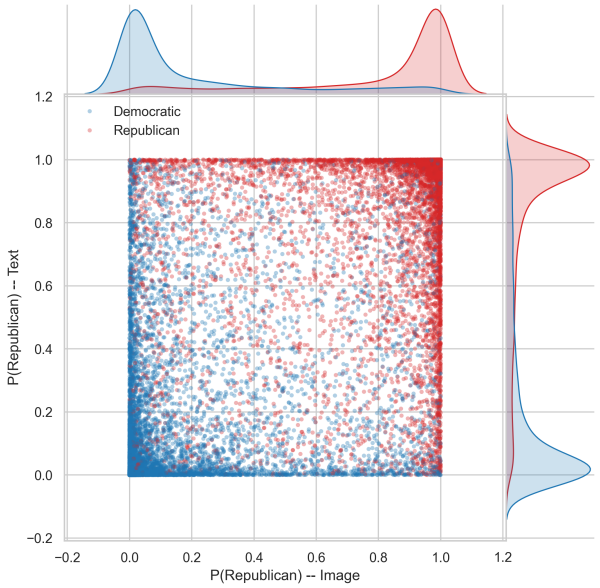
# An Image is Worth 1000 Words? Partisanship and Attention in Videos

Supplemental Appendix

Andrea Ciccarone
ETH Zürich

# A Additional tables and figures

**Figure A1:** Predicted Republican Probability: Image vs. Text Models




*Notes:* Each point represents one political ad. Horizontal and vertical axes show the predicted Republican probability from the image and text models, respectively. Points are colored by ground-truth party label.

**Table A1:** Emotion Vocabulary

| Category | Tokens |
|---|---|
| Admiration | admiration, respect |
| Adoration | adoration, devotion |
| Aesthetic Appreciation | aesthetic, beauty, gorgeous |
| Amusement | amusement, humorous, funny |
| Anger | anger, enraged, furious |
| Anxiety | anxiety, uneasy, worried |
| Awe | awe, awestruck, wonder |
| Boredom | boredom, bored, uninterested |
| Calm | calm, serene, tranquil |
| Confusion | confusion, confused, perplexed |
| Contempt | contempt, scorn, disdain |
| Contentment | contentment, satisfied, fulfilled |
| Craving | craving, desire, yearning |
| Disgust | disgust, repulsed, nauseated |
| Empathic Pain | empathy, sympathy, sorrow |
| Entrancement | entranced, captivated, mesmerized |
| Excitement | excitement, exhilarated, thrilled |
| Fear | fear, terrified, frightened |
| Gratitude | gratitude, thankful, appreciative |
| Guilt | guilt, guilty, remorse |
| Horror | horror, horrified, appalled |
| Interest | interest, intrigued, curious |
| Joy | joy, joyful, ecstatic |
| Nostalgia | nostalgia, nostalgic, wistful |
| Pride | pride, proud, triumphant |
| Relief | relief, relieved, reassured |
| Romantic Love | love, loving, romantic |

*Notes:* Emotion categories are adapted from the taxonomy of Cowen and Keltner (2017). Tokens are the natural-language entries used to represent each category in the joint CLIP vocabulary; the same dictionary is used for images and text. These categories are distinct from the broader emotion buckets used to summarize CTEs.

**Table A2:** Topic Vocabulary

| Category | Tokens |
|---|---|
| Economy/Jobs | economy, jobs/unemployment, outsourcing, minimum wage, manufacturing/construction, trade, trade: china, financial services, financial reform, retirement, union |
| Education | education |
| Environment | energy/environment, oil, oil-anti, green energy, global warming, coal-pro, coal, oil-pro |
| Foreign/Defense | international affairs, china, national defense, defense/aerospace, iraq/afghan war, veterans affairs |
| Governance | corruption, supreme court, term limits |
| Healthcare | healthcare, medicare, pro-aca, anti-insurance, anti-obama plan, anti-ahca, anti-aca, health insurance reform, prescription drugs, prescription drugs: cost, prescription drugs: anti-industry, coronavirus |
| Housing | housing/home ownership |
| Immigration | immigration, immigration: anti, immigration: pro |
| Infrastructure | transportation, telecommunications |
| Interest Groups | food/agriculture, aarp: 50+ voters, aarp mention |
| Political Campaigns | anti-trump, pro-trump, anti-biden, pro-biden, anti-clinton, pro-clinton, anti-obama message, pro-obama message, anti-sanders, pro-sanders, campaign finance reform, call to action, impeachment |
| Public Safety | public safety, gun control, anti-gun control, pro-gun control, terrorism |
| Social Issues | social issues, abortion, women's rights, drugs, civil rights, opioids/opiates, faith/religion, guns, birth control, human rights |
| Tax/Budget | taxes, tax reform, budget/government spending, social security |

*Notes:* Topic tokens derive from issue metadata in the political-ad training corpus. Issues appearing in at least 1% of advertisements enter the vocabulary. The categories organize these terms for interpretation; image and text representations use the same dictionary.

**Table A3:** Concept Aggregation for the CTE Analysis

*Panel A. Emotions*

| Aggregation group | Constituent concepts |
|---|---|
| Anger/Disgust | anger; enraged; furious; disgust; repulsed; nauseated; contempt; scorn; disdain |
| Fear/Anxiety | fear; terrified; frightened; anxiety; uneasy; worried |
| Sadness/Empathy | empathy; sympathy; sorrow |
| Calm/Relief | calm; serene; tranquil; relief; relieved; reassured |
| Joy/Amusement | joy; joyful; ecstatic; amusement; humorous; funny |
| Pride/Confidence | pride; proud; triumphant |
| Awe/Inspiration | awe; awestruck; wonder |
| Nostalgia/Tradition | nostalgia; nostalgic; wistful |
| Desire/Motivation | craving; desire; yearning |
| Love/Warmth | love; loving; romantic; adoration; devotion |
| Curiosity/Interest | interest; intrigued; curious |
| Gratitude/Hope | gratitude; thankful; appreciative; admiration; respect |
| Confusion/Fatigue | confusion; confused; perplexed; boredom; bored; uninterested |
| Horror/Shock | horror; horrified; appalled; guilt; guilty; remorse; entranced; captivated; mesmerized; excitement; exhilarated; thrilled; aesthetic; beauty; gorgeous; contentment; satisfied; fulfilled |

*Notes:* The table reports the concept-to-group mapping used in Figure 11. The same mapping is applied to both models. Group-level effects are weighted averages of individual concept effects, with weights equal to the number of videos affected by each concept's removal.

**Table A3:** Concept Aggregation for the CTE Analysis (continued)

*Panel B. Topics*

| Aggregation group | Constituent concepts |
|---|---|
| Healthcare | healthcare; healthcare: medicare; healthcare: support for ACA; healthcare: opposition to insurance companies; healthcare: opposition to Obama plan; healthcare: opposition to AHCA; healthcare: opposition to ACA; healthcare: insurance reform; prescription drugs; prescription drugs: cost; prescription drugs: opposition to industry; COVID-19 pandemic |
| Economy/Jobs | economy; jobs and unemployment; jobs and unemployment: outsourcing; manufacturing and construction; minimum wage; trade; trade: China; financial services; financial reform; retirement; labor unions |
| Tax/Budget | tax policy; tax reform; government budget and spending; social security |
| Social Issues | social issues; social issues: abortion; social issues: women's rights; social issues: drugs; social issues: civil rights; social issues: opioids; social issues: birth control; faith and religion; guns: support for gun control; guns: opposition to gun control; gun control; guns: general gun messaging; human rights |
| Public Safety | public safety; terrorism; defense and aerospace |
| Environment/Energy | energy and environment; energy and environment: oil; energy and environment: opposition to oil; energy and environment: green energy; energy and environment: global warming; energy and environment: coal; energy and environment: support for oil; energy and environment: support for coal |
| Immigration | immigration; immigration: opposition; immigration: support |
| Foreign/Defense | international affairs; international affairs: China; national defense; Iraq and Afghanistan wars; veterans affairs |
| Political Campaigns | anti Trump messaging; pro Trump messaging; anti Biden messaging; pro Biden messaging; anti Clinton messaging; pro Clinton messaging; anti Obama messaging; pro Obama messaging; anti Sanders messaging; pro Sanders messaging; calls to action; campaign finance reform; impeachment |
| Governance | political corruption; supreme court; term limits |
| Education | education |
| Housing | housing and home ownership |
| Infrastructure | transportation; telecommunications |
| Agriculture/Int. Groups | food and agriculture; AARP: 50+ voters; AARP mention |

*Notes:* The table reports the concept-to-group mapping used in Figure 11. The same mapping is applied to both models. Group-level effects are weighted averages of individual concept effects, with weights equal to the number of videos affected by each concept's removal.

**Figure A2:** Distribution of non-zero concepts across image and text modalities, fixing similarity across modalities

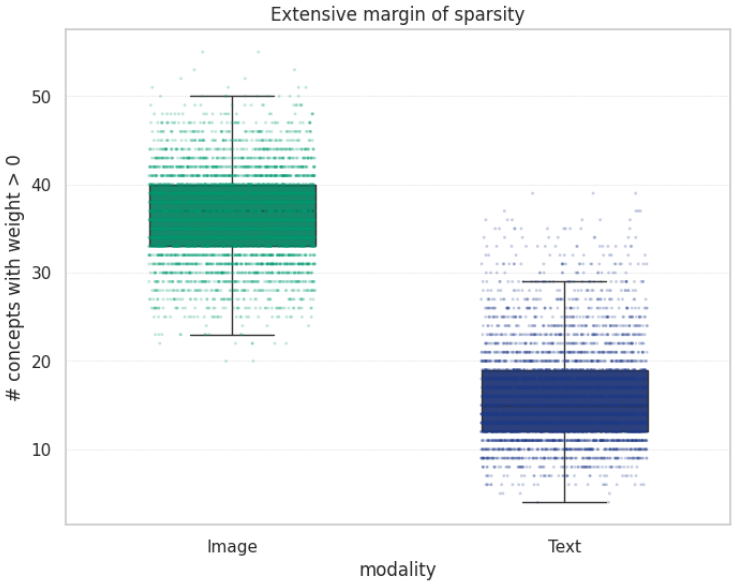


**Figure A3:** Cumulative distribution of the number of active concepts with nonzero weight per video.

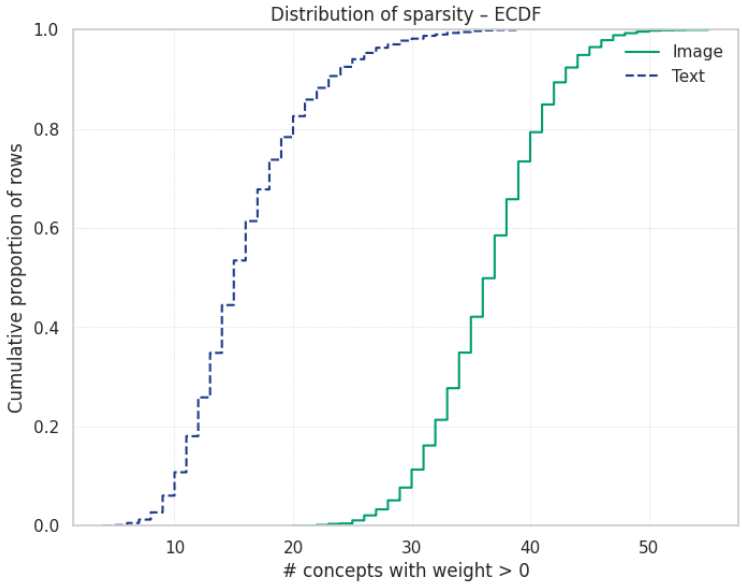
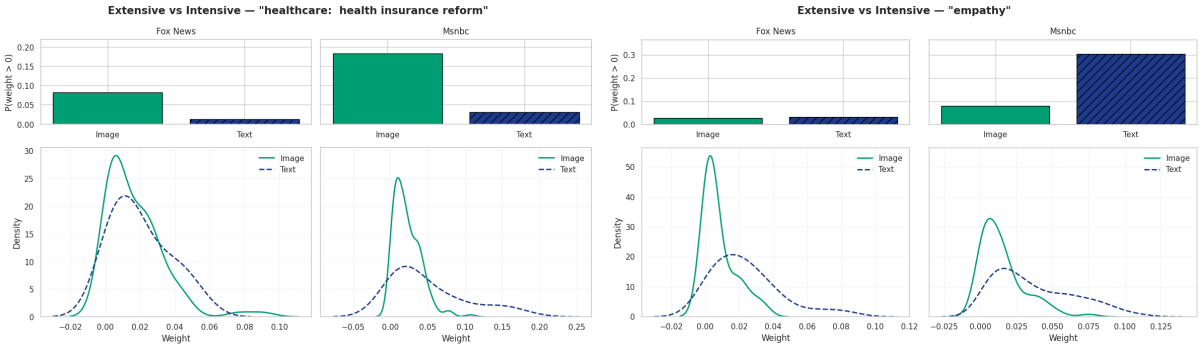
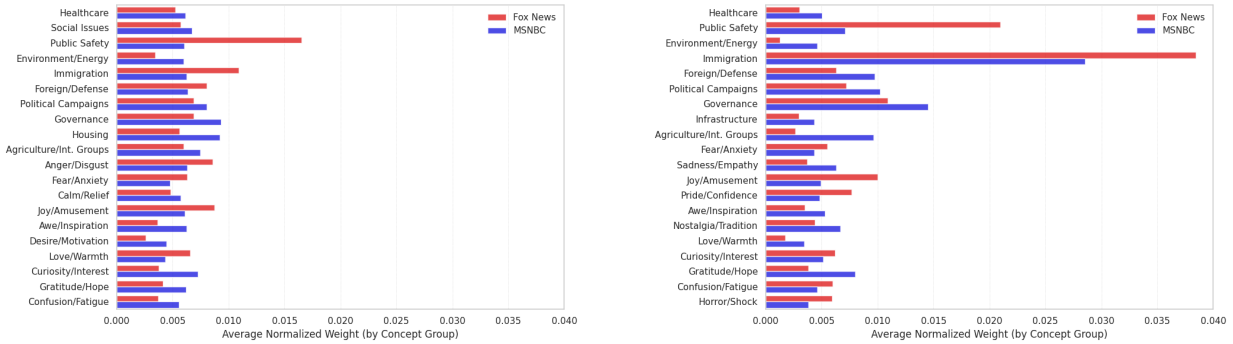


**Figure A4:** Extensive and intensive margins for two example concepts: Health Insurance Reform and Empathy.




*Notes:* The top panels report the share of news videos in which the concept has a strictly positive weight in the sparse decomposition (extensive margin). The bottom panels plot the distribution of weights when the concept is present (intensive margin), separately for image and text modalities and by channel.

**Figure A5:** Average Concept Weights in Image and Text Representations of YouTube News Videos

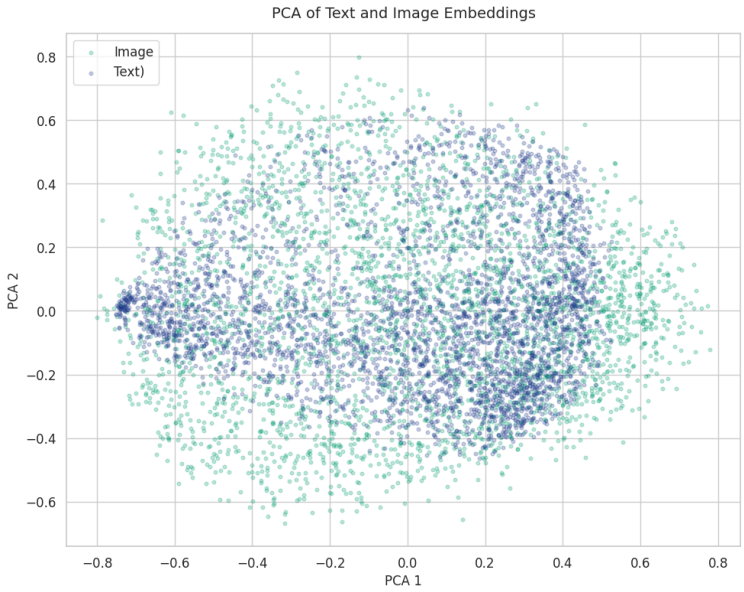



**(a)** Average Weights in Images

**(b)** Average Weights in Text

*Notes:* Each bar reports the average normalized weight of concepts in the joint SpLICE representation, aggregated at the group level (emotion or topic) and averaged across videos from Fox News (red) and MSNBC (blue). Buckets are ordered according to their position in the joint vocabulary. Weights represent the share of each modality's representation explained by a given conceptual group.

**Figure A6:** Principal component structure of image and text embeddings across all videos.




*Notes:* The figure shows the first two principal components of the original image and text embeddings for all videos in the dataset. The lack of visible clustering across modalities is expected given the mean-centering of embeddings. This projection provides a reference for the movement observed in Figure 10.

**Table A4:** TV News Sample Summary (All Channels)

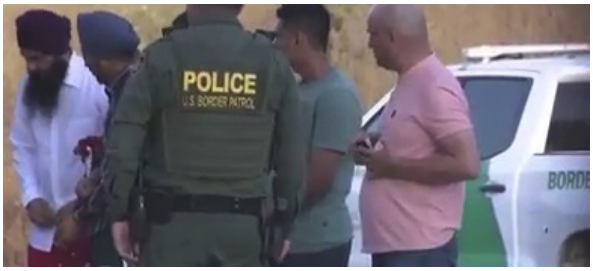
| Channel | # Videos | Avg Words /Video | Avg Images /Video | Avg Views /Video |
|---|---|---|---|---|
| ABC News | 372 | 1,675.1 | 118.5 | 184,200 |
| CBS News | 493 | 936.1 | 80.0 | 34,792 |
| CNN | 595 | 1,281.5 | 105.3 | 467,424 |
| Democracy Now! | 368 | 1,787.7 | 147.8 | 178,746 |
| FOX 4 Dallas-Fort Worth | 67 | 1,113.6 | 104.7 | 82,781 |
| Fox News | 955 | 1,092.9 | 91.1 | 350,153 |
| KTLA 5 | 34 | 816.7 | 75.3 | 34,213 |
| MSNBC | 947 | 2,060.5 | 104.2 | 182,954 |
| NBC News | 455 | 1,627.9 | 110.2 | 116,952 |
| PBS NewsHour | 495 | 879.9 | 96.8 | 73,203 |
| The Next News Network | 96 | 2,040.2 | 151.8 | 52,658 |
| WGN News | 279 | 1,033.8 | 80.2 | 3,586 |
| **Total** | **5,156** | **1,408.6** | **102.99** | **146,805** |

*Notes:* Averages are per video. Channel upload date ranges are reported in the Appendix along with scraping and preprocessing details.

**Figure A7:** TV news selection: defining the "image." The left panel shows a typical broadcast frame that mixes story content and on-screen graphics; the right panel shows the content-focused crop used for embedding. The procedure is applied uniformly across channels and videos.

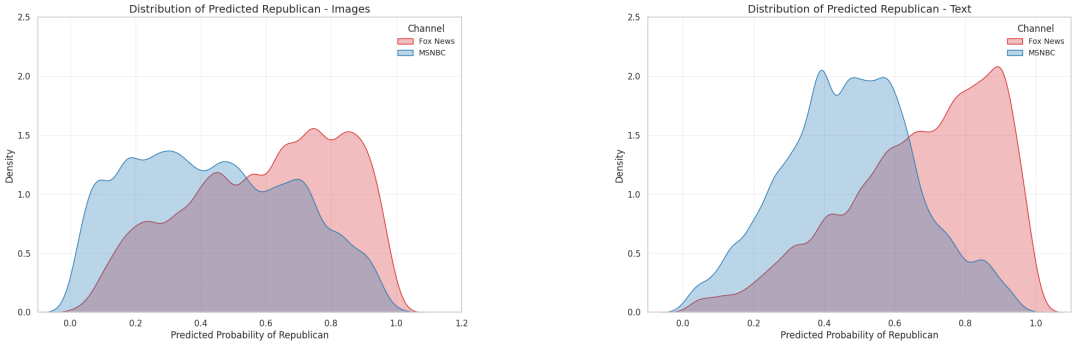



Raw TV news frame with graphics (banners, ticker, split screen).



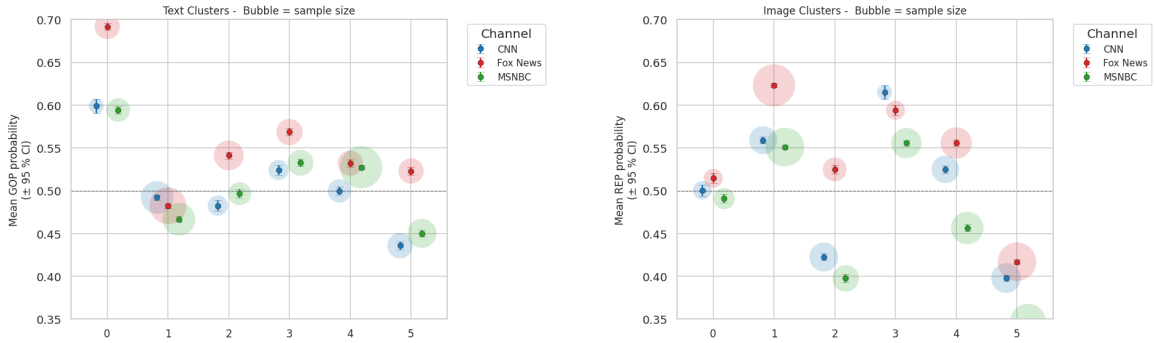

On-scene crop retained via line detection; overlays removed.

**Figure A8:** Distribution of predicted Republican probabilities for Fox News and MSNBC videos.



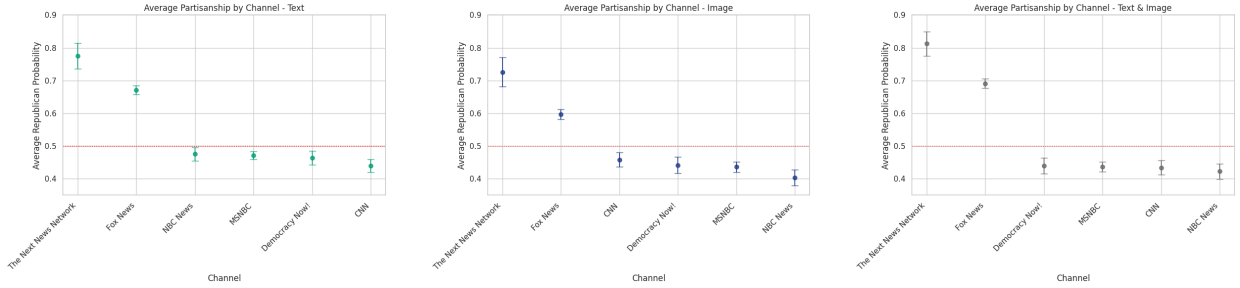

*Notes:* Each panel shows the kernel density of predicted probabilities from the image-only and text-only classifiers, respectively.

**Figure A9:** Average predicted Republican probability by channel and content cluster.



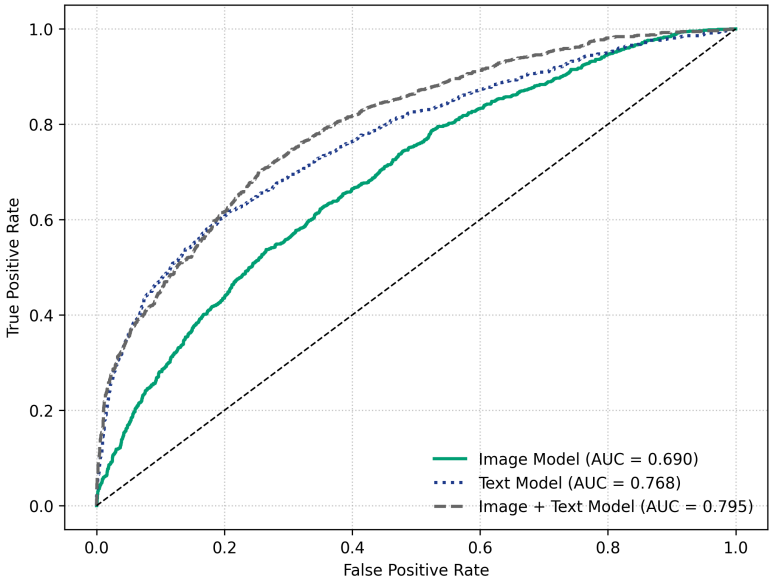

*Notes:* Each point shows the mean predicted Republican probability for a cluster of semantically similar text (left) or image (right) embeddings, with 95% confidence intervals. Bubble size reflects the number of observations within each cluster. Predictions are obtained from the ad-trained classifier applied to the news video corpus.

**Figure A10:** Average predicted Republican probability by channel at the video level, for the 3 models.




Text only. Image only. Text & Image.

**Figure A11:** ROC curves for image-only, text-only, and pooled models at the video level.




*Notes:* The validation sample contains 3,305 videos from six channels. Fox News and The Next News Network are coded as Republican-leaning; MSNBC, CNN, NBC News, and Democracy Now! are coded as Democratic-leaning.

**Figure A12:** Average word-equivalent per image, $\phi(K)/K$, as exposure increases. The value at $K = 1$ is identified by direct interpolation along the text AUC curve. Early images are especially informative.

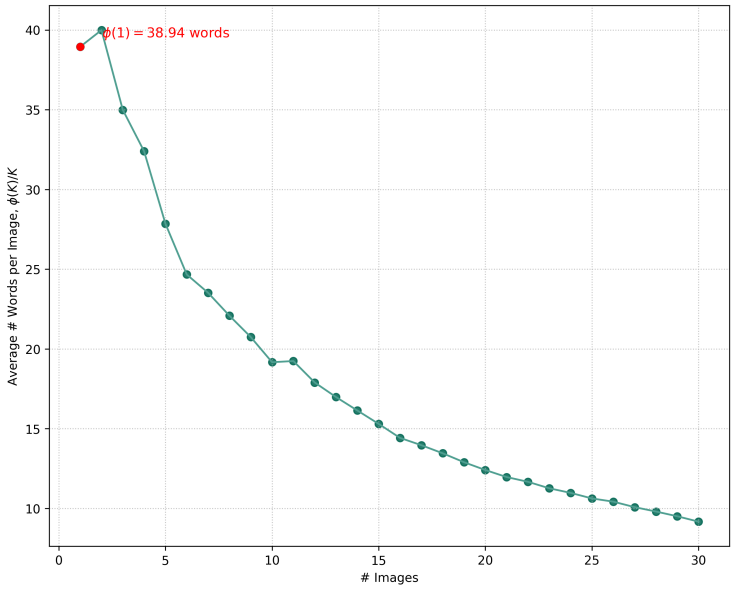
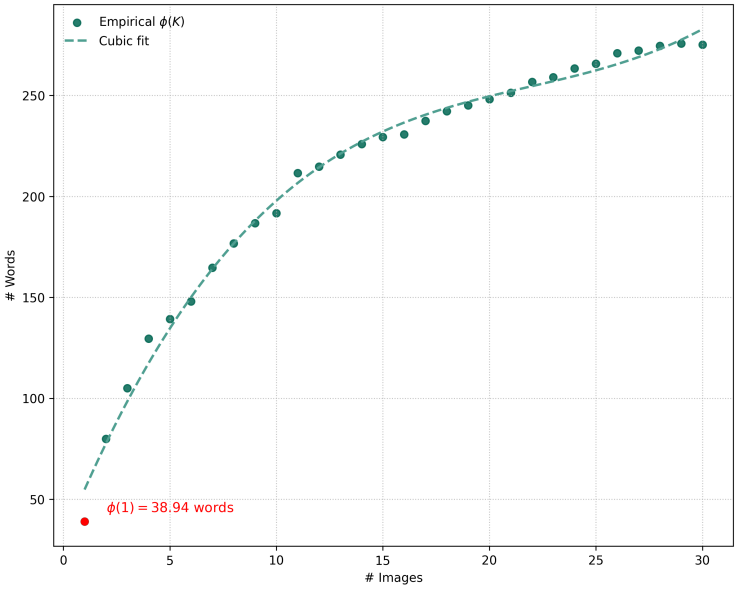



**Figure A13:** Estimated words equivalence of images.

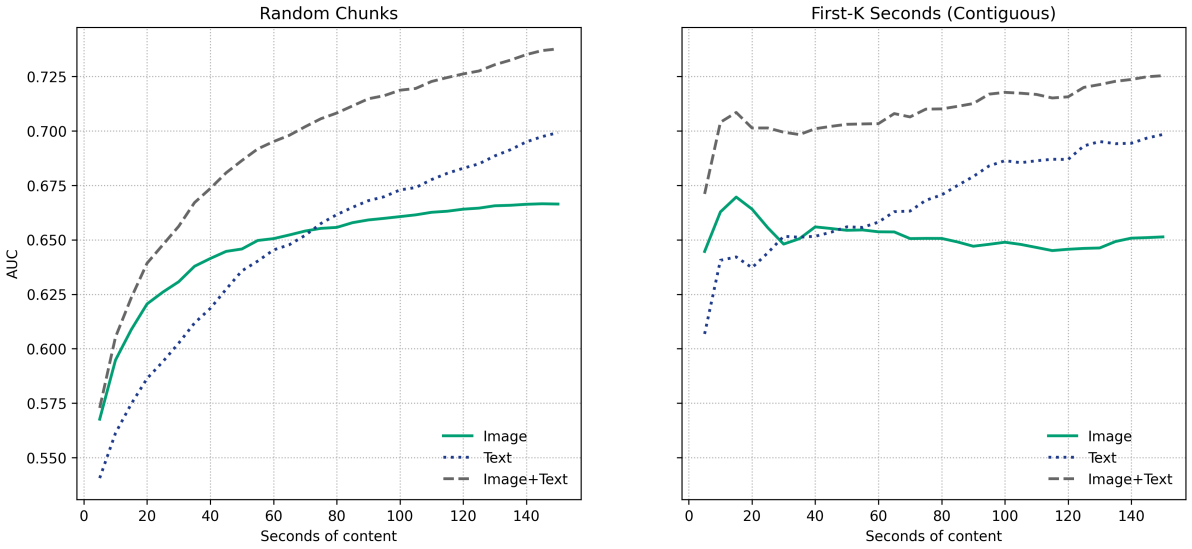



*Notes:* The figure plots the empirical mapping between the number of images and their equivalent transcript length in words, based on matching the AUC of the image model to that of the text model. For one image, direct interpolation along the text curve gives $\phi(1) = 38.94$ words. The dashed cubic is included only as a smooth summary of the empirical mapping; its intercept is not used to estimate $\phi(1)$.

## A.1 Robustness: Contiguous Exposure and the All-Topics Corpus

The main analysis uses the immigration corpus and samples chunks at random. Random sampling measures the partisan information contained in a fixed amount of fragmented exposure. Figure A14 therefore holds the corpus, models, and exposure grid fixed and changes only the sampling rule. The left panel reproduces the random-sampling result; the right panel aggregates the first $k$ contiguous seconds of each clip, approximating a viewer who starts at the beginning and then stops. The substantive result is unchanged. Images dominate text at short exposures, while text accumulates more gradually and eventually overtakes images. The opening image advantage is, if anything, stronger under contiguous exposure.

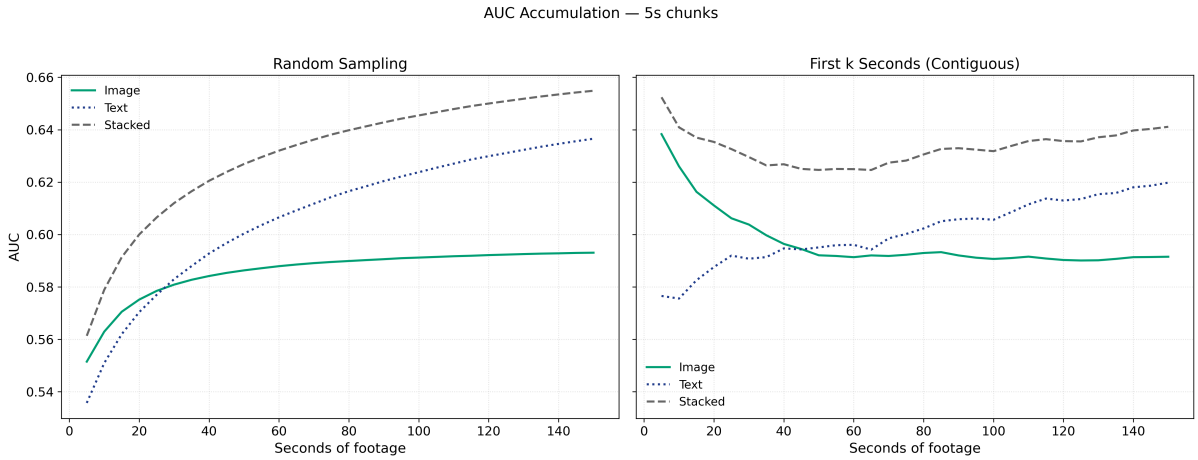
**Figure A14:** Partisan-information accumulation in the immigration corpus: random versus contiguous exposure.




*Notes:* AUC of the image-only, text-only, and stacked models as a function of seconds of content, estimated on the same 1,895 Fox News and MSNBC videos used for Figure 7. Left: five-second chunks sampled at random across the video. Right: the first $k$ contiguous five-second chunks from the start of the clip. The second panel changes only the sampling rule, providing a direct sequential-exposure robustness test for the main result.

Figure A15 then tests whether the pattern generalizes beyond immigration. It re-estimates both sampling designs on the 7,460 videos with matched image and transcript embeddings from the broader all-topics collection of 7,678 Fox News and MSNBC uploads. Predictive accuracy is lower throughout, as general partisanship is harder to read than immigration, but the ordering is unchanged: images lead through the first 20 seconds under random exposure and the first 45 seconds under contiguous exposure; text first overtakes at 30 and 50 seconds, respectively.

**Figure A15:** Partisan-information accumulation in the all-topics corpus: random versus contiguous exposure.

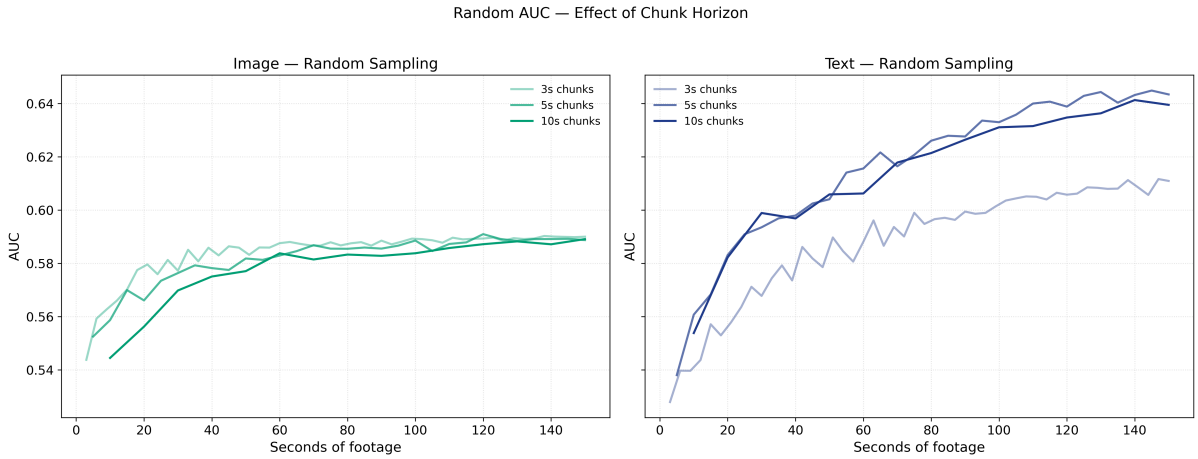



*Notes:* AUC of the image-only, text-only, and stacked models as a function of seconds of footage, estimated on the matched all-topics sample ($N$ = 7,460). Left: mean AUC across 200 random draws. Within each draw, image and text use the same randomly ordered five-second chunks, and estimates at longer exposure lengths add chunks to the shorter-exposure sample. Right: the first $k$ contiguous seconds from the start of the clip.

## A.2 Robustness to Chunk Length

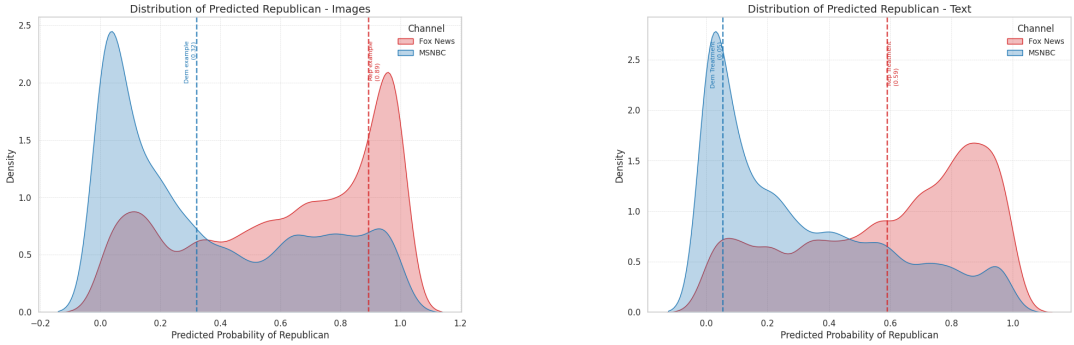
The limited-attention analysis in Section 3.3 aggregates each video into five-second chunks, pairing one frame with its matched transcript segment. This appendix verifies that the accumulation results are not sensitive to that choice. I re-embed the all-topics corpus at chunk lengths of three, five, and ten seconds and recompute the AUC accumulation curves under random sampling (Figure A16). The image curve is essentially invariant to chunk length: the visual partisan signal accumulates at the same rate regardless of how much context each frame is paired with. The text curve is nearly identical at five and ten seconds and only slightly lower at three seconds, where segments contain too few words (roughly nine) to carry partisanship reliably—consistent with the paper's central mechanism that text requires sustained exposure to accumulate signal. The five-second window used in the main text is therefore representative, and the linear aggregation of embeddings is not driving the results.

**Figure A16:** AUC accumulation under different chunk lengths (all-topics corpus).



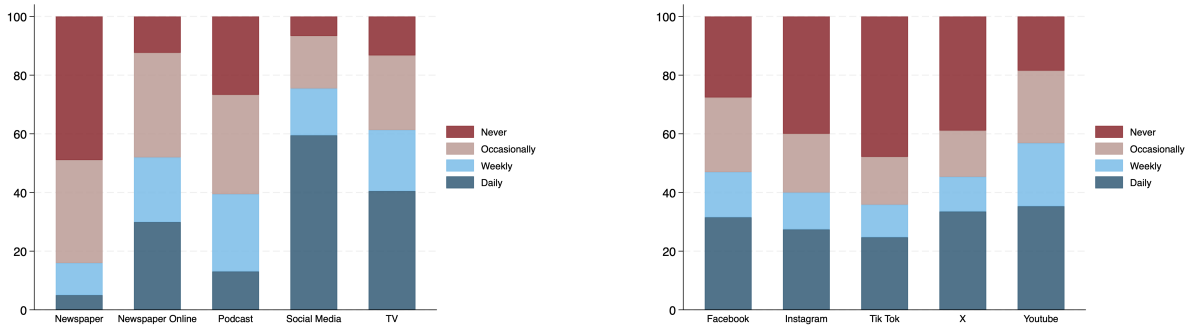

*Notes:* AUC of the image-only (left) and text-only (right) models as a function of seconds of footage, for chunk lengths of three, five, and ten seconds, under random sampling on the all-topics Fox News and MSNBC corpus. Image accumulation is invariant to chunk length; text accumulation is nearly identical at five and ten seconds and slightly lower at three seconds, where segments carry too few words.

**Figure A17:** Distribution of predicted Republican probabilities for Fox News and MSNBC videos.




*Notes:* Each panel shows the kernel density of predicted probabilities from the image-only and text-only classifiers, respectively. The distribution are re-estimated within the immigration-only clusters using a partisanship prediction model trained only on immigration-focused political ads.

**Figure A18:** News Consumption Frequency by Platform




*Notes:* The left panel shows the frequency of news consumption across broad media platforms (TV, social media, newspapers, podcasts). The right panel shows the frequency of news consumption across individual social media platforms. Social media and television are the most widely used sources overall. YouTube is the dominant platform for news consumption within social media, followed by X and Facebook. Data from original survey ($N = 4000$).

**Table A5:** Balance Across Treatment Groups

| | Full Sample | | | |
|---|---|---|---|---|
| | Mean | S.D. | F-Test | P-Value |
| Female | 0.497 | (0.500) | 1.272 | 0.260 |
| White | 0.754 | (0.431) | 1.448 | 0.182 |
| Mixed | 0.058 | (0.234) | 1.613 | 0.127 |
| Black | 0.110 | (0.314) | 0.890 | 0.513 |
| 18–24 years | 0.070 | (0.255) | 1.011 | 0.421 |
| 25–34 years | 0.238 | (0.426) | 0.845 | 0.550 |
| 45–64 years | 0.422 | (0.494) | 1.254 | 0.269 |
| College or More | 0.611 | (0.488) | 1.480 | 0.170 |
| Full-time Employed | 0.553 | (0.497) | 0.807 | 0.581 |
| Newspaper | 0.026 | (0.158) | 1.304 | 0.244 |
| Newspaper Online | 0.256 | (0.436) | 0.499 | 0.836 |
| TV | 0.307 | (0.461) | 0.641 | 0.722 |
| Social Media | 0.537 | (0.499) | 0.580 | 0.772 |
| Podcast | 0.113 | (0.317) | 2.570 | 0.012 |
| NY Times | 0.107 | (0.309) | 0.537 | 0.807 |
| CNN | 0.124 | (0.330) | 1.611 | 0.127 |
| MSNBC | 0.074 | (0.262) | 0.578 | 0.774 |
| Newsmax | 0.029 | (0.169) | 2.060 | 0.044 |
| Facebook | 0.238 | (0.426) | 2.908 | 0.005 |
| Twitter | 0.206 | (0.404) | 0.634 | 0.728 |
| Instagram | 0.183 | (0.386) | 0.532 | 0.811 |
| TikTok | 0.190 | (0.392) | 0.605 | 0.753 |
| YouTube | 0.289 | (0.453) | 1.455 | 0.179 |
| Top Issue: Healthcare | 0.149 | (0.356) | 0.552 | 0.795 |
| Voted (2024) | 0.918 | (0.275) | 1.113 | 0.351 |
| Voted for Trump (2024) | 0.469 | (0.499) | 1.914 | 0.063 |
| N | 3430 | | | |

*Notes:* The table reports mean characteristics and balance tests across the eight experimental treatment groups in the final analysis sample. Each row displays the sample mean, standard deviation, and the $F$-test and $p$-value for equality of means across treatments. Three of the 26 joint tests have $p < 0.05$. $N = 3{,}430$. All news-consumption indicators refer to consumption at least weekly.

**Table A6:** Perceived Partisanship of Images and Text

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| **Panel A. Image partisanship** | | | | | | |
| Image (Rep) | 0.485*** | 0.414*** | 0.253*** | 0.340*** | 0.647*** | 0.505*** |
| | (0.045) | (0.045) | (0.068) | (0.064) | (0.059) | (0.062) |
| Text (Rep) | 0.150*** | -0.044 | 0.098 | -0.140** | 0.196*** | 0.056 |
| | (0.045) | (0.045) | (0.067) | (0.064) | (0.061) | (0.063) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |
| **Panel B. Text partisanship** | | | | | | |
| Image (Rep) | 0.201*** | 0.121*** | 0.095 | 0.089 | 0.270*** | 0.151** |
| | (0.047) | (0.046) | (0.071) | (0.066) | (0.062) | (0.066) |
| Text (Rep) | 0.183*** | -0.019 | 0.073 | -0.184*** | 0.284*** | 0.164** |
| | (0.047) | (0.047) | (0.070) | (0.067) | (0.063) | (0.066) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |

*Notes:* The outcomes measure the perceived Republican leaning of the images and transcript separately and are standardized over the full sample. Higher values indicate stronger perceived Republican leaning. Image (Rep) and Text (Rep) indicate Republican rather than Democratic content. OLS regressions include both modality indicators and demographic, political-attitude, and media-consumption controls, separately by exposure length and respondent group. Long and short clips last approximately two minutes and 30 seconds, respectively. Robust standard errors appear in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A7:** Effect on Increase Border Patrol

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| Image (Rep) | -0.045 | -0.027 | -0.046 | -0.025 | -0.030 | -0.027 |
| | (0.034) | (0.034) | (0.040) | (0.037) | (0.052) | (0.056) |
| Text (Rep) | 0.170*** | 0.009 | 0.077* | 0.033 | 0.250*** | -0.019 |
| | (0.035) | (0.034) | (0.041) | (0.037) | (0.052) | (0.055) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |

*Notes:* The outcome is support for hiring significantly more border-patrol agents, standardized over the full sample; higher values indicate greater support. Image (Rep) and Text (Rep) indicate Republican rather than Democratic content. OLS regressions include both modality indicators and demographic, political-attitude, and media-consumption controls, separately by exposure length and respondent group. Long and short clips last approximately two minutes and 30 seconds, respectively. Robust standard errors appear in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A8:** Effect on Sympathy

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| Image (Rep) | -0.112*** | -0.076* | -0.102* | -0.083 | -0.114** | -0.067 |
| | (0.039) | (0.040) | (0.056) | (0.052) | (0.055) | (0.061) |
| Text (Rep) | -0.062 | 0.023 | -0.048 | -0.024 | -0.071 | 0.070 |
| | (0.040) | (0.040) | (0.055) | (0.054) | (0.055) | (0.061) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |

*Notes:* The outcome is self-reported sympathy, standardized over the full sample; higher values indicate stronger sympathy. Image (Rep) and Text (Rep) indicate Republican rather than Democratic content. OLS regressions include both modality indicators and demographic, political-attitude, and media-consumption controls, separately by exposure length and respondent group. Long and short clips last approximately two minutes and 30 seconds, respectively. Robust standard errors appear in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A9:** Effect on Anger

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| Image (Rep) | 0.093* | 0.136*** | -0.049 | 0.172*** | 0.183*** | 0.094 |
| | (0.048) | (0.046) | (0.066) | (0.060) | (0.068) | (0.071) |
| Text (Rep) | 0.003 | 0.021 | 0.032 | 0.018 | -0.052 | 0.011 |
| | (0.048) | (0.046) | (0.065) | (0.060) | (0.068) | (0.072) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |

*Notes:* The outcome is self-reported anger, standardized over the full sample; higher values indicate stronger anger. Image (Rep) and Text (Rep) indicate Republican rather than Democratic content. OLS regressions include both modality indicators and demographic, political-attitude, and media-consumption controls, separately by exposure length and respondent group. Long and short clips last approximately two minutes and 30 seconds, respectively. Robust standard errors appear in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A10:** Effect on Disgust

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| Image (Rep) | 0.162*** | 0.158*** | 0.044 | 0.132** | 0.242*** | 0.185** |
| | (0.048) | (0.046) | (0.066) | (0.059) | (0.068) | (0.073) |
| Text (Rep) | -0.004 | 0.033 | -0.018 | 0.066 | -0.023 | -0.008 |
| | (0.048) | (0.046) | (0.064) | (0.058) | (0.068) | (0.073) |
| Observations | 1748 | 1682 | 788 | 869 | 960 | 813 |

*Notes:* The outcome is self-reported disgust, standardized over the full sample; higher values indicate stronger disgust. Image (Rep) and Text (Rep) indicate Republican rather than Democratic content. OLS regressions include both modality indicators and demographic, political-attitude, and media-consumption controls, separately by exposure length and respondent group. Long and short clips last approximately two minutes and 30 seconds, respectively. Robust standard errors appear in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A11:** Effects of Cross-Partisan Pairings (RD versus DR)

| | All respondents | | Republicans | | Democrats | |
|---|---|---|---|---|---|---|
| | Long | Short | Long | Short | Long | Short |
| **Anti-immigration** | | | | | | |
| RD vs DR | -0.100** | -0.021 | 0.008 | 0.020 | -0.180*** | 0.013 |
| | (0.042) | (0.041) | (0.055) | (0.047) | (0.060) | (0.063) |
| Observations | 874 | 831 | — | — | — | — |
| **Charity choice** | | | | | | |
| RD vs DR | 0.028 | -0.126** | -0.034 | -0.154** | 0.114 | -0.067 |
| | (0.063) | (0.063) | (0.073) | (0.074) | (0.090) | (0.102) |
| Observations | 874 | 831 | — | — | — | — |
| **Negative emotions** | | | | | | |
| RD vs DR | 0.106 | 0.157** | -0.026 | 0.108 | 0.204** | 0.194* |
| | (0.070) | (0.068) | (0.095) | (0.101) | (0.104) | (0.107) |
| Observations | 874 | 831 | 402 | 420 | 472 | 411 |

*Notes:* The estimates compare videos where visual and textual partisanship are misaligned (Republican text with Democratic images, RD, versus Democratic text with Republican images, DR). Rows report each outcome for the full sample and by respondent party. Dashes indicate sample sizes not reported in the source panels. Outcomes are standardized over the full sample. Regressions include demographic, political-attitude, and media-consumption controls. Robust standard errors are in parentheses; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# B Partisanship Model Details

**Details of CLIP embedding representation.** Let $\mathbf{x}_j^{\text{img}} \in \mathbb{R}^{d_i}$ and $\mathbf{x}_j^{\text{txt}} \in \mathbb{R}^{d_t}$ represent the $j$-th image frame and text snippet extracted from advertisement $A$. Define the CLIP image and text encoders as mappings $f : \mathbb{R}^{d_i} \to \mathbb{R}^d$ and $g : \mathbb{R}^{d_t} \to \mathbb{R}^d$, respectively.

**Assumption 1.** *The embedding representing an entire advertisement $A$ can be approximated by the average of the embeddings of its individual constituent frames or snippets:*

$$\mathbf{x}_A^{mod} \approx \frac{1}{N} \sum_{j=1}^{N} f^{mod}(\mathbf{x}_j^{mod}), \quad mod \in \{img, txt\},$$

*where $N$ is the number of frames or snippets, and $f^{mod}$ denotes the modality-specific CLIP encoder.*

The main limitation of this assumption is that it ignores sequence effects and higher-order interactions among elements of the ad. Appendix A.2 shows that the accumulation results are robust to chunk lengths of three, five, and ten seconds.

A challenge in multimodal analysis is the modality gap: image and text embeddings tend to occupy different regions of the shared space (Liang et al. (2022)). Without adjustment, this limits their direct comparability and can bias pooled estimates. I address this in three steps: (i) normalize each ad's embeddings to lie on the unit sphere; (ii) compute the average embedding across all ads for each modality; (iii) subtract the modality-specific global mean from each embedding and re-normalize.

For ad $A$ in modality mod $\in \{\text{img}, \text{txt}\}$, with normalized embeddings $\bar{\mathbf{x}}_A^{\text{mod}}$ and global mean $\boldsymbol{\mu}_{\text{mod}}$, the adjusted embedding is

$$\tilde{\mathbf{x}}_A^{\text{mod}} = \frac{\bar{\mathbf{x}}_A^{\text{mod}} - \boldsymbol{\mu}_{\text{mod}}}{\|\bar{\mathbf{x}}_A^{\text{mod}} - \boldsymbol{\mu}_{\text{mod}}\|_2}, \quad \text{mod} \in \{\text{img}, \text{txt}\}.$$

This centering ensures that both modalities share a common reference point, enabling meaningful comparison and integration in subsequent analyses. Intuitively, this adjustment controls systematic differences between how CLIP encodes text and images, so that any remaining variation in the embeddings reflects content rather than the modality itself.

# C Experiment Text & Image Treatments

## Text Treatment.

**Democratic Text:**

*"» Right now a federal appeals court in New Orleans is hearing arguments again about a controversial new Texas immigration law. This law, which the court has currently put on hold, would let the state arrest and deport migrants for illegally crossing the border. The State and Federal government have been in and out of court for months over this, arguing about whether the state should have that power. let's get right to my homeland security correspondent. I've been following this law through so many twists and turns, what should I know about this hearing today?*

*» So, I've been listening to the oral arguments and there's something you should know, because Texas is actually changing the way they said this law would work. Just as you pointed out, they've always said they would arrest and deport them. Their plan was to walk migrants back into Mexico and change the territory on which these migrants stood. Now, they're arguing, before this judge panel in the 5th circuit, that they would simply turn these migrants back over to the custody of the United States by taking them to a port of entry on the U.S. side, and handing them over to customs and border protection. That's similar to something Texas already has been doing for a number of years, after they charged someone for trespassing on Texas state property. And so it really left the justice department, arguing on behalf of the federal government, in a strange situation as judges then started questioning them: well, could this stand? Would this be okay, if Texas did it this way? They simply pointed to the language that was already in the law and then the arguments to say: that's not how they said they would do this. So, a lot of confusion going on inside this courtroom right now. It seems that texas might need to rework the way the law is worded and the justice department would need more time to figure out how they would respond to a law that looked very different than what they set out in the first place."*

**"Republican" Text:**

*"» The showdown between texas and the biden administration, now going one step higher on the legal ladder. appeals court set to hear oral arguments about a law that allows the lone star state to arrest and deport illegals, only weeks after a lower court put that law on hold. Where to now? my correspondent live in new orleans to cover this today. Good morning.*

*» Oral arguments expected to go about an hour, though they could go a little bit longer given this issue, very controversial, this will be a three-judge appeals panel here in new orleans. The reason I're in new orleans talking about a texas immigration law is because new orleans fifth court of appeals covers the state of texas. Now, the fifth circuit Court of Appeals here in downtown new orleans is a traditionally conservative appeals court here in the united states. One of the most in the united states, if not the most. I won't know exactly how the court is leaning until these arguments are underway. You may remember last month, the supreme court allowed the Texas law that allows local Texas law enforcement to arrest anyone who is suspected of crossing the border illegally to be enforced, and to arrest those people. Just hours later a lower court put a*

*pause, or a stay in place. The justice department will be here today, fighting to make this law unconstitutional arguing that immigration control is an entirely federal power and U.S. border patrol agents alone control the border and make arrests. Governor Greg Abbott in Texas says that the law is needed because of the state's sovereignty from an invasion of illegal immigrants bringing crime and drugs. He says the border patrol agents are understaffed and overwhelmed and the State should be able to make arrests. Abbott and his team says that the border patrol not doing their job, so he wants to do his job by having the state enforce this law. I're not going to have a final decision today, they're not gonna rule from the bench. But no matter what happens today in New Orleans I expect to fight in Washington back at the steps of the Supreme Court on this very issue."*

## Image Treatment.

**Figure C1:** Image Treatments




**(a)** Democratic Images




**(b)** Republican Images

# D Preprocessing Details

**Text.** I standardize and represent video transcripts in four steps.

(i) Normalization. Transcripts are lowercased, control characters and simple markup are removed, and excess whitespace is collapsed. To reduce caption noise, I drop common non-lexical fillers (e.g., *um*, *uh*, *erm*, *like*, *you*, *know*). Immediate word repetitions generated by automatic speech recognition (ASR) are de-duplicated.

(ii) Punctuation restoration. Because ASR captions often lack sentence boundaries, I restore punctuation with a pre-trained restoration model. Processing is performed in short, fixed-length slices to respect the model's token limits; slices are then rejoined and sentence starts capitalized.

(iii) Alignment to frames. Let $N$ denote the number of retained frames for the video after image filtering. I partition the cleaned transcript into $N$ contiguous segments of (approximately) equal word count. This yields a one-to-one alignment between text segments and frames without relying on timestamps.

(iv) Representation. Each segment is tokenized and embedded with the CLIP text encoder, producing a segment-level vector used in chunk-based analyses. I also compute a video-level average text embedding for aggregation.

(v) Quality control. I exclude videos with missing or empty transcripts or with zero retained frames. For included videos I enforce $N_{\text{text}} = N_{\text{frames}}$ and report counts of exclusions and alignment failures.

Design rationale. I intentionally avoid lemmatization, stemming, and broad stopword removal beyond the filler list. Because CLIP uses subword tokenization, preserving surface forms maintains fidelity to the encoder's pretraining distribution and avoids altering token boundaries in ways that may distort embeddings.

**Images.** I process extracted frames to isolate the primary content region and suppress broadcast overlays (tickers, lower-thirds, pillarboxes).

(i) Structural detection. Each frame is converted to grayscale and lightly smoothed. I compute edges and detect straight lines using standard computer-vision primitives (Canny edges and a probabilistic Hough transform), restricting attention to near-horizontal and near-vertical lines. Detected lines are bucketed by screen half (top/bottom for horizontals; left/right for verticals) to exploit typical broadcast layouts.

(ii) Crop boundary selection. From each bucket I select the interior-most candidate to propose top/bottom and left/right crop boundaries. To avoid degenerate crops, candidate boundaries must be separated by at least 10% of the corresponding image dimension; if candidates are absent or too close, I fall back to the full extent on that axis. Bounds are clipped to the image size.

(iii) Subregion selection. The resulting boundaries partition the frame into up to four rectangles. I retain the largest rectangle as the primary content crop. When a second rectangle's area is at least 90% of the largest (e.g., split-screen interviews), I also retain it. If no valid partition is produced, the original frame is kept.

(iv) Representation and QA. Cropped images are embedded with the CLIP image encoder using its standard preprocessing (resize and center-crop as implemented by the encoder). I compute segment-level embeddings and a video-level average image embedding. Corrupted frames are skipped; processing is idempotent, and logs record exclusions and fallbacks.

Design rationale. I favor geometric cropping based on detected layout rather than inpainting or heavy denoising. This removes consistent studio graphics while keeping images close to the encoder's training distribution, limiting the risk of artifacts that could contaminate embeddings.

# E A Framework for Understanding Video Processing

**Setup.** A viewer of partisan type $g \in \{\text{Rep}, \text{Dem}\}$ watches a video of length $\tau$. I model two key elements:

- **Effective Exposure**: Effective exposure is $s = \tau A$, where $A \in [0, 1]$ denotes the fraction of attention devoted to the video.
- **Modalities**: Videos contain images (I) and text (T), each carrying partisan cues and factual content

**Stage 1: Engagement.** Viewers must first decide whether to maintain engagement with the video. Engagement depends on two sources of utility:

$$U^g(s) = \omega^g \cdot \mathbb{E}[\text{Information}] + \eta^g \cdot E(s)$$

where $\omega^g$ captures the value of learning content and $\eta^g$ captures the value of emotional arousal. Emotional arousal evolves with exposure according to:

$$E(s) = e\,(1 - e^{-\rho s})\,e^{-\delta^g s}$$

The first term captures initial emotional rise; the second term $e^{-\delta^g s}$ captures adaptation: emotions fade with prolonged exposure, consistent with the psychology literature.

Viewers remain engaged if $U^g(s)$ exceeds the cognitive cost of sustained attention.

**Stage 2: Information processing.** Conditional on engagement, viewers absorb information from both modalities according to:

$$I_m(s) = \lambda_m\,(1 - e^{-\alpha_m s})$$

where $\alpha_m$ governs the speed of information delivery and $\lambda_m$ is the maximum information capacity. Based on my empirical evidence:

- $\alpha_I > \alpha_T$: Images deliver information quickly, text slowly
- $\lambda_T > \lambda_I$: Text has higher total information capacity

**Predictions.** Combining the two stages. Suppose we have to groups:

- Republicans: High $\eta^{\text{Rep}}$ (emotion-driven), low $\omega^{\text{Rep}}$ (less content-motivated), high $\delta^{\text{Rep}}$ (rapid adaptation)
- Democrats: Low $\eta^{\text{Dem}}$ (less emotion-driven), high $\omega^{\text{Dem}}$ (content-motivated), low $\delta^{\text{Dem}}$ (slow adaptation)

this would imply that Republicans disengage in long videos as emotions adapt ($e^{-\delta^{\text{Rep}}s} \to 0$), while Democrats maintain engagement through informational interest.

Republicans:

- **Short exposure**: High emotions $\implies$ engaged $\implies$ respond to fast-delivering images
- **Long exposure**: Emotions adapted $\implies$ disengaged $\implies$ no response to either modality

Democrats:

- **Short exposure**: Remain engaged $\implies$ respond to fast-delivering images
- **Long exposure**: Remain engaged $\implies$ respond to both images and slow-accumulating text

# F Survey Questionnaire and Instructions

The questionnaire and participant instructions are reproduced below. Repeated video instructions are shown once for each exposure length. Contact details in the consent form are omitted from this public reproduction. Treatment materials appear in Appendix C.

## Consent and survey instructions

Thank you for agreeing to take this survey. It is very important for the success of our research project that you complete the survey until the end, once you have started. This survey should take (on average) about 10 minutes to complete.

The survey requires the ability to watch a short video clip - please make sure you will be able to do this before you begin.

**Will I get compensated?**

You will be paid $2.00 to complete this survey.

**What are the features of the lottery payment in the survey?**

As part of the study, there will be a lottery where participants can earn up to $25 in addition to the base compensation for participation. Survey participants do not need to complete the entirety of the survey or remain in the study to be entered into the drawing for $25. The probability of winning the lottery is approximately 0.01%. If you win the lottery, you will be notified at the conclusion of the survey and will be paid within 2 months.

**Why is this study being done?**

Researchers at [institution omitted for review] are conducting a study about media, politics and perception. Your participation in this research study is voluntary.

**What will happen to you if you are in the study?**

If you volunteer to participate in this study, the researcher will ask you to do the following: answer a few questions about your demographics and political/social views and watch a news video clip.

**Are there any risks?**

There are no anticipated risks or discomforts. You can choose whether or not you want to be in this study, and you may withdraw your consent and discontinue participation at any time. We will not collect any information that can identify you in our research.

**What about confidentiality?**

All of your responses are completely confidential, so you may answer freely and openly. We will not ask you for any personally identifiable information such as your name.

**Do you have to be in this study?**

You can choose whether or not you want to be in this study, and you may withdraw your consent and discontinue participation at any time. Whatever decision you make, there will be no penalty to you, and no loss of benefits to which you were otherwise entitled.

**Whom may I call if I have questions?**

If you have any questions, comments or concerns about the research you can talk to one of the researchers. Please contact: [researcher name omitted for review] at [researcher email omitted for review]. If you have questions about your rights as a research subject, or you have concerns or suggestions and you want to talk to someone other than the researchers, you may contact the [institutional review board] by phone at [telephone number omitted for review] or by email at [review board email omitted for review].

**Statement of Consent**

I have read this information sheet and the research study has been explained to me. I agree to be in the research study described above. A copy of this information sheet will be provided to me. By agreeing to participate, I have not given up any of the legal rights that I would have if I were not a participant in the study.

*Response options:* Yes, I consent; No I do not consent.

## Eligibility and demographics

**1.** This survey requires the ability to maintain a strong internet connection and the capacity to watch video content on Youtube. Does your device have this capability?
*Response options:* Yes; No.

**2.** Are you a U.S. Citizen?
*Response options:* Yes; No.

**3.** What is your gender?
*Response options:* Male; Female; Other/Prefer not to answer.

**4.** What is your age?
*Response options:* 18-24 years old; 25-34 years old; 35-44 years old; 45-54 years old; 55-64 years old; 65-74 years old; 75 or older.

**5.** Which category best describes your highest level of education?
*Response options:* Less than High School; High School Degree / GED; Some College; College degree or higher.

**6.** What is your current employment status?
*Response options:* Full-time employee; Part-time employee; Self-employed or small business owner; Not employed.

**7.** Were you born in the United States?
*Response options:* Yes; No.

**8.** In politics, as of today, do you consider yourself:
*Response options:* Republican; Democrat; Independent; Other.

**9.** Did you vote in the last presidential elections?
*Response options:* Yes; No.

*Shown to respondents who voted in the last presidential election.*
**10.** In the last presidential election, you supported:
*Response options:* Kamala Harris; Donald Trump; Chase Oliver; Jill Stein; Robert F. Kennedy Jr.

*Shown to respondents who did not vote in the last presidential election.*
**11.** Even if you did NOT vote, please indicate the candidate that you were most likely to have voted for or who represents your views most closely
*Response options:* Kamala Harris; Donald Trump; Chase Oliver; Jill Stein; Robert F. Kennedy Jr.

## Background views

**12.** Thinking now about fire arms and guns – in general, do you feel that the laws covering the sale of firearms should be made more strict, less strict or kept as they are now?
*Response options:* More strict; Less strict; Kept as they are now; No opinion.

**13.** Thinking now about abortion – do you think abortions should be legal under any circumstances, legal only under certain circumstances or illegal in all circumstances?
*Response options:* Legal under any circumstances; Legal only under certain circumstances; Illegal in all circumstances; No opinion.

**14.** Thinking now about healthcare – do you think it is the responsibility of the federal government to make sure all Americans have healthcare coverage, or is that not the responsibility of the federal government?
*Response options:* Yes, it is the government's responsibility; No, it is not the government's responsibility; No opinion.

**15.** Thinking now about taxes – do you consider the amount of federal income tax you have to pay too high, about right or too low?
*Response options:* Too high; About right; Too low; No opinion.

**16.** Which of the following issues is most important to you?
*Response options:* Taxes; Healthcare; Abortion; Gun Control.

## News consumption

**17.** How often do you read or watch the news through the following outlets?

| | Never | Occasionally | Weekly | Daily |
|---|---|---|---|---|
| Newspaper - Print | ○ | ○ | ○ | ○ |
| Newspaper - Online | ○ | ○ | ○ | ○ |
| TV | ○ | ○ | ○ | ○ |
| Social Media | ○ | ○ | ○ | ○ |
| Podcast | ○ | ○ | ○ | ○ |

**18.** How often do you read or watch news through the following outlets?

| | Never | Occasionally | Weekly | Daily |
|---|---|---|---|---|
| Fox News | ○ | ○ | ○ | ○ |
| The New York Times | ○ | ○ | ○ | ○ |
| CNN | ○ | ○ | ○ | ○ |
| MSNBC | ○ | ○ | ○ | ○ |
| Newsmax | ○ | ○ | ○ | ○ |

**19.** How often do you read or watch news through the following social media platform?

| | Never | Occasionally | Weekly | Daily |
|---|---|---|---|---|
| X (Twitter) | ○ | ○ | ○ | ○ |
| Instagram | ○ | ○ | ○ | ○ |
| Facebook | ○ | ○ | ○ | ○ |
| TikTok | ○ | ○ | ○ | ○ |
| YouTube | ○ | ○ | ○ | ○ |

## Video instructions

**Long videos**
Please watch carefully the entire ~2-minute video in the next page. You can continue only after watching the full clip.
You will be asked some questions about the video.

**Short videos**
Please watch carefully the entire ~30-seconds video in the next page. You can continue only after watching the full clip.
You will be asked some questions about the video.

## Video checks

**20.** Did you watch the entire video?
*Response options:* Yes; No, there was a technical problem; No, I skipped part of the video.

**21.** What was the video about?
*Response options:* Immigration; Healthcare; Gun Control; Abortion.

**22.** What kind of images were shown in the video?
- Scenes of migrants and/or U.S. border-patrol officers at the Texas-Mexico border
- Studio shots of the news anchor and in-studio guests talking
- Charts and graphics about immigration in the U.S.

**23.** What was the main legal dispute being argued?
- Whether the U.S. Border Patrol has enough funding to hire more agents
- Whether the State of Texas may arrest people it suspects of crossing the border illegally
- Whether the Texas Governor can deploy the National Guard outside Texas
- Whether New Orleans police must assist federal agents in immigration arrests

## Immigration views

**24.** Do you favor or oppose the following proposal?
Hire significantly more border patrol agents.
*Response options:* Strongly Favor; Favor; Oppose; Strongly Oppose; No Opinion.

**25.** Do you favor or oppose the following proposal?
Allow the U.S. military to assist immigration raids inside the U.S.
*Response options:* Strongly Favor; Favor; Oppose; Strongly Oppose; No opinion.

**26.** Do you favor or oppose the following proposal? Deport all undocumented immigrants living in the U.S. back to their home countries.
*Response options:* Strongly Favor; Favor; Oppose; Strongly Oppose; No Opinion.

**27.** Do you favor or oppose the following proposal?
Deport all undocumented immigrants with a criminal record living in the U.S. back to their home countries.
*Response options:* Strongly Favor; Favor; Oppose; Strongly Oppose; No Opinion.

**28.** Do you think the number of immigrants from foreign countries who are permitted to come to the United States to live should be increased or decreased?
*Response options:* Increased a lot; Increased a little; Left the same as it is now; Decreased a little; Decreased a lot.

**29.** Do you agree or disagree the following statement?
Immigrants living in the U.S. commit a large or disproportionate share of violent crimes in the U.S.
*Response options:* Strongly Agree; Agree; Disagree; Strongly Disagree; No Opinion.

## Emotional responses

**30.** Below, you will see a list of emotions. We ask you to rate the intensity with which you experienced each emotion during the survey until now.
Please make your assessment on a scale from 1 to 7, where 1 indicates very low intensity and 7 indicates very high intensity.
*Scale labels:* 1; 2; 3; 4; 5; 6; 7.
*Items:* Anger; Disgust; Fear; Sympathy; Sadness; Indifference.

**31.** To what extent do you feel you can relate to the circumstances of the immigrants at the border? (0 = not at all, 10 = very much)
*Scale labels:* Not at all; Very much so.
*Scale labels:* 0; 1; 2; 3; 4; 5; 6; 7; 8; 9; 10.
*Item:* I relate to the circumstances.

**32.** In a few words, please describe your experience watching the video
*Open response.*

## Lottery and donation decision

You will be entered a cash lottery to win up to $25. The probability of winning the lottery is approximately 0.01%.
You will be notified at the end of the survey if you won the lottery. Lottery payments will be made within 2 months. Survey participants do not need to complete the entirety of the survey or remain in the study to be entered into the lottery.

**33.** If you win the lottery, you can choose a prize of $25 dollars or you can donate a portion of this amount to The American Civil Liberties Union (ACLU) Immigrants' Rights Project.
This project is dedicated to expanding and enforcing the civil liberties and civil rights of immigrants and to combating public and private discrimination against them.
For more information, you can visit: https://www.aclu.org/issues/immigrants-rights
*Response options:* I wish to donate a portion of the prize to the charity; I do not wish to donate.

*Shown to respondents who chose to donate a portion of the prize.*
**34.** How much of the $25 prize would you like to donate?
*Scale labels:* 0; 5; 10; 15; 20; 25.
*Item:* Give donation.

## Perceived partisanship

**35.** The video you just watched was taken from a TV News report (originally broadcast on April 3, 2024). On the U.S. political spectrum, how would you classify the video?
*Response options:* Strongly Democratic/Liberal; Moderately Democratic/Liberal; Neutral; Moderately Republican/Conservative; Strongly Republican/Conservative.

**36.** On the U.S. political spectrum, how would you classify each element of the video?
*Scale labels:* Strongly Democratic; Moderately Democratic; Neutral; Moderately Republican; Strongly Republican.
*Scale labels:* -100; -80; -60; -40; -20; 0; 20; 40; 60; 80; 100.
*Items:* The images and visuals shown; The words and information spoken; The speaker's tone of voice.

## Optional additional viewing

*After short videos.*
**37.** You have finished the main part of the study. You now have the option to watch the remaining 80-second segment of the clip.
Watching is not required and carries no extra payment. Whether or not you watch, you will still receive the full $2.00 participation fee.
*Response options:* Watch the rest of the clip; Skip and continue.

*After long videos.*
**38.** You have finished the main part of the study. You now have the option to watch an additional 80-second segment.
Watching is not required and carries no extra payment. Whether or not you watch, you will still receive the full $2.00 participation fee.
*Response options:* Watch the rest of the clip; Skip and continue.

*After either exposure length, shown to respondents who chose "Skip and continue."*
**39.** If you were offered a cash bonus to watch the extra 80 seconds, what is the minimum amount you would accept to watch?
*Response options:* $0.10; $0.25; $0.50; $0.75; $1.00 or more.

## Closing questions and completion

**40.** In a few words, please state whether you believe the video you watched was biased or not, and please explain why.
*Open response.*

**41.** Please use this box to mention any comment/feedback you may have on the survey
*Open response.*

**Message to lottery winners**
Congratulations! You won the lottery prize. We will submit a bonus payment within the next 2 months.
Thank you for taking the time to complete this survey. Your response has been recorded.
Click to end the survey and confirm completion.

**Message to other participants**
Thank you for taking the time to complete this survey. Your response has been recorded.
Click to end the survey and confirm completion.

## References

**Cowen, Alan S, and Dacher Keltner.** 2017. "Self-report captures 27 distinct categories of emotion bridged by continuous gradients." *Proceedings of the national academy of sciences*, 114(38): E7900–E7909.

**Liang, Victor Weixin, Yuhui Zhang, Yongchan Kwon, Serena Yeung, and James Y Zou.** 2022. "Mind the gap: Understanding the modality gap in multi-modal contrastive representation learning." *Advances in Neural Information Processing Systems*, 35: 17612–17625.