

Witnessing, Condemning, Debating: The Multiple Roles of Facebook in Sorcery Accusation-Related Violence in Papua New Guinea

Miranda Forsyth, Anurug Chakma and Joe Barak

Abstract

Facebook is the dominant digital public sphere in Papua New Guinea (PNG), and sorcery accusation-related violence (SARV) is amongst the country's most serious forms of collective violence. This article presents the first systematic empirical analysis of how the two intersect. We draw on 507 posts from 24 public pages and groups collected between 2013 and 2024. Posts were coded for stance and analysed using Firth logistic regression and structural topic modelling, supplemented by qualitative analysis of two high-engagement comment threads. We find an apparent paradox. Public-facing Facebook discourse is overwhelmingly anti-SARV: 83 per cent of posts oppose the violence. Yet this apparent consensus is largely produced by a relatively small group of high-volume government, NGO and advocacy actors, rather than reflecting the distribution of opinion amongst Facebook users more generally. Stance also varies significantly across actor categories, with church pages in our dataset markedly less likely to produce oppositional content, despite strong church opposition to SARV offline. At comment level, discourse is considerably more contested. A particularly important position accepts the reality of sorcery while rejecting violence against accused persons, complicating any advocacy approach that would assume disbelief in sorcery to be a necessary basis for opposing SARV. Across the dataset, Facebook emerges not simply as a medium through which harmful narratives circulate, but simultaneously as an incident-reporting system, advocacy infrastructure, a stage for official public statements and an arena for public debate. These findings demonstrate the importance of distinguishing institutional messaging from audience discourse and suggest that effective digital strategies against SARV must engage seriously with the plural religious, cultural and epistemological frameworks through which sorcery and violence are understood.

Witnessing, Condemning, Debating: The Multiple Roles of Facebook in Sorcery Accusation-Related Violence in Papua New Guinea

Miranda Forsyth, Anurug Chakma and Joe Barak

Miranda Forsyth is a professor at the School of Regulation and Global Governance at the Australian National University.

Anurug Chakma is a research fellow at the School of Government and International Relations at Griffith University.

Joe Barak is a researcher at the Papua New Guinea National Research Institute.

Forsyth, M. Chakma, A. Barak, J. 2026. Witnessing, Condemning, Debating: The Multiple Roles of Facebook in Sorcery Accusation-Related Violence in Papua New Guinea, *Development Policy Centre Discussion Paper 122*, Crawford School of Public Policy, The Australian National University, Canberra.

This research was funded by Australian Research Council Future Fellowship FT230100612.

Please send any comments to miranda.forsyth@anu.edu.au.

The Development Policy Centre is a research unit at the Crawford School of Public Policy, The Australian National University. The discussion paper series is intended to facilitate academic and policy discussion. Use and dissemination of this discussion paper is encouraged; however, reproduced copies may not be used for commercial purposes.

For more information on the Development Policy Centre, visit devpolicy.crawford.anu.edu.au

1. Introduction

In Papua New Guinea (PNG), Facebook functions as the dominant digital public sphere: 71.6 per cent of all social media users in the country are on Facebook (Statcounter 2026), creating a shared space where news circulates, communities debate, and events are witnessed and contested in real time. This digital landscape has profound implications for a country facing serious and persistent levels of violence, including sorcery accusation-related violence (SARV). SARV involves individuals being accused of practising sorcery or witchcraft, typically in response to sickness, death or misfortune, and subsequently subjected to community ostracisation, torture, displacement and killing (Forsyth et al 2025). As one interviewee in our study stated, "Sorcery gives people hope, it provides an explanation, provides a solution. But this is where it gets dangerous".

Researchers and human rights observers have increasingly noted the potential for social media to shape violence dynamics in PNG. When a mass killing took place in Enga Province in February 2024, graphic images of a truck piled with dead bodies circulated on WhatsApp, first alerting people across the country to the violence. Similarly, when riots spread through Port Moresby and other cities in January 2024, Facebook posts and WhatsApp messages played a demonstrable role in their propagation (Barak, Chakma and Forsyth, 2024).

How, then, do these digital dynamics operate in relation to SARV, arguably the most extreme form of collective violence in PNG today? In a previous pilot study (Barak, Chakma and Forsyth, 2024), we studied patterns in YouTube comments on videos involving SARV. This research revealed complex patterns in the interactions between people inside and outside of PNG that raised questions about the role of international and diaspora actors and a striking proliferation of "spiritual warfare" discourse (rhetoric that blended Pentecostal Christian beliefs with justifications for violence against those accused of sorcery). However, that pilot study was necessarily limited in its methods and scope. This article presents a more systematic and comprehensive empirical analysis focused on Facebook specifically.

We ask three related core research questions:

1. Actor opposition: Who actively opposes SARV on Facebook in PNG, and how does this opposition vary across actor categories, including government entities, non-government organisations (NGOs), religious organisations, media, and public forums?
2. Discursive framing: How do dominant themes and discursive frames surrounding SARV differ by stance - support, opposition, or neutrality - and by actor category?

3. Platform functions: More broadly, beyond the simple question of whether Facebook promotes or opposes SARV, what other roles does the platform play, and what does this mean for efforts to reduce SARV through digital communications?

To our knowledge, this study represents the first peer-reviewed systematic empirical analysis of social media and SARV in any Pacific context. Our findings confirm the “ambivalent platform” dynamic noted by human rights advocate Leo Igwe in Nigeria (Igwe, 2026) whilst also revealing dynamics specific to PNG: the dominance of state and civil society actors in organised opposition to SARV, significant variation across actor categories, and the contested and multifaceted character of grassroots public debate that is visible at the level of individual comments but largely invisible in institutional posts. This article also contributes to the emerging literature on online misinformation in PNG (Yangin, 2023) and the Pacific more generally.²

The article is structured as follows. Section two provides background on SARV in PNG and on the social media and digital media landscape. Section three describes the data and methods. Section four presents our findings. Section five discusses the implications of our findings and outlines policy recommendations. Section six concludes the paper.

2. Background

2.1 SARV in PNG: Scale, Characteristics and Policy Context

SARV encompasses the violence - physical, psychological, economic, and social - perpetrated against individuals accused of practising sorcery or witchcraft. There are different narratives about sorcery circulating in PNG, the most common form in the Highlands involves *sanguma* (spiritual attack, often framed as the accused consuming the vital essence of others), whilst in coastal and island provinces, *posin* (sorcery done through deliberate practices, often using people’s hair or other body parts) is more prevalent. The accusation process commonly involves initial suspicion following a death or misfortune, gossip that circulates, and the crystallisation of suspicion by an authority figure such as a diviner or *glasman* who identifies the *sanguma* (Forsyth et al., 2024).

SARV is not new, but its scale is widely underappreciated (Forsyth et al., 2025). Analysis of newspaper reports from 1996 to 2022 identified 557 reported incidents, with 694 victims killed - figures that represent only a fraction of actual violence, given that only eight per cent of violent incidents and 17 per cent of

² The field of Pacific digital research has developed rapidly, marked by milestones such as the inaugural Pacific Disinfo Conference in 2024 (see Pacific Disinfo Network, 2024: <https://disinfopacific.org/>).

incidents leading to death are captured in national media (Forsyth et al., 2024). SARV impacts victims of all genders and ages, though gendered patterns are strong and vary significantly by province. In the Highlands, and particularly the province of Enga, women are the majority of victims; in the Autonomous Region of Bougainville, the majority are men; in the province of Jiwaka, accusations typically target both a husband and wife. Violence against female victims disproportionately involves torture, burning, sexual violence and captivity; killing is slightly more prevalent in incidents with male victims (Forsyth et al., 2024). These figures make clear that SARV is neither a predominantly female problem, as is sometimes assumed, nor a predominantly rural or Highlands one: incidents have been reported in all provinces and in the National Capital District.

The harm of SARV extends well beyond the immediate incident. Accusation functions, as one survivor described, like “a birthmark that can never be removed”, bringing stigma, displacement, ongoing fear, and often also condemning children of those accused to lifetimes of shame and potential future accusation (Forsyth et al., 2024).

PNG's legal response to SARV has strengthened over time. The *Sorcery Act 1971*, widely regarded as having legitimised aspects of sorcery belief, was repealed in 2013. The Criminal Code (*Criminal Code Act 1974*) now includes section 299A (wilful murder on account of sorcery accusation) and the *Glasman Act 2022* (section 520e of the *Criminal Code Amendment Act*), which significantly strengthened the legal framework. Prosecutions have increased over time, though they remain far below the actual scale of violence (Forsyth, 2021).

2.2 Facebook and the Digital Media Landscape in PNG

Papua New Guinea had 2.57 million internet users (24.1 per cent of the population) and 1.30 million social media user identities (12.2 per cent) at the start of 2025 (DataReportal, 2025). Facebook overwhelmingly dominates the digital landscape (DT Institute, 2026). A 2026 national scoping study of technology-facilitated gender-based violence similarly found that Facebook functions as a quasi-internet for many users, and identified it as the primary platform for the perpetration of technology-facilitated abuse (Osborne, 2026). As Dwyer and Minnegal (2020: 243) observed in their study of a regional Facebook group in PNG during the COVID-19 period, the platform “provides far more than a medium of contact between ‘friends’. It facilitates the spread of local and national news, and discussion of that news, more widely than may be achieved by print, radio and television”.

Connectivity, however, is limited and uneven. As Foster (2023; 2025) shows in the most sustained scholarly analysis of PNG's digital infrastructure, connectivity is “tenuous”, spatially concentrated in urban centres, temporally variable, and

constrained by high data costs and limited electricity access. Approximately two-thirds of Facebook activity originates from Port Moresby and Lae (Pacific Security College, 2024). Rural areas, where 86 per cent of the population lives and where SARV is particularly prevalent, are often only intermittently connected. The DataReportal statistics cited above should therefore be understood as reflecting an urban-weighted, predominantly male social media population (approximately 62 per cent of Facebook's advertising-reachable users in PNG identifies as male [NapoleonCat, 2024]).

Facebook is not the only relevant social media platform in PNG; TikTok and WhatsApp also command significant user bases, particularly amongst younger demographics (UNICEF PNG, 2024). However, our focus on Facebook reflects its clear dominance in news sharing and civic information, the availability of open public data for research purposes, and insights gained from our prior pilot work.

The adequacy of platform governance in PNG generally is a broader topic than can be covered here. Facebook's content moderation capacity in Tok Pisin and Papua New Guinean languages is widely assessed as inadequate (DT Institute, 2026; Tapia, 2024). Security expert Meg Tapia argues that automated moderation systems and human reviewers fail because they do not cater to the region's linguistic complexities or cultural contexts (Tapia, 2024; see also Sinpeng et al., 2021). The PNG government's 2018 attempt to suspend Facebook over concerns about fake accounts and graphic content, and its invocation of anti-terrorism legislation to order a 24-hour shutdown in March 2025, citing misinformation, tribal killings, hate speech, and incitement to violence (ABC News Australia, 2025), indicate that the state recognises the platform's role in enabling real-world violence. This regulatory friction provides a significant contextual factor for the analysis that follows.

This expansion of connectivity has outpaced the development of media literacy. The dynamics of SARV narrative spread in PNG have previously been described using a contagion framework (Forsyth and Gibbs, 2020): harmful narratives transmit through social networks, with spread shaped by population susceptibility, available transmission mechanisms, and the presence of "contagion entrepreneurs", such as the *glasman* referred to earlier, who actively disseminate and amplify harmful content. These narratives often take the form of "causal stories" that explain deaths, illness and other misfortunes through reference to sorcery (Forsyth and Gibbs, 2022). As discussed below, we hypothesise that Facebook's architecture, specifically algorithmic amplification, visible public engagement metrics, and frictionless share mechanics, functions as an efficient transmission mechanism for these causal narratives.

3. Data and Methods

3.1 Data Collection

We developed the dataset for this study by extracting 67,400 posts from more than 30 public Facebook groups and pages over the period 2013–2024, using the ESUIT data extraction tool.³ Pages and groups were selected to include SARV-specific forums, news pages, development and civil society organisations, and provincial and national community discussion groups, covering Enga, Hela and Madang provinces, as well as national-level pages. This quantitative dataset was supplemented with semi-structured interviews with representatives from two civil society organisations in 2025 to examine their use of Facebook. The organisations were selected based on their direct involvement in SARV advocacy and policy implementation. Human research ethics approval for these interviews was granted by Australian National University.⁴

We filtered this initial dataset using the keywords *sorcery*, *sanguma*, *SARV*, *magic*, and *evil spirits*, producing 660 posts from 24 public Facebook pages and groups. We then manually reviewed all remaining posts to remove irrelevant entries, producing a final cleaned dataset comprised of 507 Facebook posts substantively addressing SARV. Our search did not identify any Evangelical or Pentecostal churches posting on SARV; only Catholic Church related pages and the Melanesian Institute (covering a range of denominations) appeared. Whilst this represents a boundary constraint on the dataset, the absence of Pentecostal institutional discourse is an important empirical finding in itself.

The dataset covers publicly accessible content only. Private Facebook groups, personal timelines, and encrypted platforms such as WhatsApp fall outside the study's scope. The keyword-based filtering strategy using terms such as **sorcery**, **sanguma**, **SARV**, **magic**, and **evil spirits**, may have introduced a sampling bias towards posts opposing SARV, since these terms are often used in public, institutional and advocacy-oriented discussions that frame SARV as a human rights, criminal or social problem. As discussed below, content circulating in less visible spaces is likely to differ substantially from the institutional discourse examined here.

3.2 Stance Coding and Actor Classification

Two members of the research team (Chakma and Barak) independently assigned each post a stance label - *support*, *neutral*, or *oppose* - concerning SARV. Inter-

³ On Facebook, "pages" are public profiles managed by official entities (such as government agencies, media outlets, churches or NGOs) to broadcast announcements and posts to followers. Groups are collective spaces designed for peer-to-peer discussion, where any member can create a post or start a discussion.

⁴ Protocol 2024/0653.

coder reliability was substantial (Cohen's $\kappa = 0.70$), with 91.5 per cent raw agreement across the 507 posts. Disagreements were resolved through discussion between both coders.

Each page and group was also assigned to one of seven actor categories based on its stated purpose, organisational affiliation, and posting patterns: Advocacy, Church, Government Committee, Media, NGO, Public Forum, and United Nations (UN).

3.3 Quantitative Analysis

To address the first research question, we conducted a descriptive analysis of stance distributions across pages and actor categories. We then used Firth logistic regression (Firth, 1993) to identify predictors of an oppositional stance towards SARV, controlling for total post engagement (log-transformed combined reactions, comments and shares). Firth's penalised maximum likelihood approach is specifically suited to this dataset, given the small number of posts coded as supportive of SARV, where standard logistic regression would produce biased estimates. Advocacy groups serve as the reference category. Because posts are nested within 24 Facebook pages/groups and actor category is defined at the page/group level, we additionally conducted page-level robustness analyses and a sensitivity analysis excluding the single Government Committee page.

3.4 Structural Topic Modelling

To address the second research question, we applied Structural Topic Modelling (STM) (Roberts et al., 2014) to the post content. Text was preprocessed to remove punctuation, URLs, stopwords, and noise. Model selection diagnostics, evaluating held-out likelihood, residuals, semantic coherence and exclusivity, indicated that a 15-topic solution best balanced model quality and interpretability. Topics were labelled by examining the highest-scoring FREX words (terms that are both frequent within a topic and relatively exclusive to it) in conjunction with a sample of representative posts for each topic.

3.5 Qualitative Analysis of Comment Threads

To capture dynamics not visible in post-level data, we analysed two comment threads from recent public Facebook posts relating to specific SARV incidents in October and December 2025. These threads, totalling approximately 85 comments, were selected because they represented comments posted publicly in response to recent, real-time, named SARV incidents and attracted substantial engagement. We conducted thematic analysis of the comment data, identifying recurrent positions, popular framings, and underlying patterns of argument.

The artificial intelligence tool Claude Code (Anthropic) was used during the qualitative analysis phase to identify cross-cutting patterns across the comment

data. To support thematic extraction, Claude’s prompts were structured iteratively using responsive analytical instructions. All computational outputs were then manually cross-checked and verified against our own qualitative coding to ensure analytical rigour. It was further used during manuscript preparation to assist with structuring and scoping the presentation of the findings. All analytical interpretations, critical evaluation and final drafting are the responsibility of the human authors.

All data analysed in this study are publicly accessible. No individual commenters are identified by name in this article.

4. Results

4.1 Who Is Posting and With What Stance?

Three features of the stance data stand out. First, public-facing Facebook discourse in our dataset is overwhelmingly opposed to SARV. Of the 507 posts analysed, 422 (83.2 per cent) were coded as oppositional, 75 (14.8 per cent) as neutral, and only 10 (2.0 per cent) as supportive. This apparent consensus, however, is strongly shaped by who is doing the posting. The Government committee page, together with NGO and advocacy pages generated 421 posts - 83.0 per cent of the entire dataset - and accounted for 368 of the 422 oppositional posts (87.2 per cent). The 83 per cent figure should therefore be understood primarily as evidence of the dominance of anti-SARV institutional actors within the public Facebook spaces captured by our dataset, rather than as a measure of the distribution of opinion amongst Facebook users in PNG more generally.

Table 1: Dataset Composition and Stance Distribution.

Actor category	Total posts (N)	Oppose n (%)	Neutral n (%)	Support n (%)
Government committee	72	72 (100.0)	0 (0.0)	0 (0.0)
NGO	216	184 (85.2)	32 (14.8)	0 (0.0)
Advocacy	133	112 (84.2)	17 (12.8)	4 (3.0)
Church	27	13 (48.1)	13 (48.1)	1 (3.7)
Media	14	11 (78.6)	1 (7.1)	2 (14.3)
Public forum	25	16 (64.0)	6 (24.0)	3 (12.0)
United Nations	20	14 (70.0)	6 (30.0)	0 (0.0)
Total	507	422 (83.2)	75 (14.8)	10 (2.0)

Note: Percentages are calculated within actor categories. *N* denotes the total number of posts; *n* denotes the number of posts in each stance category.

Second, stance varies considerably by actor type. Opposition was universal amongst the 72 posts from the Government Committee page and very high amongst NGOs (85.2 per cent) and advocacy organisations (84.2 per cent). Media pages were also predominantly oppositional (78.6 per cent). By contrast, only 48.1 per cent of Church posts were coded as oppositional, with an equal proportion coded as neutral. Public Forums were 64.0 per cent oppositional, while UN pages were 70.0 per cent oppositional. Neutral posts were particularly associated with organisations whose Facebook activity included information dissemination, service reporting and program announcements. Such posts often documented SARV cases or organisational responses without themselves taking an explicit position on the legitimacy of sorcery accusations or associated violence.

Posting activity was also highly concentrated within a relatively small number of high-volume institutional actors. Amongst the most prolific producers of anti-SARV content were the PNG Permanent Parliamentary Committee on Gender Equality, Women's Empowerment and Gender-Based Violence, the Melanesian Institute, PNG Tribal Foundation and PNG Against Violence. These actors repeatedly framed SARV through the language of human rights, criminal accountability and government responsibility. Thus, the strong anti-SARV character of the dataset reflects not simply the aggregation of many independent voices, but the sustained digital activity of a relatively small institutional advocacy ecosystem.

Third, explicit support for SARV was exceptionally rare within these public spaces. Only ten posts (2.0 per cent) were coded as supportive. These appeared sporadically across a small number of groups and pages, including EHP Forum, *Sorcery Killings IN PNG* and media pages, and often arose in the context of reporting or discussing particular incidents rather than as unambiguous advocacy of violence. This low prevalence should not be read as evidence that pro-SARV views are equally uncommon amongst Facebook users in PNG. Our dataset consists of publicly accessible pages and groups and is weighted towards institutional and advocacy actors; content circulating through personal accounts, private groups and encrypted platforms falls outside its scope. As the exploratory comment analysis in Section 4.3 suggests, positions that are largely absent from the public-facing post dataset become visible when analysis moves to individual users' comments.

We tested the descriptive differences between actor categories using Firth logistic regression, with opposition to SARV as the outcome, Advocacy actors as the reference category, and post engagement included as a control. The post-level analysis broadly confirmed the descriptive pattern. Government Committee posts had substantially higher estimated odds of opposition than Advocacy posts (OR =

24.79, $p < .001$), although this estimate requires considerable caution because the category consists of a single parliamentary committee page and all 72 of its posts were oppositional. NGO and Media posts did not differ significantly from Advocacy posts. Church, Public Forum and UN posts were significantly less likely than Advocacy posts to be oppositional (Church OR = 0.15, $p < .001$; Public Forum OR = 0.26, $p = .007$; UN OR = 0.28, $p = .028$). Higher post engagement was also positively associated with an oppositional stance (OR = 1.37, $p < .001$).

Table 2: Firth Logistic Regression Predicting Oppositional Stance Towards SARV Across Actor Categories.

Predictor (Actor category)	Coefficient (β)	Std. Error (SE)	Odds Ratio (OR)	95% CI for OR	p-value
Intercept (Advocacy ref.)	1.02	0.29	2.77	[1.58, 5.02]	< .001
Church	-1.88	0.46	0.15	[0.06, 0.37]	< .001
Government committee	3.21	1.44	24.79	[3.31, 3172.39]	< .001
Media	-0.8	0.67	0.45	[0.13, 1.92]	.256
NGO	-0.32	0.33	0.73	[0.38, 1.38]	.335
Public forum	-1.36	0.49	0.26	[0.1, 0.69]	.007
United Nations	-1.28	0.55	0.28	[0.1, 0.86]	.028
Engagement (log transformed)	0.31	0.09	1.37	[1.14, 1.66]	< .001

Note: Advocacy is the reference category. Odds ratios and 95% confidence intervals are exponentiated from the Firth logistic regression coefficients. Confidence intervals and p-values are based on profile penalized likelihood.

Because the 507 posts are nested within only 24 Facebook pages and groups, these post-level estimates cannot be treated as if every post were an entirely independent observation. We therefore conducted additional page-level and sensitivity analyses. These produced a more conservative picture. When posts were aggregated to the page/group level, the lower likelihood of opposition amongst Church pages remained statistically significant, whereas the differences for Public Forums and UN pages no longer reached the conventional 0.05 significance threshold. The Government Committee effect could not be reliably estimated at page level because the category contained only one page. As a further check, we re-estimated the post-level model after excluding the Government Committee page; the coefficients for the remaining actor categories and engagement were virtually unchanged. These analyses strengthen confidence in

the finding concerning Church pages but suggest greater caution in interpreting the Public Forum and UN differences as general actor-category effects (see **Tables 3 and 4** in Appendix).

The Church finding is particularly noteworthy, but it requires careful interpretation. The Church pages represented in the dataset were Catholic, and their relatively low level of explicit opposition does not imply that Catholic organisations or leaders are indifferent to, let alone supportive of, SARV. On the contrary, prominent Catholic leaders have been amongst PNG's most consistent public opponents of sorcery accusation-related violence. Rather, the finding points to a gap between strong offline church advocacy and the content circulating through the Catholic Facebook pages captured in our dataset, which tended more towards pastoral, devotional and informational material than explicit anti-SARV advocacy.

The absence of other church voices is also significant. Our search identified no Pentecostal or Evangelical institutional pages posting on SARV, despite the prominence of these denominations in PNG and the existence of substantial offline church engagement with the issue. Ten major denominations have, for example, adopted the *National Churches Strategic Plan to Address Sorcery Accusation Related Violence*, committing themselves to collective action against SARV (Australian High Commission, n.d.). Taken together, these findings suggest that a substantial body of church opposition to SARV is not being translated into the public digital sphere captured by our dataset. Given the reach and authority of churches within PNG communities, this represents an important gap in the current anti-SARV Facebook ecosystem.

Overall, then, the central finding of this section is not simply that 83 per cent of Facebook posts oppose SARV. Rather, the apparent anti-SARV consensus is largely a product of the institutional production of public Facebook discourse. A relatively small group of government, NGO and advocacy actors generates much of the visible content and has successfully established a strongly condemnatory public-facing narrative. At the same time, the variation between actor categories - and especially the relative digital absence of church opposition - shows that this institutional consensus is uneven. As the following sections demonstrate, it also differs substantially from the more plural and contested positions visible when analysis moves from organisational posts to the topics they discuss and, especially, to individual users' comments.

4.2 How Is SARV Framed? Evidence from the STM

The STM identifies 15 topics, which we interpret using FREX terms and representative posts and organize into five broad online discourses (**Table 5**). The prevalence of these topics varies across stance categories and over time, respectively (**Figures 1-2**).

Table 5: Structural Topic Model (STM) clusters and keywords.

Cluster/Theme	Topics	Top FREX words (frequency & exclusivity)
1. Incident reporting & witnessing	T2: Police response & torture cases	T2: <i>rescu, enga, facebook, tortur, commission, post, polic</i>
	T8: Village killings & criminal incidents	T8: <i>three, exist, husband, bodi, suspect, villag, told</i>
2. Religious & cosmological discourse	T3: Religious beliefs & <i>sanguma</i> cosmology	T3: <i>sanguma, sick, christian, sorcer, god, believ</i>
	T7: Church & local religious leadership	T7: <i>wabag, land, sister, hunt, ain, mean, bishop</i>
	T13: Church statements & moral framing	T13: <i>cathol, societi, church, bishop, archbishop</i>
3. Advocacy & awareness	T1: Media/policy communications	T1: <i>brief, polici, film, octob, video, australia</i>
	T4: False accusations & justice narratives	T4: <i>fals, real, littl, tortur, foundat, lutz, child</i>
	T6: Community outreach & tribal prevention	T6: <i>senisim, tribal, pasin, partnership, thank</i>
	T11: Storytelling, awareness & mobilisation	T11: <i>stori, relat, violenc, share, activ, respect</i>
4. Institutional & governance responses	T5: Institutional capacity & GBV coordination	T5: <i>djag, ask, wit, dfcdr, respond, money, want</i>
		T12: <i>committe, hear, gew, parliamentari, parliament</i>

	T12: Parliamentary inquiries & GBV governance	T15: <i>abel, crimin, crime, hon, glasman, make</i>
	T15: Legal reform & criminalisation debates	
5. Development programming & services	T9: Program implementation & dev plans	T9: <i>tvi, llcap, pdia, kundiawa, development</i>
	T10: International partnerships & global campaigns	T10: <i>qut, undp, queensland, pacif, yumisanapstrong</i>
	T14: Survivor support & service provision	T14: <i>femili, trauma, sexual, servic, child, abus</i>

Incident reporting and witnessing (Topics T2, T8).

This cluster involves posts reporting specific SARV events, such as rescues, torture cases and killings, often drawing on police or NGO accounts. Representative content includes posts circulating news reports of specific incidents: identifying perpetrators, naming victims, and detailing community responses. Incident-reporting topics are visible across the earlier years of the dataset, although their prevalence varies considerably by year. These topics occur across stance categories, reinforcing that incident reporting does not itself constitute either endorsement or condemnation of SARV (**Figure 1**).

Religious and cosmological discourse (Topics T3, T7, T13).

This area encompasses posts engaging with questions of belief: what *sanguma* is, whether it is ontologically real and how Christianity relates to it. Topic T3 loads heavily on terms including *sanguma, sick, christian, sorcer, god, believ*, reflecting the contested theological terrain around sorcery in PNG. Posts range from those asserting the reality and danger of sorcery from a spiritual warfare perspective (framing sorcery as demonic activity that must be combatted through prayer and faith) to those arguing that sorcery belief is incompatible with Christian faith and that violence is unjustifiable. Both oppositional, neutral and supportive posts appear here, reflecting genuine diversity of SARV positioning within Christian communities.

Advocacy and awareness (Topics T1, T4, T6, T11).

Covering community outreach, peacebuilding, storytelling and mobilisation content, this cluster consists of predominantly oppositional and neutral posts. Representative posts in this cluster frequently come from advocacy groups and NGOs. Topic T11, loading on *stori, relat, violenc, share, activ, respect, read, sarv, workshop*, represents a deliberate use of Facebook as an awareness-raising and coalition-building platform. Representative posts announce workshops, share program reports and circulate survivor stories carefully framed in a clear human rights register.

Institutional and governance responses (Topics T5, T12, T15).

This cluster includes parliamentary inquiries, gender-based violence (GBV) governance, legal reform and criminalisation debates, drawing on official statements, policy documents and legislative reports. These topics have become considerably more prominent in several years after 2020. Predominantly associated with government committee pages and major NGOs, they reflect a broader shift from incident documentation to formal institutional advocacy (discussed further in Section 5).

Development programming and survivor services (Topics T9, T10, T14).

This category covers program implementation, survivor support and frontline service provision. Femili PNG's service announcements exemplify this category, alongside posts announcing Safe House services, Village Court magistrates' training and school education initiatives. The distribution of these topics across stance categories and over time is shown in **Figure 1**.

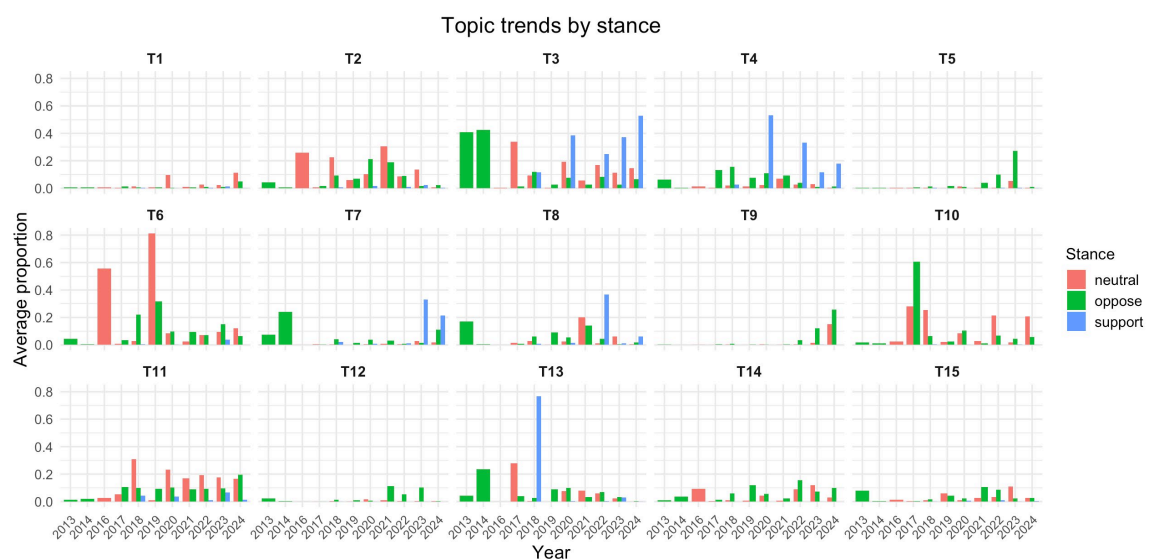


Figure 1: Topic prevalence by stance and year. Graph produced by Authors.

The temporal pattern across the dataset is particularly informative:

1. Early discourse was characterised by incident reporting and religious and *sanguma* cosmology. Facebook functioned as a space for witnessing and discussing violence, with religious framings providing the dominant interpretive lens for either condemning or making sense of incidents.
2. Governance, legal reform, and survivor services topics have become substantially more prominent over time (**Figure 2**).

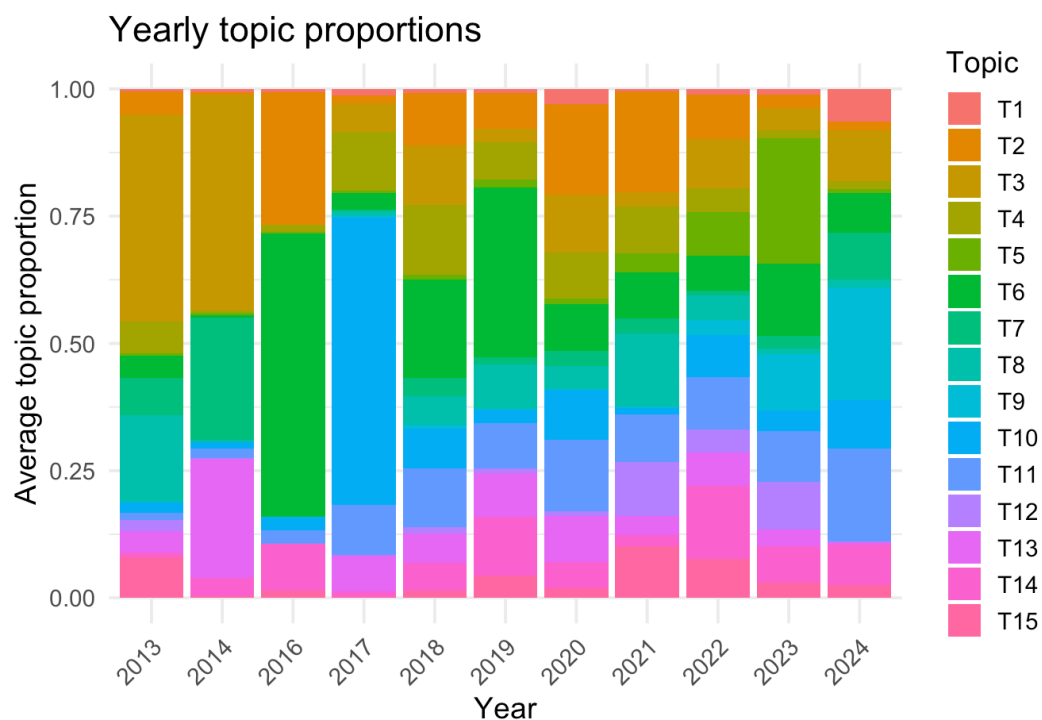


Figure 2: Yearly topic proportions. Graph produced by Authors.

It is also possible, although it cannot be tested with our data, that pro-SARV discourse has partly migrated away from public pages to private Facebook pages and encrypted WhatsApp groups during this period, potentially contributing to the relative predominance of anti-SARV actors in the public Facebook sphere.

In sum, the STM identifies considerable variation in the prevalence of different discursive functions across both stance categories and time. Incident-reporting topics occur across all stance categories, consistent with the fact that reporting an incident does not in itself indicate endorsement or condemnation. Religious and cosmological topics likewise span oppositional, neutral and supportive posts, illustrating the contested role of religious and sorcery belief in online debate. Advocacy, governance and service-delivery topics become increasingly visible in

later years, although the temporal pattern is uneven rather than a simple linear progression. Taken together, the results suggest a diversification of Facebook discourse over time, with the public discussion expanding beyond incident reporting and cosmological interpretation to encompass institutional responses, legal reform, advocacy and service provision.

4.3 What Comment-Level Discourse Reveals

To contextualise the topic modelling patterns, this section examines the targeted sample of 85 comments from two high-engagement threads. Given this focused sample, these qualitative findings are exploratory and intended to generate hypotheses about how key actors articulate SARV discourses in practice.

Post-level analysis captures the positions of organisations and page administrators, whereas comment threads reveal how individuals respond to, engage with, and negotiate these issues in real time. The picture at this level is considerably more complex, and in several respects, more challenging for current advocacy approaches.

Our qualitative analysis of approximately 85 comments posted by Facebook users from two public incident-response threads identified five recurrent positions.

1. Sorcery is real, but violence is wrong.

Amongst commenters who engage directly with the question of whether sorcery is real, affirmation outnumbers denial by roughly two to one, and the most common position amongst them holds that sorcery is real but violence is nonetheless wrong. We argue this is the most analytically significant position for policy. Representative comments include:

- *“Sorcery and sugar sick [diabetes] are real”.*
- *“Sanguma is real in Papua New Guinea... sorcery related killings has become a norm in the last 50 years, and it won't stop”,*

Many commenters who do not contest the reality of sorcery draw a distinction between sorcery as a social phenomenon requiring a spiritual or traditional response, and violence as an unacceptable form of that response. For instance, one stated:

- *“Its real... accept it na deal with it... kilim man won't change the problem... go LOTU [church]”.*

These comments are made alongside replies debating non-violent responses. Another commenter simultaneously affirmed sorcery beliefs, expressed concern about deaths caused by the devil, whilst also rejecting violence, stating:

Underworld exists and those who practiced it live amongst us. The devil is real and it also exists and lives in human form amongst us. We just cannot deny the facts! On the other end we should also be advocating on mass killings...

This stance illustrates the layered, non-binary moral reasoning that characterises grassroots positions on SARV.

Although our sample of two threads is small, the pattern of significant levels of belief in sorcery being publicly expressed is echoed elsewhere. In a qualitative review of 461 audience comments on collaboratively produced anti-SARV videos circulated via TikTok and YouTube, Wolfram (2026) identified 22 comments that affirmed that sorcery is real and continues to cause harm. Whilst a larger number (53) rejected sorcery belief, the appearance of comments asserting sorcery's reality even in response to explicitly anti-SARV content suggests that this position is a recurrent feature of comment-level discourse in PNG, rather than an artefact of our thread selection.

Moreover, recent national survey data also corroborates this insight at scale. In a 2026 Melanesian Institute survey of 2,000 Papua New Guineans across all provinces, 68 per cent said they personally believe sorcery exists and is real, whilst 89 per cent said SARV is not morally right and 97 per cent said it is not Christian (Melanesian Institute, 2026). Read together, these figures mean that a majority of respondents simultaneously believe in sorcery and reject the violence.⁵

2. Western law cannot solve a spiritual or cultural problem.

A related but distinct position frequently expressed by individual members of the public commenting on media reports and institutional updates questions the appropriateness of state legal responses to SARV, framing sorcery as a spiritual and cultural matter outside the jurisdiction of the post-colonial legal system. Representative comments include: *"White man's law cannot solve sanguma... trust me, it won't work"* and *"Our ancestors lived in this land since time eternity... whiteman and western religion just came recently"*. This position reflects a deeper epistemological question about whether PNG's post-colonial legal framework can address a phenomenon understood by many as operating in an entirely different ontological register. It has implications for how to frame advocacy for legal reform,

⁵ The Melanesian Institute surveyed 2,000 Papua New Guineans (aged 18–86, approximately half women) across all provinces in May 2026, by mobile telephone in Tok Pisin. The survey instrument distinguished explicitly between belief in sorcery, the acceptability of accusations, and the acceptability of violence. We note that only headline figures have been published; a telephone-based sample likely over-represents the connected population; and the Institute is itself an anti-SARV advocacy organisation whose Facebook page appears in our dataset.

suggesting that an exclusive emphasis on criminal penalties may resonate with some audiences whilst appearing irrelevant or illegitimate to others.

3. Demands for punitive legal action.

Commenters opposing SARV violence often demand harsher, highly retributive state legal penalties, most commonly reinstatement of the death penalty. Comments such as *"Reintroduce Death Penalty. Enough of verbal diarrhea"* and *"Death Penalty must given"* are representative. Whilst aligned with opposition to violence, this heavily punitive orientation sits in direct tension with victim-centred human rights and restorative approaches advocated by civil society organisations. Its prevalence indicates substantial, genuine public anger with perceived impunity.

4. Civic accountability demands directed at politicians.

Commenters frequently use thread space to challenge politicians who post symbolic condemnations of SARV without following through with corresponding legislative or executive action. *"Are you lobbying anything regarding this as a parliamentarian or are you just condemning on Facebook?"* is representative of the frustration expressed by such commenters. This suggests that Facebook serves a civic accountability function extending beyond advocacy and information dissemination into active political contestation. A recurring thread theme involved comments about how embarrassing it is for PNG citizens to have people in other countries witnessing their fellow citizens behaving in this violent way. Example expressions of this sense of shame include posts such as: *"Animalistic behaviours online, the world is watching"* and *"this kind of video makes us look bad in the eye of the world and white people ask what we do in PNG."*

5. Direct incitement to violence.

Whilst rare in the public institutional dataset (approximately 1–2 per cent of posts), content that encourages, rationalises or celebrates SARV does circulate. We identified posts directly inciting violence against women accused of sorcery, referring to them as "weeds" to be removed, and sharing images of female survivors in states of undress or restraint whilst inviting others to locate and harm them. Specific examples from the comment threads include:

"Sorcery is real many people die without sick. Only way to stop that sorcery business in our society is to kill them."

"Governor u mangi Sepik ya and you definitely know that sanguma & sorcery is real in Sepik. These sangumas & sorcerers have taken away the lives of too many people so please governor let people apply jungle justice to those sorcerers and sangumas."

"The whole nation knows that Sorcery is real but our leaders have failed to address this issue. One Sorcery is killing 20 people and goes unnoticed but if the community kill one, we apply the unjust law."

The 1–2 per cent figure in this institutional dataset should be understood as a floor, rather than a true estimate of the prevalence of such content. Much of this content likely circulates more extensively in private groups and encrypted platforms outside of this study's scope.

SARV as cultural identity: a cross-cutting dimension..

A dimension cutting across all five topic clusters is whether SARV is constitutive of PNG or Melanesian cultural identity. Positions range from strong rejection to qualified affirmation:

- Rejection as culture: Those rejecting SARV as cultural identity frequently describe it in temporal and developmental terms, labelling it a "primitive" or "stone age" mentality incompatible with modern PNG citizenship.
- Affirmation as culture: Those defending its cultural status argue that sorcery belief is intrinsic to Melanesian lifeworlds, irreducible by law or external human rights advocacy. Representatives of this position make comments such as:

Papua New Guinea we are melanesian and sorcery is part of our culture. Yupla wanem kain Christian country taim police, army, politicians, lawyer, magistrate, judge na even ol pastor to sa poretim poisin sanguma stap... maski lo pretend... come on"
(Translation: You call yourselves a Christian country but police, army, politicians, lawyers, magistrates, judges, and even pastors are themselves engaged in sorcery - stop pretending).

And:

Even though we have Sorcery Laws in place, but we must also understand that Sorcery is deeply Rooted in our society from generation to generation, we may execute and prosecute perpetrators but it will not solve the problem, it's here to stay.

These comments do not straightforwardly endorse SARV. Rather, they are making an epistemological and cultural claim that sorcery is ontologically real and culturally embedded in ways that legal reform cannot address. In other words, resistance to the human rights framing of SARV is a grounded claim about cultural legitimacy and the limits of post-colonial law, rather than simple ignorance or indifference.

4.4 Images, Video, and the Visual Dimensions of SARV Online

A significant dimension of SARV's digital dynamics is not captured by text-based analysis alone and warrants separate discussion: the circulation of images and video, often showing severe physical abuse and prolonged torture. Our data collection and broader monitoring identified a substantial volume of such visual content on Facebook. This includes images of victims in states of undress, restraint or severe injury and photographs and videos of violence actively being perpetrated or its immediate aftermath. Furthermore, there is also content that appears deliberately staged to communicate power and impunity.

The role of this visual content in SARV dynamics is likely substantial. As Mukherjee (2020) argues, mobile witnessing, understood as the viewing and sharing of visual documentation of violence, does not merely record events; it actively constitutes a witnessing community that can inadvertently confer social legitimacy upon them. Fujii's (2021) concept of violent display captures a closely related dynamic: the deliberate public staging of violence as a social performance designed to communicate messages about community power, the consequences of accusation, and the reach of retribution. Facebook's public-feed sharing mechanics functions as an efficient amplifier of these dynamics, extending violent display beyond the immediate physical scene *and* recruiting passive witnesses who may not otherwise engage with such content into the spectacle.

Key informant interviews conducted for a recent national scoping study corroborate these dynamics: stakeholders reported that the online circulation of SARV torture imagery is normalising the violence and contributing to its spread into provinces where accusations previously did not result in violence, and identified an emerging practice of perpetrators using social media and messaging services to demand electronic ransom payments for the release of captured victims (Osborne, 2026).

The consequences of widespread visual circulation of these images can create severe practical and ethical risks, even when shared with explicit condemnatory intent. Images of survivors shared to condemn violence, raise alarm, or pressure authorities for emergency intervention can expose survivors to renewed harm by circulating their appearance, location and alleged accusation specifics to new audiences. This tension between the need to post to ensure people are rescued and to pressure for accountability of perpetrators on the one hand, and the risks of surveillance and targeting on the other, is one of the most urgent practical problems in the contemporary SARV digital landscape.

Our text-based dataset does not allow systematic analysis of visual content, an important limitation of the current study. Future research combining multi-modal image and video analysis with text-based STM methods would advance

understanding of SARV's digital dynamics. Nevertheless, available qualitative evidence strongly suggests that images and video are likely central to how Facebook mediates, amplifies and shapes SARV dynamics in PNG.

4.5 Facebook as a Reporting Platform

A key informant interview with a prominent SARV activist provided us with significant insight into the ways members of the community routinely use Facebook to bypass what they perceive as ineffective official channels to report SARV incidents and exert pressure on public authorities to act. She noted:

When victims use this practice, known as 'digital vigilantism', it shows that people want accountability when they cannot rely on official channels. When victims or witnesses think the police will not help or may act corruptly, Facebook gives them a way to make their case public and hard to ignore. . . . In some cases, video evidence shared on Facebook has sparked public outrage so significant that it has forced official responses. For example, a video of police viciously beating three men in Port Moresby emerged on social media and drew widespread condemnation, illustrating how Facebook can serve as a watchdog against both criminal and state misconduct.

Posting reports of cases on Facebook is a strategy born of necessity. It is particularly attractive in contexts such as PNG, where official reporting channels are difficult to access, slow and widely believed to be corrupt and unreliable. Facebook functions as a valuable, if double-edged, arena for public advocacy, rapid incident exposure and civic oversight.

This practice is corroborated by the recent national scoping study discussed above, which found that women use the public nature of Facebook to seek 'public justice' when they fear the formal system will fail them, whilst noting that survivors who do so risk prosecution for defamation under the Cybercrime Code Act 2016 (Osborne, 2026).

5. Discussion

Six main insights emerge from our multi-method analysis, detailing how Facebook functions within the broader socio-legal and advocacy landscape of PNG.

5.1 Platform Evolution and Professionalisation of Digital Strategy

First, Facebook has evolved from a basic reporting tool into a multi-functional advocacy platform. It simultaneously serves incident documentation, awareness-raising, institutional signalling, survivor support, and governance advocacy. The

temporal shift documented in our dataset, from incident reporting and religious cosmology in earlier years towards governance, legal reform, and survivor services from 2020 onwards, reflects a growing professionalisation of digital strategy within the anti-SARV movement. This represents a change in content and also in Facebook's function. Early posts focused on making the issue of SARV visible to wider audiences. More recent posts are far more diverse – they mobilise support for specific legislative and institutional responses, hold politicians publicly accountable, and build coalitions amongst civil society organisations.

This evolution is consistent with patterns documented in other advocacy contexts where civil society organisations develop sophisticated digital communications strategies over time (Dwyer and Minnegal, 2020). As Bennett and Segerberg (2012: 760) have observed social media-based “connective action” relies on “the formative element of ‘sharing’: the personalization [of messages] that leads actions and content to be distributed widely across social networks”. In this way, they argue that this form of connective action can ultimately catalyse collective action (Bennett and Segerberg 2012: 745).

5.2 Institutional Dominance, Asymmetric Moderation and Survivor Risk

Our second insight is that in its organised, institutional discourse, Facebook in PNG is primarily an anti-SARV space. The 83 per cent opposition figure identified in the dataset is produced by a relatively small number of high-volume institutional actors - government committee pages, NGOs and human rights advocacy groups, operating what amounts to a coordinated digital anti-SARV campaign. Their digital communications capacity deserves deliberate support from donors and government, including dedicated funding for content creation, social media management and strategy production, rather than focusing solely on funding offline program delivery.

As one activist noted:

Facebook accounts run by law enforcement or justice agencies, including the National Department of Education or UN Systems, should keep raising awareness and calling for action. When strong institutions join together and speak as one, their voice is more powerful.

However, this institutional focus highlights a critical risk as well: the lack of standardised protocols regarding survivor identification. Even well-intentioned posts condemning SARV can inadvertently expose survivors to further harm if they identify individuals by name, image, or location. Civil society organisations, platform moderators and social media users need clear, accessible guidance on reporting SARV incidents that prioritises victim safety without increasing

vulnerability. This should be developed with survivors' organisations and shared across the advocacy network.

As observed earlier, the low prevalence of “supportive” posts in the public dataset (2 per cent) almost certainly understates the extent of pro-SARV content online. The pro-SARV posts identified reflect only public institutional spaces, which are more subject to potential basic content moderation. Content encouraging or inciting SARV circulates predominantly in private groups and encrypted channels, which are more difficult to access for research purposes and, in practice, prove difficult to moderate. Scholars and activists working in the SARV space are often incidentally exposed to such private material, informing broader understanding of its nature and extent of spread but often making it more challenging to capture.

As a result of the relatively small number of public posts concerned with supporting SARV, we were not able to test our hypotheses about the role played by Facebook in directly amplifying accusations and offline violence in real time. We know this phenomenon happens in other parts of the world; for example, The Media Diversity Institute (2022) has documented how African media can be complicit in witchcraft-accusation related violence through content that presents accusations as fact, normalising the violence and reaching audiences without editorial mediation. Newell (2021) has also shown how internet culture and occult belief systems mutually constitute each other in Côte d'Ivoire. Investigating these private spaces remains an urgent priority for future research in the PNG context. However, given the need to do this through private Facebook groups, it will require a specialised ethics protocol and approach.

5.3 Worldview Pluralism and the Disjunct Between Institutional and Public Discourse

Our third insight is how comments-level discourse reveals a wide diversity of perspectives on SARV, and that Facebook is serving as a forum in which these diverse viewpoints are being aired and debated. This disjunct between the rather uniform, secular and human rights-based framing of institutional posters and the complex, nuanced and pluralistic perspectives of individual commenters is a key finding. The dominance of opposition in organised Facebook discourse reflects who is posting, not necessarily who is reading. Organisational posting volume reflects those with the resources and motivation to shape the narrative but is not equivalent to audience persuasion. An advocacy ecosystem generating large volumes of condemnatory content may signal norms to already-sympathetic audiences more than it changes the minds of those who hold nuanced views that “sorcery is real, violence is wrong” or more dangerous, entrenched views that violence is acceptable. To be effective, state and civil society digital campaigns must bridge the gap between acknowledging local ontological realities whilst

firmly reinforcing community norms against violence, extrajudicial harm and torture.

Earlier scholarship on worldview pluralism in PNG (Forsyth and Gibbs, 2020) argues that people frequently hold multiple worldviews - magical, Christian, and biomedical - simultaneously. Comment data provide strong empirical support for this: the most prevalent community-level response is not the denial of sorcery or the endorsement of violence, but a morally complex acceptance of sorcery's reality paired with a firm rejection of violence as an appropriate response ("sorcery is real, but violence is wrong"). Effective advocacy must work *with* this ontological plurality. Campaigning that assumes sorcery belief is merely irrational or will be abandoned through education or legal pressure alone, is likely to bypass a substantial portion of its intended audience.

Effective advocacy may be achieved by engaging several integrated approaches:

- Emphasising that violence does not make communities safer or resolve underlying problems;
- Drawing on Christian frameworks alongside legal ones;
- Using trusted insider voices rather than external authority figures; and
- Framing accountability in terms of community wellbeing rather than external legal obligation.

A pluralistic advocacy approach may be better positioned to reach audiences who remain unengaged by secular human rights messaging. Wolfram's (2026) collaborative video project in Goroka offers an early demonstration of such an approach in practice: short-form anti-SARV videos co-produced with young people, performed in Tok Pisin and circulated via TikTok and YouTube, accumulated over one million views across platforms within roughly ten months, reaching an audience that was predominantly male (approximately 78 per cent), aged 18–34, and located within PNG.

This framework has a number of important practical ramifications. First, digital literacy campaigns alone are insufficient responses to SARV narratives on Facebook. Addressing the underlying beliefs and social conditions that legitimise accusations and violence is also essential. This resonates with the findings from Banaji et al. (2019) who studied WhatsApp and its role in facilitating lynching in India. The authors concluded that pre-existing prejudices and beliefs, rather than digital illiteracy, drove the sharing of harmful content.

A further ramification of engaging in a meaningful way with online opinion plurality is that more is needed than a standalone online messaging strategy. Here, we refer to emerging Global South scholarship that stresses the necessity of combining offline and online discussions, ideally incorporating community-driven

approaches where a variety of voices can be engaged and emphasis is placed in talking through particular local grievances (Ong and Jackson, 2026).

Emerging scholarship from India further suggests that locally grounded media literacy initiatives may be more substantially effective than national programs (Medeiros and Singh, 2020). These initiatives rely on tailored messaging and the mobilisation of trusted community voices. Participants at the first Disinfo Pacific Conference similarly highlighted the value of “hyper-local responses that combine social listening, training, and engaging trusted voices [that include] [e]lders, religious leaders, and health professionals ... uniting to counter the spread of false information and promote accurate messaging” (Tapia, 2024). These insights have relevance for addressing SARV through both online and offline campaigns.

Importantly, none of these recommendations imply that there is no role for messaging that asserts that sorcery does not exist. In separate work, we explore how such messaging is also important and can be particularly effective when communicated by trusted voices (Forsyth et al., forthcoming).

5.4. The Digital Absenteeism of Church Leadership

The fourth insight concerns the gap between church social media activity and church leadership advocacy. The church pages in our dataset are amongst the least likely to post content opposing SARV, in contrast to the firm and consistent anti-SARV positions taken by Catholic bishops, priests and other church leaders offline. No Pentecostal or Evangelical pages appear in the dataset at all. The lack of their voice is significant because churches are amongst the most trusted institutions in PNG communities. In addition, many church leaders already speak from within the “sorcery is real, violence is wrong” register that our comment data suggest is a common starting point for community audiences, and is a register that secular human rights messaging does not reach.

The two constituencies call for different approaches. For the mainstream churches the task is largely one of capacity: few congregations maintain Facebook pages, and few pastors address SARV on their own. Supporting faith-based organisations to develop digital communications capacity, and working with prominent church figures to record and circulate statements through existing church networks, would extend the reach of anti-SARV messaging at modest cost. Pentecostal and Evangelical spaces present a harder problem, since they frequently operate through spiritual warfare frameworks that can inadvertently validate sorcery accusations. Rather than expecting these groups to adopt legalistic or human rights discourse, ecumenical networks might engage them on their own theological terms, framing anti-violence as a Christian duty - a potentially wide-reaching pathway into congregations where SARV risks are highest.

5.5. Structural Boundaries of Platform Governance and Content Moderation

Our findings also highlight the consequences of limitations of relying on automated content moderation and centralised social media governance in complex Global South environments. Major platform architectures lack the linguistic capacity and contextual nuance required to parse Tok Pisin, local dialects or subtler forms of digital accusation and harassment. Consequently, state-level interventions, such as temporary platform bans or threats of digital censorship, are blunt instruments that fail to curb harmful narrative flows. Rather than depending entirely on platform-led algorithmic oversight, national governance efforts might also consider supporting community-led digital monitoring networks and local civil society moderators who best understand the cultural registers of digital SARV discourse.

Another governance limitation is the legal framework around SARV. The *Glasman Act* amendments, whilst significantly strengthening criminal penalties for making sorcery accusations, contain no provision addressing the use of technology or social media to incite violence or distribute SARV-related content (Osborne, 2026). The evidentiary framework presents a different problem: the Evidence (Amendment) Act 2016 inserted a comprehensive electronic evidence regime into the Evidence Act, expressly defining data to include video and audio and providing that admissibility may not be denied solely because evidence is in electronic form. Yet practitioners continue to report that courts require evidence in printed form and are increasingly reluctant to admit digital material whose authenticity might be contested in the era of generative AI (Osborne, 2026).

6. Conclusion

To conclude, this article has examined how Facebook is used in relation to SARV in PNG, drawing on systematic quantitative and qualitative analysis of 507 posts collected over 2013–2024 and comment threads collected Oct-Dec 2025. Empirically, this is the first peer-reviewed systematic analysis of Facebook content relating to SARV in PNG, and to our knowledge, the first in any Pacific context. We have documented a decade-long trajectory in which connective action, initially centred on witnessing and religious cosmological framing, has progressively hardened into coordinated governance advocacy, driven by a small number of institutional actors operating with sustained digital communications capacity.

This trajectory offers a contribution to theories of connective and collective action in the Global South: in low-resource, high-violence contexts, the shift from

witnessing platform to governance infrastructure does not emerge spontaneously from platform affordances, but requires deliberate institutional investment.

Our findings also speak to critical debates in platform governance. PNG illustrates the result of asymmetric moderation: a public sphere where organised civil society succeeds in dominating institutional discourse whilst harmful content, particularly visual content and private-group discussion, circulates largely unchecked due to structural gaps in platform language capacity and cultural context knowledge. This is not a uniquely PNG problem, but rather an instance of a broader global failure of platform accountability in non-Western linguistic contexts.

Taken together, our findings show that Facebook is neither simply a driver of SARV nor a straightforward tool for preventing it. Public-facing discourse is overwhelmingly anti-SARV, but this apparent consensus is largely produced by institutional actors and should not be mistaken for public opinion. Comment-level discourse reveals a more plural landscape in which belief in sorcery can coexist with rejection of violence. Effective digital strategies must therefore do more than amplify condemnatory messages: they must engage with the religious, cultural and epistemological frameworks through which sorcery is understood, while strengthening trusted voices and linking online advocacy with offline prevention.

References

ABC News Australia (2025, 25 March). Papua New Guinea shuts down Facebook. <<https://www.abc.net.au/news/2025-03-25/facebook-shut-down-in-papua-new-guinea-in-test/105092564>>

Australian High Commission. Papua New Guinea (no date). *Churches join forces to combat sorcery accusation-related violence*. <<https://png.highcommission.gov.au/pmsb/916.html>>

Banaji, S., Bhat, R., Agarwal, A., Andreou, N., & Maheli, R. (2019). *WhatsApp vigilantes: An exploration of citizen reception and circulation of WhatsApp misinformation linked to mob violence in India*. LSE Department of Media and Communications.

Barak, J., Chakma, A., & Forsyth, M. (2024). *The relationship between social media and violence in Papua New Guinea*. United States Institute of Peace.

Bennett, W.L., & Segerberg, A. (2012) The logic of connective action. *Information, Communication & Society*, 15:5, 739-768, <https://doi.org/10.1080/1369118X.2012.670661>.

DataReportal (2025, 8 November). *Digital 2025: Papua New Guinea*. <<https://datareportal.com>>

DT Institute / DT Global (2026). *Facebook dominates the information landscape in Papua New Guinea*. <<https://www.dt-institute.org/facebook-dominates-the-information-landscape-in-papua-new-guinea/>>

Dwyer, P., & Minnegal, M. (2020). COVID-19 and Facebook in Papua New Guinea: Fly River Forum. *Asia & the Pacific Policy Studies*, 7(3), 233-246. <https://doi.org/10.1002/app5.312>.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27-38. <https://www.jstor.org/stable/2336755>.

Forsyth, M., & Gibbs, P. (2020). Contagion of violence: The role of narratives, worldviews, mechanisms of transmission and contagion entrepreneurs. *International Journal for Crime, Justice and Social Democracy*, 9(2), 37-59. <<https://doi.org/10.5204/ijcjsd.v9i2.1217>>

Forsyth, M., & Gibbs, P. (2022). Causal stories and the role of worldviews in analysing responses to sorcery accusations and related violence. *Foundations of Science*, 27, 773-784.

Forsyth, M., Kipongi, W., Hukula, F., Gibbs, P., & Losoncz, I. (2021). Sorcery accusation-related violence in PNG, Part 5: Incidents and victims. *In Brief* 2021/5.

Forsyth, M. (2021). Prosecution of sorcery accusation related violence in PNG: What more is required? *Policy Brief 21. Development Policy Centre, Crawford School of Public Policy, The Australian National University*. https://devpolicy.org/publications/policy_briefs/PB21-Prosecution-of-SARV-in-PNG-Aug-2021.pdf

Forsyth, M., Kipongi, W., Barak, J., Malala, E., Kopel, E., & Losoncz, I. (2024). *Putting data around intergroup violence and sorcery accusation-related violence in Papua New Guinea*. Discussion Paper 24-002. United States Institute of Peace. <<https://www.usip.org/sites/default/files/2024-07/dp24-002-putting-data-around-intergroup-violence-png.pdf>>

Forsyth, M., Gibbs, P., Kiru C., Kipongi, W., Barak, J., Kopel, E., Ealedona, L., Malala, E., Lutz, A., & Kamya, P. (2025). *No one should stand alone: Report on SARV in the Highlands*. UNFPA. <<https://png.unfpa.org/sites/default/files/pub-pdf/2025-08/No%20One%20Should%20Stand%20Alone%20-%20Report%20on%20SARV%20in%20the%20Highlands.pdf>>

Forsyth, M., Lutz, A., & Gibbs, P. (forthcoming). Conspiracy theories and sorcery accusations as forms of social contagion in PNG and the USA and how to address them. (Forthcoming as part of an edited collection emerging from the ASAO session on conspiracy theories 2025 and 2026.)

Foster, R.J. (2023). Tenuous connectivity: time, citizenship, and infrastructure in a Papua New Guinea telecommunications network. *The Asia Pacific Journal of Anthropology*, 24(2), 91-115. <https://doi.org/10.1080/14442213.2023.2177330>.

Foster, R.J. (2025). *Uneven connections: A partial history of the mobile phone in Papua New Guinea*. ANU Press.

Fujii, L.A. (2021). *Show time: The logic and power of violent display*. Cornell University Press.

Igwe, L. (2026). Social media, witchcraft, and violence in post-colonial Nigeria. Abstract for Conference on Witchcraft Accusations and Human Rights, Malawi, 2026.

Medeiros, B. & Singh, P. (2020). Addressing misinformation on Whatsapp in India through intermediary liability policy, platform design modification, and media literacy. *Journal of Information Policy*, 10, 276-298. <https://doi.org/10.5325/jinfopoli.10.2020.0276>.

Media Diversity Institute (2022). African media is complicit in deadly witchcraft violence. <<https://www.media-diversity.org/african-media-is-complicit-in-deadly-witchcraft-violence/>>

Melanesian Institute (2026, August). *What Papua New Guineans think about sorcery accusation related violence (SARV) and what the government should do about it*. Goroka: Melanesian Institute. <<https://www.melanesianinstitute.network/reports/what-png-thinks-about-sarv-2026.pdf>>

Mukherjee, R. (2020). Mobile witnessing on WhatsApp: Vigilante virality and the anatomy of mob lynching. *South Asian Popular Culture*, 18(1), 1–18. <https://doi.org/10.1080/14746689.2020.1736810>

NapoleonCat. (2024). Facebook users in Papua New Guinea. <<https://napoleoncat.com/stats>>

Newell, S. (2021). Hackers of the heart: digital sorcery and virtual intimacy in Côte d'Ivoire. *Africa: Journal of the International African Institute*, 91(4), 661–685. <https://doi.org/10.1017/S0001972021000449>.

Ong, J.C. & Jackson, D. (2026). What counterdisinformation funders got wrong: Global South critiques of the field. *Information, Communication & Society*, published online 06 Apr 2026, 1–18. <<https://doi.org/10.1080/1369118X.2026.2647357>>

Osborne, K. (2026). Technology facilitated gender-based violence in Papua New Guinea: A national scoping study. Australia Papua New Guinea Law and Justice Partnership (APLJ).

Pacific Security College. (2024). PNG's digital divide. <<https://pacificsecurity.net/blog/pngs-digital-divide/>>

Roberts, M.E., Stewart, B.M., Tingley, D., Lucas, C., Leder-Luis, J., Kushner Gadarian, S., Albertson, B., & Rand, D.G. (2014). Structural topic models for open-ended survey responses. *American Journal of Political Science*, 58(4), 1064–1082. <https://doi.org/10.1111/ajps.12103>.

Tapia, Meg. (2024, 27 November). The Pacific disinformation playbook. *The Interpreter*, Lowy Institute. <<https://www.lowyinstitute.org/the-interpreter/pacific-disinformation-playbook>>

Sinpeng, A., Martin, F.R., Gelber, K., & Shields, K. (2021). Facebook: Regulating hate speech in the Asia Pacific. University of Sydney / R2P Asia Pacific. <[https://r2pasiapacific.org/files/7099/2021 Facebook hate speech Asia report.pdf](https://r2pasiapacific.org/files/7099/2021%20Facebook%20hate%20speech%20Asia%20report.pdf)>

Statcounter. (2026). Social media stats in Papua New Guinea [by month and year]. <<https://gs.statcounter.com/social-media-stats/all/papua-new-guinea>>

UNICEF Papua New Guinea [UNICEF PNG] (2024). *Navigating the digital minefield: Online safety challenges and solutions in Papua New Guinea*. <<https://www.unicef.org/png/stories/navigating-digital-minefield>>

Wolffram, P. (2026). Performing change: Anti-violence video production with communities in Papua New Guinea. *Visual Anthropology*, 39(1-3), 37-52. <https://doi.org/10.1080/08949468.2026.2681460>.

Yangin, R. (2023, 9 August). Media freedom, accountability and online misinformation in Papua New Guinea. *In Brief* 2023/4. <<https://openresearch-repository.anu.edu.au/server/api/core/bitstreams/ce63c320-0586-4b20-b026-aca1575a37bd/content>>

Appendix

Table 3: Page-level robustness analysis.

Predictor	β	SE	p_value
Intercept	1.67	0.31	<.001
Church	-1.75	0.60	0.009
Government Committee	19.31	3374.42	0.995
Media	-0.38	0.91	0.685
NGO	0.08	0.40	0.853
Public Forum	-1.10	0.63	0.098
UN	-0.83	0.71	0.260

Note: Quasi-binomial model estimated at the Facebook page/group level ($N = 24$), with Advocacy as the reference category. Posts were aggregated into oppositional and non-oppositional counts within each page/group. The estimated dispersion parameter was 1.71. The Government Committee coefficient is unstable because this category contains a single page with 72 oppositional and no non-oppositional posts.

Table 4: Sensitivity analysis excluding the Government Committee page.

Predictor	Full Firth model β	Excluding Government Committee β
Intercept	1.018	1.018
Church	-1.883	-1.883
Media	-0.802	-0.802
NGO	-0.315	-0.315
Public Forum	-1.356	-1.356
UN	-1.275	-1.275
Engagement	0.312	0.312

Note: The sensitivity model excludes the single Government Committee page and its 72 posts ($N = 435$). The two models were estimated separately; the coefficients are unchanged to three decimal places because all 72 Government Committee posts are oppositional, so that category's indicator accounts for them entirely and they carry no information about the other predictors. The estimates differ by at most 0.00006, which is not visible at three decimals.