What Makes A Video Radicalizing? Identifying Sources of Influence in Extremist Videos

Lin Ai, Yu-Wen Chen, Seoyoung Kweon, Yuwen Yu, Julia Hirschberg, Sarah Ita Levitan

2025

Goals of Our Research

- Increase our understanding of online radicalization efforts by analyzing right- and left-leaning groups' videos in social media
    - What sort of viewers do these videos appeal to
    - What methods are employed in producing more persuasive videos
    - How can we use multimodality features to identify radical videos computationally
    - How might new, potentially influential extremist groups be identified
- Identify potentially useful de-radicalizing videos

Related Work

- Radicalization detection in social media using **textual features**
  - [Hartung et al. (2017)](#) attempt to identify right-wing extremist content in German Twitter profiles
  - [Hofmann et al. (2022)](#) leverage network structure of Reddit forums to detect polarized concepts
  - [Lopez-Sancez et al. (2018)](#) and [Araque and Iglesias (2020)](#) develop methods to identify radicalizing content in Twitter

Related Work

- Radicalization detection in **multimodal content**
    - [Bermingham et al. (2009)](#) detect radicalization in Jihadist YouTube videos using social network analysis and sentiment
    - [Ribeiro et al. (2020)](#) collect 330,925 YouTube videos to identify radicalizing pipelines for far-right groups
    - [Ai et al. (2021)](#) identify multimodal features of far-right and far-left groups that makes them more popular and more persuasive

Prior Work

- Collected a large number of videos for 5 extremist groups from YouTube and BitChute and extracted metadata and multimodal features
- Used # comments and likes to identify video popularity changes over time
- Trained stance detection models to identify stance (pro-, anti-group, or neutral) of a given video using video title and description

Group	QAnon	Antifa	Proud Boys	Oath Keepers	BLM
Total Videos	5,924	17,242	1,326	589	6,592
Comments	478,358	1,685,374	824,132	82,817	852,861
Avg Likes	255.88	575.82	7,051	790	23,217
Avg Dislikes	50.9	41.23	589	86	1,433
Avg Views	13,967	24,722	207,777	45,576	893,561

QAnon – Background

- QAnon
  - Originated in 2017
  - One of the prime conspiracy-based radicalization groups [1, 2]
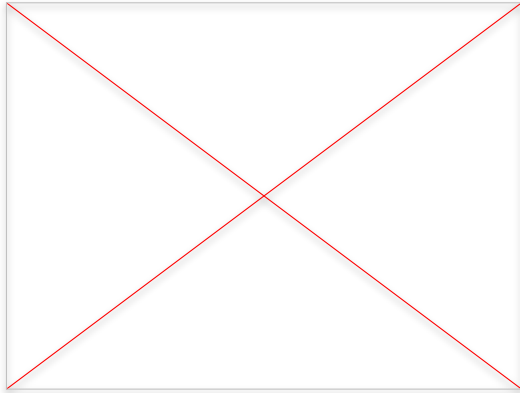
[1] Amarasingam, Amarnath, and Marc-André Argentino. "The QAnon conspiracy theory: A security threat in the making." *CTC Sentinel* 13.7 (2020): 37-44.
[2] Garry, Amanda, et al. "QAnon conspiracy theory: examining its evolution and mechanisms of radicalization." *Journal for Deradicalization* 26 (2021): 152-216.
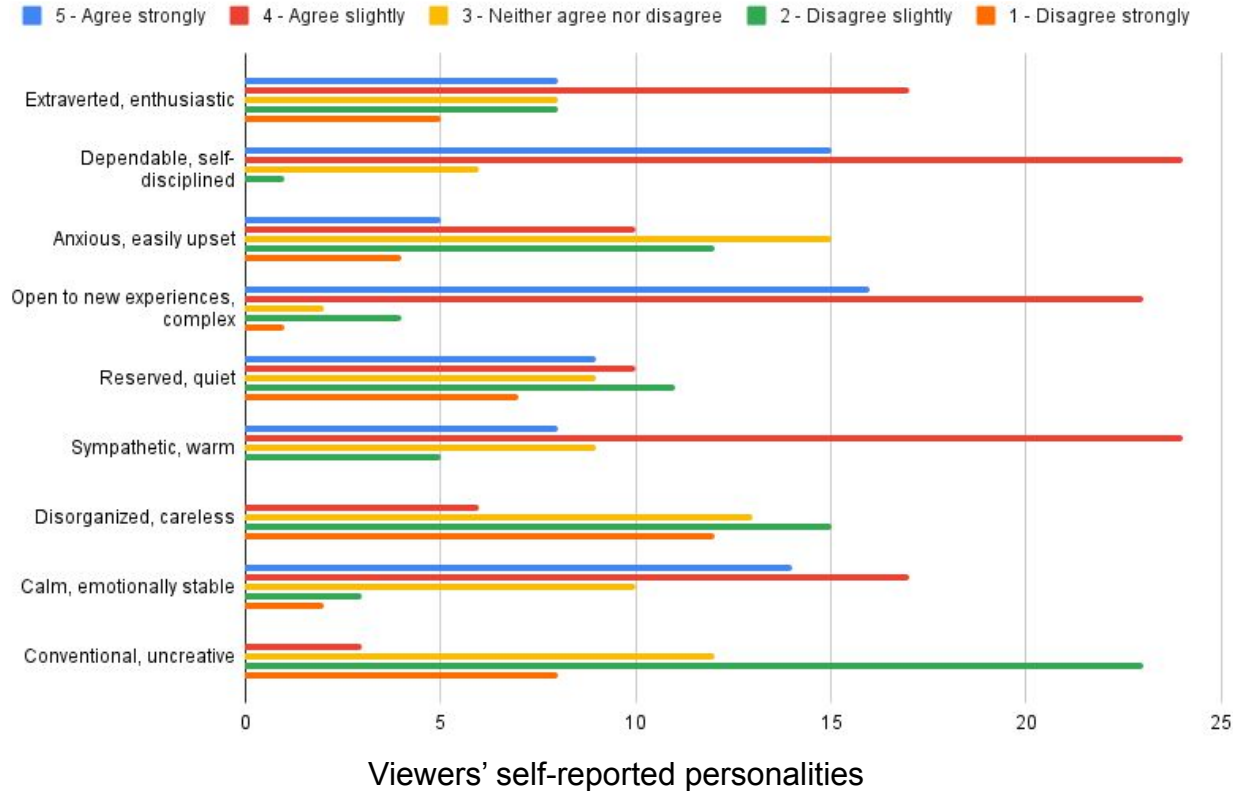
QAnon – Questionnaire

- Little study has computationally analyzed QAnon related videos
  - How these videos drag viewers into the process of radicalization?
  - Who the videos are the most appealing to?
- We designed a comprehensive questionnaire to answer three research questions:
  - **RQ1:** What viewer traits, such as personality traits and media consumption, are associated with their video preferences?
  - **RQ2:** What video characteristics, such as speaker traits, video quality, and arousing emotions, are correlated with viewers' perception?
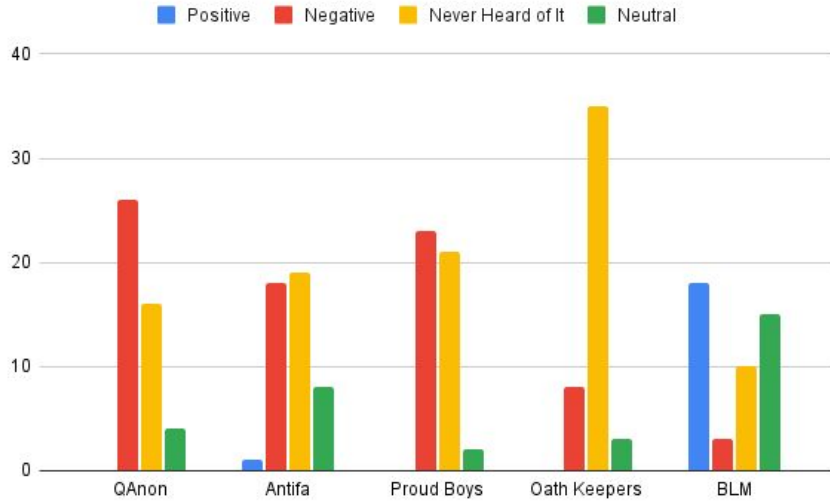 - **RQ3:** What multimodal features are correlated with viewers' perception?

QAnon – Viewer Traits

- We selected 3 pro- and 3 anti-QAnon videos
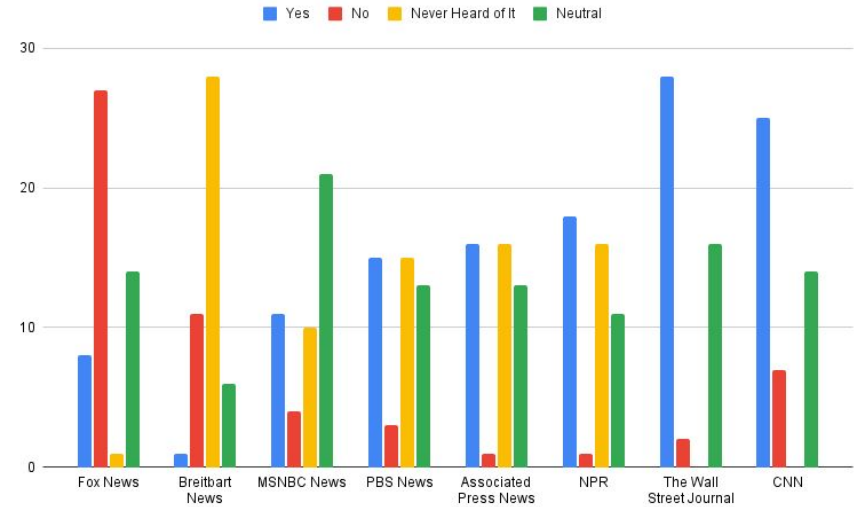- We recruited 46 participants, primarily college students

QAnon – Viewer Traits

Viewers' self-reported personalities
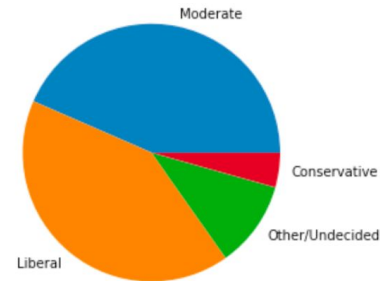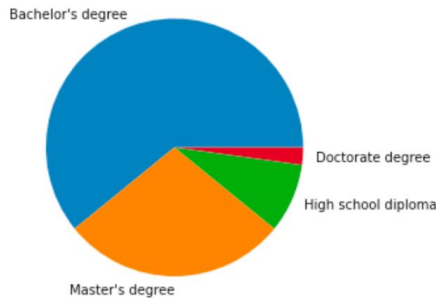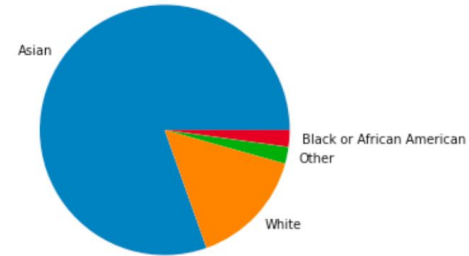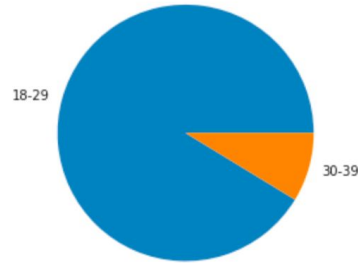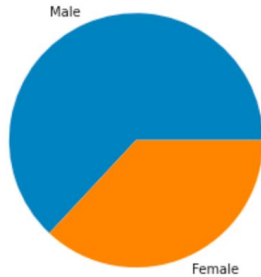
QAnon – Viewer Traits

Viewers' opinion on radical groups

Viewers' opinion on media sources

QAnon – Questions

1. Did you understand the video?

☐ Yes
☐ No

2. Do you think the video was professionally produced with good quality?

☐ Yes
☐ No

3. Who do you think the video was trying to appeal to?: _____

4. Was there any violence displayed in the video?

☐ Yes
☐ No

5. Was there any music in video?

☐ Yes
☐ No

6. Did any of the following objects appear in the video? Choose all that apply.

☐ Guns
☐ Swords
☐ Other Weapons
☐ Flags
☐ Symbols of the Group
☐ None of the Above

7. How likely do you think it is that the people in the video will become involved in the following actions?

	Not at All Likely	Not Much Likely	Undecided	Somewhat Likely	Very Much Likely
Protests	☐	☐	☐	☐	☐
Violence	☐	☐	☐	☐	☐
Illegal Acts	☐	☐	☐	☐	☐

8. Would you associate the following traits with this video?

	Yes	Neutral	No
Boring (could you pay attention the whole time or not?)	☐	☐	☐
Lively (was it energetic? Ex. was there music?)	☐	☐	☐
Persuasive (were you convinced by the content?)	☐	☐	☐
Trustworthy (did you trust the content?)	☐	☐	☐
Logical (was there a structured argument or data presented?)	☐	☐	☐

9. Is the video's stance positive, negative, or neutral towards the group?

☐ Positive
☐ Negative
☐ Neutral

10. Do you think this speaker demonstrated any of the following characteristics?

	Yes	Neutral	No
Charismatic	☐	☐	☐
Confident	☐	☐	☐
Eloquent	☐	☐	☐
Enthusiastic	☐	☐	☐
Intelligent	☐	☐	☐
Convincing	☐	☐	☐
Tough	☐	☐	☐
Charming	☐	☐	☐
Angry	☐	☐	☐

11. Did you enjoy watching the video?

	1	2	3	4	5	
Not at All	☐	☐	☐	☐	☐	Very Much

12. What emotions did you feel when you watched the video? Check all that apply.

☐ Happiness
☐ Sadness
☐ Surprise
☐ Fear
☐ Disgust
☐ Anger
☐ Confused

13. Which part of the video was most impactful? (Give the approximate timestamps.) Enter N/A if not applicable.: _____

14. Give a short description (a sentence) of the most impactful part of the video you listed above. Enter N/A if not applicable. : _____

15. Do you think any of the content in the video makes a valid point?

	1	2	3	4	5	
Not at All	☐	☐	☐	☐	☐	Very Much

16. Would you take any of the following actions after watching this video? Check all that apply.

☐ Like the video
☐ Dislike the video
☐ Post a supporting comment under the video
☐ Post a criticizing comment under the video
☐ Share the video with friends, families, or on social media platforms
☐ Search for similar videos
☐ Learn more about the group
☐ Consider joining the group
☐ Non of the Above

QAnon – Evaluation Metrics

- **Enjoyment Score**
  - How much viewers enjoy watching each video
- **Content Score**
  - Whether viewers think a video is persuasive, trustworthy, logical, and professionally created
- **Actions Score**
  - posting a criticizing comment [score -2]
  - disliking the video [score -1]
  - liking the video [score 1]
  - posting a supporting comment [score 2]
  - considering joining the group [score 3]

QAnon – Significant Viewer Traits (RQ1)

Enjoyment on Pro-QAnon Videos			Enjoyment on Anti-QAnon Videos		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Opinion_CNN	0.329	2.55E-02	Opinion_Antifa	0.368	1.19E-02
Opinion_WSJ	0.298	4.40E-02			
Content of Pro-QAnon Videos			**Content of Anti-QAnon Videos**		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Opinion_Fox	0.487	5.92E-04	Researved	0.339	2.13E-02
Opinion_NPR	-0.376	1.00E-02			
Opinion_AP	-0.33	2.53E-02			
Actions after Pro-QAnon Videos			**Actions after Anti-QAnon Videos**		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Opinion_OathKeepers	0.37	1.14E-02	Disorganized	0.318	3.12E-02
Opinion_Fox	0.358	1.45E-02	Sympathetic	-0.317	3.21E-02
Opinion_CNN	0.298	4.42E-02			

QAnon – Significant Video Characteristics (RQ2)

Enjoyment on Pro-QAnon Videos			Enjoyment on Anti-QAnon Videos		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Validness	0.999	2.34E-02			
Persuasive	0.997	4.52E-02			

Content of Pro-QAnon Videos			Content of Anti-QAnon Videos		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
			Disgust	-0.998	4.40E-02
			Boring	-0.998	4.40E-02

Actions after Pro-QAnon Videos			Actions after Anti-QAnon Videos		
Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
			Trustworthy	1	1.14E-02
			Lively	-1	1.67E-02

QAnon – Multimodal Features

- Textual features: LIWC [3], Grievance Dictionary [4], VADER [5]
- Acoustic features
    - OpenSmile features (pitch, intensity, etc)
    - SpeechBrain's emotion detection model [6]
- Visual features
    - Pre-trained FER (facial expression recognition) model [7]
    - Clarifai's weapon detector model [8]
- Data pre-processing
    - Transcripts extracted using the Google Speech-to-text service [9]
    - IPU level segmentation

[3] Pennebaker, James W., et al. *The development and psychometric properties of LIWC2015*. 2015.
[4] van der Vegt, Isabelle, et al. "The grievance dictionary: Understanding threatening language use." *Behavior research methods* 53.5 (2021): 2105-2119.
[5] Hutto, Clayton, and Eric Gilbert. "Vader: A parsimonious rule-based model for sentiment analysis of social media text." *Proceedings of the international AAAI conference on web and social media*. Vol. 8. No. 1. 2014.
[6] Ravanelli, Mirco, et al. "SpeechBrain: A general-purpose speech toolkit." *arXiv preprint arXiv:2106.04624* (2021).
[7] Facial-Expression-Recognition.Pytorch
[8] Clarifai weapon detector
[9] Google Speech-To-Text

QAnon – Significant Multimodal Features (RQ3)

	Enjoyment on Pro-QAnon Videos			Enjoyment on Anti-QAnon Videos		
	Feature	Corr	p-value	Feature	Corr	p-value
Textual	sentiment	0.205	3.37E-02	interrog	-0.165	3.95E-02
	weaponry	-0.387	3.84E-05			
	violence	-0.324	6.71E-03			
Acoustic	HNR	0.87	5.67E-34	max intensity	-0.832	3.02E-41
	mean pitch	0.738	1.26E-19	mean intensity	-0.928	9.96E-41
	mean intensity	-0.713	7.18E-18	max pitch	-0.336	1.84E-05
Visual	happy	0.259	1.05E-05	fear	-0.23	1.54E-04
	sad	-0.226	1.27E-04	surprise	-0.169	5.79E-03
	has_weapon	-0.243	1.33E-04	disgust	-0.156	1.08E-02

QAnon – Significant Multimodal Features (RQ3)

	Content of Pro-QAnon Videos			Content of Anti-QAnon Videos		
	Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Textual	weaponry	-0.387	3.84E-05	time	-0.285	3.16E-04
	power	-0.328	5.54E-04	friend	0.213	7.65E-03
	violence	-0.324	6.71E-04	female	0.164	4.09E-02
Acoustic	HNR	0.87	5.76E-34	min intensity	0.676	3.58E-22
	mean pitch	0.738	1.26E-19	HNR	-0.311	7.67E-05
	mean intensity	-0.713	7.18E-18	mean intensity	0.179	2.56E-02
Visual	happy	0.259	1.09E-05	angry	0.482	7.06E-17
	sad	-0.226	1.29E-04	fear	-0.123	4.51E-02
	has weapon	-0.243	1.36E-04	has weapon	0.141	2.13E-02

QAnon – Significant Multimodal Features (RQ3)

	Actions after Pro-QAnon Videos			Actions after Anti-QAnon Videos		
	Feature	**Corr**	**p-value**	**Feature**	**Corr**	**p-value**
Textual	weaponry	-0.387	3.84E-05	time	-0.266	7.95E-04
	power	-0.328	5.54E-04	friend	0.195	1.46E-02
	violence	-0.324	6.71E-04	insight	-0.161	4.42E-02
Acoustic	HNR	0.87	5.76E-34	min intensity	0.687	3.77E-23
	mean pitch	0.738	1.26E-19	HNR	-0.372	1.74E-06
	mean intensity	-0.713	7.18E-18	max intensity	-0.164	4.09E-02
Visual	happy	0.268	5.20E-06	angry	0.429	2.43E-13
	has weapon	-0.251	8.16E-05	fear	-0.179	3.35E-03
	sad	-0.231	8.92E-05	has weapon	0.125	4.24E-02

QAnon – Conclusions

- Collected a corpus of videos for 5 extremist groups
- Designed a comprehensive questionnaire
- Answered three questions
    - **RQ1:** What viewer traits are associated with their video preferences?
        - Viewers prefer pro- videos: positive impression to right-leaning news/groups
        - Viewers prefer anti- videos: positive impression to left-learning groups, reserved
    - **RQ2:** What video characteristics are correlated with viewers' perception?
        - Validness, persuasiveness, trustworthiness have positive impact to viewers' perception
    - **RQ3:** What multimodal features are correlated with viewers' perception?
        - Pro- videos: weapons and violence, higher intensity and sad speaker faces have negative impact to viewers' perception
        - Anti- videos: friend and weapon and angry speaker faces have positive impact
- Accepted at IC^{2}S^2 2023!

Ongoing Work

- Define techniques/strategies being used by extremist groups in different phases of radicalization
- Annotate videos with these techniques/strategies to score the radical level of the videos
- Define and annotate strategies of de-radicalizing videos
    - Former group members detaching from the group
    - Intra-group conflicts

Thank you!
Questions?