projects.

Freeman notes that we are left in an overwhelming state of loneliness, paralysis, depression and recklessness, feeling conspired against, and nostalgic for the past. Unable to alter the assault on meaning, we retreat into hyper-subjectivity, striving for self-coherence amidst the incoherence.

Seeking relief, we try to possess and consume representations of reality rather than seeing reality itself. But this only reinforces the loop of degraded images and myths. Mistrust spreads as the uninformed public is turned against the very systems meant to serve them.

These are helpful frames from philosophy and literary theory to explain complex, interrelated, networked social media phenomena related to disinformation in general, and ongoing representations of violence.

What is happening is both an engineered paralysis of our cognitive ability to discern truth, as well as a more algorithmic amplification of psychosis, where the most extreme takes on the conflict are given visibility, and primacy over nuance, critical reflection, experience and expertise.

The result is what we noted earlier: The industrial-grade production, and diffusion of content related to the ongoing atrocities, including GHGV, and disinformation, has completely outpaced institutional frameworks, regulations, laws and responses in Aotearoa New Zealand.

TDP research framing

Our research presents real, present and growing challenges to Aotearoa New Zealand's social cohesion based on this content's ubiquity in domestic information environments and disinformation ecologies.

In our report on 26 October, we stressed Aotearoa New Zealand's Jewish, Palestinian, and Muslim communities needed:

“to be forewarned, and proactively provided with (1) strategic communications around accelerated hyper-partisan affective polarisation, and echo chamber/confirmation bias effects, (2) crisis communication frameworks, in the event of offline, stochastic terrorism, and (3) guidance on what will be a step-change in Jewish-Muslim relationships globally, which obviously will map domestically.”

We note that since 7 October, TDP has provided our research to relevant community groups, civil society organisations, and agencies, through established protocols.

Our observations are anchored to ‘cognitive warfare,’ expanding on the theoretical frames around irrealism, and spectacle. Our research presents cognitive warfare as a synergistic blend of cyber, information and electronic warfare designed to manipulate mass behaviour and thought.

This research surpasses traditional propaganda methods and incorporates sophisticated techniques, such as hyper-personalised disinformation, to exploit people's cognitive biases and heightened emotional states, particularly in crisis situations like stochastic terrorist attacks.

This weaponisation of psychology and technology aims to sow long-term divisions and control societal narratives by presenting 'violent othering' as necessary and urgent, thereby poisoning the well for all forms of activism and advocacy.

Our research also underscores the need for an 'all-of-society' and 'all-of-government' approach to mitigate these threats.

We emphasise that community-oriented strategies must be urgently devised and enduringly supported, especially in the context of escalating geopolitical tensions that already have significant domestic implications both online, and offline.

Mirroring what the 9/11 Commission in the US stated¹¹, the *Guardian* newspaper noted:

“A British intelligence expert suggested Israel had suffered a “failure of imagination”. In 2001, there had been another one, said Alex Younger, a former head of MI6. “9/11 classically was that. The assumption was... that it essentially wasn't possible.”¹²

Our research argues absence of clear, public guidance, and disclosures around Aotearoa New Zealand's national security responses will add to anger, anxiety, and insecurities felt by communities.

The lack of any public-facing commentary makes it impossible to know if, and how, intelligence services are treating the domestic fallout of the violence in relation to, and beyond national security imperatives.

We are particularly concerned that stochastic terrorism early into the new government's term may define how it approaches social cohesion, and other key recommendations of the Christchurch Commission's Report.

This is in addition to whatever may surface during the on-going Coronial Inquiry into the March 2019 massacre. TDP is particularly mindful that this year's NZ SIS report flagged concern about Iranian operations, and operatives in Aotearoa New Zealand¹³.

¹¹ “The most important failure was one of imagination. We do not believe leaders understood the gravity of the threat. The terrorist danger from Bin Ladin and al Qaeda was not a major topic for policy debate among the public, the media, or in the Congress. Indeed, it barely came up during the 2000 presidential campaign.” The 9/11 Commission Report, <https://www.govinfo.gov/content/pkg/GPO-911REPORT/pdf/GPO-911REPORT-24.pdf>.

¹² Seven days of terror that shook the world and changed the Middle East, <https://www.theguardian.com/world/2023/oct/14/seven-days-of-terror-that-shook-the-world-and-changed-the-middle-east#:~:text=A%20British%20intelligence%20expert%20suggested,Younger%20told%20BBC%20Radio%204.>

¹³ China, Russia and Iran interfering in New Zealand's democracy, SIS says, <https://www.stuff.co.nz/national/politics/132727240/china-russia-and-iran-interfering-in-new-zealands-democracy-sis-says>

If Iran enters the conflict, directly and/or through Hezbollah, this may have immediate, and lasting repercussions domestically.

Social cohesion and the permacrisis

Representations of violence since October 7 are stress-testing the social cohesion of Aotearoa New Zealand in unprecedented ways.

Aotearoa New Zealand is now moving from dealing with polycrises alongside other liberal democratic states, (emerging from, or accelerated by the invasion of Ukraine by Russia in early 2022), into a state of permacrisis:

- The conflict is being used by extremist groups globally, and domestically to promote their ideologies and recruit new members. Both far-right and Islamist extremist groups will actively exploit the violence to radicalise individuals. The Islamophobic ideology that motivated the Christchurch shooter flourishes unfettered today in disinformation ecologies, alongside antisemitism which is an everyday content, and commentary signature which on Twitter/X is algorithmically rewarded.
- Online, both forms of hate noted above cite perceived threats by the "other" group and use dehumanising language. Given the country's diverse (visible, and religious) minorities, including recently resettled refugees from Syria, Afghanistan, and elsewhere, the conflict has the potential to inflame extremism domestically if leveraged effectively by these groups to sow or entrench division through disinformation, as well as GHGV's constant, daily flood serving identity motivated violence extremism (IMVE), and white supremacist violent extremism (WSVE).
- The Israel-Palestine conflict is also being used by state and non-state actors to shape public opinion through coordinated influence operations. Foreign states like Iran and Russia have an interest in portraying Israel negatively while bolstering (or remaining silent on) Hamas or Hezbollah. Their state-sponsored media and social media activity, aim to sway perspectives on the conflict in a context where NZ SIS has already confirmed operations aimed at undermining Aotearoa New Zealand's liberal democratic firmament.
- Such influence operations, which TDP has studied over the past two years, and which grew at pace ahead of the general election in 2023, make it much harder for the public to discern truth. They are also 'constant campaigns', in that the disinformation is produced at pace, every day, and not just in the lead-up to an electoral moment.
- It is critical to map known influence networks, detect inauthentic activity, and build societal resilience against propaganda relating to the conflict, which accelerates the truth decay which is already well underway in Aotearoa New Zealand. Increased media literacy and exposure of foreign influence attempts are urgently needed to support critical thinking and mitigate harm.
- The violent conflict's effects, especially if it expands to regional actors/states, will extend beyond social cohesion to include possible economic impacts for

Aotearoa New Zealand. This will increase socio-political volatility. Alongside this, online content and commentary will seek to leverage disenchantment, disconnect, and anti-establishment or anti-government sentiment for more parochial, violent ends.

- For example, prolonged or protracted violent conflict will cause fluctuations in global oil prices, which flow through to domestic fuel costs, and everything else. The cost of living of a core general election campaign issue across all the major parties. Financial markets may also react negatively to geopolitical instability caused by the conflict.
- Resulting economic stresses, like rising living costs or drops in investments, will impact social stability/cohesion. Economic precarity leaves populations more vulnerable to narratives blaming others, including immigrants – adding to disinformation narratives already well-established. This strategic scapegoating, alongside a rise in populism and nationalist rhetoric, can further fray social relations, with disinformation online acting as an accelerant.

HIGH-LEVEL RECOMMENDATIONS

1. Multilateral efforts spanning technology, law, research, education, civil society and government are necessary to curb the production, spread and impact of GHGV online. Platforms must proactively moderate GHGV's seeding and spread. If self-regulation by platforms continues to critically fail in the manner it has since 9 October, government regulations along the lines of proposed Safer Online Services and Media Platforms (SOSMP)¹⁴ will be more urgently required.
2. Public advisories, media literacy programs, law enforcement training and civil society-led counter-narratives must be urgently undertaken to help curb violent content's production and spread. This must be grounded in transparent, persistent research access to platform data and international. Some of these actions can – and should - be brought about by the government, but others will require bilateral and multilateral collaboration, including with private industry.
3. Government, community, and faith leaders must model nuance and acknowledge the complexity of current events, even if this is politically inexpedient. Their example and speech can shift culture and responses online and offline. This impact of modelling has been well-established over many decades in dangerous speech research. Media literacy programs and standards also help.
4. We believe critical citizens, leveraging empathy against simplistic narratives which seek to farm outrage can be one of the strongest safeguards to social cohesion. Education and engagement are paramount to equipping people with these skills.
5. While respecting rights and first principles, Aotearoa New Zealand's security agencies need to closely monitor far-right online activity and threats, embracing new, complex, fluid, transnational, multi-modal network dynamics which now go beyond what they have trained around and have insights on (see "streetlight effect" below).

¹⁴ <https://www.dia.govt.nz/safer-online-services-media-platforms-consultation>, DPMC

6. An early recommendation in our analysis of the conflict was for Aotearoa New Zealand's intelligence services and others to urgently expand their ability to directly engage with Arabic content, without relying on online, machine translations. There's a strong argument to focus on a humanities-based approach to issues presented by the conflict's shadow in Aotearoa New Zealand, embracing the deconstruction of narratives. This includes critical evaluation of how they are told or presented, what they exclude, where they originated, how they were imported, why, and by whom.
7. Public education around extremist appropriation of causes is critical to prevent radicalisation. Interfaith initiatives which emphasise shared values can counter these divide-and-rule tactics.
8. Responsible rhetoric is required from all leadership, including civil society. Politicians and activist groups must work to identify when the support of either position in the violence is being used to further antisemitism or Islamophobia. A collaborative, multi-media, offline, and online strategic communications effort – led by communities, urgently scaled up, and helped by government agencies - can deny (or significantly stymie) extremists the oxygen to incite harm, and inflame anger.
9. Progressive civil society groups, interfaith alliances and human rights organisations can proactively counter extremist voices by spreading unifying narratives that emphasise shared values, highlight common humanity, and prevent the violent othering of minority communities.
10. Law enforcement agencies must remain vigilant to the possibility of attacks or hate crimes targeting Muslim, Jewish, Palestinian or other minority communities, which may be provoked by the spread of inflammatory propaganda and disinformation by groups exploiting foreign conflicts.

CONCLUDING REMARKS

There's an old tale that contains an important truth about human nature. A man searching for his lost keys is looking under a streetlight even though he knows he dropped them elsewhere. When asked why he doesn't search where he dropped them, the man replies "Because the light is better here."

This parable illustrates a common tendency to seek solutions in convenient places, and observable spaces, rather than where the answer likely lies. Social scientists refer to this as the "streetlight effect" – the focusing of efforts on easily examined spaces, while neglecting obscured but relevant areas.

The man's search for his lost keys exemplifies universal patterns of human behaviour and reminds us that the most meaningful discoveries often exist outside the familiar glow of what we already know.

It also finds expression in mathematician Abraham Wald's study of Allied Forces planes during World War II, and what many at the time thought would help them fight better – which was to strengthen the sections with the most bullet holes. However, Wald noted the bullet holes indicated where planes could be hit, and still make it

home - not the most vulnerable areas. He proposed that they should armour areas without bullet holes because planes hit in those places never made it back. This became known as Wald's Paradox - seeming evidence that should strengthen a hypothesis can actually invalidate it if the sample is incomplete. By considering the "missing" data (the aircraft that didn't return), Wald helped the right conclusions to be drawn.

This research note's authorship was led by Dr Hattotuwa and fundamentally informed by life experiences which provide unique insights into truth decay and disinformation. That this background is now so critical to the analysis of Aotearoa New Zealand's disinformation ecologies, especially after 9 October, is a very worrying sign.

Given the relative absence of comparable situated, lived, and research experience in Aotearoa New Zealand, one danger is that the country's institutions, including intelligence services, only rely on what they know and the places they usually gather intelligence, to strengthen what they already believe. This can lead to the unimagined being realised without preparation, or meaningful responses.

As our research attests every day, the complex, real-time, and emotionally charged social media conversations around the conflict and underlying context make it near impossible to introduce nuance without being perceived as biased.

On social media, attempting to provide context about Israel's security concerns and actions inevitably angers some who feel this ignores Palestinian suffering. Meanwhile, criticising Hamas' tactics elicits accusations of justifying Israel's disproportionate force. Every perspective is interpreted as propaganda by the opposing side. Rage, grief, and hurt dominate. There's little to no space to reflect, and connect.

This is cognitive warfare. Strategic, well-resourced outrage actors seed division in order to benefit from accelerated affective polarisation, and the realisation of end states which have nothing to do with Israel or Palestine, including in countries like Aotearoa New Zealand. This is a socio-technological incarnation of the 'fog of war'.

What can one meaningfully do, to retain the elasticity of hope?

Perhaps Naomi Klein has a moral answer¹⁵. An excerpt from a recent article of hers provides a meaningful foundation for the design of strategic communications urgently and enduringly needed in Aotearoa New Zealand (online, and offline, including through political leadership) to fight the fallout of the conflict in Israel and Palestine and protect the country's communities, increasingly fearful, "from the shafts of strife and war," through "bonds of love".

The Israeli state's current murderous levelling of Gaza is the latest, unspeakably horrific manifestation of this ideology, and there will be more in the coming days. The responsibility for these crimes of collective punishment

¹⁵ In Gaza and Israel, side with the child over the gun, <https://www.theguardian.com/commentisfree/2023/oct/11/why-are-some-of-the-left-celebrating-the-killings-of-israeli-jews>

rests solely with their perpetrators and their financial and military backers abroad. But we all have to figure out how to make it stop.

So how do we confront this violent ideology? For one thing, we can recognize that when Israeli Jews are killed in their homes and it is celebrated by people who claim to be anti-racists and anti-fascists, that is experienced as antisemitism by a great many Jews. And antisemitism (besides being hateful) is the rocket fuel of militant Zionism.

What could lessen its power, drain it of some of that fuel? True solidarity. Humanism that unites people across ethnic and religious lines. Fierce opposition to all forms of identity-based hatred, including antisemitism. An international left rooted in values that side with the child over the gun every single time, no matter whose gun and no matter whose child.

Kia hua ko te pai.

Dr Sanjana Hattotuwa
Kate Hannah

Originally sent to relevant groups and agencies on 3 November 2023