

# Framing the Citizenship Amendment Act: A Multimodal Analysis of Short-Video Discourse and the Construction of Belonging in Digital Spaces

Namit Hans<sup>1\*</sup>, Shikha Rai<sup>2</sup>

<sup>1</sup>Research Scholar SOJNMS, IGNOU Affiliations

<sup>2</sup>Director (SOJNMS,IGNOU)

## ARTICLE INFO

### \*Correspondence:

namit.hans@gmail.com  
Research Scholar  
SOJNMS, IGNOU

### Dates:

Received: 21-04-2026  
Accepted: 24-05-2026  
Published: 30-06-2026

### Keywords:

CAA, Framing  
Analysis, Short Videos,  
Democracy, Social  
Media Discourse

### How to Cite:

Hans,N., Rai, S. (2026)  
Framing the Citizenship  
Amendment Act: A  
Multimodal Analysis of  
Short-Video Discourse  
and the Construction  
of Belonging in Digital  
Spaces. *MediaSpace:*  
*DME Journal of*  
*Communication*,7(1),85-97.  
doi: 10.53361/dmej.  
v7i01.08

## Abstract

The rapid rise of digital media has led to a fundamental alteration of the contemporary public discourse, and short-video platforms have emerged as critical sites hosting reconstruction of complex socio-political issues within the digital sphere. This study presents a framing analysis of short-video content related to the Citizenship Amendment Act (CAA), analysing how diverse and competing interpretive frameworks influence the audience's understanding of the issue and the overall mediation of social identity. This research is situated within a larger PhD project by the researcher, which examined 15 prominent topics in India, including the focus of this study - CAA. The 10 analysed videos, a subset of total 100 videos, were selected because the topic (CAA) serves as a strong case study for competing narratives and high emotional intensity, thereby influencing identity construction of the audience.

Findings of the study reveal a highly fragmented discursive landscape. The Legal and Constitutional Frame anchors the entire discourse in formal logic and institutional validity of the act. On the other hand, the Humanitarian Frame sees how unique platform characteristics of the short-video format, like personal storytelling and facial expressions, are used to evoke empathy and highlight the impact on lived experience of citizens. The Security Frame, conversely, puts an emphasis on need for administrative regulation and national interest, while the Identity and Cultural Frame focus on the collective memory and symbolic belonging aspects of the issue. The study demonstrates that short-videos bring together visual, linguistic, and auditory elements to create 'hybrid' frames, playing a crucial role in influencing social cohesion and formation of a collective identity.

## INTRODUCTION

The contemporary landscape of digital media or new media has gone through a major structural change in the last decade. The rise of short-video platforms as one of the biggest arenas of political communication has transformed public discourse in many ways. In the past, legacy media platforms like television channels, newspapers, or broadcast institutions maintained strict gatekeeping

## MediaSpace:

*DME Journal of Communication*  
e-ISSN: 2583-035X

© Delhi Metropolitan Education, 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>.

of information. However, the ease of access created by cheap smartphones and democratization of internet access has worked well in removing this monopoly over gatekeeping (Chadwick, 2013). In this new structure, micro-video platforms like YouTube Shorts and Instagram Reels are not only limited to entertainment but are also becoming central to dissemination of political messages and news.

However, the short format of these videos aimed towards quick consumption introduces several constraints to what is communicated and how it is communicated. Unlike traditional media content that allowed for in-depth discussion related to nuances of topics, short-videos compress the issues in simple, digestible nuggets of information. The design of short-video platforms is such that it rewards rapid scrolling, immediate stimulation of senses, and extraordinary brief engagement. All of this is, in effect, against the idea of carrying out any nuanced discussion on a complex issue. Furthermore, digital discourse operates within the constraints of attention economy in which attention of the audience is treated as a scarce resource which everyone competes for aggressively in order to monetize it (Goldhaber, 1997). Within this structure, media creators are incentivized to capture attention of the viewers through highly condensed, visually aggressive, and emotionally charged content instead of putting the effort into carrying out a detailed, objective analysis.

In this context, the introduction and subsequent implementation of Citizenship Amendment Act (CAA) provide a strong and relevant opportunity for investigating how these platform dynamics shape the discourse and/or narrative related to highly contentious issues like these. At the time of its introduction, due to the nature of the proposed legislation, CAA led to an unprecedented reaction from the public creating extreme polarization among the citizens in both offline and online spaces. The debate surrounding the issue did not restrict itself to simple legal or policy related deliberations. However, it went further and touched upon sensitive questions of national identity, historical trauma of partition, constitutional secularism, and belonging as a citizen of a post-colonial state.

The CAA issue which was connected to the

complex realities of international human rights frameworks and constitutional law was stripped of its nuance and flattened into binary tropes fitting for platform dynamics of short-videos. Everyone, including independent creators, political commenters, political parties, and digital media platforms, utilized the short-videos by creating emotional narrative around the issue in order to achieve strong algorithmic visibility. In this paper, the researcher looks at the media artifacts isolating them from the audience reception and response in order to assess the narratives portrayed and utilized in the videos. The researcher applies a multimodal analysis technique in order to dissect the audio-visual techniques used within the video and how it worked towards constructing public meaning.

## Objectives and Research Questions

The research is aimed at addressing the primary objective of understanding the narrative and structural framing strategies employed within short-videos related to CAA. The researcher has intentionally focused only on the site of production, i.e. the videos, and not on the site of consumption and reception (audience comments and engagement) to specifically decode the visual and auditory grammar used by digital creators with the goal of compressing complex issues like CAA into a bite-sized, polarizing narrative.

- This objective is addressed by systematically focusing on three distinct research questions:
- Research Question 1 (RQ1) - What are the dominant narrative and structural frames constructed within short-form digital videos to interpret and communicate the legislative realities of the Citizenship Amendment Act?
- Research Question 2 (RQ2) - How do digital creators synthesize multimodal affordances, like typography, audio, visual editing, and cultural symbols, to construct these frames within a temporal constraint of short-videos?
- Research Question 3 (RQ3) - In what ways does this multimodal framing strip away statutory nuance to redefine the parameters of national belonging and identity in the digital public sphere?

## LITERATURE REVIEW

This research has been situated within the broader academic discourse related to digital communication, attention economy, visual display of suffering, and historical context of political communication especially in the Global South.

### Attention Economy and Algorithmic Relevance

There has been an epistemological shift in how publics engage with and process political information because of the transition from long-form broadcast media to short-form vertical videos on social media. As internet allowed more and more people to become a participant in the discourse, early theorists argued that this will allow highly deliberative and egalitarian public sphere. However, over the years, it has been observed that the commercialization of social media platforms has led to rise of surveillance capitalism, putting healthy civic engagement in the backseat. The concept of attention economy was first articulated by Michael Goldhaber (1997) who argued that in a society where information is available in abundance, the resource that becomes scarce is human attention. Therefore, as the digital platforms evolved, they competed for harvesting this resource and their algorithms were mathematically designed to prioritize content that kept the audiences hooked through high-arousal emotional states including fear, anger, awe, and moral outrage (Berry & Sobieraj, 2014).

On the other hand, Tarleton Gillespie (2014) states that the algorithmic architecture relies heavily on politics of 'algorithmic relevance'. This is because algorithms cannot be seen as neutral mathematical formulas. Instead, they are deeply personalized because they decide which content should become more visible and what should remain hidden based on the audience engagement metrics. When we see this within the context of journalism, it shows that the algorithmic architecture has led to the rise of ambient journalism. In such a system, news is consumed in fragmented, continuous streams like in the short-videos rather than through a dedicated, deliberate reading as it was done in the past (Hermida, 2010).

At the time of data collection for this study, short-video platforms restricted the length of a video to not more than 90 seconds. However, most creators restricted themselves to less than a minute. Now, since the creators have such a short-time to capture the user's attention and prevent them from scrolling down, they are forced to adopt techniques that are not rooted in objectivity. Instead, objectivity is purposefully replaced with highly partisan and intense framing. Zeynep Tufekci (2013) argues that working within a competitive attention economy forces creators and political actors to adopt micro-celebrity tactics. This helps them gain visibility, and in doing so, brings together digital entertainment and political deliberation. This shows that digital discourse has focused itself on triggering emotional resonance instead of trying to primarily educate and inform.

### Multimodality, Visual Politics, and Spectacle of Suffering

When we look at the difference between the visual-heavy digital communication and the traditional media, we understand that the latter has historically relied on text-heavy and spoken rhetoric. This shift invites analysis of the short-video content with a fresh perspective because images tend to possess a unique cognitive authority. According to Kress and van Leeuwen (2006), visuals operate with a uniquely specific grammar that is independent of language. They argue that visuals can bypass the rational and critical filters of the human brain, while creating immediate affective connections.

This is further connected to multimodality within digital networks. Multimodality, which combines moving images with audio, text, pacing, colors, and graphics compounds the visual power of short-videos. Research has shown that this combination of elements enables video creators to construct highly persuasive artifacts (Chadwick, 2013). These digital artefacts also communicate complex ideas and political stances through shared and contextual visual vernacular like memes and short-videos (Highfield and Leaver, 2016). Within the context of analysing political communication, it becomes even more crucial to take into account the aesthetics of visual framing. According to Lilie Chouliaraki (2006),

media often aestheticizes trauma and creates a 'spectatorship of suffering' in order to evoke strong moral response from the audience. When we see how this aestheticization operates within short-videos, we see that they are condensed into a simplistic narrative of victim and perpetrator. This allows for ideological manipulation through controlling of the narrative.

## **Indian Public Sphere and Nationalism in Digital Spaces**

When we look at these broadly global digital trends from an Indian context, it is important to place them within an informed socio-historical framework. In the last few years, India has seen a rapid spread of digital networks fuelled by cheap mobile internet and smartphones. Hundreds of millions of citizens who were not part of the larger digital discourse in the past are now participating in meaning-making process simultaneously. However, this change is highly complex because of the diversity of Indian society on the lines of language, religion, and socio-economic backgrounds.

Therefore, the discourse related to CAA needs to be understood keeping in mind this deeply stratified nature of Indian society. Within this structure, 'digital nationalism' has been on the rise as political actors use extreme speech through memes and short-videos in order to consolidate majoritarian identities (Sahana Udapa, 2019). According to Udapa, this digital nationalism is often hidden behind entertainment where engaging content is used to further exclusionary political rhetoric. In India, the introduction of the CAA had immediately led to volatile information war on the internet with large media houses competing against independent creators and influencers who shared their opinions in highly sensitive language. Within this, short-video platforms became a battlefield where everyone tried to exert their influence and to push their idea. This makes the multimodal framing of CAA related short-videos a highly relevant inquiry into media studies as well as how digital framing is used to negotiate boundaries of national identity.

## **THEORETICAL FRAMEWORK**

This study analyses the short-videos by bringing together classical framing theory and contemporary frameworks of multimodal visual analysis. This approach where two theories are used, provides strong scaffolding for interpretation of what is being said and how it is constructed visually.

### **Entman's Framing Paradigm**

According to Robert Entman (1993), framing involves selecting specific aspects of perceived reality and highlighting them to make them salient within a context. Further, he proposes that framing performs four distinct functions of defining the problem, isolating the underlying causal agents, making moral judgements regarding the actors involved, and suggesting specific programmatic remedies. In this context, when a digital creator focuses on a complex topic like CAA, he or she is forced to act within the constraints of the platform. Therefore, they have to make important framing choices. In a 60 or 90-second video, the creator cannot show the entire history of the issue or provide an overall legal perspective. Instead, they focus on a fragmented perspective which is highly curated. In order to do this, the creator must decide in advance the victim and the perpetrator, and they should provide the ultimate moral standpoint related to the issue. All of this has to be communicated to the viewer within seconds. Entman's four-stage paradigm can be utilized to understand how creators frame the same legislative issue and construct diverse political realities.

### **Multimodal Visual Grammar**

However, as we conduct the framing analysis, it needs to be considered that the traditional framing literature was constructed primarily for television and print media where text-heavy and/or spoken dialogue based artifacts were prioritized (Scheufele, 1999). When we conduct framing analysis within the context of short-videos, we need an expansive vocabulary which accounts for the multimodal nature of these platforms. Thus, the framework developed by Kress and van Leeuwen (2006)

becomes crucial where the concept of 'visual grammar' is explained. According to them, images, colours, spatial arrangements, camera angles, and much more is used to construct social meaning at par with language.

As per Kress and Van Leeuwen, visual design fulfils three meta functions, namely representational, interpersonal, and compositional. The first meta function shows how the subject is portrayed in relation to other objects within the frame. The second meta function tells us how relationship between the viewer and the image/text is established, often through manipulation of eye contact and camera angles. The third meta function dictates the placement of objects on the screen to guide the attention of the viewer creating a careful hierarchy of elements.

When we see this framework within the context of short-videos, it is observed that they involve all three meta functions. For example, moving images are layered with background music, text, etc. This means that the viewer, as he or she watches the visuals, is guided by certain music to feel specific feelings and emotions. Similarly, by placing two unrelated or related visuals within a split screen, the creator pushes the viewer to connect the two instantly and construct the intended meaning through this spatial arrangement. Therefore, this framework enables us in dissecting the digital artefacts frame by frame, understanding the aesthetic choices as not random occurrences but as being inherently political.

## RESEARCH METHODOLOGY

The research aims to examine the visual and narrative construction of CAA related short-videos. To achieve this, the researcher has used a rigorous qualitative and multimodal analysis. Unlike quantitative analysis that focuses on how algorithms work and look at the engagement with the videos, this qualitative approach is suitable to look at the anatomy of the videos being analysed. The epistemological stance of this research is that the short-videos as media texts do not just reflect reality but are used to actively construct it.

## Collection of Sample and Finalizing the Corpus

The larger sample from which the videos for this research have been derived are part of a larger institutional PhD research which is aimed at analysing the manifestations of outrage in digital discourse. The data was collected from December 2023 to May 2024, focusing on highly relevant window of six months characterized by important events like upcoming General Elections, ongoing State Elections, and heightened nationalistic rhetoric owing to the events like introduction of CAA. To ensure the videos collected for the larger corpus were free of any selection bias, the researcher used a multi-step process of corpus finalization. The initial corpus of more than 1000 videos were selected by looking at the videos on two major platforms, YouTube Shorts and Instagram Reels, by focusing on subject-specific hashtags for each topic like #CAA, #CitizenshipAmendmentAct, and #IndianPolitics. The researcher selected at least two topics for each month based on their newsworthiness.

In the second stage, the initial corpus was manually analysed by the researcher to remove any redundancies, videos that were removed by the creators, or videos that used narrative styles that were not relevant for the analysis of the study. This provided a second corpus of total 329 videos which was then run through a python-based random sampling algorithm to eliminate any human bias. The rigorous computational sampling provided a final dataset of 100 videos. The dataset included videos from fifteen distinct topics, including the 10 videos related to Citizenship Amendment Act (CAA). While the sample size of 10 videos is quantitatively small, it is relevant for the multimodal in-depth qualitative analysis of the study. The frame-by-frame analysis of the 10 videos provides a sufficiently dense pattern of recurring frames analysed by the researcher.

## Multimodal Coding and Analytical Procedure

To conduct the in-depth analysis of the 10 videos, the researcher used a structured qualitative framework to evaluate the narrative and multimodal

components. Each video was analysed frame-by-frame with the researcher often pausing, playing back, and choosing slower speed to carefully identify the patterns. The framework primarily focused on the following multimodal parameters:

- Visual framing – Including use of camera angles, colour grading, and spatial arrangement.
- Kinetic typography – How the text overlays are placed, their colour and size, and the animation used to anchor narrative meaning.
- Auditory framing – Use of background music, sound effects, and cadence, pace and emotional tone of the narrator's voice.
- Symbolic repertoire – The use of partisan visual symbols, religious and political iconography, historical elements, and national flags.

The researcher looked at how these multimodal elements were used together or combined to create recurring narrative frames. The analysis found that the discourse around CAA in the short-videos was not something that occurred randomly. Rather, it was crafted carefully around the four identifiable and dominant framing categories as highlighted in the subsequent sections.

## Ethical Considerations

To ensure that the analysis of highly polarizing content shared in public domain was done in an ethical manner, the researcher only looked at the videos that were available openly on the platforms and were allowed for mass consumption. The researcher did not take into account any private interactions, closed-group materials, or direct messages with the creators. Further, the researcher did not identify the video creators or uploaders to maintain their anonymity and focused entirely on the content of the videos. Each video was coded with a combination of the topic initials (CAA) and the number assigned to the video (001, 002, etc), to ensure there were no overlaps in the analysis and the videos were arranged structurally. The researcher also maintained high flexibility by acknowledging biases to make sure the analysis remained rooted in the established theories and frameworks that are the pillars of this study.

## ARCHITECTURE OF THE FOUR FRAMES

The qualitative analysis conducted by the researcher found recurring patterns within the videos used by the creators to communicate meaning to the audience. The four identified frames are as follows:

### Legal and Constitutional Frame

The first frame employed by the creators is the legal and constitutional frame which deliberately



**Figure 1:** Juxtaposition used by the creator to make political statement without providing detailed context of the statement.

removes any interpretive depth from the issue and instead utilizes aggressive, kinetic typography. Within this category, the creator often assumes the role of an absolute legal authority on the subject. In doing so, the creator focuses on the aesthetic of the law rather than explaining the legal nuances. Within the video, text overlays are used to highlight constitutional articles, previous court rulings, etc. without providing any detailed context to the viewers. Most of the times, such text is highlighted using bold text and high-contrast colour (Figure 1).

This technique of using legal text as an overlay in a video which contains highly partisan language helps the creator to situate their argument within a so-called objective constitutional context. In doing so, the creator simulates intellectual rigour within the visual grammar when there is not enough opportunity to actually engage with the topic intellectually because of the time limitations set by the format. Often, the appearance and disappearance of the text is faster than what the viewer can process forcing the viewer to accept it as a legal backing of the argument being made without actually processing it by themselves. For example, in the video 002-CAA, the narrator looks directly into the camera while speaking with an intense expression on their face. With this, their arguments display certain declaration and clarity without any hesitation, conveying to the viewer that there is no room for ambiguity in the argument being presented.

Entman (1993) argues that effective framing is a result of defining the problem clearly and providing a definite remedy for it. The legal frame shows how this technique is used effectively across the videos. One side of the ideological spectrum uses legal language to frame CAA as a necessary intervention to enforce historical justice. For example, in video 004-CAA, creator makes use of unrelated visuals of government seals and archival parliamentary broadcasts to frame the issue as having institutional legitimacy and, thus, an unyielding state authority backing it. These visuals are further combined with high contrast colour grading to evoke strong emotions in the viewer.

On the other hand, those opposing the implementation of the act utilize the same visual

framing to portray it as a sign of constitutional collapse. Split screen technique is used where unrelated videos of marginalized communities facing police action along with a black-and-white framing is done to create a dramatic effect. This turns a highly complex legal issue into a spectacle with binary voices, portraying each side as either absolutely right or wrong. Therefore, the law does not remain a subject of deliberation but an aesthetic to legitimize the ideological and political view of the video creator.

## **Humanitarian Rescue Frame**

The legal frame utilizes constitutional language to create an illusion of authority of the narrator or the creator. However, the humanitarian frame lets go of this objective attempt at meaning-making and targets the emotional responses of the viewer. In this frame, the creator does not focus on any legal or constitutional arguments. Instead, they rely solely on the raw element of human suffering. The goal of the frame is to make the viewer look at the issue as not an administrative procedure, but as a question of moral right or wrong. In doing so, the video creator tries to show that the state should act like a benevolent saviour.

The multimodal techniques used by the creators in this frame rely on visuals of historical memory and intense cinematic expression. To achieve this, archival footage is used to portray trauma, historical persecution, displacement, and religious bias. All of this is done along with the help of desaturated images, and slow, deliberate cross-fading of visuals. Here, the creators use the technique of 'spectatorship of suffering' as highlighted by Chouliaraki's (2006). Through this, the aesthetic of pain and suffering is used to elicit a moral and political reaction from the viewer aiming to push them towards political allegiance of the creator. For example, video 006-CAA uses a quotation of a non-Indian athlete regarding the treatment of minorities in their country (Figure 4). This decontextualized use of a quote does not just support the visual information being displayed in the video but also actively creates an emotional reality for the viewer Kress and van Leeuwen (2006). At the same time, the creator uses a melancholic background music



**Figure 2:** High-arousal editing techniques used by the creator to create an impression of threat and/or collective anxiety.

throughout the video to create a sense of tragedy in the minds of the viewer.

The visual grammar used in this frame constructs a victim-perpetrator binary where distressed faces of marginalized minorities are zoomed into with the aim of pulling the viewer into an intimate relationship with the subject of perceived persecution. And then suddenly, the narrative voiceover shifts in tone, the melancholic music makes way for a triumphant crescendo, and the visual transitions abruptly from historical black-and-white to vibrant, high-definition colour. In video 006-CAA, the creators utilize the split-screen format to place the images of political

leaders along with rescued people intending to manipulate the emotions of the viewer. The framing shows the issue as a civilizational rescue of historically marginalized people and any criticism of the policy as cruel and apathetic. Therefore, the frame creates the impression that questioning the act through a legal standpoint is going against the humanitarian values.

## The Security and Threat Frame

In the humanitarian frame, the creator utilizes empathy as an emotion to communicate with the audience. However, in this frame, the creator makes use of aesthetics of fear, anxiety, and vulnerability. Within this context, the issue is not portrayed as aimed towards rescuing those persecuted, but as an urgent need to protect the national borders from infiltration and demographic changes. The frame calls for an urgent intervention from perceived existential threat. This can be viewed from Entman's (1993) model and the concept of problem definition. In the frame, the focus shifts from the marginalized victim to a dangerous infiltrator, demanding their immediate expulsion.

The primary technique used to create this emotional resonance is hyper-kinetic editing that intentionally creates a sensory overload for the viewer (Figure 2). For example, in video 010-CAA, the creator uses jarring, rapid, and disorienting jump cuts while the entire video is flooded with shaky, handheld smartphone footage. This footage shows border crossings and crowded urban spaces. The destabilizing visuals are a deliberate choice made by the creator intending to signal a loss of institutional control over the infiltration. It also portrays that the country is currently facing a crisis that is almost out-of-control and unmanageable. At the same time, the video makes use of aggressive, flashing borders and red colour grading. The red colour is primarily used as an indicator of threat or danger and it is embedded in the videos to simulate a language of emergency.

Just like the visuals, the auditory techniques used within this frame are also aggressive. The creators employ high-tempo, electronic beats or heavy percussion tracks to influence the viewer's mood and perception. It creates a panic-inducing pace



**Figure 3:** Cultural symbolism used by the creator to project cultural identification as a marker of national identity.

forcing the viewer to feel the anxiety as intended by the creator. Like legal frame uses constitutional language in text overlays, in this frame the creators use unverified statistics and facts related to infiltration and its economic impact on the nation. Often, the text is capitalized to create a strong impression in the minds of the viewer. Another important technique in this frame is the use of digital cartography. For example, in video 010-CAA, the creator uses animated maps with aggressive red arrows on top of it showcasing illegal infiltration as a physical and militaristic invasion of international borders. The viewers are bombarded by a mix of

shaky visuals, aggressive music, and the rhetoric of invasion in order to frame the issue as an emergency defence against an organized attack completely eliminating any legal nuance from the topic.

## Identity and National Belonging Frame

The final and ultimate frame what short-videos display is the identity and national belonging frame. While this frame is comparatively abstract, it is also arguably the most potent of all four. This is because the frame does not dwell onto the technicalities of legal frame or the statistical posture of the security frame. Instead, in this frame, video creators ask a single, all-encompassing question about who belongs to the nation legitimately? Within this context, the short-video becomes a potent virtual border around which the insiders and outsiders are placed. The insiders are considered part of the nation while the outsiders are seen as infiltrators.

Particularly, this frame uses multimodal strategies that are rooted in cultural symbolism. On screen, you see visuals of historical monuments, religious iconography, and traditional garments, all aimed towards making the distinction between the insider and the outsider (Figure 3). Here, citizenship becomes not a legal issue but rather a shared, cultural mythology and rhetoric. For example, in video 005-CAA, graphics related to the country's partition in 1947 are used to justify the current policies. On the other hand, the editing techniques rely heavily on juxtaposition like using a split screen where, in one frame, the creator showcases those who are the insiders and are engaged in patriotic activities, while on the other side, they place those who are supposed outsiders and are engaging in culturally alien activities or disruptive behaviour. This juxtaposition is used to create a strong contradiction between the two parties, with one being pitted against the other has culturally and civilizational different.

Further, the narration or vocal performance used in this frame is highly theatrical. The narrators do not use an objective language or cadence while talking about the issue. Instead, they adopt a highly dramatized tone or rhythm to deliver their monologue. Narrators raise big issues like historical



**Figure 4:** Split screen technique used to provide decontextualized visuals of a foreign athlete while comparing it with CAA implementation.

betrayals, legacy of a civilization, and the need to maintain cultural purity. The background music or audio supports this framing by setting up a dramatic mood using cultural anthems or sound of traditional instruments. For example, in video 009-CAA, the entire video consists of visuals of mass rallies along with a background music consisting of religious slogans and military style soundtracks. By using this combination of audio and visuals, the creator immediately connects with the viewer and communicates a communal message through emotional appeal. This bypasses the cognition of

the viewer and they are pulled into the narrative. Lastly, in this frame, the creator projects that national belonging cannot be a legislative issue. Rather, it is an issue of cultural alignment. The sixty-second video narrates to the viewer who should be considered a citizen and who shouldn't be through a highly distilled discourse surrounding a very complex issue.

## DISCUSSION

As observed, the in-depth qualitative analysis of the 10 videos provides a deep perspective into the four major frames used by creators to narrate meaning to the viewers. This analysis also provides a foundation for answering the three research questions, and to see how the findings connect with the theoretical framework proposed at the beginning.

### Collapse of Deliberative Discourse: Addressing RQ1 and RQ3

First, we should look at the RQ1 regarding identification of dominant frames and RQ3 regarding elimination of statutory nuance. It has been observed in the analysis of the videos that narrative framing of short-videos related to Citizenship Amendment Act (CAA) systematically strips away any deliberative depth from meaning-making. In all the four frames identified above, namely legal, humanitarian, security, and identity frames, any nuance is eliminated and creators employ aggressive simplification to connect with the audience. This is also because of the limitations of the short-video where the creators cannot practically explain the legal, constitutional, historical, or cultural depth of the issue at hand in a span of 60 or 90 seconds. Therefore, creators regularly make use of hybrid techniques of framing, bringing together more than one frame in a video to narrate a complex meaning to the audience about legal complexity and identity-based anxieties parallelly (Chadwick, 2013).

As observed by the researcher, even when legal and constitutional explanation is used, it is not done to make the audience aware or educate them. Instead, it is done as a visual technique to create an impression of objective clarity while pushing a highly political rhetoric. Therefore, it can be safely

said that the short-video discourse operates through a structural flattening where a complex issue of citizenship is not evaluated institutionally or seen as a human rights issue. Instead, it gets transformed into a polarized debate around historical justice and cultural belonging. This means that the digital public sphere loses any faculty of hosting a rational, critical deliberation.

What happens instead is that the discourse is dominated by polarized, volatile, and manipulative messaging. The idea of national belonging is packed into easily digestible nuggets of saviour vs perpetrator, and citizen vs infiltrator. The legal questions related to citizenship are sidelined and the cultural and ideological identity takes precedence in the debate. Thus, by eliminating any nuance, the techno politics of such framing threatens the inclusive inquiry needed to function in a democratic society.

## **Weaponization of Multimodal Aesthetics: Addressing RQ2**

The second research question was aimed at understanding how creators make use of multimodal techniques to construct the frames as identified above. This, especially, within the limitations posed by short-video format. The findings of the study align with the theoretical framework proposed by Kress and van Leeuwen (2006) which show that the format of short-videos makes any nuanced argumentation impossible. This is why, the creators need to heavily rely on the multimodal visual grammar of the short-videos in order to maximize the emotional response from the audience and causing an immediate arousal.

This is observed in the use of hyper-kinetic editing cuts, text overlays with high contrast, polarized visuals in split-screens, and aggressive and dramatic audio tracks. Through a combination of these elements, the creators carefully engineer a narrative. The analysis of the 10 videos also reveals that the use of audio in framing is perhaps an under-observed phenomenon as it is used to immediately create a certain emotional impression in the minds of the audience while they actively engage with the visuals on the screen. For example, a transition from melancholic music to aggressive beats can shift the

understanding of the viewer as they make sense of the visuals, leading them to either empathize with someone or opposing someone.

The combination of these techniques is a potent tool for immediate psychological arousal and tilts human attention towards moralized indignation, collective anxiety, or cultural pride. By making use of these techniques, video creators hack the engagement metrics and algorithm to gain attention of the audience. This shows how aesthetics of short-videos are weaponized aggressively by the creators where the deliberation of traditional media is replaced by emotional display and high-arousal content of short-videos. In such an environment, the aesthetic impact becomes stronger than the truth value of the content.

## **Algorithmic Affordances and Framing**

The four frames identified by the researcher and the overall analyses of the videos also provides a clear support for the four-step paradigm and three meta functions proposed by Entman (1993) where the creators provide a clear outline of the problem, propose a strong remedy, and make use of visual techniques to guide the attention of the viewers. The authoritative tone used by the narrator along with juxtaposition of text and graphics makes a viewer believe that the message is rooted in objective analysis of the issue. The strong solutions provided by the narrators also create an impression of authority, thereby lending support to the framing.

These framing strategies should also be situated within the context of the platform affordances or short-videos because they do not operate through the same rules of social media. In short-video platforms, anyone can discover content whether they are following the account or not. Infinite scroll is a key feature of these platforms where algorithms decide what kind of content to push towards the audience based on their engagement history (Gillespie, 2014). The four frames identified in this research are particularly a potent tool to hack the algorithm.

This is because the framing leads to high-arousal state making the audience view the content in a sustained manner and wanting for more. The jarring

cuts of security frame keep the viewer hooked while the strategically placed emotional crescendos of humanitarian frame in the middle of the video stop the viewer from scrolling without watching the video till the end. The symbols used in identity frame use resonance as a tool to connect with the viewer immediately and the legal frame gives them a sense of being educated. Therefore, the multimodal techniques highlighted in this study should not be seen as mere artistic choices, but as a potent tool of political and ideological propaganda. These techniques give an edge to the creator in a highly competitive algorithmic environment shaping the discourse not through deliberation but through the mathematics of machine learning.

## CONCLUSION

This study provides a rigorous analysis of the multimodal techniques used in short-videos related to Citizenship Amendment Act (CAA). By solely focusing on the site of production and conducting a careful frame-by-frame analysis of each video, the study has highlighted the socio-technical mechanisms behind framing of complex topics like CAA. The findings clearly illustrate that the architecture of short-video platforms favours high-arousal engagement over in-depth deliberations. The creators rely on four dominant frames, i.e. Legal and Constitutional frame, Humanitarian Rescue frame, Security and Threat frame, and Identity and National Belonging frame. Through the use of these frames, highly complex issue like CAA is turned into bite-sized information that is digestible for the masses. The techniques employed by the creators bypass cognitive engagement and target the emotional arousal of the audience.

In conclusion, this study highlights that short-videos are not just a neutral tool of information distribution. Instead, they are highly ideological and are used to create meaning for the audience. In today's digital age, the debate around complex constitutional and legal issues does not only happen in parliamentary chambers of large media houses where the information is carefully gatekept. Rather, the debate happens in the open arena of social media platforms where meaning is continuously

constructed in a shared manner and communicated instantly. Powerful symbolism, audio, graphics, etc are used to connect with the audience at an emotional level fundamentally rewiring how democratic deliberation happens.

## LIMITATIONS OF THE STUDY

While this study provides a detailed analysis of the videos and highlights the dominant frames, it is also important to acknowledge the limitations. First, the study only focuses on the production of the information and excludes the reception and circulation through likes, comments, and engagement data. This way, the study focused on a specific aspect of short-videos. Future studies can look at a more holistic impact of this meaning-making by analysing the response of the audiences and how they engage with such content. Second, the sample used to conduct the analysis consisted of 10 videos which allowed for an in-depth qualitative analysis, but it also provides a highly localized view of a vast digital ecosystem where thousands of videos are produced and shared. Future studies can focus on a larger dataset to include quantitative analysis as well.

Further, future research can also combine the qualitative methods with new-age quantitative measurement techniques like Natural Language Processing (NLP) to study the comments section to understand the actual impact of this framing. There is also a need to understand and compare how this framing operates across platforms like Facebook, WhatsApp, etc. By diving deep into this analysis, researchers can provide robust solutions and suggestions to policymakers, educators, and social media networks to be able to navigate the complexities of digital age and allow for a healthy democratic deliberation.

## REFERENCES

- Berry, J. M., & Sobieraj, S. (2014). *The outrage industry: Political opinion media and the new incivility*. Oxford University Press.
- Chadwick, A. (2013). *The hybrid media system: Politics and power*. Oxford University Press.
- Chouliaraki, L. (2006). *The spectatorship of suffering*. SAGE Publications.

- Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication*, 43(4), 51-58. <https://doi.org/10.1111/j.1460-2466.1993.tb01304.x>
- Gillespie, T. (2014). The relevance of algorithms. In *Media technologies* (pp. 167-194). MIT Press.
- Goldhaber, M. H. (1997). The attention economy and the net. *First Monday*, 2(4). <https://doi.org/10.5210/fm.v2i4.519>
- Hermida, A. (2010). Twittering the news: The emergence of ambient journalism. *Journalism Practice*, 4(3), 297-308. <https://doi.org/10.1080/17512781003640703>
- Highfield, T., & Leaver, T. (2016). Instagrammatics and digital methods: Studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, 2(1), 47-62.
- Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design* (2nd ed.). Routledge.
- Scheufele, D. A. (1999). Framing as a theory of media effects. *Journal of Communication*, 49(1), 103-122. <https://doi.org/10.1111/j.1460-2466.1999.tb02784.x>
- Tufekci, Z. (2013). "Not this one": Social movements, the attention economy, and microcelebrity networked guerrilla journalism. *American Behavioral Scientist*, 57(7), 848-870.
- Udupa, S. (2019). Nationalism in the digital age: Fun as a metapractice of extreme speech. *International Journal of Communication*, 13, 3143-3163.

