A MULTILINGUAL ANALYSIS OF YOUTUBE COMMENTS ON SCIENCE COMMUNICATION VIDEOS

By

Jonus Cottrell

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Journalism – Master of Arts

2022

ABSTRACT

This thesis aimed to fill a gap in the literature regarding the consumption of YouTube science videos as a form of science communication and engagement. This was done by conducting an automated sentiment analysis of comments posted to YouTube videos published by popular science communication channel, Kurzgesagt. The two videos used for analysis covered the production and consumption of meat and the production and consumption of dairy and were published in English, Spanish, and German, which provided an opportunity to compare expressed sentiment for identical content across three different target audiences. Results showed that for English and German audiences, expressed sentiment was higher for the video covering the production and consumption of meat. There were no statistically significant differences in expressed sentiment for the two videos published in Spanish. For the video covering the production and consumption of meat, the most negative expressed sentiment was found in the English-speaking audience. For the video covering the production and consumption of milk, the most positive expressed sentiment was found in the Spanish-speaking audience. Findings suggest that there are some differences in engagement across publication languages, but further research needs to be done to account for the effect of exogenous variables. Results align with previous research that suggests YouTube is a useful platform to facilitate audience engagement.

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION 1

CHAPTER 2: LITERATURE REVIEW AND RESEARCH QUESTIONS 4

CHAPTER 3: METHOD 16

CHAPTER 4: RESULTS 20

CHAPTER 5: DISCUSSION, LIMITATIONS, AND CONCLUSION 27

BIBLIOGRAPHY 33

CHAPTER 1: INTRODUCTION

With the emergence of blogs, alternative news sites, messaging apps, and social media platforms, journalists and news organizations have a diminished gatekeeping role regarding what information becomes news in the public (Boy, Bucher & Christ, 2020). No longer are the days when consumers rely on newspapers, broadcast media, or even National Geographic magazines to get their information about the world. Individuals now have the option to search online for any topic that interests them via a multitude of digital communication platforms such as YouTube, TikTok, Facebook, Twitter, and Instagram.

One of the most popular digital communication sites is YouTube, a video-sharing platform with over 2 billion monthly users. It has also become a popular place to consume news, making it an important platform for journalism scholars to research (Pew Research Center, 2020).

A study by Pew Research Center (2020) showed that 26% of U.S. adults use YouTube as a news source, and of those that do, 72% say that it is "an important way" or "the most important way" they keep up with news. Additionally, most people who consume news on YouTube do not see misinformation, political bias, censorship, or the tone of news discussions to be "very big" issues for the platform (ibid., 2020).

Although consumers may not recognize political bias as an issue, a memo published by the Oxford Internet Institute analyzed coronavirus news on YouTube and found that highly politicized content facilitated the most engagement in the form of comments (Au, Howard, and Marchal, 2020).

Many news consumers don't appear to differentiate between news organizations and independent news producers while using YouTube, and they say they use the platform because it provides access to sources and opinions that are excluded from mainstream media (ibid., 2020). Simultaneously, the YouTubers that create news content find it important to differentiate themselves from mainstream news (Lewis, 2018).

In their research, Pew Research Center (2020) focused primarily on news organizations and independent news producers, but also described a third category of news producers on YouTube that they refer to as “other organizations.” In their analysis, only 9% of news producing channels on YouTube belong to this category. These channels produce news but have a clear affiliation with an external entity that is not a news organization (ibid., 2020).

One popular channel that falls into this “other organizations” category is Kurzgesagt, a science communication channel based in Munich, Germany that publishes content in English, Spanish, and German. Kurzgesagt publishes content covering several science topics relevant to space, technology, biology, philosophy, physics, and society. This thesis aims to analyze how consumers are engaging with Kurzgesagt’s YouTube content by comparing post-video comment sections on identical science videos published in multiple languages. The goal of this thesis is to analyze what differences in emotional engagement exist between videos published in different languages that communicate scientific information about animal agriculture’s impact on climate change – a topic that evokes strong emotion among consumers and has historically received a lack of coverage from mainstream media. The results of this thesis imply that journalists and science communicators looking to facilitate engagement with multiple audiences in different languages should expect those audiences to have differing levels of emotional engagement – even when the content delivered is identical. However, further research is needed to account for exogenous variables.

For this thesis, emotional engagement is defined as using language that expresses positive or negative sentiment in YouTube's comment section. It is important to note that while this study focuses on language, there are likely other sociological factors that could explain differences in emotional engagement regarding animal agriculture – such as age, political orientation, gender, or religious affiliation.

The literature review (Chapter 2) for this study is organized into six sections. First, I discuss traditional mainstream media's lack of coverage on animal agriculture's relationship with the environment and why this might be the case. Second, I give an overview of journalism's transition from traditional media to digital media. Third, I discuss environmental science communication on YouTube. Fourth, I give a brief overview of research that explores differences in engagement with science between cultures. Fifth, I introduce the outcome variable — emotional engagement. Finally, I discuss the popular science communication channel Kurzgesagt and the research questions relevant to this study.

Following the literature review, I discuss the methods (Chapter 3) used for analysis and the results of the analysis (Chapter 4).

Finally, the thesis closes with a discussion of the research process, limitations, ideas for future research, and a conclusion (Chapter 5).

CHAPTER 2: LITERATURE REVIEW AND RESEARCH QUESTIONS

Animal Agriculture, Climate Change, and Mainstream News Coverage

The animal agriculture industry plays a significant role in the ongoing climate crisis. It is estimated that 14.5% of human-induced emissions come from the animal agriculture sector, and this percentage is expected to grow as global meat consumption increases alongside human population growth (Food and Agriculture Organization of the United Nations, 2006, 2011, 2018; Godfray et al., 2018; Kristiansen, Painter, and Shea, 2020). Animal agriculture has also been linked to deforestation, ocean acidification, biodiversity loss, and health risks in humans (Poore and Nemecek, 2018).

Even though there is an abundance of research showing these links and offering alternatives – such as switching to a plant-based diet – there appears to be a lack of public awareness. In a survey of 12,000 participants across twelve countries, Bailey et al. (2014) found that "across all the emissions sectors asked about in the survey, recognition of the livestock sector as a contributor to climate change was markedly the lowest" (p. 18) and "one-quarter of respondents overall stated that meat and dairy production contributes either little or nothing to climate change" (p.19).

One proposed reason for this awareness gap is the lack of attention given to the link between animal agriculture and climate change by mainstream media. In one quantitative media content analysis of UK and US elite media from 2006 to 2018, results showed that "continuously low media attention" was given to animal agriculture's contribution to greenhouse gas emissions (Kristiansen, Painter, and Shea, 2020). In an earlier content analysis of leading newspapers from Spain and Italy, Almiron and Zoppeddu (2014) found that only 1.4% (Spain) and 3.6% (Italy) of articles on climate change published between 2006 and 2013 mentioned the impacts of meat consumption. The results of these content analyses are in line with previous research which found that only 2.4% of climate change articles in US newspapers mentioned food, farming, or agriculture (Neff, Chan, and Smith, 2009).

There have been several reasons proposed by researchers for the lack of media attention devoted to animal agriculture's role in climate change. Almiron (2020) proposes that journalism's deontological codes don't allow for objective communication of animal agriculture and suggests that mainstream media ethics be rewritten entirely before journalists can accurately portray the human relationship with animal agriculture.

Neff, Chan, and Smith (2009) take a less radical stance, suggesting that individual journalistic interest may play a role since, in their analysis, many climate change articles within a publication are written by a single journalist. They also suggest that media editors may view food consumption to be an individual choice and therefore less newsworthy. Neff, Chan, and Smith (2009) also point to the conflicts between environmental advocacy groups and industry as difficult to navigate and potentially alienating to portions of the public.

There is also evidence to suggest that the topic of animal agriculture is controversial, and evokes high levels of emotion, which may be why mainstream media avoid the conversation. Kaul, Schröger, and Humm (2020) examined YouTube videos on this topic and reported that the video that received the most negative comments and in-depth engagement discussed the connection between the consumption of animal products and climate change. Similarly, a content analysis of tweets surrounding the Intergovernmental Panel on Climate Change's Special Report on Climate Change and Land found that meat consumption was the most controversial topic, with high levels of toxicity and increasingly polarized narratives (Sanford et al., 2021).

Since mainstream media have historically devoted little attention to animal agriculture's impact on climate change, individuals have had to seek out other sources of information on the topic, such as blogs and social media. As mentioned in the introduction, survey data suggests that people use YouTube to seek out perspectives excluded from mainstream media sources (Pew Research Center, 2020). Therefore, the YouTube videos chosen for this study cover animal agriculture and its environmental impacts to get insight into how consumers engage with content regarding this underreported topic.

Journalism and Digital Media

News consumption has changed drastically since the days when newspapers, radio, and television were the only channels to consume news. In those days, there was a unidirectional flow of media content in which a select few provided daily information to the public. With the emergence of the internet, individuals were given the opportunity to become producers of information by developing their own websites, but the flow of information remained unidirectional. This phenomenon is referred to as Web 1.0 (Alejandro, 2010).

Internet users are now operating in what is called Web 2.0. The Web 2.0 model differs from Web 1.0 due to its "openness, organization, and community" (ibid., p. 5). Internet users in Web 2.0 are not only consuming information published on the internet, but can comment on, redistribute, and create their own content.

The barrier to entry is also lower than it has ever been. Members of the public no longer need a printing press, a broadcasting company, or extensive coding knowledge to publish their own ideas. It is free to comment on and redistribute content and increasingly cheap for an individual to produce content of their own. With over 6 billion smartphones worldwide (Statista, 2022), roughly 83% of the global population can participate in the new media landscape. This is astonishing considering the extreme barriers to participation only a few decades ago.

For journalism, this means that wealthy media conglomerates are no longer the sole gatekeepers of public information. Citizen journalists have already changed the way journalism is conducted, using platforms like Twitter and Facebook to spread information about protests, discrimination, and wars quicker than mainstream media outlets. For example, in July 2009, early news of the Bali bombings was published by a Twitter user before being reported by mainstream media outlets (Agence France Presse, 2009; Alejandro, 2010). The opportunity for citizens to conduct their own journalism means that the public no longer rely solely on mainstream media to shape their worldview.

Of course, the emergence of new technologies and a participatory media landscape brings new challenges. Social media platforms have been used in mass disinformation campaigns, where many users participate in the redistribution of false information regarding politics, health, and several other topics. Recent examples include the misinformation surrounding COVID-19 (Ferrara, Cresci, and Luceri, 2020) and the 2016 and 2020 U.S. Presidential elections (Chen et al., 2021; Tucker et al., 2018).

During the 2016 U.S. Presidential election campaign, teenagers in Macedonia produced disinformation articles about Donald Trump and Hillary Clinton across at least 100 sites, earning tens of thousands of dollars (Marwick and Lewis, 2017; Subramanian, 2017; Tucker et al., 2018). Although purely motivated by financial reasons, these disinformation campaigns had long-lasting effects on the political landscape.

Similarly, conspiracy theories flourish on social media and are often perpetuated by mainstream media or political leaders (Byford, 2011; Marwick and Lewis, 2017; Tucker et al., 2018; Wiggins, 2017). An example of this would be the relationship between Alex Jones' website Infowars and Donald Trump, where Trump "frequently amplified conspiracy theories"

(Tucker et al., 2018, p. 26) that had been perpetuated among Jones and his audience (Marwick and Lewis, 2017).

There have also been concerns regarding the role of social media algorithms pertaining to disinformation and radicalization. YouTube has been a target of study due to its recommended video algorithm, which is responsible for 70% of total watch time on the platform (Hao, 2019; Solsman, 2018). Some scholars have found evidence that YouTube's recommended algorithm is a pathway for radicalization of ideas (Alfano et al., 2020; Ribeiro, 2020), while others argue that the algorithm discourages users from engaging with radical content (Hosseinmardi et al., 2021; Ledwich and Zaitsev, 2019). Either way, YouTube videos are making an impact on the media and information environment that the global public now operates within.

As mentioned in the introduction, over a quarter of U.S. adults use YouTube as a news source and, for the majority of those that do, it is "an important way" or "the most important way" they consume news and information (Pew Research Center, 2020).

Science Communication on YouTube

The internet has not only changed the way that the public consumes news and information regarding politics, but also news and information regarding science (Velho, Mendes, and Azevedo, 2020). YouTube is a particularly important component of the spread of online science information due to the platform's 2 billion active monthly users (Cooper, 2019; Velho, Mendes, and Azevedo, 2020).

Research into popular science videos on YouTube is relatively new and there is no standard approach to conducting research on the platform. A popular science video can be defined as a "short video that focuses on the communication of scientific contents for a broad audience on the Internet" (Morcillo, Czurda, and Trotha, 2015, p. 1).

One of the first analyses conducted on YouTube science videos was carried out by Morcillo, Czurda, and Trotha (2015), who sought to identify the most popular science video channels, produce a typological study on aesthetic and narrative trends, and provide context for network analysis. Although their sample included a couple of news channels, such as Euronews Knowledge and Northwestern News Center, most of the channels analyzed were non-journalistic in nature. Their analysis found that while there are popular videos in languages other than English, the global list is dominated by the English language and that most science communication videos can be placed into one of three categories: (1) documentary, (2) animation, or (3) reportage. Their analysis also found that many of the science videos contained components found in professional productions, such as special FX, montages, studio lights, external sound devices, and reoccurring intro and outro sequences. Morcillo, Czurda, and Trotha (2015) concluded that the most significant aspect of their analysis is that many YouTubers are storytelling experts – a description often applied to journalists.

Other research into YouTube science videos has explored how video features and video metrics affect their popularity on the platform. Welbourne and Grant (2015) conducted a content analysis of 390 videos across 21 professionally generated YouTube channels and 18 user-generated YouTube channels and found that user-generated content is much more popular than professional content. Their analysis also showed that having a regular commentator and delivering information at a rapid pace resulted in higher video views.

A complementary study conducted by Velho, Mendes, and Azevedo (2020) analyzed 441 videos from a group of Brazilian science communication channels that were a part of the ScienceVlogs Brasil project. Their analysis explored how other video features and metrics affect video popularity: video theme, video format, number of comments, video age, channel productivity, and the channel responsible for uploading. Velho, Medes, and Azevedo (2020) determined that the most popular science videos in their sample were vlogs, animated videos, and group conversations that discuss interdisciplinary themes targeted toward a broader audience. They also found that newer videos with more likes and comments performed better than older videos with less user engagement (ibid., 2020).

Additionally, science communication videos that include dramatic questions, changes in narration techniques, and evoke emotional arousal perform better on the platform than those that do not (Huang and Grant, 2020).

In another study, 26 science YouTubers who were actively involved in the ScienceVlogs Brasil project were surveyed to analyze their sociodemographic data, their relationship with science communication, their relationship with YouTube, and strategies for communicating science on YouTube (Velho and Barata, 2020). Survey results showed that the majority of the science YouTubers were men between the ages of 18 and 35, were in the process of completing or already had completed a higher education degree, and worked in education (ibid., 2020).

The top motivation for these YouTubers to communicate science was "the perceived need of the population to be 'educated' about scientific issues" (ibid., 2020, p. 8), but they were also motivated by the need to combat misinformation, getting the lay public interested in science and academia, and the joy that communicating science brings them.

In a semantic analysis of user comments posted to YouTube videos covering climate change, Shapiro and Park (2014) found that the science content discussed in the videos was politicized by viewers in the post-video comment section. This was the case even when the video did not link climate change and politics. Shapiro and Park (2014) concluded that the specific content of the video, as well as the facts presented in the video, had little connection to the post-video discussion.

Similarly, an analysis of comments posted to videos from the #EarthOvershootDay campaign on YouTube found a lack of in-depth engagement with the scientific content discussed in the videos (Kaul, Schrögel, and Humm, 2020). The #EarthOvershootDay campaign was a collaboration between influencers on the platform, a German education initiative, the World Wide Fund For Nature, and the Robert Bosch Foundation. Due to the lack of engagement and low views relative to other videos published on the participating channels, Kaul, Schrögel, and Humm (2020) concluded that cooperating with influencers might not be the most effective way to communicate science on the platform.

Dubovi and Tarak (2020), however, found evidence of knowledge construction in the post-video comment sections from leading science communication channels on YouTube. After an analysis of 1,530 comments, they concluded that YouTube has potential to be a rich domain for informal learning about science concepts (Dubovi and Tarak, 2020).

Previous studies have explored science video production, motivations, popularity, consumption, and post-video engagement. While there has been research into non-English language YouTube videos, most of the content analyzed has been in English. To my knowledge, there has not been a study conducted that compares science videos published on YouTube in multiple languages. This thesis aims to fill that gap in the literature by analyzing the sentiment expressed within post-video comment sections for identical science videos published to YouTube in multiple languages.

Language and Social Media Use

As mentioned above, there has been little research analyzing YouTube users who engage with the platform using non-English languages. However, there is some previous research that suggests that language matters when it comes to social media. In one computational analysis of 62 million tweets, researchers found that there were cross-language differences in how users engaged with the platform, namely through features such as mentions, replies, and hashtags (Hong, Convertino, and Chi, 2021).

In another analysis, researchers compared how English, Spanish, and Portuguese audiences engaged with Zika-related information on Facebook and Twitter and found meaningful differences in audiences' use of social media platforms regarding expressed blame and topics discussed (Wirz et al., 2018). Wirz et al. (2018) also highlighted the importance of multilingual approaches in future communication research. This thesis aims to add to the literature regarding multilingual approaches in communication research.

Engagement with Science Between Cultures

Off social media, there have also been studies that compare emotional engagement with (Fleer at al., 2016) and attitudes toward science across cultures (Allum et al., 2008). While studying early childhood engagement with science among refugees, Fleer et al. (2016) found that it was challenging to explain Western science concepts using words in the Dinka language, but that discussing everyday routines through a scientific lens amplified emotional engagement with science.

In another study, after conducting a meta-analysis of public knowledge and attitudes about science, Allum et al. (2008) found that while there are some differences in attitudes toward science across cultures, most of the difference can be attributed at the individual level, rather than the country-level. Both studies were interested in engagement with science among groups who speak different languages or reside in different countries.

Additionally, scholars in the field of science communication recognize the disproportionate amount of English content and have made arguments for more inclusive science communication that involves non-English perspectives (Márquez and Porras, 2020).

Since this thesis is concerned with videos discussing the production and consumption of meat and dairy products, it is also important to consider how people who speak different languages engage with this specific topic and not just science in general.

For example, one survey showed that Hispanic/Latino Americans reported that meat was more important to their culture's traditional foods than non-Hispanic White Americans (Ellithorpe et al., 2021). It is possible, then, that these feelings may impact engagement with science communication that includes discussions regarding the production and consumption of meat. This thesis does not aim to answer if this is why emotional engagement differs, only to assess if differences in emotional engagement exist between audiences.

Automated Text Analysis and Emotional Engagement

While social media is becoming an increasingly popular way for the public to engage with science, research on how this engagement occurs is still in its infancy (Dubovi and Tabak, 2021). For starters, "engagement" does not have an agreed upon conceptualization in academic literature, making it difficult to pursue as a research avenue (ibid. 2021). While behavioral engagement can be operationalized through actions on social media such as viewing, sharing, and liking, emotional engagement is more difficult to quantify.

Emotional engagement is typically studied on social media through computational techniques that infer comment sentiment and emotional intensity (Dubovi and Tabak, 2021; Huang and Grant 2020; Mohammed 2016). Other research utilizes a qualitative approach, where researchers look at the context surrounding the comment in more depth (Dubovi and Tabak, 2021; Gasper et al. 2016). While a qualitative approach has the advantage of detecting nuance and contextual cues, quantitative analysis “offers greater consistency and reliability” (Dubovi and Tabak, 2021, p.762).

In an analysis of 89,000 comments posted to trending science videos on YouTube, Dubovi and Tabak (2021) sought to analyze the frequency of positive and negative comments as well as how eight different emotions were expressed: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. Results showed that there were significantly fewer negative comments than positive or neutral comments, and that trust was the most commonly expressed emotion (ibid. 2021).

This thesis aims to extend this line of research by analyzing whether emotional engagement with popular science videos on YouTube varies between the comment sections for identical science content published in different languages.

For this thesis, emotional engagement refers to expressed sentiment within a comment section. Expressed sentiment is calculated using an automated process that counts the positive and negative words present in each comment. This is explained in more depth in the method section (Chapter 3).

Kurzgesagt

One of the most popular science channels on YouTube is Kurzgesagt – In a Nutshell, a science communication channel with nearly 18 million subscribers and 2 billion views. It has even been recommended in *Nature* as “bite-size science for beginners,” with the author stating that “Kurzgesagt’s creators have science communication down to a tee” (Kareh, 2021). Their animated science videos are created and published by a design company based in Munich, Germany.

Kurzgesagt publishes animated science communication videos on three YouTube channels – In a Nutshell, Dinge Erklärt, and En Pocas Palabras – where videos are published in English, German, and Spanish, respectively.

Kurzgesagt's website says that the goals of their YouTube channels are to share knowledge, make a positive impact on the world, and tell a story using the facts (Kurzgesagt, nd). These goals are similar to those shared by journalists.

Many videos are published on all three channels, but some are not. The reasons for this are not explicitly stated but could be due to factors such as time and labor spent to translate videos or perceived lack of cultural relevance. When videos are published to multiple channels, the animations are identical and the script is translated but presents the same content in the same order.

For this analysis, I will only be looking at select videos published to all three channels. The content of the videos chosen for analysis discuss the science surrounding the production and consumption of animal products – a topic historically given a lack of coverage by mainstream media.

Research Questions

Based on the literature discussed above, I have proposed the following research questions:

RQ1: What differences in emotional engagement, operationalized by expressed sentiment, exist between comments posted to identical scientific content published to YouTube in English, Spanish, and German?

RQ2: What differences in emotional engagement, operationalized by expressed sentiment, exist between comments posted to scientific content published to YouTube covering different topics in the same language?

CHAPTER 3: METHOD

Sampling

Videos chosen for this analysis had to meet two requirements: the video must be available on all three Kurzgesagt channels and the content of the video must pertain to the production and consumption of animal products. Only two videos met these requirements, so six videos total were used for further analysis. The URL for each video can be found in the Appendix. Publishing data for each video is available in Table 1.

Table 1: Publication data for sample of videos used for analysis.

Channel	Published	Title	Views	Comments
In a Nutshell	Sep 30, 2018	Why Meat is the Best Worst Thing in the World	11,425,760	78,070
In a Nutshell	Jan 26, 2020	Milk. White Poison or Healthy Drink?	16,410,355	46,070
Dinge Erklärt	Jan 24, 2019	Fleisch – Das leckerste Übel der Welt	1,640,067	9,946
Dinge Erklärt	Oct 23, 2019	Milch – So ungesund ist sie wirklich	2,350,510	8,334
En Pocas Palabras	Sep 9, 2020	¿Por qué la carne es la mejor peor cosa del mundo?	1,020,924	6,128
En Pocas Palabras	Aug 11, 2021	Leche: ¿veneno blanco o bebida saludable?	1,235,386	3,196

All comments were extracted from each video using the YouTube API before conducting further analysis. However, there is a discrepancy between the total number of comments the YouTube platform displays and the total number of comments that the YouTube API extracts. For example, the video regarding meat production published in Spanish displayed a total of 6,282 comments, but the YouTube API only pulled 4,936 comments (as of April 1, 2022). This was the case even after multiple attempts to pull more comments. Therefore, there is either a discrepancy between the live site and the API data or there are comments that have been deleted from the platform but still show up in the counter. This was true for each of the comment sections used for analysis.

A much higher quantity of English comments was collected during this step due to the English channel having more views than the German and Spanish channels.

The YouTube API allows users to extract a limited number of units per day, so the data was collected over the course of multiple days. The comments posted to the German and Spanish videos were collected on the first day of data retrieval, while the comments posted to the English videos were collected on the second and third days of data retrieval.

Once all possible comments were pulled from the video platform, the data for each individual comment section were separated into three groups: (1) initial period, (2) middle period, and (3) recent period. This was done by dividing the number of comments into three equal sections as opposed to choosing comments before or after certain dates. This route was chosen because the videos were all published at different times, so different dates would've had to be chosen for each comment section, even if the videos were covering the same topic. This stratification was done to analyze how the conversation evolves over time regarding sentiment within each individual comment section. Although arbitrary, the time-series analysis was conducted to provide more insight into the expressed emotional engagement for each video. Limitations of this approach are discussed in Chapter 5.

Data Preprocessing

Since comments were collected in three languages (English, German, and Spanish), translation from Spanish and German into English was done before conducting further analysis. To achieve this, the document translator function of the Google Translate API was used. The Google Translate API has been shown to be an effective way to translate multilingual data sets into a monolingual data set for comparative analysis (de Vries, Schoonvelde, and Schumaker, 2018; Hase et al., 2021; Reber, 2018; Windsor et al., 2019).

The data were translated into English for two reasons: (1) machine translations perform best when translating to and from English since it is the lingua franca of the internet (de Vries, Schoonvelde, and Schumaker, 2018) and (2) English is my native language, making it easier to conduct further analysis.

After translation, but prior to analysis, data cleaning took place. Data cleaning, or data preprocessing, is conducted to remove any noise from the text data to get the most relevant data set possible. This includes converting characters to lowercase, conducting tokenization, and removing stop words, punctuation, and usernames. Tokenization is the process of breaking down a text string into smaller units, or tokens, such as words and phrases. Stop words are words that should be removed because their inclusion would create noise within the data set. These include words that are incredibly common for the language being analyzed, such as ‘the’ or ‘and’. This process was conducted using the tidytext package in R (Robinson and Silge, 2016, 2017).

Emotional Engagement

To assess the level of emotional engagement, an automated sentiment analysis was conducted, also using the tidytext R package, in a manner similar to a previous analysis of comments posted to popular science videos conducted by Dubovi and Tabak (2021).

While Dubovi and Tabak used the Syuzhet R package to report sentiment polarity and categorize data into emotional types, the tidytext R package was used for this thesis. Both packages rely on the NRC Word-Emotion Association Lexicon (Mohammad and Turney, 2013) to categorize text data into eight emotional types: anger, anticipation, disgust, fear, joy, sadness, surprise, or trust. Text data is also categorized into positive or negative expression of sentiment. The NRC lexicon has been utilized in several studies and found reliable (Dubovi and Tabak, 2021; Ragupathy and Maguluri, 2018; Widyaningrum et al., 2019; Yoon et al., 2017).

For this analysis, an extra column was created called “expressed sentiment,” which calculated the sum of the positive and negative scores for each individual comment. Additionally, the eight emotional types were filtered out because there are multiple languages involved in this analysis and there was likely meaningful emotional information lost during translation.

Analysis

Descriptive statistics for each video were computed using the filter, summary, and emmeans functions in R. Repeated analysis of variance (ANOVA) tests were conducted to explore the differences in expressed sentiment between the comment sections of the Kurzgesagt videos and the expressed sentiment over time within the individual comment sections.

CHAPTER 4: RESULTS

A sentiment analysis was conducted on 126,619 comments taken from the comment sections of six YouTube videos published by the animation and design studio, Kurzgesagt. Of the 126,619 comments analyzed, only 82,474 comments expressed sentiment when using the NRC lexicon. An example of a comment which expresses sentiment and one that does not express sentiment according to the NRC lexicon can be seen in the comment section for the English video covering the production and consumption of meat: “Urgh … Feel sick to my stomach watching this and thinking about this” (expressed sentiment) and “Nice bobs burgers reference” (no sentiment). These 82,474 comments were used for further analysis. Figure 1 shows the distribution of comments between the videos.

Figure 1: Total comments with expressed sentiment.

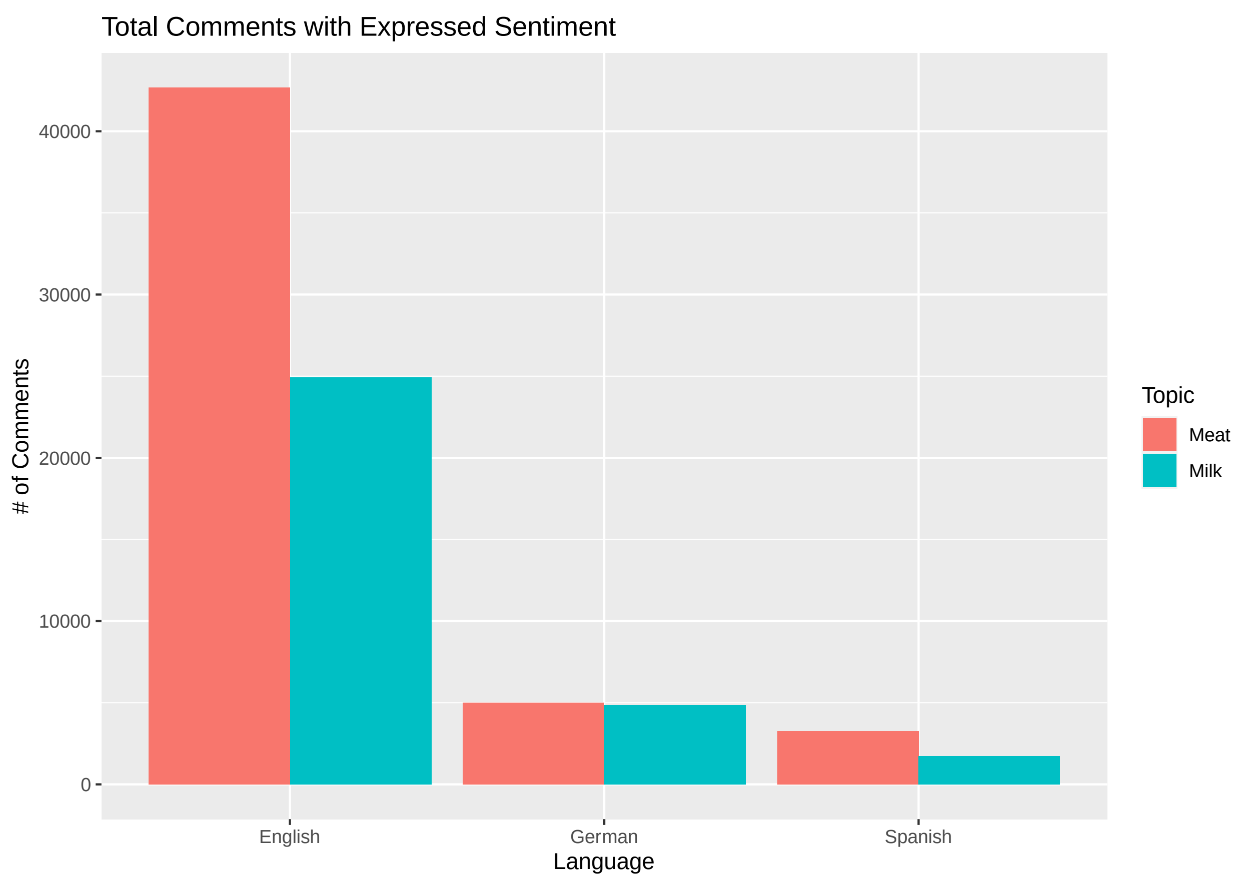
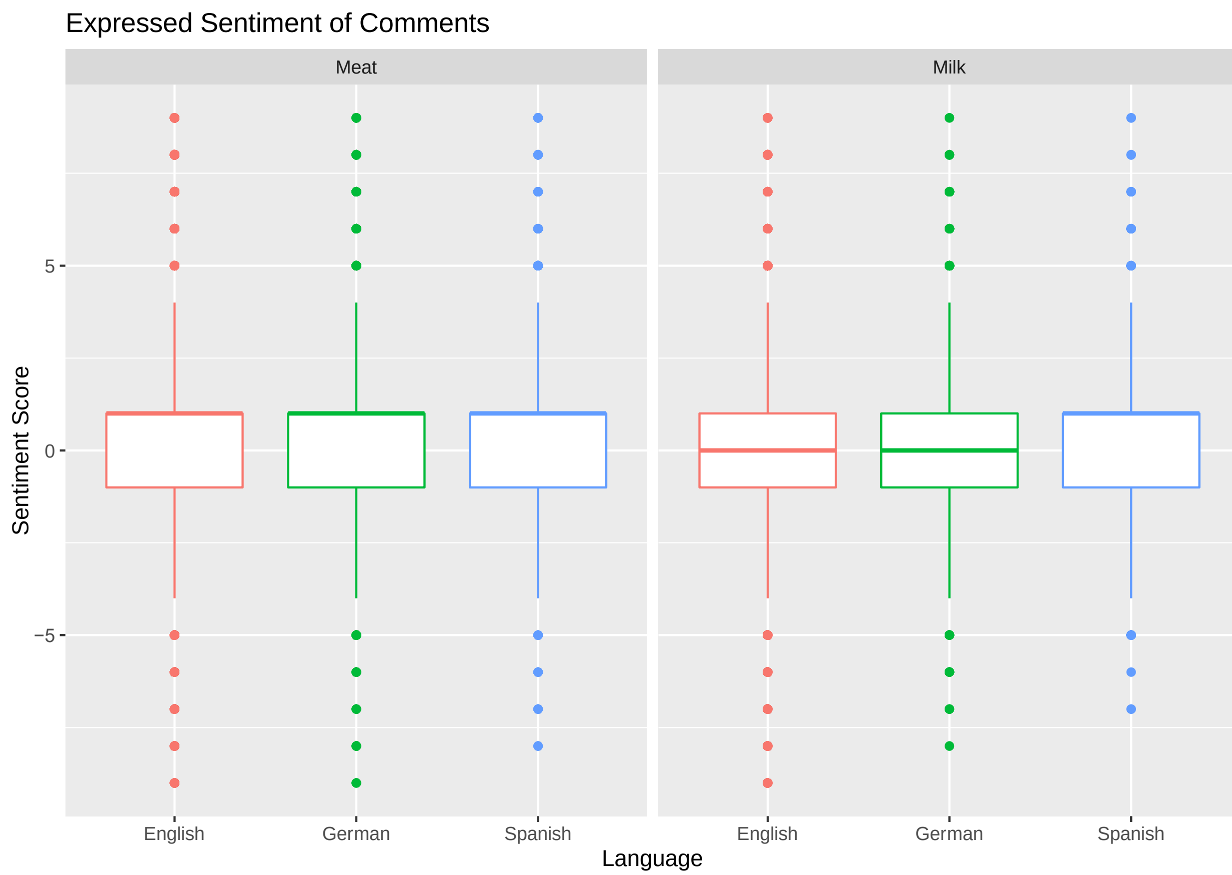


Figure 2 shows a boxplot of the sentiment scores of all six YouTube videos, with outliers cropped out for clarity.

Figure 2: Boxplot of expressed sentiment scores for videos.



English Videos

The English video covering the production and consumption of meat had a mean expressed sentiment of 0.48 ± 0.01 (SD = 2.25), with a minimum and maximum sentiment of -25 and 64, respectively. The English video covering the production and consumption of milk had a mean expressed sentiment of -2.45e-3 ± 0.01 (SD = 1.88), with a minimum and maximum sentiment of -75 and 39, respectively.

After conducting a t-test, results showed that the mean expressed sentiment for the comments posted to the meat video was significantly higher than the mean expressed sentiment for the comments posted to the milk video $t(59{,}697) = 29.70$, p-value < 0.001. Therefore, on average, more positive comments were posted to the English meat video than the English milk video.

Within the comment section of the English video covering the production and consumption of meat, the mean expressed sentiment for the initial, middle, and recent periods were 0.47 ± 0.02, 0.55 ± 0.02, and 0.41 ± 0.02, respectively. The middle period had significantly higher mean expressed sentiment than both the initial period, $t(42{,}658) = 3.14$, p-value = 0.005, and recent period, $t(42{,}658) = 5.09$, p-value < 0.001. There was no statistical difference between the initial and recent periods. Therefore, during the middle period of commenting, more positive expressions were posted than during the initial or recent periods of commenting. However, the number of positive expressions posted during the initial and recent periods of commenting were more similar.

Within the comment section of the English video covering the production and consumption of milk, the mean expressed sentiment for the initial, middle, and recent periods were 0.02 ± 0.02, 0.01 ± 0.02, and -0.05, respectively. The initial period had significantly higher mean expressed sentiment than the recent period, $t(24{,}940) = 2.40$, p-value = 0.04. There was no statistically significant difference between the initial and middle periods or middle and recent periods. Therefore, during the initial period of commenting, more positive expressions were posted than in the most recent period of commenting. However, the number of positive expressions posted during the initial and middle periods and the middle and recent periods of commenting were more similar.

German Videos

The German video covering the production and consumption of meat had a mean expressed sentiment of 0.61 ± 0.03 (SD = 2.30), with a minimum and maximum expressed sentiment of -27 and 26, respectively. The German video covering the production and consumption of milk had a mean expressed sentiment of -0.04 ± 0.03 (SD = 1.83), with a minimum and maximum sentiment of -18 and 24, respectively.

After conducting a t-test, results showed that the mean expressed sentiment for the comments posted to the meat video was significantly higher than the mean expressed sentiment for the comments posted to the milk video $t(9{,}525) = 15.51$, p-value < 0.001. Therefore, on average, more positive comments were posted to the German meat video than the German milk video.

Within the comment section of the German video covering the production and consumption of meat, the mean expressed sentiment for the initial, middle, and recent periods were 0.68 ± 0.06, 0.57 ± 0.06, and 0.58 ± 0.06, respectively. There were no statistical differences in the mean expressed sentiments between the three periods on the German meat video. Therefore, there was a comparable number of positive expressions posted during each of the time periods.

Within the comment section of the German video covering the production and consumption of milk, the mean expressed sentiment for the initial, middle, and recent periods were -0.22 ± 0.05, -0.04 ± 0.05, and 0.15 ± 0.05, respectively. The recent period had significantly higher mean expressed sentiment than both the initial period, $t(4{,}866) = 5.87$, p-value < 0.001, and middle period, $t(4{,}866) = 2.98$, p-value = 0.01. Additionally, the middle period had a significantly higher mean expressed sentiment than the initial period, $t(4{,}866) = 2.90$, p-value = 0.01. Therefore, the middle period had a higher number of positive expressions than the initial period, and the recent period had a higher number of positive expressions than both the middle and initial periods.

Spanish Videos

The Spanish video covering the production and consumption of meat had a mean expressed sentiment of 0.62 ± 0.04 (SD = 2.27), with a minimum and maximum expressed sentiment of -14 and 37, respectively. The Spanish video covering milk had a mean expressed sentiment of 0.55 ± 0.05 (SD = 2.05), with a minimum and maximum expressed sentiment of -7 and 18, respectively.

After conducting a t-test, results showed that there was no statistical difference between the mean expressed sentiment of comments posted to the meat video and the milk video, $t(3,856) = 1.10$, p-value = 0.27. Therefore, on average, the expressed sentiment of comments posted to the two Spanish videos analyzed was comparable.

Within the comment section of the Spanish video covering the production and consumption of meat, the mean expressed sentiment for the initial, middle, and recent periods were 0.48 ± 0.07, 0.45 ± 0.07, and 0.94 ± 0.07, respectively. The recent period had significantly higher mean expressed sentiment than both the initial period, $t(3,249) = 4.81$, p-value < 0.001, and middle period, $t(3,249) = 5.02$, p-value < 0.001. There was no statistical difference between the mean expressed sentiment for the initial and middle periods. Therefore, during the recent period of commenting, more positive expressions were posted than during the initial or middle periods of commenting. However, the number of positive expressions posted during the initial and middle periods of commenting were more similar.

Within the comment section of the Spanish video covering the production and consumption of milk, the mean expressed sentiment for the initial, middle, and recent periods were 0.53 ± 0.09, 0.32 ± 0.09, and 0.79 ± 0.08, respectively. The recent period had significantly higher mean expressed sentiment than the middle period, $t(1{,}733) = 3.93$, p-value < 0.001. There was no statistical difference between the mean expressed sentiment for the initial and recent periods. Therefore, during the recent period of commenting, more positive expressions were posted than during the middle period of commenting. However, the number of positive expressions posted during the initial and recent periods of commenting were more similar.

Meat Videos

Comparing the mean expressed sentiment of comments posted to the Kurzgesagt meat video across the three languages, the English video had a significantly lower mean expressed sentiment than both the German video, $t(50{,}923) = 4.02$, p-value < 0.001, and Spanish video, $t(50{,}923) = 3.61$, p-value < 0.001. There was no statistical difference between the mean expressed sentiment of comments posted to German and Spanish videos. Therefore, the expressed sentiment of comments posted to the English video covering the production and consumption of meat was less positive overall than those posted to the German or Spanish video. However, the expressed sentiment of comments posted to the German and Spanish videos were more similar.

Milk Videos

Comparing the mean expressed sentiment of comments posted to the Kurzgesagt milk video across three languages, the Spanish video had a significantly higher mean expressed sentiment than both the English video, $t(31{,}545) = 11.93$, p-value < 0.001, and German video, $t(31{,}545) = 11.24$, p-value < 0.001. There was no statistical difference between the mean expressed sentiment of comments posted to the German and English videos. Therefore, the expressed sentiment of comments posted to the Spanish video covering the production and consumption of milk was more positive overall than those posted to the German or English video. However, the expressed sentiment of comments posted to the German and English videos were more similar.

CHAPTER 5: DISCUSSION, LIMITATIONS, AND CONCLUSION

The primary goal of this thesis was to contribute to the literature regarding consumer engagement with digital science content published in multiple languages. This was done by conducting a sentiment analysis of 126,619 comments from YouTube videos published in three languages by science communication channel Kurzgesagt.

Expressed sentiment was analyzed both between comment sections and within comment sections. Within comment sections, the expressed sentiment for each individual video was compared over three time periods: initial, middle, and recent. For this thesis, the time periods were determined by splitting the comment section into three equal parts. This was done to see if there were any consistencies in how the emotional engagement evolved over time. Although there were some statistically significant differences within comment sections and between time periods, no meaningful patterns emerged across topic or language.

The results of the sentiment analysis showed that there were significant differences in the mean expressed sentiment for different languages (RQ1), but average sentiment for all videos was close to neutral, regardless of language or topic (RQ2).

In this sample, the English video covering the production and consumption of meat evoked more negative emotional engagement on average than the Spanish or German equivalent. This video also had the highest number of comments, so the results are in line with previous research suggests that more comments lead to more negativity (Dubovi and Tabak, 2021; Thelwall, Sud, and Vis, 2011). The Spanish video covering the production and consumption of milk evoked more positive emotional engagement on average than the English or German equivalent (RQ1). As suggested by previous research (Hall et al., 2020), these differences in engagement with information regarding the consumption of food may be connected to the cultural relevance of the food being discussed. Future research is needed to explore this potential connection in more detail.

In this sample, science videos discussing the production and consumption of meat evoked more positive expressions of sentiment than science videos discussing the production and consumption of milk for the English and German channels. This suggests that these audiences feel more negative about the discussion highlighting the links between the dairy industry and environmental and human health. The audience for the Spanish channel did not respond more positively to one video over the other, suggesting that the topics of meat and dairy evoke similar levels of emotional engagement (RQ2).

Interestingly, the milk videos for each language received more views but fewer comments than the respective meat videos for each language. The reason for this is unknown but previous research suggests that it could be due to the highly controversial nature of discussing meat consumption (Kaul, Schrögel, and Humm, 2020; Sanford et al., 2021). It is possible that the controversial nature of the topic means less people are inclined to view the video but are more inclined to comment if they do view the video.

While this thesis did not explicitly analyze expressions of trust, a qualitative look at the comments shows that both expressions of trust and claims of misinformation are present. For example, one comment from the English video regarding the production and consumption of meat reads, "But by only showing the harm of the meat industry they make everyone think a plant based diet is superior by default. Just look at this comment section. There's already so much negativity aimed at the meat industry. By omission Kurzgesagt has spread misinformation."

The results of this thesis are in line with previous research, which shows that YouTube is a useful platform to facilitate engagement with scientific content (Dubovi and Tabak, 2021). Journalists and science communicators can use YouTube to engage with a diverse audience from around the world. Additionally, the results of this research suggest that journalists and science communicators looking to publish science communication in multiple languages on YouTube should expect their audiences to emotionally engage with the videos at differing levels. This is important as scholars are calling for more inclusive science communication that includes non-western and non-English perspectives (Márquez and Porras, 2020).

Limitations and Future Research

As with any research, this thesis is not without its limitations. An early design of this thesis was focused on Twitter engagement, but the Twitter API would not allow me to pull the data required for my research questions. Therefore, I moved on to design a study for a different platform: YouTube.

The first limitation in the YouTube design was the initial sample size for comments posted to videos in different languages. Since Kurzgesagt's English channel receives many more views and comments than the Spanish and German counterparts, more English comments were analyzed for the final study. In future research, this limitation could be mitigated by designing a survey or lab experiment that doesn't rely on pre-existing data. However, researchers would lose the advantage of analyzing users and comments unprompted and in their natural settings.

Additionally, Kurzgesagt's videos were published at different times in different languages, sometimes by as much as two years. It is possible that events that occurred during the lag time between publishing could play a role in how YouTube audiences engage with the content that they consume. It would be beneficial for future researchers to mitigate this limitation by choosing videos that are published closer to one another, ideally at the same time.

This analysis also only included comparison between three languages: English, Spanish, and German. Future researchers could expand the relevancy of this work by including comments posted to science communication published in more or other languages. Similarly, future research could analyze comments posted to videos covering a wider array of science topics, instead of solely relying on videos covering the environmental and health risks associated with industrial animal agriculture. These could come from other videos published to Kurzgesagt's channels, or videos from a variety of channels on YouTube. It could also be fruitful to research comments posted to science communication on other popular social media channels, such as Instagram and Facebook.

As mentioned above, the milk videos for each language received more views but fewer comments than the respective meat videos for each language. Future research could test behavioral engagement based on the topic but should also be aware of the limitations of using comments, views, and likes as measures of engagement. A primary limitation would be the prevalence of bots on social media who can inflate likes, comments, and views.

The next limitations were ones of technology, including the YouTube API and Google Translate. Relying on the YouTube API allowed limited control over the number of comments available for the initial sample. The number of comments pulled using the YouTube API was not consistent with the number of comments on the live site for any of the comment sections analyzed for this study. This discrepancy was caused for an unknown reason but occurred for each video. Therefore, unless the YouTube site or the API fixes the bug, future researchers will likely run into a similar issue.

Google Translate made the methods utilized in this thesis possible, but still provided a limitation: because machine translation was used, I cannot be certain of the accuracy of the content of all the comments. This limitation could be overcome by analyzing comment sections in languages that the researcher is proficient in.

This thesis was concerned with emotional engagement, but future studies could research further into behavioral and cognitive engagement or the co-construction of knowledge, following the methods outlined by Dubovi and Tabak (2020, 2021).

Initially, this thesis was going to not only analyze the results of a sentiment analysis, but also the topics extracted from the comment sections after conducting topic modeling. This step of the method was dropped because it was too advanced for my level of skill with R.

However, future research could explore similar questions with more advanced methods that could lead to results with more practical implications. For example, topic modeling could provide useful information for researchers and science communicators, alike. It could also be lucrative to analyze the levels of expressed trust in scientific information between audiences who engage with content published in different languages to understand how best to communicate trustworthy science to diverse audiences.

Regarding expressions of trust, future research should explore the source, frequency, and validity of misinformation claims that are posted to science content on YouTube and what implications these claims may have for the public understanding of science.

Conclusion

In conclusion, this thesis explored how audiences engage with professionally generated science content on YouTube covering a topic that is underreported by mainstream media, namely the scientific risks associated with animal agriculture industries. Although the videos used in this study were not published by a traditional journalism outlet, the videos provided information about newsworthy topics on a platform that many adults use to consume news and make decisions about the world around them, making it relevant to the academic field researching digital media and journalism.

The results of this study suggest that journalists and science communicators interested in facilitating higher levels of positive engagement with science among multiple social media audiences who speak different languages should consider the cultural relevancy of their topic. Additionally, it could be beneficial to create multiple versions of their content, each one tailored to a specific audience. However, future research is needed to understand how to do this effectively.

The method utilized was novel in that it compared videos that already existed on the platform in question and were published in multiple languages, namely English, Spanish, and German. However, the method borrowed heavily from Dubovi and Tabak's (2020) analysis of engagement from comments. The attempt to compare emotional engagement with science content across different languages provides future studies with a starting point to conduct similar research.

BIBLIOGRAPHY

Agence France Presse. (2009, December 15). *“2009, the year of Twitter and Facebook"*. Retrieved March 12, 2022.

Alejandro, J. (2010). *Journalism in the age of Social Media*. Reuters Institute for the Study of Journalism. Retrieved March 12, 2022, from https://reutersinstitute.politics.ox.ac.uk/our-research/journalism-age-social-media

Almiron, N. (2020). The “animal-based food taboo.” climate change denial and deontological codes in journalism. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.512956

Almiron, N., & Zoppeddu, M. (2014). Eating meat and climate change: The Media Blind Spot—a study of Spanish and Italian press coverage. *Environmental Communication*, *9*(3), 307–325. https://doi.org/10.1080/17524032.2014.953968

Alfano, M., Fard, A. E., Carter, J. A., Clutton, P., & Klein, C. (2020). Technologically scaffolded atypical cognition: The case of YouTube’s Recommender System. *Synthese, 199*(1-2), 835–858. https://doi.org/10.1007/s11229-020-02724-x

Allum, N., Sturgis, P., Tabourazi, D., & Brunton-Smith, I. (2008). Science knowledge and attitudes across cultures: A meta-analysis. *Public Understanding of Science*, *17*(1), 35–54. https://doi.org/10.1177/0963662506070159

Au, H., Howard, P. N., & Marchal, N. (2020). Coronavirus news and information on YouTube. *Health*, *1*(1), 0-3.

Bailey, R., Froggatt, A., & Wellesley, L. (2014). In *Livestock-Climate Change's forgotten sector: Global public opinion on meat and dairy consumption: Research paper*. London; The Royal Institute of International Affairs.

Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003) Latent Dirichlet allocation. *Journal of Machine Learning Research, 3,* 993-1022.

Boy, B., Bucher, H.-J., & Christ, K. (2020). Audiovisual Science Communication on TV and YouTube. how recipients understand and evaluate science videos. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.608620

Byford, J. (2011). *Conspiracy theories a critical introduction*. Palgrave Macmillian.

Chen, E., Chang, H., Rao, A., Lerman, K., Cowan, G., & Ferrara, E. (2021). Covid-19 misinformation and the 2020 U.S. presidential election. *Harvard Kennedy School Misinformation Review*. https://doi.org/10.37016/mr-2020-57 de Vries, E., Schoonvelde, M., & Schumacher, G. (2018). No longer lost in translation: Evidence that google translate works for comparative bag-of-words text applications. *Political Analysis*, *26*(4), 417–430. https://doi.org/10.1017/pan.2018.26

Dorrance Hall, E., Ma, M., Azimova, D., Campbell, N., Ellithorpe, M., Plasencia, J., Chavez, M., Zeldes, G. A., Takahashi, B., Bleakley, A., & Hennessy, M. (2020). The mediating role of family and cultural food beliefs on the relationship between family communication patterns and diet and health issues across racial/ethnic groups. *Health Communication*, *36*(5), 593–605. https://doi.org/10.1080/10410236.2020.1733213

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

Dubovi, I., & Tabak, I. (2021). Interactions between emotional and cognitive engagement with science on YouTube. *Public Understanding of Science, 30*(6), 759–776. https://doi.org/10.1177/0963662521990848

Dustin Welbourne, & Will J Grant. (2015). *What makes a popular science video on YouTube*. The Conversation. Retrieved March 12, 2022, from https://theconversation.com/what-makes-a-popular-science-video-on-youtube-36657

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure INR. *Journal of Statistical Software*, *25*(5). https://doi.org/10.18637/jss.v025.i05

Ferrara, E., Cresci, S., & Luceri, L. (2020). Misinformation, manipulation, and abuse on social media in the era of covid-19. *Journal of Computational Social Science, 3*(2), 271–277. https://doi.org/10.1007/s42001-020-00094-5

Fleer, M., Adams, M., Gunstone, R., & Hao, Y. (2016). Studying the Landscape of Families and Children's Emotional Engagement in Science across Cultural Contexts. *International Research in Early Childhood Education*, *7*(1), 122-141.

Gaspar, R., Pedro, C., Panagiotopoulos, P., & Seibt, B. (2016). Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Computers in Human Behavior*, *56*, 179–191. https://doi.org/10.1016/j.chb.2015.11.040

Hao, K. (2019, September 27). *YouTube is experimenting with ways to make its algorithm even more addictive*. MIT Technology Review. Retrieved March 12, 2022, from https://www.technologyreview.com/2019/09/27/132829/youtube-algorithm-gets-more-addictive/

Hase, V., Mahl, D., Schäfer, M. S., & Keller, T. R. (2021). Climate change in news media across the globe: An automated analysis of issue attention and themes in climate change coverage in 10 countries (2006–2018). *Global Environmental Change*, *70*, 102353. https://doi.org/10.1016/j.gloenvcha.2021.102353

Hong, L., Convertino, G., & Chi, E. (2021). Language Matters In Twitter: A Large Scale Study. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), 518-521. Retrieved from https://ojs.aaai.org/index.php/ICWSM/article/view/14184

Huang, T., & Grant, W. J. (2020). A good story well told: Storytelling Components That Impact Science video popularity on YouTube. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.581349

Kareh, K. M. (2021). Bite-size science for Beginners. *Nature Physics*, *17*(8), 874–874. https://doi.org/10.1038/s41567-021-01320-1

Kaul, L., Schrögel, P., & Humm, C. (2020). Environmental science communication for a young audience: A case study on the #EarthOvershootDay campaign on YouTube. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.601177

Kristiansen, S., Painter, J., & Shea, M. (2020). Animal Agriculture and climate change in the US and UK Elite Media: Volume, responsibilities, causes and solutions. *Environmental Communication*, *15*(2), 153–172. https://doi.org/10.1080/17524032.2020.1805344

Ledwich, M., & Zaitsev, A. (2020). Algorithmic extremism: Examining YouTube's Rabbit Hole of Radicalization. *First Monday*. https://doi.org/10.5210/fm.v25i3.10419

Lewis, R. (2019). "This is what the news won't show you": YouTube creators and the reactionary politics of micro-celebrity. *Television & New Media*, *21*(2), 201–217. https://doi.org/10.1177/1527476419879919

Lucas, C., Nielsen, R. A., Roberts, M. E., Stewart, B. M., Storer, A., & Tingley, D. (2015). Computer-Assisted Text Analysis for Comparative politics. *Political Analysis*, *23*(2), 254–277. https://doi.org/10.1093/pan/mpu019

Márquez, M. C., & Porras, A. M. (2020). Science communication in multiple languages is critical to its effectiveness. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.00031

Marwick, A., & Lewis, R. (2017). *Media manipulation and disinformation online*.

Mohammad, SM (2016) 9—sentiment analysis: Detecting valence, emotions, and other affectual states from text. In: Meiselman, HL (ed.) *Emotion Measurement*. Cambridge: Woodhead Publishing, pp. 201–237.

Munoz Morcillo, J., Czurda, K., & Robertson-von Trotha, C. Y. (2016). Typologies of the Popular Science Web Video. *Journal of Science Communication*, *15*(04). https://doi.org/10.22323/2.15040202

Neff, R. A., Chan, I. L., & Smith, K. C. (2009). Yesterday's dinner, Tomorrow's weather, today's news? US newspaper coverage of food system contributions to climate change. *Public Health Nutrition*, *12*(7), 1006–1014. https://doi.org/10.1017/s1368980008003480

Poore, J., & Nemecek, T. (2018). Reducing Food's environmental impacts through producers and consumers. *Science*, *360*(6392), 987–992. https://doi.org/10.1126/science.aaq0216

Published by S. O'Dea, & 23, F. (2022, February 23). *Smartphone users 2026*. Statista. Retrieved March 12, 2022, from https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/

Reber, U. (2018). Overcoming language barriers: Assessing the potential of machine translation and topic modeling for the comparative analysis of multilingual text corpora. *Communication Methods and Measures*, *13*(2), 102–125. https://doi.org/10.1080/19312458.2018.1555798

Ribeiro, M. H., Ottoni, R., West, R., Almeida, V. A., & Meira, W. (2020). Auditing radicalization pathways on YouTube. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*. https://doi.org/10.1145/3351095.3372879

Sanford, M., Painter, J., Yasseri, T., & Lorimer, J. (2021). Controversy around climate change reports: A case study of twitter responses to the 2019 IPCC report on land. *Climatic Change*, *167*(3-4). https://doi.org/10.1007/s10584-021-03182-1

Shapiro, M. A., & Park, H. W. (2014). More than entertainment: YouTube and public responses to the science of Global Warming and climate change. *Social Science Information*, *54*(1), 115–145. https://doi.org/10.1177/0539018414554730

Silge, J., & Robinson, D. (2016). Tidytext: Text mining and analysis using tidy data principles in R. *The Journal of Open Source Software*, *1*(3), 37. https://doi.org/10.21105/joss.00037

Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly.

Solsman, J. E. (2018, January 10). *Ever get caught in an unexpected hourlong YouTube binge? thank YouTube AI for that*. CNET. Retrieved March 12, 2022, from https://www.cnet.com/tech/services-and-software/youtube-ces-2018-neal-mohan/

Stocking, G., Kessel, P. van, Barthel, M., Matsa, K. E., & Khuzam, M. (2021, November 16). *Many americans get news on YouTube, where news organizations and independent producers thrive side by side*. Pew Research Center's Journalism Project. Retrieved March 12, 2022, from https://www.pewresearch.org/journalism/2020/09/28/many-americans-get-news-on-youtube-where-news-organizations-and-independent-producers-thrive-side-by-side/

Subramanian, S. (2017, February 17). *The Macedonian teens who mastered fake news*. Wired. Retrieved March 12, 2022, from https://www.wired.com/2017/02/veles-macedonia-fake-news/

T.S, A., Batra, H., & Ramachandran, M. (2018). Comparative analysis of machine learning algorithms on IVR data. *EAI Endorsed Transactions on Energy Web*, *5*(18), 154814. https://doi.org/10.4108/eai.12-6-2018.154814

Thelwall, M., Sud, P., & Vis, F. (2011). Commenting on YouTube videos: From Guatemalan Rock to el big bang. *Journal of the American Society for Information Science and Technology*, *63*(3), 616–629. https://doi.org/10.1002/asi.21679

Tucker, J., Guess, A., Barbera, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., & Nyhan, B. (2018). Social media, political polarization, and political disinformation: A review of the scientific literature. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3144139

Velho, R. M., & Barata, G. (2020). Profiles, challenges, and motivations of science Youtubers. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.542936

Velho, R. M., Mendes, A. M., & Azevedo, C. L. (2020). Communicating science with YouTube videos: How nine factors relate to and affect video views. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.567606

Widyaningrum, P., Ruldeviyani, Y., & Dharayani, R. (2019). Sentiment analysis to assess the community's enthusiasm towards the development chatbot using an appraisal theory. *Procedia Computer Science*, *161*, 723–730. https://doi.org/10.1016/j.procs.2019.11.176

Wiggins, B. E. (2017). Navigating an immersive narratology. *International Journal of E-Politics*, *8*(3), 16–29. https://doi.org/10.4018/ijep.2017070102

Windsor, L. C., Cupit, J. G., & Windsor, A. J. (2019). Automated content analysis across Six languages. *PLOS ONE*, *14*(11). https://doi.org/10.1371/journal.pone.0224425

Wirz, C. D., Xenos, M. A., Brossard, D., Scheufele, D., Chung, J. H., & Massarani, L. (2018). Rethinking social amplification of risk: Social media and Zika in three languages. *Risk Analysis*, *38*(12), 2599–2624. https://doi.org/10.1111/risa.13228

Yoon, S., Parsons, F., Sundquist, K., Julian, J., Schwartz, J. E., Burg, M. M., Davidson, K. W., & Diaz, K. M. (2017). Comparison of Different Algorithms for Sentiment Analysis: Psychological Stress Notes. *Studies in health technology and informatics*, *245*, 1292.

YouTube. Kurzgesagt. (n.d.). Retrieved March 12, 2022, from https://kurzgesagt.org/youtube/