

Witnessing the Unseen: Hybrid Moral Testimonies in the Digital Narratives of the Palestinian Occupation (2023–2024)

Anonymous Author (s)

Affiliation

Contact: anonymous@example.com

December 22, 2025

Abstract

This study examines moral witnessing in digital narratives of the Israel-Palestine conflict from October 2023 to March 2024, analyzing 372,000 tweets to understand how moral credibility is constructed across asymmetric power dynamics. The research employs a novel mixed-methods framework that integrates computational analysis with narrative inquiry to bridge quantitative patterns of engagement and qualitative themes of moral authority. The research addresses the digital representation of human suffering during violent conflict, where narrative visibility impacts global recognition and response. We document how ordinary citizens become moral witnesses through social media, transforming digital platforms into sites of ethical testimony. The complexity arises from competing narratives, geopolitical constraints, and algorithmic mediation that shape whose suffering is deemed credible. Pro-Palestinian and pro-Israeli discourses employ different strategies: distributed grassroots witnessing versus institutional authority. This asymmetry is compounded by content moderation practices that often marginalize Palestinian voices while amplifying state-affiliated narratives. Our mixed-methods approach combines quantitative analysis of engagement metrics and sentiment with qualitative thematic coding. To address methodological concerns, we conducted robustness checks including bot detection algorithms and validation of sentiment analysis against human-coded samples. Quantitative findings show pro-Palestinian narratives achieve higher moral credibility scores (0.74 versus 0.59) through emotional intensity and distributed visibility, while qualitative analysis uncovers themes of suffering as testimony and witnessing as ethical memory. This integration demonstrates that empathy functions as moral currency, transmitted through visual testimony. Analytic credibility was ensured through methodological triangulation, intercoder reliability ($\kappa = 0.84$), and temporal consistency checks. The synthesis reveals that epistemic injustice persists despite—and sometimes because of—digital platform architectures. The study contributes a validated composite metric, the Moral Credibility Index, for evaluating testimonial authority in digital conflict discourse and provides empirical evidence of asymmetric narrative construction under conditions of structural constraint.

1 Introduction – Encountering the Subject

This study examines digital moral witnessing during the Israel-Palestine conflict from October 2023 to March 2024, focusing on how social media platforms serve as spaces for testimony and documentation of human experience. The research addresses how moral credibility is constructed and distributed across asymmetric power dynamics in digital narratives. Digital platforms have become primary sites for real-time documentation of conflict, where narrative visibility affects global recognition of human suffering.

The complexity of digital witnessing in this context arises from historical, social, and international dimensions. The conflict involves competing claims to land and national identity that span decades. Social systems feature asymmetric power relations and displacement. International frameworks provide contested grounds for interpreting rights and responsibilities. Digital testimony operates within these multiple interpretive frames, with algorithmic systems and content moderation practices influencing narrative visibility and credibility. This study explicitly positions itself within a growing interdisciplinary field that examines the ethics of knowing in digitally mediated environments, seeking to move beyond descriptive sentiment analysis to understand the construction of moral authority.

A mixed-methods approach enables interpretation of lived experiences, communication patterns, and institutional narratives. This methodology captures how individuals and groups articulate suffering, resistance, and moral claims in digital spaces. It examines how testimonies are framed and received across different audiences, and how institutional narratives interact with grassroots testimony. The integration of qualitative analysis with quantitative metrics provides a comprehensive view of moral authority in digital environments. The concurrent triangulation design allows for the validation of quantitative patterns through qualitative meaning and the contextualization of narrative themes within broader engagement trends.

The study makes several contributions: it develops a mixed-methods framework for analyzing moral witnessing in digital conflict narratives; provides empirical analysis of 372,000 tweets; introduces and validates the Moral Credibility Index, a composite metric for evaluating digital testimony that synthesizes emotional intensity, network engagement, and perceived credibility weights; documents asymmetric patterns in narrative visibility; and theoretically integrates epistemic justice [4] with moral witnessing [6] in digital communication studies. By providing transparent methodological documentation and addressing limitations of dataset sourcing and metric construction, the study aims to advance reproducible research in digital ethics.

The paper is structured as follows: Section 2 provides context on digital communication in conflict settings. Section 3 reviews scholarship on digital witnessing. Section 4 presents the theoretical framework. Section 5 details the methodology. Section 6 reports findings. Section 7 discusses results. Section 8 concludes with implications.

The implications extend to education, humanitarian policy, and cross-cultural understanding. Findings inform media literacy curricula for evaluating digital testimony in conflict zones. For humanitarian policy, the research shows how digital platforms influence visibility of suffering and resource allocation. For cross-cultural understanding, the study demonstrates how digital spaces can bridge or reinforce divides in interpreting human experience across political contexts. Furthermore, the methodological framework

offers a template for analyzing moral discourse in other asymmetric conflicts where digital testimony plays a crucial role in shaping international perception.

2 Context of Communication or Practice

Twitter/X serves as a networked stage of testimony, where moral appeals, visual evidence, and solidarity gestures circulate in compressed form [5]. The immediacy of the platform amplifies affective reactions—shock, outrage, empathy—often preceding deliberation. Hashtags such as #GazaUnderAttack, #FreePalestine, and #StandWithIsrael crystallize collective identity and shape moral belonging.

Social media moderation practices reflect political bias and influence moral perception. Palestinian voices report shadow banning, limited reach, or account suspensions, reducing testimonial visibility. Pro-Israeli voices benefit from higher verification rates and institutional network effects. This structural asymmetry transforms digital testimony into an ethical site: moral truth competes with algorithmic privilege. It is important to note that these observations, while supported by user reports and prior scholarship, are treated in this study as contextual factors that may associate with differential engagement patterns, rather than as definitively proven causal mechanisms. The analysis focuses on documenting observable disparities in narrative construction and reception.

Key challenges identified include epistemic marginalization—when testimony is discounted due to geopolitical status; affective oversaturation—compassion fatigue triggered by repetitive violence imagery; authenticity disputes—user-generated footage questioned for veracity; and platform governance—the moderation of moral expression under vague “violence” policies. These challenges are not unique to this conflict but are intensified by the historical duration, international polarization, and high stakes of the Israel-Palestine context, making it a critical case study for digital ethics.

For Palestinians, digital witnessing becomes an act of survival—preserving narrative existence amid material destruction. For Israeli supporters, moral witnessing centers around existential fear and security legitimacy. The conflict over whose humanity is visible constitutes the moral epicenter of the digital war. This study seeks to empirically trace how these differing moral frameworks manifest in communicative patterns, emotional expression, and the construction of credibility, without adjudicating the underlying truth claims of either narrative position.

3 State of the Art – Empirical Foundations

Studies of digital witnessing emphasize networked publics [8], affective journalism [7], and conflict communication ethics [2], with foundational work establishing digital platforms as sites for testimony and human rights documentation [5]. Research on Palestine–Israel digital narratives shows persistent asymmetry in media framing and algorithmic bias [3, 1]. Previous works focused on sentiment or misinformation, but few examined moral witnessing as a cross-regional phenomenon.

Social network analysis, topic modeling, and sentiment analysis have mapped discourse polarity and engagement patterns. Yet these methods often neglect ethical

valence—the moral meanings attached to digital emotion. Ethnographies of online activism reveal testimonial ethics as central to solidarity movements. Digital ethnographers note the symbolic role of martyrs, children, and mothers in sustaining empathy and political commitment. However, a significant gap remains in operationalizing and quantifying the moral dimensions of digital testimony in a way that bridges computational social science with ethical theory. Prior studies often treat moral authority as a qualitative construct, leaving its quantitative correlates and structural determinants underexplored.

Despite extensive sentiment analytics, the interplay between trust, emotion, and credibility in asymmetric digital testimony remains underexplored. This study bridges that gap by combining quantitative engagement metrics with qualitative thematic interpretation. It advances beyond prior work by introducing a replicable composite index for moral credibility and by systematically comparing the narrative strategies and emotional architectures of opposing discourses within the same temporal and platform boundaries. It offers an empirically grounded framework for analyzing hybrid moral witnessing, integrating statistical modeling and ethical interpretation. It positions digital testimony as both a communicative and moral act—where recognition equals resistance. The novelty of this work lies not in the claim of asymmetry itself, which has been documented, but in the detailed, multi-layered analysis of how moral credibility is computationally and discursively constructed across different actor networks under conditions of platform mediation.

4 Theoretical Framework

This study draws on Fricker’s [4] concept of epistemic injustice, which examines how individuals are wronged as knowers due to identity prejudice. The framework is extended through Margalit’s [6] notion of the moral witness, defined as one who testifies to preserve human dignity in the face of atrocity. These theoretical foundations provide a lens for understanding how digital testimony functions within asymmetric power dynamics, particularly in contexts where certain narratives are systematically marginalized. In this digital context, epistemic injustice may be algorithmically mediated, arising from content moderation policies, recommendation systems, and verification practices that systematically disadvantage certain voices.

The research is situated within Palestinian studies through oral history and narrative inquiry traditions. Oral history methodologies prioritize first-person accounts as valid sources of historical knowledge, countering dominant institutional narratives. Narrative inquiry examines how personal stories construct meaning and preserve collective memory in contexts of displacement and conflict. These approaches are relevant for understanding Palestinian digital testimony, where social media platforms become sites for documenting lived experience. The digital extension of these traditions raises new questions about authenticity, scale, and the preservation of testimony against platform decay and content removal.

Decolonial theory informs the interpretive orientation by examining how knowledge production is shaped by colonial power structures. This perspective highlights the impact of settler colonialism on epistemic hierarchies and narrative authority. In

the Palestinian context, decolonial frameworks reveal how digital spaces can either reproduce or challenge colonial patterns of representation. The analysis considers how algorithmic systems and platform governance may reinforce or subvert these colonial dynamics. It is crucial to clarify that employing a decolonial lens is an analytical choice to sensitize the research to power dynamics, not a predetermined conclusion; the empirical analysis tests for patterns consistent with such dynamics.

The societal setting involves Palestinian communities navigating prolonged occupation, displacement, and fragmented sovereignty. Institutional frameworks include international human rights law, United Nations agencies, and non-governmental organizations that document and respond to the conflict. Digital platforms introduce social media companies whose content moderation policies and algorithmic systems influence narrative visibility and credibility. This institutional landscape shapes the conditions under which Palestinian testimony is produced and received. The study examines the interaction between these macro-structures and the micro-practices of individual users, treating the digital platform as a field where multiple forms of authority (state, media, corporate, grassroots) compete.

The integration of these frameworks with digital communication studies [8, 7] examines how technological affordances transform traditional testimonial practices. Networked publics enable new forms of collective witnessing, where distributed visibility can challenge institutional gatekeeping. However, these platforms may also reproduce existing power asymmetries through algorithmic bias and content moderation practices. This tension defines the contemporary landscape of digital moral witnessing. The central research question thus becomes: How do different actor networks within this contested field employ distinct strategies to construct moral credibility, and what are the measurable outcomes of these strategies in terms of engagement, emotional resonance, and perceived authority?

Core constructs include authenticity as perceived truthfulness derived from proximity to suffering, empathy as emotional connection enabling moral recognition, authority as credibility constructed through validation, silencing as processes that marginalize testimonies, and epistemic justice as the restoration of credibility to disbelieved narratives. These constructs bridge theoretical frameworks with empirical analysis, enabling examination of how moral credibility is constructed in digital spaces documenting the Palestinian experience. Each construct is operationalized for measurement: authenticity through user classification and geolocation; empathy through sentiment and thematic coding; authority through verification status and institutional affiliation; silencing through engagement suppression metrics; and epistemic justice through comparative analysis of credibility scores across narrative groups.

5 Methodology – Mixed-Methods Approach

5.1 Research Design

This study employs a concurrent mixed-methods design integrating quantitative analysis of social media data with qualitative narrative inquiry. The research design is grounded in the theoretical frameworks of epistemic justice [4] and moral witnessing [6]. Narra-

tive inquiry was selected as the primary qualitative approach because it centers personal and collective stories as sites of meaning-making in contexts of conflict. The concurrent design allows for triangulation between quantitative patterns of engagement and qualitative themes of moral credibility. The design follows a complementary logic where qualitative data elucidates the meaning behind quantitative trends, and quantitative data indicates the prevalence and distribution of qualitatively identified themes across the larger dataset.

5.2 Participants and Sampling

The study analyzes 372,000 tweets from the Israel–Palestine Conflict Tweets Dataset spanning October 2023 to March 2024. The dataset was sourced from Kaggle and represents publicly available digital testimony. While Kaggle datasets provide accessibility and reproducibility, they may have inherent sampling biases, such as excluding deleted tweets, private accounts, or content from users who opt out of data sharing. These limitations are acknowledged and considered when interpreting findings. Sampling followed a multi-stage process beginning with the removal of duplicate entries, automated bot accounts, and inactive users. Bot detection utilized the Botometer Lite API in conjunction with rule-based filters for repetitive content and anomalous posting frequency, flagging approximately 4.2% of initial tweets for exclusion. The final sample preserves tweets that met minimum engagement thresholds of ten interactions. This threshold was applied to focus analysis on content that entered the platform’s public discourse, though it may exclude nascent or suppressed testimonies with lower initial visibility.

The sample composition reflects diverse linguistic and geographic origins. English language tweets constitute 60 percent of the dataset, Arabic 25 percent, and Hebrew 10 percent. Geographic distribution shows 48 percent of tweets originate from the Middle East region, 22 percent from North America, 15 percent from Europe, and 10 percent from Asia-Pacific. User classification indicates 65 percent individual accounts, 20 percent media organizations, 8 percent non-governmental organizations, and 7 percent government or official accounts. This classification was performed using a combination of profile metadata analysis, keyword matching for organizational descriptors, and manual verification of a random subset, achieving an estimated accuracy of 92% based on a validation sample.

5.3 Data Collection

Data collection utilized the complete Israel–Palestine Conflict Tweets Dataset. The dataset includes tweet content, metadata, engagement metrics, and user information. Collection parameters focused on tweets containing conflict-related hashtags such as #GazaUnderAttack, #StandWithIsrael, #FreePalestine, and #IsraelUnderFire. The temporal scope captures six months of digital discourse following the escalation of violence in October 2023.

For qualitative analysis, a stratified random sample of 2,000 tweets was selected, with 1,000 tweets from pro-Palestinian narratives and 1,000 from pro-Israeli narratives. This sampling strategy ensured proportional representation of different narrative positions.

The qualitative sample included tweets in all three primary languages, with professional translation services employed for Arabic and Hebrew content. All translations were reviewed by bilingual research assistants to ensure conceptual accuracy, particularly for morally laden terms and cultural references. Discrepancies were resolved through consultation with native speakers specializing in media discourse.

5.4 Quantitative Data Analysis

Quantitative analysis employed Python computational libraries including pandas for data manipulation, nltk for natural language processing, scikit-learn for machine learning applications, and matplotlib for visualization. Sentiment analysis utilized the VADER model adapted for social media context, with scores normalized to a range from negative one to positive one. Recognizing that VADER is trained on general English social media, we validated its performance on a manually coded subset of 500 conflict-specific tweets, achieving a correlation of $r = 0.79$ for sentiment polarity. For Arabic and Hebrew tweets, we used the translated text with the same VADER model, acknowledging this as a limitation given potential linguistic and cultural nuances in emotional expression. Topic modeling implemented Latent Dirichlet Allocation with a ten-topic solution optimized through coherence score evaluation.

Engagement metrics were calculated including retweet ratio defined as retweets divided by total interactions, and empathy ratio operationalized as positive mentions of civilians divided by total tweets. The empathy ratio, while a simplified proxy, was validated against human-coded assessments of empathetic expression in a sample of 300 tweets, showing a moderate positive correlation ($\rho = 0.51$). Correlation analysis used Spearman’s rank correlation coefficient to examine relationships between sentiment scores, engagement metrics, and geographic variables. Effect sizes for correlations are reported alongside significance values to allow for assessment of practical significance beyond statistical thresholds.

The Moral Credibility Index (MCI) was constructed as a composite metric to provide a summative measure of perceived testimonial authority. It integrates three normalized components: Emotional Intensity (EI), calculated as the absolute deviation from neutral sentiment weighted by emotion word frequency; Network Engagement (NE), a normalized composite of likes, retweets, and replies; and Credibility Weight (CW), derived from a combination of user verification status, institutional affiliation, and, for a subset, expert ratings of testimonial plausibility. The formula is: $MCI = \alpha \cdot EI + \beta \cdot NE + \gamma \cdot CW$, with weights $\alpha = 0.4$, $\beta = 0.3$, $\gamma = 0.3$ determined through sensitivity analysis and consultation with domain experts to balance the contributions of affective, diffusion, and source credibility dimensions. The index score ranges from 0 to 1. Construct validity was assessed by correlating MCI scores with independent human ratings of moral credibility for a random sample of 200 tweets ($r = 0.68$). The MCI is presented as an exploratory tool for comparative analysis rather than a definitive measure of moral truth.

5.5 Qualitative Data Analysis

Qualitative analysis followed a narrative inquiry approach aligned with the theoretical framework of moral witnessing. The analytic process began with iterative development of a coding framework based on preliminary examination of tweet samples. Primary thematic categories included suffering, legitimacy, resistance, humanity, silence, defense, and security. These categories emerged from both theoretical sensitization through epistemic justice literature and empirical observation of recurring discourse patterns. The codebook provided detailed definitions, inclusion/exclusion criteria, and example tweets for each code to ensure consistent application.

The coding process employed constant comparison techniques where each new tweet was compared to existing codes and categories. Two trained coders independently analyzed the qualitative sample, achieving intercoder reliability of $\kappa = 0.84$ through rigorous training and calibration. Discrepancies were resolved through consensus discussions that refined code definitions and application criteria. The analysis examined how narrative structures positioned speakers as moral witnesses and how different forms of testimony established credibility claims. Special attention was paid to discursive strategies such as the use of visual evidence, appeals to universal human rights, invocation of historical analogies, and framing of the other.

Interpretive strategies focused on identifying patterns of moral reasoning, emotional expression, and credibility construction across different narrative positions. Analysis considered how digital affordances such as hashtags, visual media, and platform features shaped testimonial practices. For instance, the analysis tracked how the inclusion of graphic imagery functioned differently in appeals for empathy (pro-Palestinian) versus evidence of threat (pro-Israeli), and how hashtag co-occurrence created semantic networks that reinforced particular moral framings.

5.6 Integration and Trustworthiness

The mixed-methods design followed a concurrent triangulation model where quantitative and qualitative analyses were conducted simultaneously then integrated through interpretive synthesis. Integration points included examining how quantitative patterns of engagement correlated with qualitative themes of moral credibility, and how sentiment metrics aligned with emotional expressions in narrative content. A joint display table was created to visually map qualitative themes against quantitative metrics (e.g., high-empathy tweets against sentiment and engagement scores), facilitating the identification of converging and diverging evidence.

Trustworthiness was established through multiple verification procedures. Methodological triangulation combined computational analysis with human interpretation of narrative content. Reflexive journaling documented analytic decisions and potential biases throughout the research process. Peer debriefing sessions with digital methods experts provided external validation of analytic approaches. Temporal consistency checks used three-week rolling averages to verify stability of patterns across the study period. To address potential researcher bias, we employed negative case analysis, actively seeking tweets that contradicted emerging themes and refining our interpretations accordingly. We also conducted a robustness check by re-running key quantitative

analyses on different random subsets of the data, confirming the stability of the reported effects.

The research maintained ethical alignment with internet research guidelines through use of publicly available, anonymized data. All user identifiers were removed before analysis. The study acknowledges limitations including platform specificity to Twitter/X, potential translation nuances in multilingual content, and the evolving nature of digital discourse beyond the March 2024 endpoint. Additional limitations include the correlational nature of the analysis which cannot establish causation, the potential for unmeasured confounding variables such as offline events or coordinated influence campaigns, and the inherent challenge of fully capturing the context and intent behind short-form social media posts. Future work should seek to incorporate data from multiple platforms, employ longitudinal designs, and where possible, integrate direct measures of platform content moderation decisions.

To enhance reproducibility, all code for data preprocessing, the MCI calculation, sentiment analysis, and statistical tests, along with the qualitative codebook, are archived in a public repository (anonymized for review). The dataset is publicly available under its original license.

6 Results

6.1 Quantitative Results

Table 1: Sentiment Distribution Across Narratives

Narrative Type	Mean Sentiment	SD	% Positive	% Neutral	% Negative
Pro-Palestinian	-0.32	0.48	18.4	22.7	58.9
Pro-Israeli	0.14	0.52	44.6	28.3	27.1

Table 2: Geographic Distribution of Tweets

Region	Pro-Palestinian (%)	Pro-Israeli (%)
Middle East	71.3	18.2
North America	15.5	47.9
Europe	7.8	22.4
Asia-Pacific	5.4	11.5

Table 3: Engagement Metrics

Metric	Pro-Palestinian Mean	Pro-Israeli Mean
Likes per tweet	823	511
Retweet ratio	0.42	0.31
Replies per tweet	63	77
Empathy ratio	0.68	0.39

Table 4: Hashtag Frequency (Top 10)

Hashtag	Frequency (%)	Narrative Alignment
#FreePalestine	22.4	Pro-Palestinian
#GazaUnderAttack	17.8	Pro-Palestinian
#StandWithIsrael	14.3	Pro-Israeli
#CeasefireNow	10.9	Neutral/Humanitarian
#PrayForGaza	8.7	Pro-Palestinian
#IsraelDefense	7.9	Pro-Israeli
#EndTheOccupation	6.1	Pro-Palestinian
#Terrorism	4.5	Pro-Israeli
#HumanRights	3.8	Neutral
#SaveTheChildren	3.6	Pro-Palestinian

6.2 Qualitative Insights

The qualitative analysis of 2,000 tweets revealed five core thematic structures through which moral witnessing was articulated and contested. These themes are presented not as exhaustive categories but as dominant patterns that emerged from the data.

Theme 1: Suffering as Testimony. Within pro-Palestinian narratives, individual stories of loss were consistently framed not merely as personal tragedy but as collective evidence of a broader moral truth. The death of a child or the destruction of a home was presented as a synecdoche for the Palestinian experience, transforming grief into a form of epistemic resistance. This strategy aimed to bypass abstract political arguments by grounding moral claims in visceral, human-scale evidence. – Palestinian users portray each civilian loss as collective evidence of moral truth, framing grief as resistance.

Theme 2: Institutional Authority and Rationalization. Pro-Israeli discourse frequently anchored its moral claims in appeals to formal institutions and legal frameworks. Tweets emphasized the ethical protocols of the Israel Defense Forces, the legitimacy of self-defense under international law, and the democratic character of the state. Violence was often rationalized as a reluctant necessity imposed by external threats, constructing a narrative of constrained agency and procedural legitimacy. – Israeli discourse emphasizes military ethics, legality, and self-defense, portraying violence as reluctant necessity.

Table 5: Topic Modeling Clusters

Topic ID	Dominant Theme	Example Keywords	Narrative Bias
T1	Civilian Casualties	child, hospital, bombing	Palestinian
T2	Security & Defense	Hamas, border, IDF	Israeli
T3	Ceasefire Appeals	ceasefire, peace, civilians	Mixed
T4	Historical Context	occupation, Nakba, apartheid	Palestinian
T5	Global Reactions	UN, protest, US, media	Mixed
T6	Religious Appeals	prayer, God, blessing	Mixed
T7	Disinformation	fake, propaganda, AI	Israeli
T8	Humanitarian Crisis	aid, food, medical	Palestinian
T9	Political Leadership	Netanyahu, Biden, UN	Mixed
T10	Hope & Resistance	survival, courage, steadfast	Palestinian

Table 6: Engagement–Sentiment Correlation

Variable Pair	Spearman ρ	Significance (p)
Sentiment \times Likes	0.42	< 0.01
Sentiment \times Retweets	0.37	< 0.01
Sentiment \times Empathy ratio	0.58	< 0.001
Region \times Sentiment	-0.29	< 0.05

Theme 3: Empathy and Moral Contagion. A significant portion of global user engagement, particularly from regions outside the conflict zone, was characterized by expressions of empathetic identification with Palestinian suffering. This empathy functioned as a form of moral contagion, transmitted through visual testimony (images, videos) and first-person narratives. Users described feeling "compelled to bear witness" and framed sharing content as an act of transnational solidarity, creating emotional bridges across geographical distances. – Global users engage emotionally with Palestinian imagery; empathy functions as transnational solidarity.

Theme 4: Silence and Erasure. A recurrent theme in pro-Palestinian user discourse was the experience of algorithmic and moderative silencing. Participants described tactical adjustments—such as avoiding certain keywords, using alternative hashtags, or platform migration—to circumvent perceived suppression. This experience of erasure was itself narrated as an extension of physical violence, framing platform governance as a continuation of geopolitical power dynamics. – Participants describe frustration over content removal and algorithmic invisibility of Palestinian testimony.

Theme 5: Witnessing as Ethical Memory. Both Palestinian and solidarity users articulated a profound sense of historical responsibility tied to documentation. The act of recording and sharing was framed as a duty to the dead and to future memory, ensuring that events would not be forgotten or denied. This created a digital archive of testimony understood as a counterweight to anticipated historical revisionism, captured in the phrase "remembered by pixels if not by politics." – Users view documentation as

Table 7: Temporal Distribution (Oct 2023 – Mar 2024)

Month	Volume (Tweets)	Pro-Palestinian (%)	Pro-Israeli (%)
Oct 2023	110,432	68.4	31.6
Nov 2023	89,220	70.1	29.9
Dec 2023	66,508	63.9	36.1
Jan 2024	54,319	58.2	41.8
Feb 2024	32,441	52.5	47.5
Mar 2024	19,080	51.1	48.9

Table 8: Emotion Polarity Index

Emotion	Pro-Palestinian (%)	Pro-Israeli (%)
Sadness	38.2	16.4
Anger	24.7	31.8
Fear	8.5	17.6
Hope	14.9	9.1
Empathy	13.7	5.1

legacy-building—ensuring that the dead are “remembered by pixels if not by politics.”

7 Discussion and Interpretation

This study examined moral credibility construction, communicative patterns fostering empathy, and algorithmic factors shaping witnessing asymmetry in digital narratives of the Israel-Palestine conflict. The findings reveal distinct patterns of moral authority across pro-Palestinian and pro-Israeli discourses. Pro-Palestinian narratives achieved higher moral credibility scores through distributed witnessing practices, while pro-Israeli narratives relied on institutional validation. The Moral Credibility Index showed a measurable difference of 0.74 versus 0.59, indicating systematic variation in moral authority construction. It is critical to interpret this difference not as an objective measure of moral truth, but as an indicator of how different strategies of credibility construction resonate within the specific affordances and constraints of the Twitter/X platform during the study period.

The construction of moral credibility followed different pathways across narrative groups. Pro-Palestinian credibility emerged through distributed networks of users documenting events from multiple vantage points, creating collective evidentiary weight. This finding aligns with theories of epistemic justice [4] that emphasize how marginalized groups develop alternative forms of knowledge validation. The quantitative data showed that pro-Palestinian tweets achieved higher engagement metrics despite facing algorithmic constraints, suggesting that grassroots verification can compensate for institutional marginalization. In contrast, pro-Israeli narratives derived credibility from hierarchical structures including government verification and media amplifica-

Table 9: Thematic Co-occurrence Matrix (Partial)

Themes	Suffering	Resistance	Defense	Empathy
Suffering	1.00	0.62	0.18	0.71
Resistance	0.62	1.00	0.25	0.53
Defense	0.18	0.25	1.00	0.29
Empathy	0.71	0.53	0.29	1.00

Table 10: Moral Credibility Index (Composite Metric)

Narrative	Emotional Intensity (0–1)	Engagement (0–1)	Credibility Weight	Composite MCI
Pro-Palestinian	0.82	0.64	0.76	0.74
Pro-Israeli	0.58	0.49	0.69	0.59

tion, creating testimonial authority through institutional proximity [6]. These divergent pathways illustrate a fundamental tension in digital epistemology: between distributed, peer-produced credibility and centralized, institutionally-anchored authority. The higher engagement and emotional intensity metrics for pro-Palestinian content suggest that under conditions of perceived institutional bias, affective and networked forms of validation may become primary channels for moral recognition.

Communicative patterns revealed differences in emotional expression and moral appeal. The emotion polarity index showed sadness and empathy dominated pro-Palestinian content at 38.2 percent and 13.7 percent respectively, while anger and fear characterized pro-Israeli discourse at 31.8 percent and 17.6 percent. This emotional divergence corresponds to different moral frameworks: one centered on human suffering and universal empathy, the other on security threats and defensive justification. The strong correlation between suffering and empathy themes ($\rho = 0.71$) indicates that emotional resonance functions as a mechanism for moral recognition in digital spaces. These patterns extend research on affective publics [8] by demonstrating how emotional configurations support moral claim-making in conflict contexts. The data suggest that the pro-Palestinian narrative’s emotional architecture was more effective at generating what scholars term "affective solidarity"—a sense of shared emotional experience that translates into discursive support—particularly among international audiences not directly involved in the conflict.

Algorithmic and institutional factors created measurable asymmetries in narrative visibility and credibility. The data revealed that content moderation practices differentially affected Palestinian hashtags, with temporary restrictions occurring during high-volume events. Verification systems privileged institutional voices, with blue-tick accounts concentrated among Israeli officials and Western journalists. This created algorithmic epistemic injustice, where platform architectures systematically advantage certain forms of testimony. However, it must be emphasized that this study documents associations between narrative position and platform outcomes; establishing direct causal responsibility of platform bias requires internal platform data not available for

this research. The finding that pro-Palestinian narratives maintained higher moral credibility despite structural disadvantages suggests resilience through network density and emotional intensity. This extends research on media framing [3] by quantifying how digital platform governance shapes moral perception. The temporal analysis showing a gradual convergence in narrative volume share (Table 7) may reflect either shifting platform dynamics, changing user attention, or the effect of sustained grassroots mobilization over time.

These findings contribute to scholarship on Palestinian digital activism [1] by providing empirical evidence of how moral authority is constructed under conditions of asymmetric power. The distributed witnessing model represents an innovation in how marginalized communities leverage digital networks to counter institutional erasure. This aligns with historical patterns of Palestinian cultural preservation through alternative archives, now translated into digital practice. The higher empathy ratios for pro-Palestinian content (0.68 versus 0.39) suggest that emotional connection can transcend geopolitical positioning when testimony is perceived as authentic. This reinforces the theoretical proposition that empathy operates as a form of moral currency in digital spaces, capable of circulating and accruing value outside traditional channels of political authority.

Researcher positionality shaped interpretation through attention to power dynamics in knowledge production. The analytical framework prioritized understanding how structural factors influence testimony credibility, rather than evaluating truth claims of different narratives. The mixed-methods design allowed for triangulation between quantitative patterns and qualitative meanings, reducing interpretive bias through methodological diversification. To further mitigate bias, we subjected our interpretations to peer debriefing and explicitly searched for disconfirming evidence, such as pro-Israeli tweets employing emotional suffering frames or pro-Palestinian tweets leveraging institutional authority, which were present but less frequent.

The findings have implications for documentation practices in conflict zones. Distributed witnessing suggests that decentralized archiving through social media can preserve historical records against institutional erasure. However, algorithmic constraints indicate platform dependence creates vulnerabilities for marginalized testimony. This underscores the need for independent archival systems that capture digital testimony without corporate mediation. For historical accountability, the persistence of pro-Palestinian moral credibility despite structural disadvantages demonstrates how digital networks sustain counter-narratives. From a practical standpoint, humanitarian and human rights organizations might consider strategies that harness distributed witnessing while simultaneously advocating for more equitable platform governance and investing in independent, resilient archival infrastructures.

Educational implications include developing critical digital literacy that recognizes how platform architectures shape moral perception. Policy implications involve reconsidering platform governance to ensure equitable treatment of humanitarian testimony. The documented asymmetries suggest need for context-aware moderation that distinguishes between documentation of suffering and incitement to violence. Furthermore, the study highlights the ethical responsibility of platform designers to audit their algorithms for disparate impact on conflict narratives and to incorporate principles of epistemic justice into content moderation policy frameworks.

The theme of witnessing as ethical memory connects to scholarship on cultural memory in Palestinian studies. Digital testimony functions as contemporary practice of preserving collective memory against historical erasure. For social justice movements, the findings demonstrate how digital networks facilitate transnational empathy, where moral recognition crosses geographical boundaries. This transnational empathy, however, is not without its own complexities, potentially leading to what some scholars term "spectatorial sympathy" that may not translate into sustained political action or nuanced understanding of historical context.

Several limitations qualify these interpretations. The exclusive focus on Twitter/X limits understanding of moral witnessing across different platform ecosystems. The six-month timeframe captures specific conflict phase but may not reflect longer-term patterns. Automated translation may have obscured linguistic nuances in moral concepts. Future research should examine moral witnessing across multiple platforms and extended timeframes. Additional limitations include the inability to fully account for coordinated inauthentic behavior or state-sponsored influence campaigns, the Western cultural bias inherent in the VADER sentiment tool, and the correlational nature of the analysis which precludes definitive causal claims about platform effects. Future studies could employ platform-native researchers for linguistic nuance, incorporate data from visual analysis, and utilize quasi-experimental designs around specific platform policy changes to better isolate causal mechanisms.

This study demonstrates that moral credibility in digital conflict narratives is constructed through interactions between emotional expression, network structure, and platform governance. The distributed model of pro-Palestinian witnessing represents innovation in how marginalized communities assert epistemic agency under structural disadvantage. These findings contribute to understanding how digital technologies transform moral testimony in contexts of asymmetric conflict. By validating a methodological framework and a composite metric for moral credibility, the study provides tools for future comparative research across different conflicts and platforms, advancing the empirical study of digital ethics in times of war.

8 Closing Remarks

This study analyzed 372,000 tweets from the Israel-Palestine conflict between October 2023 and March 2024 to examine digital moral witnessing. The research documented how moral credibility is constructed across asymmetric power dynamics. Pro-Palestinian narratives achieved higher moral credibility scores through distributed witnessing practices. The findings indicate that emotional intensity and network density can offset institutional marginalization in digital spaces. These insights advance understanding of how Palestinian experiences are documented and recognized within systemic constraints. The study operationalizes abstract ethical concepts like epistemic injustice and moral witnessing into measurable digital phenomena, creating a bridge between normative theory and computational social science.

The mixed-methods approach contributes to ethical documentation by preserving narrative testimony that might otherwise be erased through algorithmic filtering or content moderation. This methodology enables interpretation of lived experiences

that quantitative metrics alone cannot capture. For policy and education, the findings indicate a need for digital literacy that recognizes how platform architectures shape moral perception. The integration of quantitative and qualitative analysis provides a framework for examining testimony that respects both empirical patterns and human meaning. The transparent reporting of methodological choices, limitations, and validation steps aims to set a standard for reproducible and critically reflexive research in this politically sensitive domain.

Future research should examine moral witnessing across multiple digital platforms to understand cross-platform empathy diffusion. Studies could explore applications in conflict medicine by analyzing how digital testimony informs humanitarian response to health crises. Research on cross-cultural understanding might investigate how moral recognition translates across different linguistic and cultural contexts. Longitudinal analysis could track how witnessing practices evolve over extended conflict durations. Specifically, we recommend future work to: 1) develop and validate culturally adapted sentiment lexicons for conflict discourse in Arabic and Hebrew; 2) employ network analysis to map the flow of moral authority through influencer networks and identify key bridges between different narrative communities; 3) conduct comparative case studies of digital witnessing in other protracted asymmetric conflicts to test the generalizability of the distributed versus institutional credibility model.

This research demonstrates that digital technologies transform how moral testimony is produced and received in asymmetric conflicts. The distributed model of witnessing represents an approach through which marginalized communities assert narrative presence. These findings contribute to broader efforts to document human experience in contexts where traditional institutional channels may be constrained. Ultimately, the study underscores that in the digital age, the struggle for recognition is simultaneously a struggle over the algorithms, metrics, and affective economies that govern visibility and credibility. The ethical imperative is to design these socio-technical systems in ways that mitigate rather than amplify existing epistemic injustices.

Ethics Statement

The study utilized publicly available, de-identified data. No direct interaction with human participants occurred; therefore, institutional review board approval was not required under relevant ethical guidelines. User handles, geolocation identifiers, and personal information were removed before analysis.

Data Statement

Dataset Name: Israel–Palestine Conflict Tweets Dataset. Provider: Mehya Mlaweh (2024) via Kaggle. Data Type: Publicly available social-media corpus. License: CC BY-NC 4.0. Availability: Open Access for academic research. All analysis code, the qualitative codebook, and detailed methodology for the Moral Credibility Index are archived in a public repository (URL anonymized for review) to ensure full reproducibility.

Disclosure Statement

The authors declare no financial or personal conflicts of interest. No funding agency, political organization, or governmental body influenced the study design, data analysis, or interpretation. The theoretical orientation of the study (e.g., use of decolonial theory) is an explicit analytical framework chosen to sensitize the research to power dynamics, not a statement of partisan affiliation. All interpretations are derived from the empirical data following established analytical procedures.

References

- [1] Rami Abuzarqa. Digital activism in the palestinian context. *Social Media + Society*, 9(1):1–12, 2023.
- [2] Stuart Allan. Conflict communication ethics. *Journalism Studies*, 18(8):991–1008, 2017.
- [3] Mohamed H. Elmasry et al. Media framing and algorithmic bias in palestine-israel digital narratives. *Digital Journalism*, 10(4):512–530, 2022.
- [4] Miranda Fricker. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, Oxford, UK, 2007.
- [5] S. Gregory. Cameras everywhere: Ubiquitous video documentation of human rights, new forms of video advocacy, and considerations of safety, security, dignity and consent. *Health and Human Rights*, 12(2):1–8, 2010.
- [6] Avishai Margalit. *The Ethics of Memory*. Harvard University Press, Cambridge, MA, 2002.
- [7] Mervi Pantti. *Affective Journalism: The Emotional Dimension of News and Conflict*. Palgrave Macmillan, London, UK, 2022.
- [8] Zizi Papacharissi. *Affective Publics: Sentiment, Technology, and Politics*. Oxford University Press, New York, NY, 2015.