International Society of the Learning Sciences

Co-Constructing (Twisted) Knowledge: Argumentation in Far-Right YouTube Comments on Brazil's War with X's Elon Musk

Carolina Sotério, Teachers College, Columbia University, cs4465@tc.columbia.edu
Renato Russo, Teachers College, Columbia University, rfr2126@tc.columbia.edu
Paulo Blikstein, Columbia University, and Instituto Tecnológico de Monterrey, paulob@tc.columbia.edu

Abstract: Social media can foster collective knowledge construction alongside communication. However, these platforms can also be hubs for disseminating distorted or "twisted" knowledge, particularly within extremist contexts. This study analyzes argument negotiation in the comments section of a far-right YouTube channel implicated in anti-democratic activities in Brazil. Using an adapted coding scheme from Dubovi and Tabak (2020), we examined 123 comments to identify patterns of argumentation and knowledge construction. Findings reveal three key insights: (1) users engage in logical reasoning that often leads to the construction of misinformation, reinforcing ideological polarization; (2) participants primarily act as consumers interacting with the influencer rather than engaging in peer-to-peer dialogue, limiting the potential for civic discourse; (3) instances of deeper argumentation reveal the potential for fostering critical engagement. These results highlight the interplay between argument negotiation and disinformation on social media, emphasizing the need for strategies to mitigate extremism and foster constructive dialogue online.

Introduction and literature review

Among the many ways social media has become part of life in society, its role in politics is remarkable. Elected officials, activists, and lay citizens all take part in online conversations concerning politics at different levels. One disturbing facet of that relationship is the association between social media and alt-right extremism. For example, evidence has been found of paths toward radicalization through YouTube's recommendation systems (Haroon et al., 2023). While echo chambers amplify the problem, we argue that distorted learning processes on social media foster attachment to self-generated theories. By "distorted," we refer to learning processes where logical reasoning reinforces preexisting ideological biases, often resulting in misinformation (Burdick & Sandlin, 2022). In this paper, we analyze evidence of collective knowledge construction in the comments section of a YouTube channel that was taken down for involvement in anti-democratic speech leading to a coup attempt in Brazil in January 2023.

Educational research has increasingly recognized social media as learning settings beyond mere communication spaces. The dynamics of social media have been compared to constructivist learning, emphasizing users' social and emotional involvement (Greenhow & Robelia, 2009). Researchers adhering to sociocultural perspectives on learning have identified how engagement in social movements involves participation in social media in ways that can be framed as participation in communities of practice (Gleason, 2013, 2016; Gleason *et al.*, 2021). Other approaches identified evidence of collective knowledge construction on Facebook (Asterhan & Hever, 2015) and TikTok (Nguyen & Diederich, 2023). Dubovi and Tabak (2020) proposed a coding scheme to identify stages in the process of knowledge construction that could be identified in YouTube comments: they put forth the proposition that comment sections transcend the "transmission model" that characterizes instructional video on social media and promote an environment of collective knowledge construction. That line of scholarship indicates that learning processes are also possible in social media.

However, research shows that not all learning on social media is beneficial: scholarly work points to the existence of "public pedagogies" – learning that happens beyond traditional settings like schools, often through media and popular culture (Hayes & Gee, 2011) – supporting the expansion of the alt-right. Researchers connect those detrimental types of learning with the work of networks of "informal miseducation" (Greenhalgh *et al.*, 2021). Other work investigates how those networks operate by instrumentalizing pedagogical and discourse practices that imply transgression and rebelliousness against "the establishment," mainstream media, and academia (Tebaldi, 2021). Burdick and Sandlin (2022) identify pedagogical practices implemented by conspiracy QAnon that distort principles of critical pedagogy and scholarly dynamics of knowledge co-construction. That line of work demonstrates how the alt-right instrumentalized learning practices in pathways that lead to concrete threats to democracy, such as the involvement of QAnon in the Capitol riots and the implication of Brazilian YouTubers in the coup attempt in January 2023.

This paper adopts a conceptual framework that draws from public pedagogy and collective knowledge construction to understand how informal learning occurs in extremist online spaces. We frame the YouTube

comment section as a learning environment shaped by discursive and pedagogical practices that often produce misinformed knowledge, what we refer to as "twisted" learning. We begin by describing how we adapted Dubovi and Tabak's (2020) coding scheme to analyze a corpus of YouTube video comments as part of a pilot study to understand argument negotiation within this context. We then present our findings, highlighting that most comments related to knowledge construction are expressions of logical reasoning that result in misinformed interpretations and ideologically driven narratives. Finally, we discuss the implications of these findings, offering insights that could inform strategies to address social media-based extremism.

Methods

Data Collection

The video was selected for its unique characteristics, including the channel's anti-democratic content, which led to its removal during the Brasilia riots on January 8, 2023 (PODER360, 2023). Second, we opted to analyze comments from a video from that channel that had the stated purpose of *explaining* a scenario — the suspension of X/Twitter in Brazil — in a setting that resembles an "informal lecture," that is, a *learning setting*. The video claimed that Brazil's Supreme Court was acting beyond its authority by suspending X/Twitter. Also, the video highlights a rhetorical question, a discursive strategy that has been associated with the politically-motivated spread of misinformation in Brazil (Russo & Blikstein, 2023; Sotério, 2022): the video background features the question "Is he the law?" in reference to Supreme Court Justice Alexandre Moraes, who decided upon X/Twitter suspension in the country.

Using the YouTube v3 API, we downloaded 5,753 comments, focusing specifically on user-to-user interactions within discussion threads, as in Dubovi & Tabak (2020). We analyzed parent comments initiating threads and their nested replies. Parent comments serve as the starting point of a conversation, while nested comments are the replies to these comments, forming a thread. This filtering process narrowed the dataset to 615 parent comments with associated replies. From this subset, a random sample of 123 comments was chosen for analysis, totaling 21 parent comments, each ranging from 1 to 46 replies (M = 5). This sampling ensured feasibility for two analysts while maintaining the sample's representativeness.

Data Analysis

The data analysis followed Bardin's (2000) content analysis steps: pre-analysis, coding and categorization, and treatment and interpretation. During *pre-analysis*, the dataset was organized, anonymized, and assessed for suitability, with sensitive or additional information marked using brackets. A coding framework, adapted from Dubovi & Tabak (2020), was developed to categorize comments based on discourse moves — the ways comments respond to others, such as by agreeing, disagreeing, or questioning (Clark & Sampson, 2008; Lu, Chiu, & Law, 2011) — and collaborative knowledge construction — the process of building shared understanding, from expressing opinions to reaching consensus (Gunawardena, Lowe, & Anderson, 1997; Lucas, Gunawardena, & Moreira, 2014). The framework included the codes listed in Table 1.

Table 1

YouTube Comments Coding Scheme for Argument Negotiation (Adapted from Dubovi & Tabak, 2020) - continues

Categories		Description	Example
Evaluation	Disagree	Directly opposes/attack a previous comment	"We're lost with these corrupt senators and congressmen"
	Counterclaim	Offers a different view without attacking	"Why isn't Facebook being sanctioned by the US government, then?
	Agree	Explicitly supports a previous comment	"Nothing more to add… exactly that"
	Question	Requests clarification or more information	"The question is… How to reverse this?"
	Claim	A standalone comment not replying to others	"Printed vote and public counting now!!!!"
Grounds	Evidence	Uses facts or data to support a point	"The term in any Bar Association is three years"
	Explanation	Clarifies ideas through context or reasoning	"X didn't go down because the Armed Forces acted quickly"

Table 1
YouTube Comments Coding Scheme for Argument Negotiation (Adapted from Dubovi & Tabak, 2020) - final

Categories		Description	Example
Evidence validation	External source	Cites outside materials like articles or websites	"This newspaper just published an article titled…"
	Authority reference	Refers to recognized institutions or experts	"What's the legal basis... Brazilian law is being violated by those who swore to uphold it"
	Personal experience	Bases argument on lived experience	"Exactly, I was there on the highway. But most people chose to trust the military instead of using our own strength…"
Emotional expressions	Negative	Rude or hostile remarks	"That venomous, treacherous snake allows everything"
	Positive	Supportive or empathetic comments	"She must never give up. She always needs our support"

During the *coding and categorization phase*, two analysts independently assigned each comment, treated as a unit of analysis, to one of the predefined categories. Each comment was assigned only one primary code based on its dominant argumentative move. It should be noted that the New/Repetition category, which indicates whether an idea is new or repeated, initially proposed by Dubovi & Tabak (2020), was excluded due to its overlap with existing categories such as Claims and Evidence, which already captured the introduction and reinforcement of ideas. This decision aligned the coding framework with the data's characteristics.

In the *treatment and interpretation phase*, the coded data were quantified and analyzed to explore the interactional dynamics of argumentation present in the comments. The results were compared with findings from existing literature on the topic. To ensure the reliability of the coding process, Cohen's kappa coefficient (K) was calculated to measure inter-rater agreement using the SPSS software. This statistical method was appropriate because the analysis adhered to the necessary conditions, including mutually exclusive categories, paired observations, and consistent independent raters. Discrepancies between coders were resolved through discussion. The resulting K of 0.933 indicated the highest level of inter-rater agreement (Landis & Koch, 1977).

Results

Claims were the most frequent argument strategy (25.2%), often used to initiate threads or respond to the video creator's pinned comment (who did not reply to any comment). These responses suggest that users frequently perceived the creator as an authoritative figure rather than engaging peers. For instance, one user wrote, *"If I were Elon Musk, I'd create another platform just for the countries that banned X."*

In contrast, agreements (8.1%), disagreements (2.4%), counterclaims (8.1%), and questions (2.4%) were less frequent. The latter two, which typically indicate more profound argumentative exchange, suggest limited dialogical interaction. For example, one user challenged the narrative by stating, *"You [video creator] barely said anything critical about [a well-known far-right politician], did you?"* A reply to this comment added another counterclaim, *"There are more important topics than that. That's all."*

Grounds such as evidence (6.5%) and explanation (13%) reflected more elaborate reasoning, though evidence validation (external sources, authority, or personal experience) was rare (1.6% each), highlighting the opinion-driven nature of the discourse. These forms of more complex reasoning often appeared in replies. One user asked, *"Can I file a complaint with PROCON [*Consumer Protection Agency*] against the STF [Supreme Court]?"*, to which another replied, *"PROCON only handles consumer issues. Filing such a complaint wouldn't be constitutional."* This exchange illustrates how users sometimes followed logical reasoning but ultimately reached flawed conclusions, such as treating government actions as consumer complaints.

Emotional expressions were common, especially positive ones (18.7%). For instance, users expressed support with comments like *"We're here for you, dear,"* while negative emotional reactions appeared less frequently (8.1%), such as in *"The spell is turning against the donkeys."*

Discussion and implications

This study analyzes argumentation in far-right YouTube comments, revealing three key findings: (1) Users often construct distorted knowledge that fuels misinformation and polarization; (2) They primarily position themselves as consumers rather than civic participants; and (3) Despite these dynamics, some interaction patterns suggest the

platform might still afford opportunities for learning.

First, *users construct knowledge, but a "twisted" version of it.* Our findings show that while YouTube users engaging in right-wing discourses attempt at logical reasoning, their knowledge construction often leads to misinformed conclusions and ideological polarization. For example, users may start with reasonable assumptions, such as recognizing PROCON as a consumer protection agency, but would draw the flawed conclusion that it could address government policy complaints (as if government policies were "products"). This surface-level reasoning relies on shortcut solutions (Walton, 2010), perpetuating misconceptions that right-wing parties exploit to reinforce ideologies, undermine institutions, and advance anti-democratic agendas (McIntyre, 2018).

The perpetuation of such twisted reasoning can be explained by psychological factors that lead to biased information processing. Individuals tend to favor information that aligns with their preexisting beliefs while avoiding worldview-challenging content to reduce cognitive dissonance (Scheufele & Krause, 2019). This helps explain the high frequency of positive emotional expressions (18.7%) and agreements (8.1%) observed in our findings, reflecting alignment with the influencer and reinforcing shared beliefs. As such, twisted knowledge construction is an expected outcome of these online discussions, since conspiracy movements intentionally manipulate pedagogical practices to distort knowledge co-construction (Burdick & Sandlin, 2022). Addressing this issue requires more than spreading corrective information; it demands integrating media literacy with targeted interventions that address the emotional and cognitive factors driving these misconceptions (Asterhan & Schwarz, 2009). Future research could distinguish between argumentation format and content, as findings suggest that logically structured discourse may still convey misinformation.

Moreover, *they may act more as consumers of products and services than as citizens*. This is evidenced by the prominence of claims (25.2%), most often directed at the influencer's pinned post. This suggests that users primarily use the comments section to engage with the content creator (offering feedback, requesting content, and showing loyalty) rather than to negotiate ideas with peers. While this marks a shift from passive information consumption to surface-level participation (Dubovi & Tabak, 2020; Gunawardena, Lowe, & Anderson, 1997), it also indicates that users perceive their involvement more as an individual consumer-seller interaction than as collective, informed civic discourse. This behavior may also reflect parasocial dynamics, where users form one-sided emotional bonds with the content creator (Paravati *et al.*, 2020), shaping how they process information and reinforce shared ideologies. Moreover, the user–influencer dynamic may affect argument quality by reinforcing power asymmetries. Future research could examine how this dynamic may shape online political argumentation.

This consumer-oriented behavior is reflected in references to agencies like PROCON and suggestions for topics the influencer should address, framing content creation as a product tailored to user preferences. It aligns with a type of framing of civic engagement on social media where users act more as consumers than active citizens (Couldry & Mejias, 2019). Additionally, this dynamic underscores — despite possible claims that online environments are more horizontal — the considerable power and influence of the "teacher" figure (the influencer) in shaping understandings that emerge from these interactions (Fischer et al., 2018).

While claims dominate the discourse, signs of deeper argumentation and inquiry processes are also present — *often associated with productive sense-making and learning opportunities*. Commenters offered grounds (19.5%), counterclaims (8.1%), evidence validation (4.8%), and questions or disagreements (2.4% each), all of which suggest potential for more meaningful engagement (Dubovi & Tabak, 2020; Gunawardena, Lowe, & Anderson, 1997). Notably, most explanations arose during non-consensual negotiations, which are strong predictors of dialectical argumentation, collective reasoning, and conceptual learning gains (Asterhan & Schwarz, 2009; Dubovi & Tabak, 2020).

Research suggests that while claims are typically the most prominent discourse moves, transitioning to deeper argumentation often requires external support or facilitation (Asterhan & Schwarz, 2016; Fischer et al., 2018). This highlights a significant opportunity within YouTube comment sections (and other similar environments) to foster meaningful learning and collective reasoning, allowing these spaces to transcend traditional hierarchical teaching models and promote democratic values. This could be supported by rethinking platform design and online community norms, as well as integrating civic education and social media literacy into curricula to help users identify misinformation, evaluate sources, and engage in critical discourse.

It should be noted that this study does not aim to make broad generalizations but focuses on a specific (but popular) YouTube video to provide insights into knowledge construction in polarized social media. Future research could explore how linguistic features and framing strategies spread "twisted" knowledge and whether these tools can promote fact-based reasoning, deeper argumentation, and democratic action. Such efforts could help leverage social media to foster meaningful discourse and knowledge co-construction.

References

Asterhan, C. S. C., & Schwarz, B. B. (2009). Argumentation and explanation in conceptual change: Indications

from protocol analyses of peer-to-peer dialog. *Cognitive Science*, 33(3), 374–400.
Asterhan, C. S. C., & Hever, R. (2015). Learning from reading argumentive group discussions in Facebook: Rhetoric style matters (again). *Computers in Human Behavior*, *53*, 570–576.
Asterhan, C. S. C., & Schwarz, B. B. (2016). Argumentation for learning: Well-trodden paths and unexplored territories. Educational Psychologist, 51(2), 164–187.
Bardin, L. (2000). *Análise de conteúdo*. Edições 70.
Burdick, J., & Sandlin, J. A. (2022). Into the UpsideDown: Conspiracy Pedagogies, Critique, and the Apocalypse of Exposure. *Journal of Public Pedagogies*, *6*, 9–23. https://doi.org/10.15209/jpp.1243
Clark, D. B., & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching*, 45(3), 293–321.
Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
Dubovi, I., & Tabak, I. (2020). An empirical analysis of knowledge co-construction in YouTube comments. *Computers & Education*, *156*, 103939. https://doi.org/10.1016/j.compedu.2020.103939
Fischer, F., Chinn, C., Engelmann, K., & Osborne, J. (Eds.). (2018). *Scientific reasoning and argumentation*. Routledge.
Gleason, B. (2013). #Occupy Wall Street: Exploring Informal Learning About a Social Movement on Twitter. *American Behavioral Scientist*, *57*(7), 966–982.
Gleason, B. (2016). Teens' Participatory Play: Digital Media Learning through Civic Engagement. In C. Greenhow, J. Sonnevend, & C. Agur (Eds.), *Education and Social Media: Toward a Digital Future* (pp. 231–238). MIT Press.
Gleason, B., He, L., & Karakaya, O. (2021). The intersection of new literacies, academic identities, and critical scholarship: Emerging learning practices of graduate students on Twitter. *International Journal of Smart Technology and Learning*, *2*(4), 250–271. https://doi.org/10.1504/IJSMARTTL.2021.118902
Greenhalgh, S., Krutka, D., & Oltmann, S. (2021). Gab, Parler, and (Mis)educational Technologies: Reconsidering Informal Learning on Social Media Platforms. *Journal of Applied Instructional Design*, *10*(3).
Greenhow, C., & Robelia, B. (2009). Old Communication, New Literacies: Social Network Sites as Social Learning Resources. *Journal of Computer-Mediated Communication*, *14*(4), 1130–1161.
Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, 17(4), 397–431.
Haroon, M., Wojcieszak, M., Chhabra, A., Liu, X., Mohapatra, P., & Shafiq, Z. (2023). Auditing YouTube's recommendation system for ideologically congenial, extreme, and problematic recommendations. *Proceedings of the National Academy of Sciences*, *120*(50), e2213020120.
Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.
Lucas, M., Gunawardena, C., & Moreira, A. (2014). Assessing social construction of knowledge online: A critique of the interaction analysis model. *Computers in Human Behavior*, 30, 574–582.
Nguyen, H., & Diederich, M. (2023). Facilitating knowledge construction in informal learning: A study of TikTok scientific, educational videos. *Computers & Education*, *205*, 104896.
McIntyre, L. (2018). Post-truth. MIT Press.
Tebaldi, C. (2021). Speaking post-truth to power. *Review of Education, Pedagogy, and Cultural Studies*, *43*(3), 205–225. https://doi.org/10.1080/10714413.2020.1729679
Paravati, E., Naidu, E., Gabriel, S., & Wiedemann, C. (2020). More than just a tweet: The unconscious impact of forming parasocial relationships through social media. *Psychology of Consciousness: Theory, Research, and Practice*, 7(4), 388–403.
PODER360. (2023, January 5). Moraes suspende redes sociais de jornalistas bolsonaristas. Poder360. https://www.poder360.com.br/midia/moraes-suspende-redes-sociais-de-jornalistas-bolsonaristas/.
Russo, R., & Blikstein, P. (2023). Just asking questions: Can a far-right president turn agentic knowledge construction into political manipulation? *Information and Learning Sciences*, 124(7/8), 197–220.
Scheufele, D. A., & Krause, N. M. (2019). Science audiences, misinformation, and fake news. *Proceedings of the National Academy of Sciences*, 116(16), 7662-7669.
Sotério, C. (2022). The provisional nature of science evidenced in times of pandemic. *Alternautas*, 8(1).
Walton, D. (2010). Why fallacies appear to be better arguments than they are. *Informal logic*, 30(2), 159-184.