Online voting behaviour, public content preferences, and content complexity of movies and digital entertainment since the late 19th century

Aristides Moustakas (✉ arismoustakas@gmail.com)

Natural History Museum of Crete <https://orcid.org/0000-0002-6334-747X>

Mohammad Firdaus

Universiti Brunei Darussalam

Research Article

Keywords: digital entertainment content, movies, computer games, content complexity, voting behaviour, multivariate statistics, timeseries analysis

Posted Date: July 9th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-692921/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Analysis of cinema and electronic art contents into quantifiable characteristics could provide insights on important elements of human behaviour such as voting behaviour, appreciation, and emotions regarding content. Analysis of movies, series, TV programs, and video games, were performed mining the IMBD database and merging datasets. Data analytics were performed including the relationship between the number of votes and rating, rating and content as timeseries, and multivariate statistical analysis between content and rating as well as production type and rating. Content complexity (i.e. the number of different genre) was also analyzed regarding rating. Results indicated that there is an overall positive relationship between the number of votes and rating for low, intermediate, and high number of votes, while very high number of votes is related with a partisan voting behaviour resulting in negative rating. Adult rated content is declining over the past three decades. Violent content is expressed via other forms of content through time, switching from war to horror. Diversity of content was the highest rated content type linked with diversity and complexity of feelings and emotions. Among all production types, video games were the highest rated ones related with integrative social behaviour with other users or addiction.

Introduction

Quantifying art is admittedly a complex issue. For movies, films and digital entertainment in particular, quality assessments by experts (e.g. in Cannes) are short lasting (Ginsburgh & Weyers, 1999) as this assessment may change in time and in general they may fail to live up to public ranking (Haan, Dijkstra, & Dijkstra, 2005). Decomposing cinema and electronic art contents into quantifiable characteristics, perhaps in a subjective (Koh, Hu, & Clemons, 2010) but consistent and reproducible way, could facilitate explaining human preferences as well as divergences between audiences and changes of appreciation over time (Ginsburgh & Weyers, 1999). In addition, investigating public appreciation on art and entertainment provides insights on important elements of human behaviour such as voting behaviour, as well as appreciation and emotions regarding content.

A question that has always interested political, social, and natural sciences is how do people vote in terms of ethics and rationality (Brennan, 2012)? Public voting or rating is a complex phenomenon and requires an interdisciplinary approach (Lago, 2019). From a political perspective voting may often merit analysis as polarized or irrational (Achterberg & Houtman, 2006; Leduc, 2002). In terms of expert knowledge, one-off voters tend to vote on popular items, while experts mostly vote for little-known possibly low-rated items (Kostakos, 2009). From a physical perspective, the distribution of votes presents scale-free behavior over several orders of magnitude (Ramos, Calvão, & Anteneodo, 2015). This points to a general underlying mechanism for the propagation of voting across potential audiences that is independent of the intrinsic features of a movie (Ramos et al., 2015) and thus facilitating understanding elements of human behaviour in relation to voting behaviour. Public rating is further complicated by the fact that voting is not necessarily secret nor simultaneous by all users; while ignorance of previous views leads to a uniform sampling of the range of opinions among a group of people, exposure of previous opinions to potential reviewers induces a trend following process which leads to the expression of

increasingly extreme views (Wu & Huberman, 2008). In addition, public voting requires some time and effort that the individual who is voting may not necessarily benefit from directly, the paradox of voting (Downs, 1957).

What are movie and digital entertainment characteristics rated highly and thus liked by humans? Good scenarios, dialogues, content, production type, plot, soundtrack, acting, directing, and the cinematography play an important role (Jakob, Weber, Müller, & Gurevych, 2009; Mehr et al., 2019; Simonton, 2002). There are many movies and digital entertainment productions released every year and performing script analysis or determining image, colour, or sound characteristics of highly rated productions is interesting but computationally unfeasible. While clearly acting plays a key role, it is also hard to quantify. Simply including the actor's name is a poor determinant as even good actors can perform lower than their standards in some cases, or have minor roles, and their names are very hard to include as a categorical variable in the analysis due to their volume and time discontinuity. The content and production type are easier to quantify (Thompson & Yokota, 2004) and online databases exist collecting those data as well as the reciprocal rating and number of votes. In addition, what is a great movie is up to a point subjective. However, the analysis of large datasets provides useful insight into both the mean values of preferences as well as the variance around the mean simply by volume (Hills, Proto, Sgroi, & Seresinhe, 2019).

Is violence and explicit adult content increasing through time in productions? In a recent report to USA Congress on the accuracy of the age content ratings and effectiveness of oversight, the Federal Communications Commission noted that the system has not changed in over 20 years (FCC, 2019). However, another recent study for the USA indicated that violent content more than doubled, in the past 10 years, both in per-episode averages and in absolute terms (PTC, 2019) and thus the age-ratings fail to reflect "content creep" - an increase in offensive content in programs with a given age-rating as compared to similarly-rated programs a decade or more ago (PTC, 2019). In addition modern technological advances facilitate more complex productions in terms of visual and special effects as well as shooting locations and techniques. Trends over time in movie content (Redfern, 2012) could potentially be explained by technological advances, which favor profitability of certain types over others (Ji & Waterman, 2010). However, this may also be translating into content complexity, by including more content categories. How is content diversity appreciated by the public? Apart from the interest that content diversity merits on its own, inclusion of more diverse content can in part incorporate violent or sexual content into a production of different major content.

We sought to quantify elements of human behaviour through digital entertainment. In specific we addressed the following questions: (1) How does public rating relate with the number of votes? (2) Is adult rated content (violent and sexually explicit content) increasing over time? (3) Are simple or complex content genre appreciated more by the public? What is the content complexity trend over time? (4) What is the highest rated by the public content and production type? For addressing these questions we performed data mining followed by data analytics of rating, number of votes, content and production type of 799,736 productions from 1874 to 2018 using the Internet Movie Database (IMDB).

Methods

Data mining

Data mining (Moustakas & Katsanevakis, 2018) regarding movies (in cinemas), TV productions (TV shows, news, TV episodes, special TV productions, sports broadcasting) and digital entertainment productions (video games, video clips, video tapes, video clips on internet resources such as YouTube, movies and series on internet resources only such as Netflix) – thereby “productions” - were retrieved from the IMDB database on Jan 25th, 2018. All the raw data are freely available at: <https://www.kaggle.com/ashirwadsangwan/imdb-dataset>

The datasets employed in the analysis included:

“title.basics.tsv.gz” which contains the following information for productions as classified in IMDB:

- Tconst: alphanumeric unique identifier of the production title
- titleType: the production type. The production type classification included: *movie, short, video, TV movie, TV series, TV mini series, TV special, TV short, and video game*. Information regarding classification of production type is provided at: https://help.imdb.com/article/imdb/discover-watch/how-do-you-decide-if-a-title-is-a-film-a-TV-movie-or-a-miniseries/GKUQEMEFSM54T2KT?ref_=helpart_nav_21#.
- In brief, production types are classified into a category based on their duration (short or not), whether they were available on cinemas or exclusively on TV, or on the internet only, and whether they were continued as series. Video games were also included.
- primaryTitle: the more popular title / the title used by the producers on promotional materials at the point of release
- originalTitle: original title, in the original language
- isAdult: - 0 for non-adult rated production; 1 for adult rated production
- startYear: represents the release year of a production. In the case of TV Series, it is the series start year. The first year entry was in 1874 and the last 2018 (till 25th Jan).
- runtimeMinutes: primary runtime of the title, in minutes
- genres (string array) – includes up to three genres associated with the title. This information was used as a proxy of production’s content. The values included: *Documentary, Action, Adventure, Animation, Biography, Comedy, Crime, Drama, Family, Fantasy, Film-Noir, Game-Show, History, Horror, Music, Musical, Mystery, News, Reality-TV, Romance, Sci-Fi, Sport, Talk-Show, Thriller, War, and Western*. The definition of each content type is provided by the IMDB database: https://help.imdb.com/article/contribution/titles/genres/GZDRMS6R742JRGAG?ref_=helpms_helpart_inline#

Data mining was performed also in the “title.ratings.tsv.gz” dataset which contains the IMDb rating and votes information for each production

- **tconst**: alphanumeric unique identifier of the production, used to match the production’s rating and number of votes with the content and production type characteristics.
- **averageRating**: weighted average of all the individual user ratings. Each individual vote spans from 1 (lowest) to 10 (highest). The average rating is calculated by adding the total rating score by all users that voted and sequentially divided by the number of users that voted. Voting can be submitted at any time (year) since the production’s release.
- **numVotes**: the number of votes by unique registered users the production has received since the date it was released.

Number of votes & Rating

There were 4,766,401 entries of productions in the first data file (title.basics.tsv.gz) whereas the second data file (title.ratings.tsv.gz) contained 799,736 productions. The two data sets were merged together using the unique alphanumeric identifier of the production. The non-rated 3,966,665 productions were discarded resulting in a total of 799,736 rated productions including the total number of votes and the mean rating. The primary and original title, the end year and the run time in the raw data were not included into the final data set as this resulted into a dataset memory size that was impractical to perform statistical analyses.

Production type

The variable **titleType** was used to infer the production type. An amendment was made to the original data by merging the production’s type “TV episode” with “TV series”; this was done because the primary title that an individual episode appears in the original IMDB dataset is different than the series title, and thus each episode would appear as an entry under “TV episode” and the series would appear as a single entry under “TV series”. As different episodes of a series receive different public ranking and number of votes we retained “TV series” and discarded “TV episode” variable type, with each episode of the same series appearing as “TV series” with its unique alphanumeric unique identifier of the production title.

Content

The initial raw data included up to three non-exclusive “genre” words that were used as proxies of content. As in the original data the genre for each production were stacked together (e.g. comedy/romance), each genre was given a unique column and clustered in a Boolean form (False= 0, True =1) so that when a particular genre was tagged on a production in our classification the result will show “1” if that genre was tagged or “0” if it was not. For example a production containing comedy/romance will result in “1” in the comedy column as well as in the romance column and “0” in all other content category columns of that row. We have added a new column termed “adult” that indicated whether the production contained adult content (“1”) or not (“0”) as classified in the IMDB dataset. We

have also added an additional column termed “diversity” indicating whether the production has a single content or if the content is more diverse containing more than one category. This was achieved by summing the number of non-exclusive content (genre) tagged in the production (e.g. comedy and romance result in diversity = 2, comedy only result in diversity = 1).

Statistical Analyses

Number of votes and rating distributions were plotted, and their mean and median confidence intervals were calculated, as well as inter-quartile intervals. Rating was plotted per production type over time. A linear regression was fit in each production type to quantify linear trends as well as a smoother line to smooth out extreme values as well as quantify potential non-linear trends over time (You, Lin, & Young, 2018).

Content was plotted over each year (time) for each content category as a percentage of the total productions that the content category was present in all the productions of that year (e.g. total comedies in 1994 / total productions in 1994). Regarding adult rated in particular, adult rated productions were quantified as a percentage of total productions of that production year across years. War and horror genre were considered as highly likely to contain violent content. Productions with crime, action, and adventure genre were also considered as potentially containing violent content.

Quantile regressions (Koenker, 2005) were used between the number of votes (dependent variable) and rating (independent variable) to quantify the relationship between low to high number of votes and rating. Regressions included the following quintiles: 1, 5, 10, 25, 50, 75, 90, 95, and 99 % of the data. Quantiles are values that partition a distribution into q subsets of (nearly) equal sizes. Quantile regression aims at estimating the regression between quantiles of the response variable and the dependent variable. Essentially, quantile regression is the extension of linear regression when the conditions of linear regression are not applicable.

Change point detection algorithms were employed to detect changes in both the mean and variance of the average number of votes, average rating, average diversity, and the average adult rated productions per year analyzed as time series. The binary segmentation test statistic (Scott & Knott, 1974) was used to detect changes in the data. Change point detection methods allow the decomposition of complex non-stationary time series into segments where the mean and variance are constant, and thus such changes can be quantified (Moustakas & Evans, 2016). Average values of rating, votes, diversity, and adult content per year were used for this analysis. Strictly speaking, the number of votes and rating are not time series as voting and thus change in rating can be submitted at any time point after the release and not just during the production year, but they are treated as time series here to quantify trends over the rating and voting of production year.

Linear mixed effects models fitted with maximum likelihood estimation to allow for comparisons between models with different fixed effects (Moustakas, Daliakopoulos, & Benton, 2019; Pinheiro & Bates, 2000) were used to investigate the relationship between the rating of the movies (dependent variable)

and the influence of content in the rating as independent variables. Content variables included: diversity, documentary, type, action, adventure, biography, comedy, crime, drama, family, fantasy, film noir, game show, history, horror, music, musical, mystery, news, reality TV, romance, sci.fi, sport, talk show, thriller, war, western, and adult rated – all as factors. Time (production year) was also included as a random effect accounting for potential temporal autocorrelation in the data as well as the fact that the number of votes and the number of productions are varying across years. Model selection was performed using the AIC.

Linear mixed effects models fitted with maximum likelihood, were also used to quantify the effect of the production type on the rating. Rating was analyzed as a dependent variable and production types as independent factors. Production type variables included: movies, TV series, short, video, TV movie, video game, TV mini series, TV special production, and TV short production. Production year was also included as a random effect. Model selection was performed using the AIC.

Results

The number of productions is exponentially increasing, with less than 1,000 productions each year till 1934, less than 5,000 per year till 1964, reaching 10,000 productions per year in 1998, while in the 21st century there are more than 20,000 per year (Fig. 1a). In terms of content, productions have been dominated by drama, and comedies accounting together close to 38% of all content of productions so far (Fig. 1b). Documentary, action, animation, crime, and adventure are following as most commonly encountered content (Fig. 1b). The lowest content includes film noir, news, musical, war, and sports (Fig. 1b). In terms of production type across all years, productions have been dominated by TV series, despite the relative recent introduction of TV, followed by movies, and short productions (Fig. 1c). Productions disseminated as a video tape only are less than 50,000 so far (Fig. 1c). The lowest number of production types includes TV mini series, TV special productions, and TV short productions (Fig. 1c). The mean number of votes is not monotonically distributed across years; some early 19th century productions receive high mean number of votes, while productions between 1900 – 1915 receive a low mean number of votes (Fig. 1d). The cumulative number of votes reached 50 million in 1973, 100 million in 1985, and 750 million in 2016 (Fig. 1d).

Public rating and number of votes

Rated productions of the late 19th century are highly rated while rating declined in the early 20th century (Fig. 2a). Overall rating is exhibiting a complex behaviour both in terms of mean as well as variance (Fig. 2a). Movies and TV series are consistently rated through years with almost no trend (Fig. 2b). Short productions and video games exhibit an increased rating through years (Fig. 2b). TV mini series, TV special productions, and TV movies are decreasingly rated through years (Fig. 2b). Video and TV short productions are showing a negative trend when seen overall through all the data (linear regression; Fig. 2b); however their rating is increasing in the last year after 2000 as exhibited by a smoother line (Fig. 2b).

The number of votes has a median value of 20 and a mean of 992.18 per rated production (Fig. 2c). However, there are productions that receive over one million votes (Fig. 2c). Despite that, the 75% of productions receive a number of votes between 9 -79 (first and third quartile), while any number of votes below 5 or above 184 are statistically considered as outliers (values outside whiskers in Fig. 2c). Rating of productions has a median value of 7.1 and a mean of 6.94 (Fig. 2d). Values of rating between 6.2 – 7.9 included 75% of the data (first and third quartile respectively); (Fig. 3d). Values below 3.7 are statistically considered as outliers while no upper value is an outlier (whiskers 3.7 – 10; Fig. 2d).

The slope between the number of votes and the resulting rating was positive for all quintiles until (including) the 90th quintile indicating an increased rating with increased number of votes (Fig. 2e; Supplementary Table 1). The 1st, 5th, and 10th quintiles had higher slopes than all other percentiles (Fig. 2e; Supplementary Table 1), indicating a strong positive relationship between public rating and the number of votes when the number of votes is relatively low. The 95th and the 99th quintiles had negative slopes (Fig. 2e; Supplementary Table 1), indicating a negative relationship between public rating and the number of votes, when the number of votes is very high.

Adult rated and violent content

Adult rated film content was < 1% until 1907; then 1% of productions contained adult rated content between 1908 – 1930, and decreased to < 1% between 1931 – 1961 (Fig. 3). Adult rated content increased dramatically from 1962 peaking at 1983 when over 6% of all productions contained adult rated content (Fig. 3). Since 1984 adult rated content is generally decreasing (with a second peak at 1990) and adult rated content is < 1% after 2009 (Fig. 3).

Violent content as deduced by war content is peaking for short time periods between 1890-1900 and 1912-1920 with over 6% of productions containing war content (Fig. 3). War content is peaking between 1939 – 1946 with over 14% war content in all productions and steeply decreasing till 1966 – 1980 with war content close to 5%, while war content is decreasing thereafter with values < 1% after year 2000 (Fig. 3). Horror content is overall increasing from 1874 to 2018, starting from horror content <1% before 1900, peaking in 1960- 1964 to 8% horror content, while in the last decade 2008 – 2018 horror content is > 5% (Fig. 4). Adventure content is overall increasing from 1874 till the decade of 1980 and decreasing thereafter, but with an overall of 8% adventure content in all productions in the last decade of 2008 – 2018 (Fig. 3). Action production content is overall increasing from 1874 to 2018 with a relative decrease between 1995 – 2005, and bouncing back to the pre-1985 levels of 15% after 2005 (Fig. 3). Crime content is increasing from 1874 – 1985, decreasing between 1986 – 2008, and increasing in the last decade of 2008 – 2008, with crime content present in 15% of all productions (Fig. 3).

Change points in time

Change points in mean and variance of average rating through years detected a segment between 1874 – 1888, 1889 – 1916, 1917 – 1951, 1952 – 1954, 1955 – 1967, 1968 – 2010, and between 2011 – 2018 (Fig. 4a). Rating of early productions was higher, followed by low rating in the early 20th century which

gradually increased to peak in the 1970's, followed by a decline, while the latest productions are the highest rated so far (Fig. 4a).

The number of votes exhibited segments between 1874 – 1886, 1887 – 1910, 1911 – 1929, 1930 – 1970, and 1971 – 2018, with votes peaking initially for production in the late 19th century, followed by a decline in the second period and increased ever since (Fig. 4b).

Adult rated content exhibited segments between 1874 – 1897, 1898 – 1916, 1917 – 1921, 1922 – 1941, 1942 – 1958, 1959 – 1960, 1961 – 1997, and 1998 – 2018 (Fig. 4c). In practice there has been virtually no adult rated content till the second segment period, rated content declined after and peaked during 1961 – 1997, while it steeply declined thereafter (Fig. 4c).

Diversity of content had segments between 1874 – 1884, 1885 – 1898, 1899 – 1927, 1928 – 1932, 1933 – 2018 (Fig. 4d). Content diversity was higher in the rated productions of the early 19th century, declined thereafter till the third period which diversity is lowest, and increased thereafter, while it remains stable from 1933 till today (Fig. 4d).

Results from mixed effects models indicated that the optimal model included all content types as factors - the most parsimonious model was the maximal model, no content type elimination was justified based on the AIC (Supplementary Table 2). The highest rated productions in terms of content included high diversity content, documentary, and western (Fig. 5a; Supplementary Table 2). Mystery, fantasy, adventure, and drama content had a positive impact on the rating (Fig. 5a; Supplementary Table 2). Music content resulted in a marginal positive effect while all other content types resulted in a negative effect (Fig. 5a; Supplementary Table 2). The worst effects were recorded in production types that included horror, thriller, news, adult, and game show content (Fig. 5a; Supplementary Table 2).

In terms of production type, results from mixed effects models indicated that the optimal model included all production types - the most parsimonious model was the maximal model, no production type elimination was justified based on the AIC (Supplementary Table 3). The highest rating derived from video games, followed by TV series, and TV miniseries (Fig. 5b; Supplementary Table 3). The lowest rated production type was movies, followed by video tape (Fig. 5b; Supplementary Table 3).

Discussion

The idea that crowd wisdom might reflect the reality better than expert opinion or any single individual, has been considered as provocative in the past (Galton, 1907; Prelec, Seung, & McCoy, 2017). However, online voting has the cultural, geographical, social, and temporal diversity of a sample size that is hard to be ignored (Sunstein, 2006; Surowiecki, 2005). That been said, the way people vote is complex depending on several factors such as fake news (Guess, Nyhan, & Reifler, 2018), social circles and networks (Galesic et al., 2018), and information gerrymandering (Stewart et al., 2019); a small number of strategically positioned individuals can influence the voting behaviour of a larger majority, particularly in cases where the larger group is undecided about its voting intentions (Stewart et al., 2019). However, strong ties are

elementary for spreading online behaviour in social networks of humans (Bond et al., 2012). Still voting behaviour exhibits universality; votes received by a candidate, rescaled by the average performance of all competitors in the same party list, has the same distribution regardless of the country and the year of the election (Chatterjee, Mitrović, & Fortunato, 2013) and a scale-free behaviour in the number of votes regarding rating of movies independently of movie attributes such as average rating, age and genre was identified (Ramos et al., 2015).

Results derived here indicate that the number of votes, as a mean property, is only weakly correlated with higher rating. But there is a subtlety to this: low and intermediate number of voting turnover results in a higher rating but a very high number of votes results in lower rating. A potential explanation for this phenomenon is partisan polarization (Gustafson et al., 2019): initially, there is low voting but the ones who do spend time to vote are the ones that saw a new and of relatively low public awareness production. Through time, voters become much more familiar with the production and partisan polarization increases significantly due to a sharp decrease in support among some groups that will vote because they disagree with the high rating of the highly voted production (Gallup, 2019). In addition, another set of people agrees with the initial high rating and although normally would not bother to vote after watching a production, will vote to oppose the negative voting trending among the other group (Gallup, 2019). Polarization between partisan groups results from selective exposure and motivated reasoning that drive strong partisans to more extreme opinions (Druckman & McGrath, 2019; Gustafson et al., 2019). Polarization may be further exacerbated by the fact that e-voting is associated with less happiness and more wariness in comparison with traditional voting (Bruter & Harrison, 2017; Cammaerts, Bruter, Banaji, Harrison, & Anstead, 2016) at least in youths.

Video games as production type contribute the highest to the mean increase in rating. Video games are popular among all generations based on content (Greenberg, Sherry, Lachlan, Lucas, & Holmstrom, 2010; Williams, Martins, Consalvo, & Ivory, 2009) and moreover currently people can easily watch stream in-live game online from social network like YouTube or Twitch (Smith, Obrist, & Wright, 2013). The reason it is very popular could potentially be explained by the fact that during the live show the viewer can communicate on the social chat platform with the host, and thus the process is communicative and interactive (Crull, Miller, Kenney, & Martin, 2007). Additionally, video game users post a walkthrough on the game in their video and many of the people watch this kind of video as it can help them finish their game or identify the rationale of other users (Petrova, Gross, & Insights, 2017). Concluding, an explanation behind the high rating of video games is their potential for belonging in a community, learning, and feeling accepted (Petrova et al., 2017; Squire, 2008, 2011). On the more negative side, video games are also addictive and associated with gambling (Nielsen & Grabarczyk, 2019; Turel & Bechara, 2019), and internet use (Ng & Wiemer-Hastings, 2005) and therefore addicted users to video games and internet are more likely to submit an e-vote.

The second highest production type to contribute to the mean is TV series. A reason behind TV series being highly rated may simply be a result of natural selection: a TV series will be discontinued if they are not popular (Esser, 2010). On the other hand, the good ones usually last for several seasons which means

the series lasts several years. This results in actors and characters to be identifiable figures (Fischer, Ekenel, & Stiefelhagen, 2011; Tapaswi, Bäuml, & Stiefelhagen, 2012) that the public may like or dislike (Tian & Hoffner, 2010) or find common attributes with their own character and behaviour (García, 2016). As the audience ages real time with the actors in TV series, watching them grow may create an additional sense of familiarity. Moreover, TV series get people in the suspense which can make it quite enthralling because the story and plot is continued across episodes. In addition, more often than not, TV series have shorter duration than movies and to that end they are also easier to fit in the daily schedule in terms of time availability.

The highest rated content was content diversity implying that productions with a more complex and diverse thematic and script are more appreciated by the public. A plausible explanation is that more diverse content is more complex, and thus able to encapsulate more sentiments and feelings (Liu, Hsaio, Lee, Lu, & Jou, 2011). Interestingly, productions from the early 19th century exhibited the highest content diversity within the same production, several combining three genres, among the movies that have been rated by at least one vote. Content diversity declined in the beginning of the 20th century to similar levels as the ones recorded today. The levels recorded today – content diversity close to two – remains constant in the past 80 years. Thus, the majority of productions so far have two content types and this is not related with technological progress. It is therefore the script that matters more rather than technologically-derived special effects that characterize content complexity.

Documentary was the highest rated content, indicating that learning in a relaxing manner, photography, exploring, nature, and narrative story-telling, themes often encountered in documentaries (Dovey & Rose, 2012; Nisbet & Aufderheide, 2009; Ruby, 1977), are highly rated among humans. Mystery and fantasy content are also highly rated on average. Both contents are related with a sense of discovery, imagination, creativity, and originality which are mostly positive feelings. Western content is also highly rated and notably is the only highly rated content that often contains violence. It seems that the dominant characteristic of western movie content, history, open landscapes, nature, wilderness, may dominate over action, war, and crime, contents that are often in western movies but when each of those contents is seen in isolation as a separate category are poorly rated having a negative effect on the mean.

Results indicated that comedy content contribute marginally negatively to the average rating, even though there are several highly rated comedies (Krutnik & Neale, 2006). This may simply be a result of the very large number of comedies produced – there are so many comedies and only some are highly rated and the rest have average or low rating. Horror and thriller content reduces rating by the highest negative coefficient. While there certainly are highly rated horror or thriller productions, the effect of their content in the mean rating indicates that the majority of people find exhibition to gore scenes to be disturbing, and only a small audience is enjoying it (Hoekstra, Harris, & Helmick, 1999). Highly sensitive people can be easily over-stimulated by their environment and they also to be more empathetic than the average person, and thus may have a more intense physiological reaction to horror (Haidt, McCauley, & Rozin, 1994; Hoekstra et al., 1999). In fact, there is evidence that young viewers who perceive greater realism in horror films are more negatively affected by their exposure to horror films than viewers who perceive the film as

unreal (Hoekstra et al., 1999). On second hand, religion may also be a factor of why horror movie are not liked by the people as horror movie can be seen as glorifying the devil (Carroll, 2003; Winstead, 2011).

News, and game show content were also low rated, but this is perhaps unsurprising as news are not watched for entertainment but for briefing, and often the news are unpleasant as bad news are more likely to be reported (Soroka, Young, & Balmas, 2015). Adult rated movies are also low rated on average indicating that exposure on strong violence or explicit sex scenes is not appreciated on average and it is linked with violent behaviour (Russell, 1980) and thus not relaxing. Game is also low rated; a game show is hard to be appreciated more than the game itself, and in fact televised sports are topping watched shows of all time (Raney, 2009). Thus, people who like sports prefer the game *per se* rather than the pre-game which has very expensive commercial time. The reality TV show content is also having a large negative effect on the rating. Since the early days of reality TV, critics have consistently attacked the genre for being voyeuristic, cheap, and sensational (Hill, 2014). Audiences watching reality TV they are not only watching them for entertainment, but are also engaged in critical viewing of the attitudes and behaviour of ordinary people in the programs, as well as the ideas and practices of the producers (McMurria, 2008). Therefore audiences are able to make distinctions between what they perceive to be good and bad reality programming.

Statement

The authors declare that this work was conducted in the absence of any conflict of interest.

References

- Achterberg, P., & Houtman, D. (2006). Why do so many people vote 'unnaturally'? A cultural explanation for voting behaviour. *European Journal of Political Research*, 45(1), 75-92. doi: 10.1111/j.1475-6765.2005.00291.x
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415), 295-298. doi: 10.1038/nature11421
- Brennan, J. (2012). *The Ethics of Voting*. Princeton University Press.
- Bruter, M., & Harrison, S. (2017). Understanding the emotional act of voting. *Nature Human Behaviour*, 1(0024), 1-3.
- Cammaerts, B., Bruter, M., Banaji, S., Harrison, S., & Anstead, N. (2016). *Youth participation in democratic life: Stories of hope and disillusion*. Springer.
- Carroll, N. (2003). *The philosophy of horror: Or, paradoxes of the heart*. Routledge.

- Chatterjee, A., Mitrović, M., & Fortunato, S. (2013). Universality in voting behavior: an empirical analysis. *Scientific Reports*, 3(1), 1049. doi: 10.1038/srep01049
- Crull, R., Miller, B., Kenney, A., & Martin, T. (2007). Social network-enabled interactive media player: Google Patents.
- Dovey, J., & Rose, M. (2012). We're happy and we know it: Documentary, data, montage. *Studies in Documentary Film*, 6(2), 159-173. doi: 10.1386/sdf.6.2.159_1
- Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, 65(2), 135-150.
- Druckman, J. N., & McGrath, M. C. (2019). The evidence for motivated reasoning in climate change preference formation. *Nature Climate Change*, 9(2), 111-119. doi: 10.1038/s41558-018-0360-1
- Esser, A. (2010). Television formats: Primetime staple, global market. *Popular Communication*, 8(4), 273-292.
- FCC. (2019). Federal Communications Commission: Implementation of the Consolidated Appropriations Act of 2019; Report on Television Ratings and The Oversight Monitoring Board. <https://docs.fcc.gov/public/attachments/DA-19-423A1.pdf>.
- Fischer, M., Ekenel, H. K., & Stiefelhagen, R. (2011). Person re-identification in tv series using robust face recognition and user feedback. *Multimedia Tools and Applications*, 55(1), 83-104.
- Galesic, M., Bruine de Bruin, W., Dumas, M., Kapteyn, A., Darling, J. E., & Meijer, E. (2018). Asking about social circles improves election predictions. *Nature Human Behaviour*, 2(3), 187-193. doi: 10.1038/s41562-018-0302-y
- Gallup. (2019). Abortion Trends by Party Identification (Gallup, 2019);.
- Galton, F. (1907). Vox populi (the wisdom of crowds). *Nature*, 75(7), 450-451.
- García, A. N. (2016). *Emotions in Contemporary TV Series*. Springer.
- Ginsburgh, V., & Weyers, S. (1999). On the perceived quality of movies. *Journal of Cultural Economics*, 23(4), 269-283.
- Greenberg, B. S., Sherry, J., Lachlan, K., Lucas, K., & Holmstrom, A. (2010). Orientations to Video Games Among Gender and Age Groups. *Simulation & Gaming*, 41(2), 238-259. doi: 10.1177/1046878108319930
- Guess, A., Nyhan, B., & Reifler, J. (2018). Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 US presidential campaign. *European Research Council*, 9.

- Gustafson, A., Rosenthal, S. A., Ballew, M. T., Goldberg, M. H., Bergquist, P., Kotcher, J. E., . . . Leiserowitz, A. (2019). The development of partisan polarization over the Green New Deal. *Nature Climate Change*, 9(12), 940-944. doi: 10.1038/s41558-019-0621-7
- Haan, M. A., Dijkstra, S. G., & Dijkstra, P. T. (2005). Expert judgment versus public opinion—evidence from the Eurovision song contest. *Journal of Cultural Economics*, 29(1), 59-78.
- Haidt, J., McCauley, C., & Rozin, P. (1994). Individual differences in sensitivity to disgust: A scale sampling seven domains of disgust elicitors. *Personality and Individual differences*, 16(5), 701-713.
- Hill, A. (2014). *Reality tv*. Routledge.
- Hills, T. T., Proto, E., Sgroi, D., & Seresinhe, C. I. (2019). Historical analysis of national subjective wellbeing using millions of digitized books. *Nature Human Behaviour*. doi: 10.1038/s41562-019-0750-z
- Hoekstra, S. J., Harris, R. J., & Helmick, A. L. (1999). Autobiographical memories about the experience of seeing frightening movies in childhood. *Media Psychology*, 1(2), 117-140.
- Jakob, N., Weber, S. H., Müller, M. C., & Gurevych, I. (2009). *Beyond the stars: exploiting free-text user reviews to improve the accuracy of movie recommendations*. Paper presented at the Proceedings of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion.
- Ji, S., & Waterman, D. (2010). Production technology and trends in movie content: An empirical study: Working Paper, Department of Telecommunications, Indiana University
- Koenker, R. (2005). *Quantile regression*. Cambridge university press.
- Koh, N. S., Hu, N., & Clemons, E. K. (2010). Do online reviews reflect a product's true perceived quality? An investigation of online movie reviews across cultures. *Electronic Commerce Research and Applications*, 9(5), 374-385. doi: <https://doi.org/10.1016/j.elerap.2010.04.001>
- Kostakos, V. (2009). *Is the crowd's wisdom biased? A quantitative analysis of three online communities*. Paper presented at the 2009 International Conference on Computational Science and Engineering.
- Krutnik, F., & Neale, S. (2006). *Popular film and television comedy*. Routledge.
- Lago, I. (2019). A Research Agenda in Elections and Voting Behavior in a Global and Changing World. *Frontiers in Political Science*, 1, 1.
- Leduc, L. (2002). Opinion change and voting behaviour in referendums. *European Journal of Political Research*, 41(6), 711-732. doi: 10.1111/1475-6765.00027
- Liu, C.-L., Hsaio, W.-H., Lee, C.-H., Lu, G.-C., & Jou, E. (2011). Movie rating and review summarization in mobile environment. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 42(3), 397-407.

- McMurria, J. (2008). Desperate Citizens and Good Samaritans: Neoliberalism and Makeover Reality TV. *Television & New Media*, 9(4), 305-332. doi: 10.1177/1527476408315115
- Mehr, S. A., Singh, M., Knox, D., Ketter, D. M., Pickens-Jones, D., Atwood, S., . . . Glowacki, L. (2019). Universality and diversity in human song. *Science*, 366(6468), eaax0868. doi: 10.1126/science.aax0868
- Moustakas, A., Daliakopoulos, I. N., & Benton, T. G. (2019). Data-driven competitive facilitative tree interactions and their implications on nature-based solutions. *Science of The Total Environment*, 651, 2269-2280. doi: <https://doi.org/10.1016/j.scitotenv.2018.09.349>
- Moustakas, A., & Evans, M. R. (2016). Regional and temporal characteristics of bovine tuberculosis of cattle in Great Britain. *Stochastic Environmental Research and Risk Assessment*, 30(3), 989-1003. doi: 10.1007/s00477-015-1140-3
- Moustakas, A., & Katsanevakis, S. (2018). Editorial: Data Mining and Methods for Early Detection, Horizon Scanning, Modelling, and Risk Assessment of Invasive Species. *Frontiers in Applied Mathematics and Statistics*, 4(5). doi: 10.3389/fams.2018.00005
- Ng, B. D., & Wiemer-Hastings, P. (2005). Addiction to the internet and online gaming. *Cyberpsychology & behavior*, 8(2), 110-113.
- Nielsen, R. K. L., & Grabarczyk, P. (2019). Are Loot Boxes Gambling? Random reward mechanisms in video games. *Transactions of the Digital Games Research Association*, 4(3).
- Nisbet, M. C., & Aufderheide, P. (2009). Documentary Film: Towards a Research Agenda on Forms, Functions, and Impacts. *Mass Communication and Society*, 12(4), 450-456. doi: 10.1080/15205430903276863
- Petrova, E., Gross, N., & Insights, G. (2017). 4 reasons people watch gaming content on youtube.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-Effects Models in S and S-PLUS*. New York: Springer Verlag.
- Prelec, D., Seung, H. S., & McCoy, J. (2017). A solution to the single-question crowd wisdom problem. *Nature*, 541(7638), 532-535. doi: 10.1038/nature21054
- PTC. (2019). Parents Television Council: A Decade of Deceit: How TV Content Ratings Have Failed Families. <https://go.parentstv.org/decades-report/documents/Decades-Report.pdf>.
- Ramos, M., Calvão, A. M., & Anteneodo, C. (2015). Statistical Patterns in Movie Rating Behavior. *Public Library of Science One*, 10(8), e0136083. doi: 10.1371/journal.pone.0136083
- Raney, A. A. (2009). The effects of viewing televised sports. *The SAGE handbook of media processes and effects*, 439-453.

- Redfern, N. (2012). Genre trends at the US box office, 1991 to 2010. *European Journal of American Culture*, 31(2), 145-167.
- Ruby, J. (1977). The image mirrored: Reflexivity and the documentary film. *Journal of the University Film Association*, 29(4), 3-11.
- Russell, D. E. (1980). Pornography and violence: What does the new research say. *Take back the night: Women on pornography*, 218-238.
- Scott, A. J., & Knott, M. (1974). A cluster analysis method for grouping means in the analysis of variance. *Biometrics*, 507-512.
- Simonton, D. K. (2002). Collaborative aesthetics in the feature film: Cinematic components predicting the differential impact of 2,323 Oscar-nominated movies. *Empirical Studies of the Arts*, 20(2), 115-125.
- Smith, T., Obrist, M., & Wright, P. (2013). *Live-streaming changes the (video) game*. Paper presented at the Proceedings of the 11th european conference on Interactive TV and video.
- Soroka, S., Young, L., & Balmas, M. (2015). Bad news or mad news? Sentiment scoring of negativity, fear, and anger in news content. *The ANNALS of the American Academy of Political and Social Science*, 659(1), 108-121.
- Squire, K. (2008). Open-ended video games: A model for developing learning for the interactive age. *The ecology of games: Connecting youth, games, and learning*, 167-198.
- Squire, K. (2011). Video games and learning. *Teaching and participatory culture in the digital age*.
- Stewart, A. J., Mosleh, M., Diakonova, M., Arechar, A. A., Rand, D. G., & Plotkin, J. B. (2019). Information gerrymandering and undemocratic decisions. *Nature*, 573(7772), 117-121. doi: 10.1038/s41586-019-1507-6
- Sunstein, C. R. (2006). *Infotopia: How many minds produce knowledge*. Oxford University Press.
- Surowiecki, J. (2005). *The wisdom of crowds*. Anchor.
- Tapaswi, M., Bäuml, M., & Stiefelhagen, R. (2012). "Knock! Knock! Who is it?" probabilistic person identification in TV-series. Paper presented at the 2012 IEEE Conference on Computer Vision and Pattern Recognition.
- Thompson, K. M., & Yokota, F. (2004). Violence, sex and profanity in films: correlation of movie ratings with content. *MedGenMed : Medscape general medicine*, 6(3), 3-3.
- Tian, Q., & Hoffner, C. A. (2010). Parasocial Interaction With Liked, Neutral, and Disliked Characters on a Popular TV Series. *Mass Communication and Society*, 13(3), 250-269. doi: 10.1080/15205430903296051

Turel, O., & Bechara, A. (2019). Little video-gaming in adolescents can be protective, but too much is associated with increased substance use. *Substance use & misuse*, 54(3), 384-395.

Williams, D., Martins, N., Consalvo, M., & Ivory, J. D. (2009). The virtual census: representations of gender, race and age in video games. *New Media & Society*, 11(5), 815-834. doi: 10.1177/1461444809105354

Winstead, A. F. (2011). The Devil Made Me Do It! The Devil in 1960s–1970s Horror Film. *Vader, Voldemort, and Other Villains: Essays on Evil in Popular Media*, 28-45.

Wu, F., & Huberman, B. A. (2008). *How public opinion forms*. Paper presented at the International Workshop on Internet and Network Economics.

You, C., Lin, D. K. J., & Young, S. S. (2018). Time series smoother for effect detection. *Public Library of Science One*, 13(4), e0195360. doi: 10.1371/journal.pone.0195360

Figures

Figure 1

Original data. a. Number of productions per year. b. Total productions per content type throughout the data 1874 – 2018. c. Total productions per production type throughout the data. d. Mean number of votes per production per year throughout the data.

Figure 2

Rating and number of votes. a. Mean rating per year across all productions. The solid red line indicates the mean, blue boxed a 50% confidence interval of the mean, and whiskers extent to the lower (25%) and upper quartile (75%) of the data. b. Rating per production type across years. Red lines indicate linear regressions to show overall trends, while black lines a smoother line to smooth out extreme values and indicate potential non-linear effects. c. Upper panel: distribution of the number of votes. Lower panel: 50% confidence interval of the mean and outliers with grey star symbols. Outliers are values outside the 75% range of the data indicated by the vertical black line whiskers. d. Upper panel: distribution of rating. Lower panel: 50% confidence interval of the mean and outliers with grey star symbols. f. Quantile regressions between 1, 5, 10, 25, 50, 75, 90, 95, and 99 % of the data regarding rating and number of votes. The

highest quantiles marked on the figure as q95, and q99, correspond to the 95 and 99% of the highest number of votes. The dotted red line indicates the least squares regression between rating and number of votes.

Figure 3

Content of productions over time. Percentage of each content present across all productions of that year.

Figure 4

Change point detection of data as time series. Change point detection methods were applied to quantify changes in the mean and variance over time. Time is referenced from the year 1874 corresponding to zero value to the 140th corresponding to year 2018. a. Change point detection over time of the average of rating per year, b. Change point detection to the average of the number of votes per year. c. Change point detection of the percentage of adult rated content per year. d. Change point detection of the average of diversity of content per year.

Figure 5

Results from mixed effects models regarding rating explained by content or production type. a. Coefficients of the mixed effects model between rating (dependent variable) and all content types as factors (independent variables). Solid bars indicate the coefficient values of that content type while red error bars indicate the coefficient's standard error. Coefficients are ranked from higher to lower value. b. Coefficients of the mixed effects model between rating (dependent variable) and production types as factors (independent variables). The defaults value in the model is movie as a production type. Solid bars indicate coefficient values while red error bars indicate the coefficient's standard error. Coefficients are ranked from higher to lower.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryinformation.docx](#)