UNIVERSITY OF WOLVERHAMPTON

School of Mathematics and Computing

Factors influencing the popularity of YouTube videos and users' decisions to watch them

David Foster MSc, MEd

A thesis submitted in partial fulfilment of the
requirements of the University of Wolverhampton
for the degree of Doctor of Philosophy

School of Mathematics and Computing

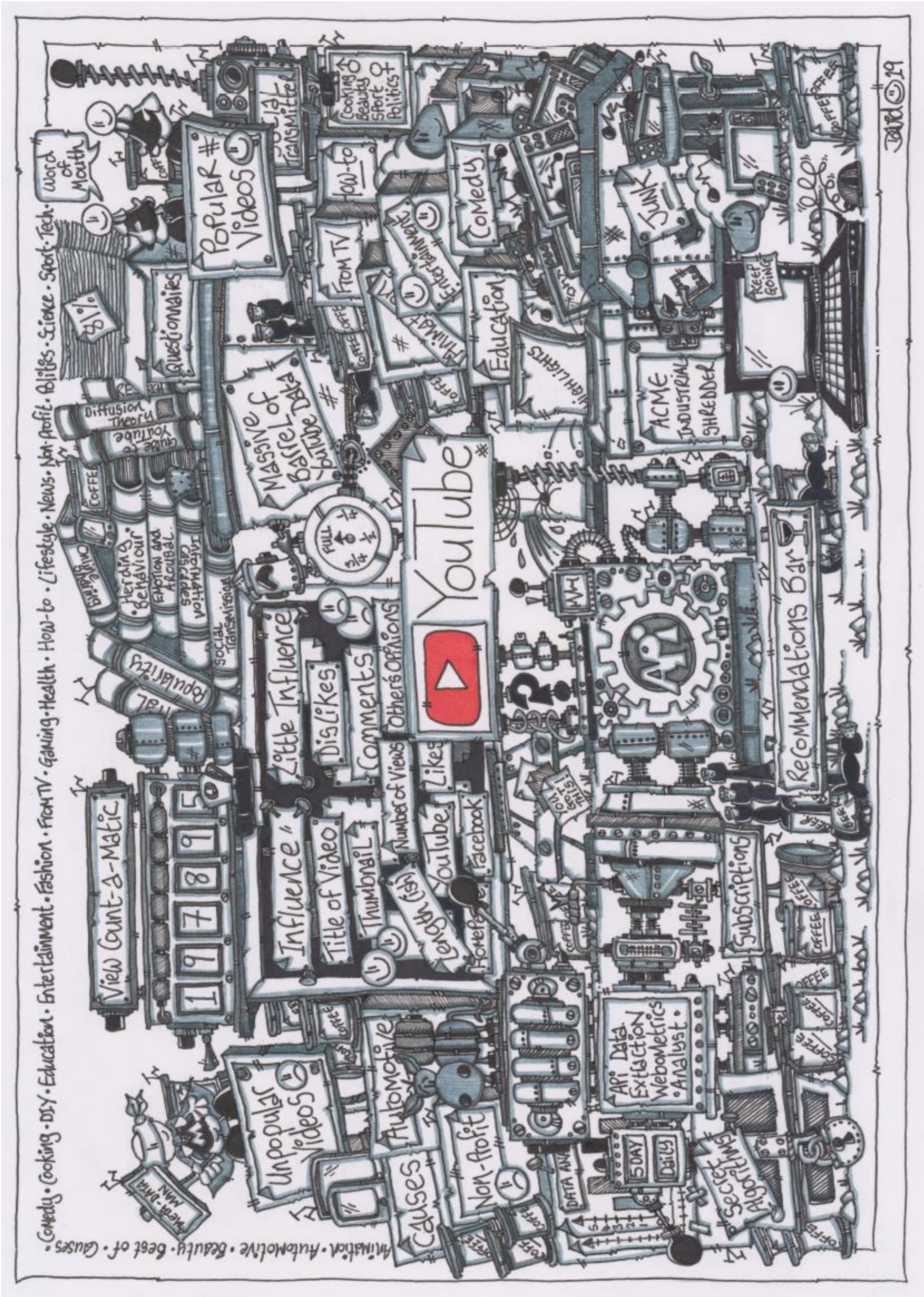
University of Wolverhampton

2020

This work or any part has not previously been presented in any form to the University or to any other body whether for the purposes of assessment, publication or for any other purpose (unless otherwise indicated). Save for any express acknowledgements, references and/or bibliographies cited in the work, I confirm that the intellectual content of the work is the result of my own efforts and of no other person.

Signature: [redacted]

Date: 7/8/20


Comedy • Cooking • DIY • Education • Entertainment • Fashion • From TV • Gaming • Health • How-to • Lifestyle • News • Not-profit • Politics • Science • Sport • Tech •
Animation • Automotive • Beauty • Best of • Causes •
Word of Mouth
81%
Questionnaires
Popular Videos
Social Transmitter
Cooking & Beauty Sport & Politics
How-to
From TV
Comedy
Education
Highlights
Junk
Keep Going
ACME Industrial Shredder
Massive Barrel of YouTube Data
Diffusion Theory
Guide to YouTube
Herding Behaviour
Emotion and Arousal
Information Cascades
Social Transmission
Popularity
View Count-a-Matic
1 9 7 8 9 5
Influence
Little Influence
Title of Video
Dislikes
Thumbnail
Comments
Number of Views
Others Opinions
Length (ish)
YouTube
Likes
Homepage
Facebook
YouTube
Full
Recommendations Bar
Subscriptions
API Data Extraction Webometrics Analyst
Unpopular Videos
Automotive
Causes
Non-profit
Secret Algorithms
5 DAY
Daily
Data An
Coffee

Abstract

YouTube has substantial impact on modern society as the second most popular website in the world. Despite its sustained popularity, little is known about which types of video are most viewed and the reasons why people choose to watch them. This research critically analyses the sample of videos provided by the YouTube API, then uses the metrics associated with these videos to help assess which types of YouTube video are popular. It also harnesses a questionnaire of mainly UK teacher education graduate YouTube users to investigate which factors influence decisions to watch YouTube videos. This was a convenience sample selected to achieve a high response rate, which it achieved (81%), minimising non-response bias. The video lists provided by the YouTube API were not random samples but contained a wide range of types of video (including both popular and unpopular), except that older videos were avoided. There were substantial differences between categories in the average properties of the videos returned and the proportion of videos returned on multiple days. The most popular categories from the YouTube metadata collected based on average view counts are varied: From TV, Best of, Animation and How-to. Cause-based video categories tended to be unpopular. Video popularity did not seem to be affected by video duration, on average. Users are more likely to interact with (comment, like, dislike) videos that are useful or supporting in some way. Videos that are interacted with more are not always more popular, with subject content affecting this relationship. In addition, high view counts associated with fewer likes, dislikes and comments per view, suggesting that indicators of popularity may not attract new viewers. The most popular categories with survey respondents were slightly different, partly reflecting their educational background (e.g., Education videos), and there were some (stereotypical) gender differences in the most popular categories. Respondents rarely believed that they were influenced by a video's popularity or evidence of other users' reactions to it when deciding to watch the video. Instead, they were most likely to be influenced by content-related factors, such as a video's title and thumbnail picture. Despite previous research showing that people can be influenced by the opinions and watching habits of others, respondents claimed to be little influenced by this. Nevertheless, they frequently reported watching videos posted to Facebook, possibly trusting the person that posted the video. Thus, despite extensive discussion of various forms of viral information spreading, content, rather than popularity, is king in YouTube, although online word-of-mouth sharing through trusted relationships is also important. The main limitations of this research are that the data used may not be representative of YouTube and all UK YouTube users overall, so the conclusions should be interpreted cautiously.

Acknowledgements

I would like to express my deepest thanks to my supervisor Professor Mike Thelwall for all of your support, help, guidance, ideas, time, patience and understanding throughout the whole process of getting this thesis completed. I feel very lucky to have had you as my supervisor and appreciate everything you have done.

I would also like to thank the wonderful Steph who without all of your support, encouragement, understanding, friendship, love and endless faith in me I would never have gotten this far.

In addition I would like to thank Erica, Neil and Rob for all of your support and for always being there for me.

I would also like to thank all those individuals, the 81%, who kindly completed my questionnaire and provided me with a wealth of data for this thesis.

Finally, I would like to thank those friends who have provided moral support, understanding and messages of encouragement throughout this process.

Table of Contents

List of Tables

List of Figures

1 Introduction

Many people watch a wide range of videos through social websites like YouTube, Vimeo, Dailymotion, Twitch and Bilibili. YouTube is the most influential video sharing site, having turned producing, watching and interacting with online videos into a mainstream activity (Khan, 2017; Duncan et al., 2014; Alloway and Alloway, 2012; Eckler and Bolls, 2011). It is also the most popular video sharing website in the world and is the second most popular website of any type, behind only Google in overall popularity according to Alexa.com in December 2019 (Alexa, 2019). In addition, individual viral YouTube videos can be influential within society and social network groups (Frasco, 2014; Broxton et al., 2013). YouTube is therefore part of the lives of many people across the globe as well as a significant influence on culture and society.

YouTube's videos have also become a resource for research. Almost all YouTube research has focused on the content of YouTube videos around a defined issue, rather than on YouTube itself, however. Most studies have analysed YouTube videos from a humanities or social sciences perspective, including for the following topics.

- Contemporary events and issues
- Social problems, behaviours and interactions
- Online behaviour
- YouTube's impact on society
- Health
- News and current affairs
- Education and learning
- Political issues
- Collective responses
- Audience partisanship
- Academic impact

(Klobas et al., 2019; Aytar et al., 2018; Kardas and Brien, 2018; Khan, 2017; Anthony et al., 2013; Snelson et al., 2012; Chenail, 2011; Lewis et al., 2011; Jang, 2011; Priem et al., 2011; Thorson, et al., 2010; Steinberg et al., 2010; Lange, 2007; Keelan et al., 2007; Cunningham and Nichols, 2008)

Despite the importance of YouTube for research and society, little is known about why individual YouTube videos are popular and which factors influence users when they are selecting videos to watch. These are important omissions for those creating and marketing on YouTube as well as for scholars seeking to understand the dynamics of the site. This thesis addresses these issues by analysing YouTube video popularity for a large collection of videos as well as by surveying a sample of users to discover the factors that they believe influence their watching decisions.

1.1 Thesis Focus

Although it is difficult to predict the popularity of a given YouTube video (Welbourne and Grant, 2016; Ahmed et al., 2013; Zink et al., 2009) this thesis investigates which video-related factors influence UK YouTube users when deciding to watch a video. There has been some research relating to video content and category popularity but no definitive answers about what makes a video popular within the YouTube website from the perspective of either the system or users.

YouTube has an Applications Programming Interface (API) that allows researchers and others to access information about its videos automatically through a computer program. This API data is the only practical way to get information about many YouTube videos and so has become the default entry to YouTube for large scale research. Little is known about the sample returned by category searches from the API, however, and so it is not clear whether it is reasonable to use it to investigate the site. Moreover, there does not appear to be any research that analyses and evaluates the sample of data extracted by the YouTube API system. This thesis investigates the YouTube API to understand its properties and any biases in the sample of videos returned by category searches in order to allow its subsequent use to investigate video popularity.

This thesis focuses on a set of category search results from the YouTube API. Comparisons between identical queries submitted at different points in time were used to try to deduce the factors that influence YouTube when selecting videos to return from category searches. This data was then used to investigate properties of videos that associate with higher attention metrics (e.g., views, likes) and to assess which categories contain the most popular videos.

A questionnaire was also used to get information about video popularity from the perspective of UK YouTube users to triangulate with the YouTube API data. YouTube users are not possible to survey from within the site and any general survey method designed to identify users, even from one country, would probably have a low response rate. This would generate a substantial non-response bias so that the results might reflect enthusiastic YouTubers rather than typical users. Thus, the questionnaire strategy instead targeted a narrow demographic of users to obtain a high response rate, minimising non-response bias. The survey used a convenience sample of UK people that work within education, mainly as teachers.

1.2 Research Questions

The primary aim is to identify which YouTube video categories are most popular and the factors that influence whether users will watch a YouTube video. In support of this, the first research question addresses the nature of the sample provided by YouTube searches via the API. This is important because it is a practical tool for generating a large sample of videos to investigate for popularity. The research questions are discussed and justified in Chapter 3.

- **RQ1:** How does the YouTube API select the sample of videos that it returns for a category search?
- **RQ2:** What is the age, length, and popularity of videos in each YouTube category and how do these vary between categories?

The following questions address the popularity of videos based on their YouTube category and use data extracted from the YouTube API category searches.

- **RQ3:** Which categories of YouTube video are the most popular?
- **RQ4:** Which categories of YouTube video do users comment on most?
- **RQ5:** How does the length, like count, dislike count and comment count of a YouTube video relate to its popularity?

The following questions investigate the popularity of videos from the user perspective and are addressed with a survey.

- **RQ6:** What are the main gender and age differences in the types of YouTube video that are the most popular?
- **RQ7:** Which factors influence the decision to watch a YouTube video for different genders and ages?

The final question draws on the two data sources (YouTube API category searches and survey) to obtain more general findings.

- **RQ8:** What influences the decision to watch a YouTube video?

1.3 Contribution

This thesis is situated within the emerging interdisciplinary literature of quantitative analysis of social media content. This literature draws on information science and computer science for methods and uses theoretical insights from media studies, psychology and sociology. The thesis generates and compares YouTube data and questionnaires to identify factors that influence the popularity of YouTube videos and decisions to watch them. The findings and conclusions of this research will be of relevance to YouTube influencers and content creators wanting to increase the popularity of their videos. It will give insights to people working in marketing, advertising, information services and political election campaigns, for example, about how they might gain an audience. Its insights into the factors behind video popularity will also help the YouTube research community by giving background information that can set the context for studies of individual topics on YouTube.

1.4 Thesis Structure

This thesis comprises ten chapters and is organised as follows.

- **Chapter 1 (Introduction):** This presents a broad overview of the research, including YouTube, the sources and focus of the research, the contribution to the field, the structure of the thesis and the research questions.
- **Chapter 2 (Literature Review):** This discusses popularity, increased online activity, sharing and online videos, the development and growth of YouTube, and a range of previous research focusing on YouTube.
- **Chapter 3 (Aims and Research Questions):** This explains the aims of the research and justifies the research questions.
- **Chapter 4 (Methods):** This explains and justifies the methods to generate and analyse the two data sets - category based searches using the main YouTube categories through its API and the development of a questionnaire to collect data from a convenience sample of web users. This chapter also discusses how the ethical considerations of the research were addressed.
- **Chapter 5 (YouTube API Category Search Result Changes Over Time):** This analyses the two different searches used to extract data from YouTube and what type of sample the website's API provides (Research question 1).
- **Chapter 6 (YouTube API Category Search Video Properties):** This analyses the whole data sample collected from the YouTube API in relation to the metrics and popularity of videos across the categories and how these vary when compared (Research question 2).
- **Chapter 7 (YouTube Video Popularity, Interactions and Metric Impact):** This analyses the whole data sample collected from the YouTube API focusing on the types of video that are most popular and have the most interaction from users. It also investigates how key metrics relate to YouTube video popularity – (Research questions 3, 4 and 5).
- **Chapter 8 (YouTube User Perspectives):** This analyses the data collected from a UK YouTube user survey in relation to video popularity and user influence in gender, age and user level (Research questions 6 and 7).
- **Chapter 9 (YouTube Video Popularity):** This compares and discusses the findings from Chapters 6, 7 and 8 in order to determine the overall key influences on user decisions to watch YouTube videos (Research question 8).
- **Chapter 10 (Conclusion):** This addresses each of the research questions and demonstrates the key findings of the research (Research questions 1 to 8).

2 Literature Review

This literature review covers three separate topics that are important for the goals of the thesis: popularity, online activity, and YouTube.

2.1 Popularity

This section discusses the concept of popularity, how things become popular and the ways in which popularity is disseminated through society and social groups. The term popularity has two different meanings, both of which are relevant to the thesis (merriam-webster.com, 2019).

- **Popular**: "frequently encountered or widely accepted".
- **Popular**: "commonly liked or approved".

This thesis is primarily concerned with the first, and more specifically the first part of the first definition, frequently encountered. Thus, **popular within this thesis primarily refers to "frequently encountered"**, although, especially in YouTube, it can be expected that this type of popularity would be closely related to the other: widely watched videos would tend to be liked by viewers. For clarity, the two meanings will sometimes be referred to as "**popularity (liked)**" and "**popularity (frequently encountered)**" and the default meaning of the term is popularity (frequently encountered) unless clear from the context or explicitly stated.

2.1.1 What is 'Popularity'?

Humans have always had a fascination with things within the world around them, having personal and specific preferences, tastes, opinions, wants, desires and social and cultural influences which lead to them making decisions about what they interact with (Akdeniz, 2014; Easley and Kleinberg, 2010). Because of these interactions some things gain greater levels of attention and could be referred to as being *popular*.

The concept of popularity (liked) is a fluid social construction within social groups and therefore can only be understood and discussed within such a context. It can relate to a person, idea or item that is liked best within a social group (Akdeniz, 2014; Easley and Kleinberg, 2010). Within a social group, popularity may reflect a socially-determined consensus or an external objective measure of attention. Nevertheless, the greater attention, opinion and activity something receives, the more popular it is deemed to be by the group (Scott and Judge, 2009). It can be difficult to agree about what is 'popular' overall at a given time due to differing preferences. Popularity (frequently encountered) can sometimes be the product of peoples' in-built nature to conform with others' ideas, rather than reflecting average personal preferences (Eger, 2015; Boyd, 2014; Cialdini and Goldstein, 2004; Banerjee, 1992).

There are many factors other than the quality of a product or artefact that influence its popularity, such as herding behaviours, word-of-mouth, information cascades, social groups, opinion makers, access and social networking (Eger, 2015; Easley and Kleinberg, 2010; Leskovec et al., 2006). Aristotle had three main principles in persuading members of society to share information with others:

- Ethos relates to the credibility of the source of the information. The greater the level of trust someone has in the origin of the information the more likely they are to engage with and share it.
- Pathos relates to the emotional appeal of the information. The greater the level of emotional engagement with the information the more likely someone will be to discuss and share it.
- Logos relates to the logical structure of the information provided. If the information is suitably supported and contains persuasive reasoning, then people are more likely to share it.

(Worthington, 2008)

2.1.2 Influencers

Society is regularly guided about what is popular (in both senses) by influential people, the media, and advertising (Perrin, 2015). People may conform to the views of opinion leaders even if they disagree with them (Eger, 2015; Takacs et al., 2014). Nevertheless, things deemed inappropriate or censured by influential figures may still be liked by a minority in reaction to the leaders, helped by attention generated by the negative publicity (Akdeniz, 2014; Brack, 2013). This can act as a symbol of defiance because some people do not want to be told what to do (Yin et al., 2012). The power of the traditional media has declined, however, through the rise of alternative influencers, including those that operate almost exclusively through YouTube (Burnett, 2013; Kattimani et al., 2010, Zink et al., 2009). The web also allows easy access to a much wider range of content than before, with fewer restrictions, and the potential to access niche content and influencers that would previously have been difficult to find (Reka et al., 1999). Because of these online developments, web users now have more choice about what to access and who to be influenced by (Yin et al., 2012).

2.1.3 Popularity as an Influencer

Popularity itself has become a powerful global force within a positive feedback loop. Popular things may get additional exposure because of their popularity, such as from "top ten" lists or news coverage. Many resource-based websites, such as Amazon, YouTube, and iTunes, provide consumers with lists of most popular, viewed or downloaded, which can have the result of influencing decisions based on the tastes, opinions or purchasing habits of others (Welbourne and Grant, 2016; Shifman, 2012). Clever manipulations of these rankings could therefore have a significant impact on how people are influenced (Bishop, 2018; Gillespie, 2015; Easley and Kleinberg, 2010).

2.1.4 Cult Popularity

Some people prefer to engage with niche culture, such as a band, hairstyle or article of clothing, *because* it is unpopular in the sense of not mainstream (Akdeniz, 2014; Burnett, 2013; Browne, 2013; Yin et al., 2012), for example to gain acceptance with a non-mainstream group (Faafat et al., 2009). Something may also develop niche or cult popularity by others imitating the behaviour of prominent non-conformist thought leaders (Yin et al., 2012). This cult popularity sometimes transitions into mainstream acceptance, often then being subsequently rejected by the initial group (Akdeniz, 2014; Browne, 2013; Yin et al., 2012). It is possible that they like to feel that they are in a small and exclusive club, where they are the only ones who have the knowledge, understanding and refined taste to appreciate something (Akdeniz, 2014; Yin et al., 2012). This elitist mentality has the tendency to lead to people criticising things because they have mainstream popularity (Browne, 2013), or, conversely, to take pride in being early followers (Browne, 2013).

A counterculture is a subculture encompassing multiple behaviours that are consciously in opposition to the mainstream (Frank, 1998). The term counterculture implies a wider set of beliefs than for cult popularity, but countercultures can result in sets of related things, including YouTube videos, maintaining a small but non-trivial audience.

Perhaps as a result of YouTube being taken over by Google and having a greater emphasis upon commercial and advertising influences, videos that are not likely to generate revenue are less likely to be promoted by the site (Meehan, 2006). Videos that have previously developed cult status due to their producer's creativity, lack of conformity and/or controversy, are less likely to be seen and may be punished by YouTube (by making them less visible through the algorithm) due to not fitting in with their preference of cultural output or not being content that is aligned with advertising demands (Bishop, 2018; Cunningham et al., 2016). Therefore, YouTube has become more mainstream, having a much greater focus on more commercial videos and content, moving away from the original open essence of the platform (Bishop, 2018; Jarboe, 2012).

2.1.5 Herding Behaviour

Social structures influence behaviour (Eger, 2015). *Herding behaviour* is the tendency to conform to the ideas, opinions and behaviours of others within a social group or network (Boyd, 2014; Faafat et al., 2009).

The web has led to increased contact with a wider range of others, creating the potential for new types of herding behaviours (Alloway and Alloway, 2012; Faafat et al., 2009). Exposure to a wide range of information may lead people to imitate others rather than making informed decisions based on personal research (Burgess and Green, 2009; Bonabeau, 2004). They may also be more influenced by their social peers than recommendations from relevant field experts (Alloway and Alloway, 2012; Chen, 2007; Lee et al., 2001). For example, the "demand for a product increases when consumers believe more people have purchased this product" (Chen, 2007, p14; see also Hanson and Putler 1996). Herding behaviour has been argued to be mediated by social learning, reducing the likelihood that it leads to negative outcomes (Toyokawa, Whalen and Laland, 2019). It is not clear whether this applies to YouTube, however, and there is a lack of empirical evidence into the contexts in which herding behaviour is particularly strong (e.g., pop music vs. DIY advice).

2.1.6 Information Cascades

A period of increased sharing of something within a group, can be referred to as an *information cascade* (Easley and Kleinberg, 2010; Banerjee, 1992). Information cascades can be amplified by social pressures to conform (De Vany and Lee, 2008) and results in temporary or permanent popularity (frequently encountered) for the cascaded information.

Diffusion through information cascades can sometimes be maximized by seeding a piece of information or a new product with key influencers (Berger and Milkman, 2010; Keller and Berry, 2003; Weimann, 1994). These are chosen for exhibiting a combination of desirable attributes that allows them to influence a disproportionately large number of others (Gladwell, 2000), either directly or indirectly via a cascade of influence (Watts, 2002).

The following are important partly counterintuitive facts about information cascades (Easley and Kleinberg, 2010).

- The cascaded information does not have to be correct.
- Cascades can be started by poorly-informed influencers.
- Cascades can be easy to stop.

Large scale information cascades produced by targeting prominent influencers are difficult to achieve (Bakshy et al., 2011). Information can tend to spread more effectively via many small-scale cascades initiated by ordinary members of a group or network (Bakshy et al., 2011). The process of effectively cascading information relies heavily on the concept of word-of–mouth interactions, whether on or off line, where people are more likely to engage with information sharing and continued sharing based on trust with others within their social group (Bakshy et al., 2011; De Vany and Lee, 2008).

2.1.7 Word-of-Mouth

Word-of-mouth refers to information spreading through personal contacts, whether online or offline (Leskovec et al., 2006). This form of diffusion, or cascading, has long been regarded as in important mechanism by which information can reach large populations, possibly influencing public opinion (Berger and Milkman, 2013; Katz and Lazarsfeld, 1955). It has the potential to support the adoption of innovations (Rogers, 1995), highlight new products within the market (Bass, 1969) or stimulate brand awareness (Keller and Berry, 2003).

Word-of-mouth is a consumer-dominated form of communication and diffusion where the distributor is apparently independent of the sources of information discussed and is perceived to be more trustworthy (Shifman, 2012; Berger and Milkman, 2010; Brown et al., 2007; Heath et al., 2001). Traditional communication theory argues that word-of-mouth can have a powerful influence on behaviour and decision making (Silverman, 2001; Money et al., 1998; Brown and Reingen, 1987; Cox, 1963). Even a single word-of-mouth message may reach a wide range of receivers in this way (Lau and Ng, 2001). The success of individual word-of-mouth connections depends on the strength of the relationship between the sender and receiver, however (Brown et al., 2007; Haythornthwaite, 1999).

Word-of-mouth grew in importance due to the greater communication power of email and the social web (Brown et al., 2007). It can have particular power within online groups based on shared cultural pursuits and frequent interactions (Dehghani et al., 2016; Perrin, 2015).

2.1.8 Diffusion Theory

Online popularity can also be viewed from the perspective of diffusion theory. Diffusion theory concerns the process leading to the successful adoption of innovations (including ideas) across a social group (Rogers, 2003). The diffusion of a new information or innovation relies heavily on someone's ability to be able to relate to and adopt the idea that is presented to them. Rogers (2003) argues that there are four main influences on the spread of a new idea: the innovation, communication channels, time and the social system. This is relevant to videos because a new video can be viewed as an innovation that has the potential to be adopted. Diffusion theory suggests a process by which video popularity can be achieved.

The initial part of the diffusion process is demonstrating the innovation and its advantages (Rogers, 2003). The individual will determine the relative advantages and benefits of adopting the innovation in economic considerations, how much social prestige it will generate, the associated convenience and the satisfaction obtained (Rogers, 2003). The innovation must also be compatible with the current values and needs of the individual. There must be adequate time and accessibility to experiment with and test the innovation.

When a person has sufficient exposure to an innovation, they will decide whether to adopt or reject it (e.g., whether to watch a video or follow a YouTube channel). The final aspect of the innovation-decision process is to look for further reinforcement of their decision from others and as a result this aspect of the process has the potential to reverse their adoption decision (Rogers, 2003), for example by unfollowing a YouTube channel or deciding to watch a video despite initially rejecting it. The quicker the innovation-decision process is completed by users the sooner the innovation will be shared and diffused throughout the group.

Once adoption has been achieved by someone, the next diffusion process stage is sharing the innovation, for example by emailing a video URL to friends or posting it to Facebook. Repeated sharing leads to mutual awareness of the innovation within a social group (Rogers, 2003). Groups based on homophilous factors are more effective at diffusing innovations within their social network structure. Homophily within social networks can therefore lead to high levels of innovation diffusion (Galeotti et al., 2013; Golub and Jackson; 2009). Nevertheless, due to the nature of homophilous groups the diffusion can be limited to the boundaries of these groups, thus confining the diffusion (Rogers, 2003; Lazarsfeld and Merton, 1954). Elements of heterophily are therefore needed to ensure that bridges are available between homophilous groups (Rogers, 2003; Lazarsfeld and Merton, 1954). Based on this, Rogers (2003) argues that ideal groups should be homphilous in variables such as education, social status, but should be heterophilous regarding ideas, information and innovations.

There are significant differences between people in their willingness to adopt new ideas. Rogers (2003) defines the following adopter categories.

- Innovators are willing to take risks and will be the first to adopt an innovation, even if there is the potential for failure;

- Early Adopters are slightly more discrete in their adoption choices and are willing to accept new ideas, will think more carefully about the wider implications;
- Early Majority have more contact with Early Adopters and as a result will adopt an innovation after a longer period than Innovators and Early adopters;
- Late Majority have a high degree of scepticism towards innovation and will wait for most of society to adopt an innovation before adopting it themselves;
- Laggards are the last to adopt an innovation as they have an aversion to change and tend to be more focused on the 'traditions' of the social group.

These terms probably do not apply to individual YouTube videos because the decision to watch one is relatively trivial compared to, for example, the decision to adopt a new technology. Nevertheless, they may be relevant to decisions to follow prominent YouTubers or to engage with genres of videos. For example, it seems reasonable to regard the early followers of now famous YouTuber PewDiePie or early users of YouTube as a source of make-up or gaming advice as innovators.

2.2 Online Activity

This section discusses uses of the web, online interactions and relationships, online videos, video popularity and online information (including videos) sharing and dissemination.

2.2.1 The Social Web

The web is a significant facet of modern life in developed nations and a global platform for distributing and accessing information (Belanger and Jordan, 2020; Wodjao, 2020; Bozkurt et al., 2018; Carbonell et al., 2018). For many years, it has been the primary information source for a substantial fraction of the world's population (Carbonell et al., 2018; Bozkurt et al., 2018; Kellner and Kim, 2010; Duffy, 2008; Huberman et al., 1998), as well as for education, work, entertainment and socialising.

The web, access technologies and social networking have also provided societies new communication channels through which popularity may spread (Ahn and Shin, 2016; Kim et al., 2013; Manyika and Roxburgh, 2011; Kattimani et al., 2010; Wright and Hinson, 2009). As part of this, users for many years have been able to upload videos to the web from mobile devices with a few simple actions (Crick, 2016; Waldron, 2012; Ashraf, 2009; Li and Bernoff, 2008). There is now an online-based society of 24-hour sharers, consumers and commenters for a wide range of purposes (Carbonell et al., 2018; Oh and Syn, 2015; Shifman, 2007; Kuipers, 2002).

The social web is particularly suited to information sharing (Khan, 2017) and hence can play an important role in the generation and transmission of popularity. Stories, videos and cultural memes that gain traction in social media can do so quickly through online social networks but can also fade quickly (Kong et al., 2018; Broxton et al., 2013; Leskovec et al., 2009; Cha et al., 2007). Memes are, "individual bits of cultural information that propagate from person to person while undergoing variation, selection, and retention" (Guadagno et al., 2010, p2313). Most web memes are developed to provide users with humour, inspiration or social commentary (Knobel and Lankshear, 2007). On YouTube, memes could be individual videos or types of video, such as the "Downfall" genre mocking Hitler.

Marketers sometimes exploit the sharing power of the social web by developing content for sharing and diffusion through online friendship groups and social networks (Sokolova and Kefi, 2020; Barry et al., 2014; Greenberg, 2010; Purcell, 2010; Southgate, Westoby and Page, 2010; Madden, 2009). For example, music videos have benefitted from a greater global audience because of the web and the distribution of materials through continued sharing and online recommendations (Henke, 2013). This process of users sending and sharing links has become an increasingly important aspect of marketing music videos to a wider audience (Dehghani et al., 2016; Henke, 2013).

Nearly 60% of people report that they frequently share online materials with colleagues, family and friends, with approximately 88% of companies using social media to advertise their products (Dehghani et al., 2016; Allsop et al., 2007).

2.2.2 Reasons for Online Sharing

Even with the global scope and easy access of social media websites, there are some pieces of information, for example videos, that are accessed, shared and forwarded to a significant extent daily, whereas others disappear into obscurity (Wu et al., 2018; Welbourne and Grant, 2016; Guadagno et al, 2013; Waldron, 2012; Burgess and Green, 2009). Online information sharing may be influenced by different factors to offline sharing, and may be affected by the accessibility of information and the methods available for sharing. In particular, broadcasting is easier online than offline, for example through posting a video link to Facebook for all friends and followers. Content is more likely to be shared if it is *funny*, generates *excitement*, is *related to the sender or receiver,* or is believed to be *useful* (Tellis et al., 2019; Wu et al., 2018; Feitosa and Botelho, 2017; Berger and Milkman, 2013; Purcell, 2010; Verhaeghe et al., 2007). An *emotional response*, whether positive or negative also increases the likelihood of someone choosing to share information (Tellis et al., 2019; Guadagno et al., 2013).

2.2.3 Online Video Sharing and Virality

Online video sharing websites (see below), such as YouTube, have helped popularise posting, watching and sharing videos as a form of entertainment (Klobas and McGill, 2019; Arthurs et al., 2018; Klobas et al., 2018; Dehghani et al., 2016; Zink et al., 2009). YouTube has had a significant impact on the increased availability and number of videos on the web, which has led to more users accessing them (Arthurs et al., 2018; Siersdorfer et al., 2010). Social networking websites, such as Facebook and Twitter, have also enabled greater functionality to share and comment on videos (Khan, 2017; Eckler and Bolls, 2011; O'Malley, 2011; Tsai, 2009; Cashmore, 2009). This, coupled with increased broadband connectivity and improved mobile networks and technologies, has provided users with easier access online videos (Khan, 2017; Mehrotra and Bhattacharya, 2017; Cunningham and Craig, 2016; Yang and Gaunt, 2015; Wang, 2015).

Traditionally, videos were distributed by large media organizations directly to consumers, so choices were limited to switching to another centralised media organisation or turning off the TV (Cunningham and Craig, 2016; Broxton et al., 2013). These organisations determined which videos were good enough to be broadcast. In doing so, they had an impact on what could become popular (Broxton et al., 2013). Some online videos do not have the quality to make them suitable for broadcast media but are still readily available to view online, providing users with greater choice and decision-making capabilities (Welbourne and Grant, 2016; Broxton et al., 2013; Halpern and Gibbs, 2013; Shifman, 2012). There are differences in length, lifespan and content of online videos compared to traditional media (Cheng et al., 2007). The social networking and commenting aspects of online video websites, such as YouTube, are important for their continuing success, and the process of linking, rating, discussing, sharing and favouriting has the potential to make videos popular in an organic fashion (Ameigeiras et al., 2012; Alloway and Alloway, 2012; Cheng et al., 2007).

When a video is shared online at a very high rate then it might be classed as having 'viral' popularity or being **viral** (Tellis et al., 2019; Kong et al., 2018; Shifman, 2012; Crane and Sornette, 2008a). These videos demonstrate the significant power of social group and network sharing. Whilst the concept of virality implies that the videos are watched at least partly because they are popular, there is also a need for interesting and/or useful content (Rubenking, 2019; Kong et al., 2018). Viral videos tend to remain popular within a social network for a short period (Tellis et al., 2019; Kong et al., 2018; Brodersen et al., 2012; Chatzopoulou et al., 2010; Cha, et al., 2008; Cheng et al., 2008). Videos can become very popular without being viral if they are watched for reasons other than sharing or over a longer period (Frasco, 2014; Broxton et al., 2013).

2.2.4 Types and Measures of Online Video Popularity

The key popularity (in the sense of *frequently encountered*) metric of an online video is the number of times it has been viewed (Kong et al., 2018; Wu et al., 2018; Park et al., 2016). Other metrics can give complementary information about popularity (in the sense of *liked*), such as user likes and dislikes (Chang, 2018; Park et al., 2016). In contrast, the number of comments may reflect the level of engagement between viewers and videos (Chang, 2018; Park et al., 2016). Popular videos tend to receive more comments, ratings and likes (Chatzopoulou et al., 2010), so engagement and the two types of popularity are all related. A grey area is the importance of time. Since online videos do not seem to reach a fixed number of daily users but instead have a period of high daily views (e.g., when they are first posted), it is not clear whether there is a fair way to compare the popularity of videos of different ages. For example, a day old video with 100,000 views might be thought to be more popular than a decade-old video with 1,000,000 views in the expectation that it would eventually surpass the total view count of the older video. Nevertheless, there is no agreed formula with which to make such comparisons. It is fair to compare the total view counts of videos with similar ages, however.

2.2.5 Video Sharing and Online Social Networks

People are more likely to watch videos when they are recommended by trusted people within their online networks (Kayumovich and Annamuradovna, 2020; Hayes et al., 2018; Oh and Syn, 2015; Kim et al., 2013; Duncan et al., 2013; Halpern and Gibbs, 2013; Waldron, 2012). Although some videos are shared thousands of times through online social networks, others are never shared (Kong et al., 2018; Nelson-Field et al., 2011a; Cha et al., 2009; Zink et al., 2009). Whilst not all widely shared videos generate many views (Shifman, 2012; Berger and Milkman, 2012; Broxton et al., 2013), they are much more likely to become popular (Burgess, 2014; Broxton et al., 2013; Guadagno et al., 2013).

Virality requires users' active participation by passing information to others (Carvalho and Gomes, 2020; Kong et al., 2018; Eckler and Bolls, 2011; O'Neil, 2010; Freeman and Chapman, 2007). Thus, viewers' willingness to share a video through relevant sharing websites, social media, email and other online platforms is an important requirement for virality (Akdeniz, 2014; Scott and Judge, 2009). The elements that encourage someone to watch a video also needs to encourage them to share it if the video is to become viral (Eckler and Bolls, 2011; Burgess, 2008).

The structures on which networks are established and built might have an influence on social transmission and what is shared (Hayes et al., 2018). The sharing of material which has, or can have, a positive emotional impact on the recipient has the potential to support the development and reinforcement of the bonds between members if a group (Rubenking, 2019; Feitosa and Botelho, 2017; Quan-Haase and Young, 2010; Barker, 2009; Peters and Kashima, 2007). The more emotionally embedded and like-minded people happen to be with the rest of the members within their social group or network the greater the chance that material and information will be spread (Feitosa and Botelho, 2017; Guadagno et al., 2013; Waldron, 2012; Karpf, 2010). Members can increase the closeness or the bonds with others within their social group by establishing greater similarities with them through shared emotional experiences (Purcariu, 2018; Hanson and Haridakis, 2008; Anderson et al., 2003).

Sharing videos and materials online can have, or can be seen to have, an impact on someone's place and positioning within their social groups or networks (Tian et al., 2011; Kellner and Kim, 2010; Lee et al., 2001). Information can provide people with a form of social currency that they can share within their social groups and networks, and the attention they receive can give them the feeling of being helpful, in the know or in some way provides them with the status of being smart in the eyes of their peers (Rubenking, 2019; Feitosa and Botelho, 2017; Huberman et al., 2009). Videos that contain useful information can encourage viewers to share from the altruistic perspective of helping others within the social group or as a tool for self-enhancement to appear knowledgeable to network peers in a particular area (Oh and Syn, 2015). People can also be looking for the approval and social validation of others through the sharing and forwarding process (Cialdini, 2009; Rafaeli et al., 2005) and want to be perceived by others in a positive light (Feitosa and Botelho, 2017; O'Neill, 2011). People also share content on the web to make sense of their experiences, to reduce dissonance with others and to deepen social connections within their group (Berryman and Kavka, 2018; Oh and Syn, 2015; Berger and Milkman, 2012; Peters and Kashima, 2007; Hall, 2003). There can also be occasions when members of a group start to copy the sharing actions and opinions of others to seek social approval. This can include sharing materials even if they are of a poor quality, do not resonate with the individual or if there is no personal emotional connection (Kuznetsov, 2006; Salganik et al., 2006; Deci and Ryan, 2000).

Generally, the original source of videos has no impact on the likelihood that people will share a video with others, except if the forwarding reinforces a derogatory opinion of the sender (Guadagno et al., 2010; Guarino, 2010). On occasions material will be forwarded that is disparaging about others or that will paint others in a negative light (Wallsten, 2011; Nahon et al., 2011; Shifman, 2007). The process of sharing can be used as a tool for building and reinforcing social walls of separation between people or groups (Kayumovich and Annamuradovna, 2020). Sharing material with negative emotional content can be used to reinforce negative feelings aimed at people who are not members of their social group or network (Peters and Kashima, 2007; Fein and Spencer, 1997). This can be used to reinforce the positive in-group feelings that established the group in the first place (Peters and Kashima, 2007; Fein and Spencer, 1997).

People are more likely to share riskier content one-to-one rather than with the whole of their social group (Guarino, 2010). They do not want to share materials that will portray a negative image within their social group and therefore there is a difference in what someone will forward depending on who will receive it and the size of that audience, and how they consider they will be perceived as a result (Wallsten, 2011; Nahon et al., 2011). Social transmission is about more than just value exchange or the development of self-presentation (Rubenking, 2019; Berger and Milkman, 2012; Berger and Schwartz, 2011). People share to provide others with entertainment, surprise and relevant content, they also want to inform others, improve their mood and provide them with things and ideas that are practically useful (Rubenking, 2019; Aytar et al., 2018; Kardas and Brien, 2018; Feitosa and Botelho, 2017;Berger and Milkman, 2012).

A stunt, shock value, surprise or some aspect of novelty can also have an impact on whether a video is watched and forwarded (Al-Rawi, 2019; Shultz, 2015; Eckler and Bolls, 2011; Bruno, 2010; O'Neil, 2010). Viewers often watch videos that fall under the umbrella of "you have got to see it to believe it" particularly in this share-all online environment in which web users feel that they have seen everything (O'Neill, 2011; Kellner and Kim, 2010). Thus, a video that is produced which has new, contemporary or original content can be a significant draw to watch it (Rubenking, 2019; Arthurs et al., 2018; Nikolinakou and King, 2018; Shifman, 2012; Van Dijk, 2009; Shifman, 2007; Kuipers, 2002). Many brands and companies have used a range of shock tactics within their advertising videos to make them become very popular or viral (Barry, et al., 2014; Henke, 2013). Shock tactics and content that startles can also have an impact on peoples' willingness to watch and share videos, but this depends on whether their social, sexual and moral codes are violated (Al-Rawi, 2019; Gaunt, 2015; Bruno, 2010; Dahl, Frankenberger and Manchanda, 2003). With people having greater access, through the web and improved technologies, to a wider and more varied range of content, what used to trigger higher levels of emotional arousal, in the form of hilarity, shock and inspiration, does not always have the same effect and therefore the baseline standard has shifted. If users are exposed to content that is either too shocking or disgusting for them then this will have a negative impact on their desire to share the video with their peers and social group (Shultz, 2015; Henke, 2013; Eckler and Rodgers, 2011; Dahl et al., 2003). Users or companies producing videos have sometimes tried too hard to achieve a significant level of popularity or viral success by making their videos too unconventional and too controversial and have instead had a negative effect on watchers (Eckler and Rodgers, 2011; Dahl et al., 2003).

2.2.6 Video Sharing Platforms

The original online video hosting website founded in 1997 was ShareYourWorld.com and provided users with the opportunity to upload video in different file formats (Crick, 2016; Plesser, 2007). Unfortunately due to issues with transcoding technologies and limitations with internet access speeds it was discontinued in 2001 (Crick, 2016; Plesser, 2007). The development and proliferation of mobile devices, the ability to easily capture videos of aspects of their life (e.g. key events, birthdays, concerts) and individuals desire to share these productions led to the rise and the exponential growth of a number of video sharing websites and platforms (such as YouTube, Dailymotion, Netflix, Vimeo, Twitch) (Mehrotra and Bhattacharya, 2017; Cunningham and Craig, 2016; Yang and Wang, 2015; Saini et al., 2012; Frey, 2007). As some of the early difficulties in developing and establishing online video sharing websites were download speeds and differences in video formats, no universal video standard (Berry, 2018). In addition, the phenomenal growth and popularity of these video sharing platforms has also been supported by the development of a range of social media based websites, such as MySpace, Facebook, Twitter, which have provided individuals' with the opportunity to share videos with their social network groups (Yang and Wang, 2015). Many video sharing websites now provide users with the option of posting and sharing their videos directly onto social media platforms. For example, YouTube has a share button that provides a range of options of social media platforms where the uploaded video can be posted.

The connections developed between platforms such as YouTube, Facebook, Twitter and Instagram have led to their continued success in the very different and competitive way in which traditional broadcasting companies have worked in the past (Cunningham and Craig, 2016). A result of the development and integration of both video sharing and social media websites has been that videos sharing platforms have morphed into more than just a repositories for getting an individual's creations online and have now become social media platforms in their own right (Mehrotra and Bhattacharya, 2017) where users can interact with, rate and comment upon videos. The growth and development seems to be cyclical in nature, as the more developments that are made in relation to social media platforms the more individuals seem to be generating videos to share (Mehrotra and Bhattacharya, 2017).

The live video streaming platform Twitch has demonstrated significant growth over the last five years particularly with the social online gaming community (Johnson and Woodcock, 2019). The website is not just a space to share videos, but has provided a social space for like-minded individuals with common interests. Twitch has provided gamers with a platform where they can view and broadcast video game content, and has had a major impact not only on the gaming industry, but the global social media ecosystem (Johnson and Woodcock, 2019). The platform has helped the promotion of games (particularly those produced by independent game developers), advertising revenue for companies and individuals, support and advice and have provided an informal forum for game reviews (Johnson and Woodcock, 2019).

Google originally had its own video platform Google Video, but this focused on searching for video content already in existence on the web and the key aspects of this was that its basis was the Google search engine (Walczyk, 2008). When Google acquired YouTube, they continued to keep the independence of each of the platforms with YouTube still focusing upon UGC and Google video being an advanced video search tool (Walczyk, 2008). Google also had a third party connection with Metacafe and presented videos from this platform (Walczyk, 2008).

AOL also had its own online video sharing websites and, like Google, had two distinction platforms. The first AOL Video focused upon finding online videos and purchasable content, and the other Uncut Video provided users with the ability to upload their creations (similar to YouTube). The advantage for AOL was its backing of the media giant Time Warner (Holahan, 2006) and the fact that its search engine provided and highlighted content from other video sharing platforms as well as its own.

Unfortunately AOL lacked some of Google Video's powerful tools and features, plus it tended to over incorporate advertising throughout its videos.

Although some video sharing platforms have developed and experienced substantial growth over the years, others have not been as successful. iFilm was a video sharing platform established in the late 90s (before YouTube) and was intended as a website to make short independent films and was a key destination of the viral online videos of the time (Walczyk, 2008). IFILM was one of the leading video streaming platforms at the time (Walczyk, 2008), but, due to the burst of the dot-com bubble (2000), iFilm suffered from significant loses and is now discontinued.

Some video sharing websites have developed in a way to establish themselves as a different type of platform to YouTube, for example Vimeo. Vimeo came slightly before YouTube and was initially known for its advances in technology, establishing an on-demand video service and was one of the first to support high definition videos, earlier than YouTube (Jaakkola, 2020). The Vimeo website has established itself as a smaller, more focused community with a greater emphasis upon providing content that is more mature, tasteful, supportive, professional and academic (Jaakkola, 2020). It does not focus upon advertising before or throughout videos, which is seen as a big positive for the website and sets it aside from other platforms such as YouTube, and provides users with the option to pay to upgrade their account (Jaakkola, 2020). Vimeo allows users to upload videos across all genres with the proviso that the content should be original and demonstrate creative expression. Generally Vimeo is seen as a platform where there is a greater possibility of your productions being seen by a wider audience and your creation having greater visibility, which is of particular interest to those wanting to become "influencers" and monetising their productions (Jaakkola, 2020; Chiang and Hsiao, 2015).

Overall, online videos have taken two key directions over the last few years (Berry, 2018). There has been an increase in real time on demand video watching through websites such as Netflix, Amazon Prime, Hulu, Disney +, BBC iPlayer and HBO Online, where some are free with others requiring payment or a subscription fee. The other types of platform focus upon the uploading of user generated content and the social network aspects of video sharing such as YouTube, Twitch and Dailymotion (Berry, 2018). The success of video sharing websites, such as YouTube, is that they have provided users with a platform where anyone with video recording device can present their video creations to a global audience (Walczyk, 2008). A significant element in terms of YouTube's particular success was the development of transcode technology which allowed/s users to upload almost any file format to YouTube (Crick, 2016). The success of online video sharing platforms also appears to be partly due to their connections to popular social media networking websites and the ability to easily and quickly share video content through those connections. The visibility of your content and the chances of it been seen by users seems to be a consistent factor in the success of a video sharing website (Jaakkola, 2020). As a substantial number of individuals are uploading videos to various sharing platforms to increase the rankings and popularity of their productions even at times through illegitimate means (Bulakh et al., 2014). Video sharing websites also need to be able to provide users with a unique aspect to their platform that sets it aside from other, for example, Twitch focuses more specifically on gaming content (Johnson and Woodcock, 2019). It would also appear that those video sharing website that are part of, or have the backing of, technology giants, media conglomerates or new corporations, such as YouTube, stand a better chance of surviving and growing particularly if they run into legal and copyright difficulties based upon content that has been uploaded to and published on their website (Jondet, 2008). YouTube, not only being owned by Google has also made key deals and partnerships with global companies such as Sony and Apple, again further establishing it presence as the most popular and significant video sharing platform (Berry 2018). Approximately 1 billion unique users per month were accessing and watching videos on YouTube, and by 2017 a substantial 3.25 billion hours of video were being watched on the platform every month (Berry 2018). Video sharing is taken for granted by society due to the developments in technology and bandwidth, and the web provides individuals with a range of options and platforms in order to achieve this quickly and easily (Berry, 2018). However, YouTube still remains the most significant platform for video sharing with a massive and world-wide user base and seeks to maintain its monopoly over this corner of the social media market (Berry, 2018).

2.2.7 Factors Influencing Online Video Popularity

There are a range of general factors, not relating to a specific platform, that have been highlighted by previous research in terms of increasing the chances of online videos being watched and, as a result, becoming popular based upon views. Popular videos need a hook to draw people to watch them (Burgess, 2008). Many popular videos are reliant on *cuteness*, *humour, happiness, sexuality, nudity, shock value* or *violence* to encourage someone to watch (Arthurs et al., 2018; Berger and Milkman, 2012; Cheng et al., 2013; Daly, 2011; Eckler and Bolls, 2011; Nelson-Field et al., 2011b; Weiner, 2011). Alternatively, a video may offer a *unique perspective, story, personality* or *style* in the way in which it is presented or in relation the content (Nikolinakou and King, 2018; Tschopp, 2014; Ahn and Shin, 2016; Ahmed et al., 2013; Thelwall, Sud and Vis, 2012; Waldron, 2012; Ashraf, 2009). *Titles* may encourage viewers to watch a video, especially if they are short (Cheng et al., 2013; West, 2011).

There appear to be substantial differences in the range of advice that is provided in terms of when the best time to upload videos to YouTube (Andress, 2020; Ferreira, 2020; Perta, 2020; Parbey, 2019; Gielen, 2015; Cheng, 2013; Robertson, 2007). During 2007 it was suggested that the best time to post videos to YouTube was during the working week (Monday to Friday) with Wednesday and Thursday being the best (Robertson, 2007). Advice in 2013 had changed and it was now recommended that videos should be uploaded on either a Monday or Tuesday, as individuals would watch them whilst at work and as a result would gain momentum throughout the week (Cheng, 2013). Weekends were a "speed bump" in terms of uploading a new video to YouTube (Cheng, online, 2013). During 2015 suggestions had again shifted with Thursday and Friday being promoted as the best days to post videos, further implying that users would be less likely to watch on Sunday and at the beginning of the week (Gielen, 2015). Parbey (2019) recommends that early afternoon during week days and late morning for weekends. Ferreira (2020) and Andress (2020) suggest Thursday and Friday as the best days to post video on YouTube. They continue by putting forward that as most users access and watch videos over the weekend Saturday and Sunday can also be effective (Ferreira, 2020 and Andress, 2020). Andress (2020) goes on to imply that Monday, Tuesday and Wednesday are the worst days, completely disagreeing with previous advice from Cheng (2013). In general terms Perta (2020) recommends that videos for business purposes are best posted between Tuesday to Thursday, with those of a non-business nature being more successful when posted Friday to Monday. In terms of other social media there is some agreement about when it is best to post, but also some differences in opinion. When posting to Facebook and Instagram the majority of current advice suggests Wednesday as the optimum day (Arens, 2020; Hearn, 2020; Powers, 2019; Vasquez, 2019). However, others recommend later in the week and the weekend for Facebook (Ferreira, 2020; Kolowich, 2019) and Friday for Instagram (Kolowich, 2019). In terms of Twitter there is more variation in when is best to post. Ferreira (2020) suggest Wednesday, Hearn (2020) and Powers (2020) recommend Tuesday and Wednesday, Arens (2020) advises Wednesday and Friday, and Kolowich (2019) highlights weekdays for business and weekends for personal content.

The *time of day* that the video is emailed or posted online can also influence its popularity (Arens, 2020; Ferreira, 2020; Hearn, 2020; Kolowich, 2019; Vasquez, 2019; Berger and Milkman, 2012). The extent to which a video is discussed in the media influences the likelihood that it will eventually become popular (Tschopp, 2014; O'Neill, 2011; Brown et al., 2007). The following types of video have also been singled out as popular.

Video length may be important, with shorter videos tending to be viewed and shared more often (Tschopp, 2014). Most online videos are short, with the average length being 20 to 40 seconds and the majority lasting under 4 minutes (Cheng et al., 2013). Videos between 16 and 30 seconds long are likely to be viewed and shared more often than those of length 15 seconds or less, or videos of 30 seconds to 1 minute (Henke, 2013). Other research has found that a popular length of video is 30 minutes with this being typical of one episode of content from TV (Rigby, et al., 2018).

Fails: A popular theme of video that has become very successful on the web are those which fall into the contemporary popular culture category of *fails* (O'Neill, 2011; O'Neil, 2010; O'Neil, 2010a; Porter and Golan, 2006). The term fail used in this way refers to a person making some sort of mistake or providing an instance of very poor performance, for example, unsuccessfully trying to complete a trick on a skateboard (Shultz, 2015). Most fail videos contain some element of a person hurting or embarrassing themselves, usually after attempting to complete some action or activity, usually one that should not be attempted in the first place (Shultz, 2015; O'Neill, 2011; O'Neil, 2010; Brown et al., 2010; Porter and Golan, 2006). Many of the videos that are viewed on websites, such as YouTube, are 'fail' related (O'Neil, 2010). This suggests that a proportion of society find the misfortunes of others entertaining (O'Neil, 2011; Porter and Golan, 2006), but this seems to be situations where the fail is determined to be deserved by the watcher (Gupta, 2015). It also suggests that people just like to see that they are not the only ones who make mistakes and an element of human vulnerability (Gupta, 2015; Shultz, 2015). There is also an element of surprise to most fail videos and a common human reaction to these unexpected, and sometimes bizarre, events is to initially laugh (O'Neil, 2010a; Verhaeghe et al., 2007). The idea of people getting things wrong also extends to the concept of bloopers or gag reels, such as actors making blunders, forgetting their lines, being unable to control fits of laughter or tripping over whilst on set, further reinforcing users' fascination with others making mistakes (Shultz, 2015; O'Neill, 2011; Verhaeghe et al., 2007). There is also some element of ego to the watching of such videos and people also enjoy feeling superior to the actions of others in the video (O'Neill, 2011). Overall, it is more the unexpected, surprise and disbelief factor that makes these types of videos funny rather than the injury of the participant of the video (Gupta, 2015; Shultz, 2015; O'Neil, 2010a).

Remixes: Remixing, or redesigning, a video that has already been popular can improve the chances of a video developing a high or possibly viral level of popularity (Henke, 2012; Bruno, 2010). Video parodies of current pop culture, in particular pop videos (Henke, 2012) or internet memes (Knobel and Lankshear, 2007), are more relevant to a wide range of viewers and tend to be accessed more regularly (Guadagno, 2013; O'Neil, 2010; O'Neil, 2011). This does depend on how well the parody is produced and the original popularity of the pop star or video being parodied (Henke, 2012; O'Neil, 2010). The quality of the remixed content and its production, and its ability to have an impact and stick in the mind of the individual also increase the chances of a video being shared (Peer and Ksiazek, 2011; Huang et al., 2011). Poor quality videos will be less appealing to viewers and could be a barrier to them accessing, or engaging with, the desired content (Tschopp, 2014; Peer and Ksiazek, 2011; Huang et al., 2011).

News: News videos can become popular if they are timely and engaging (Wu et al., 2018). Real world events can influence which videos are accessed (Abisheva et al., 2014; Pinto et al., 2013; Berger and Milkman, 2012; Van Zoonen et al., 2011; Crane et al., 2010). If a video relates to the news then it may have a greater likelihood of becoming popular (Wu et al., 2018; Abisheva et al., 2014; Berger and Milkman, 2012; Van Zoonen et al., 2011; Van Zoonen et al., 2011; Crane et al., 2010). Also, a video, not relating to the news, can have a greater chance of success and becoming popular if it is posted or emailed on a slow news day (Berger and Milkman, 2012).

Amateur: Videos may become popular for unexpectedly poor quality for their genre. This includes poor production, laughable scripting, clichéd dialogue, substandard acting and an overall cheesiness to the video (Barakat, 2014). The draw or hook can be the cringe-worthy nature of the video and as a result can be thoroughly engaging for some social groups or networks. Some users may enjoy sharing and criticising poor content (Wallsten, 2011; Nahon, et al., 2011). Such discussions and infamy may draw attention and increase the popularity of the videos.

Algorithm: All video sharing platforms use algorithms in order to determine, based upon a range of factors, which videos are promoted or shown (Bishop, 2018; Jarboe, 2012). Commercial pressures, particularly through advertising, influence what is presented to audiences, for example in lists of related videos after a user has watched their original choice. YouTube therefore has an influence in terms of what becomes popular that it exercises by modifying its algorithms to highlight specific content in line with the commercial focus and advertisers' demands (Bishop, 2018; Cunningham et al., 2016; Meehan, 2006).

2.3 Finding and watching YouTube videos

There are multiple ways in which users can encounter YouTube videos and multiple potential influences on their decisions to watch.

2.3.1 YouTube Video Search and Recommendations

As with most websites, YouTube provides a range of inbuilt systems to present users with selections of videos, for example through the front page, most popular, most viewed today or recommendation pages. Some of the videos presented by YouTube are a result of the site-wide algorithms it uses (Bishop, 2018; Crane and Sornette, 2008) and others are algorithmically selected from user preferences, viewing history and watching behaviours, for example by suggesting videos related to previously watched content (Bishop, 2018; Klobas et al., 2018; Gielen and Rosen, 2016; Zhou et al., 2010).

Although the YouTube search facility can be used to find videos, this sometimes returns only the most popular videos (Pinto et al. 2013; Zhou et al., 2010; Crane and Sornette, 2008) or content that is more commercially focused (Arthurs et al., 2018; Bishop, 2018; Jarboe, 2012). This makes more specialised or niche content harder to find or even hidden (Bishop, 2018; Cunningham et al., 2016). Nevertheless, the YouTube search facility tends to be influenced by key YouTube channel producers and videos that garner higher levels of controversy and dissent from users (Rieder et al., 2018).

The recommendations facility in YouTube is a key factor in supporting the discovery of new material. Videos at the top of the recommended videos list generated by YouTube tend to have high view counts either as a cause or effect of the recommendation (Zhou et al., 2010). Despite this ambiguity, being at the top of the recommendation list seems likely to generate substantial numbers of extra views (Zhou et al., 2010). The personalised aspect of the YouTube recommendation system suggests videos based a user's video history and YouTube channel subscriptions (Figueiredo et al., 2014). YouTube can be quite powerful through its algorithms directing users to content that they might not have otherwise found. This may tend to promote mainstream content with a general level of popularity (Bishop, 2018).

2.3.2 YouTube Video Popularity

Due to the continued uploading of new YouTube videos, each one tends to have a small window of opportunity and a low probability of becoming popular (Kong et al., 2018; Pinto et al. 2013; Guadagno et al., 2010). Moreover, video popularity is unbalanced, with most content receiving a minimal level of attention and a small proportion attracting huge numbers of views (Welbourne and Grant, 2016). This imbalance can be modelled with power-law distributions and Zipf's law (Cha et al., 2007; Figueiredo et al., 2011). Understanding what makes YouTube videos popular could be very useful information to a range of bodies such as marketers, companies, information services and researchers (Sokolova and Kefi, 2020; Kayumovich and Annamuradovna, 2020; Ahmed et al., 2013; Pinto et al., 2013). It is difficult to accurately predict the popularity of YouTube videos from their content, however

(Welbourne and Grant, 2016; Ahmed et al., 2013; Zink et al., 2009). This section reviews some known or hypothesised factors, extending the previous section about online video popularity.

A key difficulty in determining which YouTube content will become popular is that videos can follow complex journeys in how their popularity evolves (Pinto et al., 2013; Figueiredo et al, 2011; Crane and Sornette, 2008a). A factor that contributes to the different journeys that videos take to becoming popular is the many internal and external features and mechanisms of interacting with YouTube content (Törhönen et al., 2019; Pinto et al., 2013; Lerman and Hogg, 2010). Thus, a consideration that makes it more difficult to predict the popularity of YouTube videos is the number of external factors that can influence whether content is watched and then shared, for example current events happening in the news or if it is posted on a social media website (Abisheva et al., 2014; Ahmed et al., 2013). It is also a challenge to forecast how YouTube will be cascaded throughout a social group or network, which will also have a substantial impact on whether a video becomes popular (Ahmed et al., 2013). Knowing how viewers will react to or engage with different video content can also have an impact in its popularity and is particularly challenging to decipher (Welbourne and Grant, 2016; Ahmed et al., 2013).

A study investigating the relationship between opinions and popularity for a set of 306 YouTube videos found that the title, thumbnail and content of a video helped with its popularity much more than the number of likes, dislikes or the sentiment of comments (Chang, 2018).

Videos in YouTube can become viral through being shared on social media and perhaps also by being recommended in the website homepage. Virality is difficult to investigate directly on YouTube, however, because direct public sharing within the site cannot occur.

2.3.3 Sharing YouTube Videos

YouTube and other online videos are routinely shared (Rubenking, 2019; Tellis et al., 2019; Wu et al., 2018; Khan, 2017; Zink et al., 2009). The sharing of YouTube videos and the associated links within and across a range of other forms of social media can help them become popular (Khan, 2017; Abisheva et al., 2014; Jenkins et al., 2009). YouTube has made it easy for users to share videos either by linking on other websites, directly posting in other social media platforms, such as Facebook and Twitter or through pasting a URL within an email (Buzzetto-More, 2014; Freeman and Chapman, 2007). Sharing videos showing consumer responses to products are a form of electronic word of mouth communication (Bi et al., 2019). Here, opinions are expressed in videos and users share those that they agree with, with the trust relationship between the sharer and recipient being important for influence rather than direct trust in the video creator (Sokolova and Kefi, 2020; Nouri, 2018).

Views of YouTube videos can be increased because of sharing through discussion and recommendation from a range of on- and offline platforms which include email, blogs, television, links on other websites and social media, newspapers, conversation and because of being featured within other YouTube videos (Crane and Sornette, 2008). Generally, people enjoy sharing information, resources, pictures and videos within their social groups (Oh and Syn, 2015). If a social group, network or community engages with a YouTube video then there is more chance that it will be shared and continued to be viewed and shared, thus increasing the view count and popularity of the video (Crane and Sornette, 2008). There are a range of theories explaining and discussing what motivates people to share information through social means with some of the key ideas being:

- Maslow's (1946) hierarchy of five basic human needs;
- Horton's (1983) hierarchy of information needs;
- Deci and Ryan's (2000) model of intrinsic and extrinsic motivation;
- Herzberg's (2008) two factor theory – motivation and hygiene

(Oh and Syn, 2015)

In general terms, a person's motivation for sharing can fall into two key categories (Oh and Syn, 2015). The first is that they want to share knowledge, information and innovation to support the development, progress and advancement of a social group or society in the hope that others will do the same (Carvalho and Gomes, 2020; Ahmed et al., 2019; Rosenthal, 2018; Wasko and Faraj, 2005, Hall, 2003; Jarvenpaa and Staples, 2000; Ekeh, 1974). The other is that people will provide and share content and material again to support, improve and solve problems, but with the expectation of some reciprocity or reward for providing these services (Hou, 2019; Arthurs et al., 2018; Bishop, 2018; Oh and Syn, 2015; Wasko and Faraj, 2005; Hall, 2003). In sharing YouTube videos motivations tend to encompass the following:

- connecting and communicating with peers (Kayumovich and Annamuradovna, 2020;Quan-Haase and Young, 2010; Barker, 2009; Hanson and Haridakis, 2008; Joinson, 2008; Nov, 2007);
- developing a feeling of social connectivity (Khan & Vong, 2014; Liao et al., 2011; Hanson and Haridakis, 2008);
- support and community development (Oh, 2012; Hsu and Lin, 2008; Kuznetsov, 2006);
- feeling of usefulness (Rubenking, 2019; Feitosa and Botelho, 2017; Lin and Lu, 2011; Rafaeli et al., 2005; Bandura, 1997; Constant et al., 1994; Herzberg et al., 1993);
- developing their identity, personal expression and reputation (Shifman, 2012; Hanson and Haridakis, 2008; Richter and Koch, 2008; Eisentraut et al., 2001; Deci and Ryan, 2000);
- knowledge and information exchange (Carvalho and Gomes, 2020; Ahmed et al., 2019; Wu et al., 2018; Lin and Lu, 2011; Hanson and Haridakis, 2008; Richter and Koch, 2008; Kuznetsov, 2006; Rafaeli et al., 2005; Blau, 1964);
- education and learning (Arndt and Woore, 2018; Aytar et al., 2018; Kabooha and Elyas, 2018; Kardas and Brien, 2018; Nam et al., 2009; Nov, 2007; Rafaeli et al., 2005);
- entertainment and enjoyment (Arthurs et al., 2018; Klobas et al., 2018; Quan-Haase and Young, 2010; Hanson and Haridakis, 2008; Hsu and Lin, 2008; Hills et al., 2000);
- personal gain (Hou, 2019; Arthurs et al., 2018; Bishop, 2018; Oh and Syn, 2015; Deci and Ryan, 2000).

2.4 YouTube Culture

The original purpose of YouTube website was to provide a forum for sharing consumer-generated videos relating to current and well-known events (Crick, 2016; Hopkins, 2006). YouTube was subsequently bought by Google in 2006 (Welbourne and Grant, 2016). YouTube has become globally popular (Arthurs et al., 2018; Broderson et al., 2012) with almost two-thirds, 61%, of people using the web having watched and shared a video on the website (Eckler and Bolls, 2011), and with 70% of viewers being outside of the USA (Broderson et al., 2012) by 2012. YouTube is the world's second most visited website (Alexa, 2019) and the web's largest video sharing site (Alexa, 2019) by 2019, with approximately 100 hours of video content uploaded every minute by 2014 (Gaunt, 2015).

One possible reason for the original success of YouTube is that users can upload videos of any quality and in different file formats, making it easy to publish (Crick, 2016; Gill et al., 2007). Further developments in technology and improved access to the web, have made it even easier to produce and publish videos to YouTube (Dehghani et al., 2016; Gaunt, 2015; Duncan et al., 2013; Alloway and Alloway, 2012).

YouTube's progressive download technique allows users to start watching a video whilst it is being downloaded, eliminating the need to wait for the file transfer to finish (Crick, 2016; Gill et al., 2007). This was a revolutionary feature when YouTube launched because it addressed the slow internet download speeds of the time (Crick, 2016; Ameigeiras et al., 2012; Gill et al., 2007). YouTube is also an inexpensive and instantaneous way of getting your message out and sharing your experiences to a potentially global audience (Wotanis and McMillan, 2014; Kellner and Kim, 2010; Freeman and Chapman, 2007).

One of the advantages of YouTube for video producers is that traditional media barriers to getting their creations published to a global audience can be bypassed (Amarasekara and Grant, 2018; Cunningham and Craig, 2016) and as a result there is a substantial difference in the content and quality of videos (Zink et al., 2009). YouTube provides very little regulation in terms of general production quality and content of videos such as camera work skills, quality of picture and sound or the interest/popularity of the subject matter (Cunningham and Craig, 2016; Zink et al., 2009), with the exception of copyrighted or fraudulent material (Bulakh et al., 2014; Jondet, 2008; Frey, 2007), socially inappropriate images, pornography and/or illegal material. However, there are issues with the moderation processes that are employed by social media websites in order to locate and remove certain content. When people are employed to check and monitor content it can have a negative impact on the mental health and well-being of those individuals because of the material they are exposed to (Dwoskin et al., 2019). Some social media platforms, including YouTube, have used non-human means of removing ‘inappropriate’ content, such as algorithmic content moderation (Binns et al., 2017). Unfortunately, as a result of using these approaches to content moderation and filtering some levels of bias have emerged, with some content being unfairly penalised and removed, for example, material relating to LGBT issues (Binns et al., 2017; Hunt, 2017). YouTube also offers a Partner Programme which provides users with access to a range of additional resources and features to monetise videos (Google, 2019). Individuals have to agree to specific policies and guidelines in order to qualify for the scheme and the content on the YouTube channel is monitored and checked by human reviewers providing further monitoring of the appropriateness of video content (Google, 2019). In addition, based upon the commercial priorities of YouTube the platform promotes what it determines to be higher quality content that meets their advertisers demands (Bishop, 2018).

YouTube provides a different experience to broadcast television, which is one of its natural competitors. Users can be active with posting, liking, sharing and commenting on videos or passive watchers. In terms of active users, YouTube has supported a significant proportion of society moving into a post-television more interactive era, with greater control and fewer boundaries (Strangelove, 2010; Tolson, 2010; Gaunt, 2015). YouTube has provided users with the ability to:

- Choose what to watch and when
- Rate videos with likes and dislikes
- Post comments on videos
- Access similar and related videos and content on demand
- Share and link to specific content
  (Bentley et al., 2019; Jacob et al., 2018; Rigby, et al., 2018; Khan, 2017)

Videos uploaded to YouTube provide some content that is more reflective of contemporary events, issues and the culture of a community rather than the material distributed through more traditional means of broadcasting (Burgess and Green, 2009).

2.4.1 User Demographics

YouTube is most popular within the 18 to 34-year-old age range, with some of these using the platform more regularly than more traditional methods to watch videos (Blank et al., 2019; Gaunt, 2015; Perrin, 2015; Smith, 2014).

In terms of gender, YouTube has been used more regularly by males than females (Fisher and Ha, 2018; Mayoral et al., 2010; Madden, 2009; Molyneaux et al., 2008). Rainie et al. (2012) suggest that YouTube use is relatively equal across genders. Other studies suggest that females have become more prominent users (Oh and Syn, 2015; Chappell, 2012) or have shown greater interest in the platform over the past 10 years (Fisher and Ha, 2018) Research has shown that males generally have more favourable attitudes to technology than males (Cai et al., 2017), however, although both genders have similar levels of computer-based skills females are less confident (Cai, Fan, & Du, 2017). It has also been highlighted that females are more likely to use social media in general than males (Carbonell et al., 2018; Duggan and Brener, 2013). There is a significant gender imbalance in top most-subscribed channels on YouTube (Döring and Mohseni, 2019), with 20% produced by females (Lauzen, 2012; Manning and Shogan, 2012; Soares et al., 2011). Female YouTubers receive a higher proportion of negative responses to their videos than males (Wotanis and McMillan, 2014). With the developments and progress of YouTube and increase of in interactions with online videos, these gender differences could have altered over recent years. Males are less concerned with rating negatively (through dislikes) and are more likely to post comments relating to videos (Khan, 2017).

Gender differences in the types of videos commented on and watched on YouTube have been found (Amarasekara and Grant, 2019; Fisher and Ha, 2018) and in some specific contexts, including for museums (e.g., males watching military museum videos and females watching art museum videos) (Thelwall, 2018) and science education videos (Thelwall and Mas-Bleda, 2018). The topic of a science video seems to be more important than the gender of its presenter, however.

Other research (Bishop, 2018) has found gender biases within YouTube (and how its algorithm is engineered) due to its need for commercially-relevant content to meet the needs of its advertisers. The platform has developed an emphasis towards encouraging the production of content that is highly gender focused in some areas (Bishop, 2018). In addition, YouTube seems to be more focused upon males in terms of the content and producers (including vloggers) than females (Bishop, 2018). This has led to a divide between male and female content in some areas, an emphasis upon stereotypes and with the overall result being “a clear gendered bifurcation of content on the platform” (Bishop, p70, 2018).

Whilst most research about YouTube has focused on its US or Western audiences (it is banned in China), it is a global brand with different international trajectories and competitors. For example, in India, the ability for users to access local and regional content has been crucial to its success (Mohan and Punathambekar, 2019).

2.4.2 Social Networking

YouTube provides viewers with social networking opportunities (Klobas et al., 2018; Khan, 2017; Khan and Vong, 2014) and features, such as the ability to have a home page (channel), and to message and follow eachothers’ channels. Loose virtual communities can also form within YouTube by followers of popular channels communicating through comments left on videos. These virtual communities have led to people interacting with a wider range of other users within YouTube and as such are developing new personal online relationships and feel some sense of social connectedness (Alloway and Alloway, 2012). For example, some people have felt comfortable enough with their YouTube communities that they have used YouTube to come out (Tian et al., 2011; Alexander and Losh, 2010). Through its social networking features, YouTube has become a social media ecosystem fuelled by it users across the globe and has developed into a constantly evolving and complex online plurality of environments (Gaunt, 2015; Alloway and Alloway, 2012; Wesch, 2009).

2.4.3 Opinion, Debate and Emotions

As people can now publish their videos to a world-wide audience YouTube has become mainstream soap box for the publishing of opinions and commentary relating to current affairs and issues (Shapiro and Park, 2018; Van Zoonen et al., 2011; Strangelove, 2010; Van Zoonen et al., 2010). Within YouTube it is possible to leave comments and opinions relating to the videos they have watched and respond to what others have posted (Kahn, 2017; Thelwall et al., 2012). This feature provided by the YouTube platform has sometimes led to the formation of a wide variety of debates and discussions around a topic, theme or issue (Hussain et al., 2018; Thelwall et al., 2012; Van Zoonen et al., 2011; Van Zoonen et al., 2010; Jones and Schieffelin, 2009). Users have a choice in whether they decide to respond to in text, which is the case, or through a further video presentation (Kellner and Kim, 2010). Within YouTube, there are apparently organised groups that interact though comments about non-mainstream issues (e.g., conspiracy theorists: Alassad et al., 2019).

YouTube has evolved into a significant and cultural destination for the integration of social and emotional experiences (Strangelove, 2010; Gill et al., 2007). As well as hosting videos for recreation it has also become a forum for addressing and discussing contemporary issues (Al-Rawi, 2019; Van Zoonen et al., 2011; Vis and Mihelj, 2010; Van Zoonen et al., 2010). The ability to post comments and feedback relating to videos within YouTube has provided viewers with the opportunity to generate a wide range of debates around contemporary and cultural topics and issues (Thelwall et al., 2011; Strangelove, 2010; Cha et al., 2007). Some viewers, because of such debates, have created their own videos in response to those of others. The virtual YouTube community discussions reflect contemporary cultural ideas, issues, politics and trends, and is thus embedded with high levels of both unity and conflict (Strangelove, 2010). Some of these video-based online interactions and discussions can invoke significant emotion and feeling resulting in quite intense and passionate debates which are prevalent with collaboration and understanding, but at the same time can cite substantial levels of conflict and discourse (Thelwall et al., 2012). These discussions are anonymous, lack any of the structures or formalities of traditional debates and seems to have no boundaries (Thelwall et al., 2012). Due to the nature and forum of such debates the result is that an ongoing record is left of the comments and opinions that have been made (Thelwall et al., 2012). People are now more immersed and actively engaged with what they are watching than they have ever been, and there has been a metamorphosis and transition from the traditional model of a passive viewer (Bau-Franch et al., 2012; Thelwall et al., 2012; Moor et al., 2010). Some of the discussions that are initiated through interactions with videos are also continued in the offline environment further demonstrating the significant impact and influence that YouTube can have on society (Milliken et al., 2008; Milliken et al., 2008a).Those who comment on YouTube videos can remain anonymous or can give themselves an online name or pseudonym. This anonymity can provide the opportunity to ignore social norms and lead to anti-social or abusive online behaviours (Thelwall et al., 2012; Alonzo and Aiken, 2004; Freidman, et al., 2000; Burgess and Green, 2009). As a high proportion of the videos that are uploaded to YouTube are amateur, they can sometimes provide a more intimate and real context for the viewer which can lead to a greater level of empathy and fewer negative responses (Molyneaux et al., 2008). Those discussions that maintain a more negative tone, contain a greater level of undesirable remarks, are at times very abusive in nature and focus on more controversial topics tend to have longer threads of comments (Chimel et al., 2011; Sobkowicz and Sobkowicz, 2010; Stauff, 2009; Lange, 2007a). As a result, some YouTube performers will accept or ignore negative and hateful comments in relation to them and their video as the increased interaction can lead to increased popularity (Wotanis and McMillan, 2014; Lange, 2007a). Unfortunately, there is a tendency for the public to trivialise online abusive and harassing behaviours and which can have the result of discouraging viewers from sharing their videos on YouTube (Wotanis and McMillan, 2014; Citron, 2009). Videos that focus on sport seem to initiate passionate, negative and controversial discussions (Stauff, 2009), as do videos related to radicalism (Murthy and Sharma, 2019).

2.4.4 Self-Expression

YouTube is used for self-presentation, self-expression to a potentially global audience of millions of users (Berryman and Kavka, 2018; Gaunt, 2015; Shifman, 2012; Kellner and Kim, 2010; Ashraf, 2009; Wesch, 2009; Wesch, 2008; Freeman and Chapman, 2007; Grossman, 2006; Kuipers, 2002). YouTube has provided people with the opportunity to create their own personal online video profile or persona, through a channel page, which they can present to the world (Sokolova and Kefi, 2020; Alloway and Alloway, 2012; Kellner and Kim, 2010; Freeman and Chapman, 2007; Lange, 2007).

YouTube is also used to upload vlogs, diary-like regular personal videos, which have become a popular form of recreation and entertainment both for creators and viewers (Arthurs et al., 2018; Berryman and Kavka, 2018; Juhasz, 2009; Freeman and Chapman, 2007). These can also have a serious purpose, such as helping to deal with traumatic events by sharing experiences (Schuman et al., 2019). Coming out through YouTube (Lovelock, 2019) may also be a form of self-expression.

2.4.5 YouTube Influencers

YouTube celebrities are an important aspect of its popularity (Sokolova and Kefi, 2020; Nouri, 2018). This is perhaps not well understood because it does not seem to have a direct offline counterpart. A content analysis of videos in ten major YouTube channels (i.e., from YouTube celebrities) found gender differences in content (e.g., males were more likely to post gaming-related videos) but positive and negative self-disclosure were important elements of the videos (Ferchaud et al., 2018).

One of the reasons for the success of YouTube celebrities is that they may be paid for endorsements in their videos, enabling them to be professional and increasing competition (Gerhards, 2019; Bishop, 2018). They are sometimes called “influencers” because their close bond with followers of their channel formed through para-social interactions (Rihl and Wegener, 2017) translates into the power to market products (Munnukka et al., 2019). An investigation of the perception of a YouTube celebrity Misslisibell by Swedish children found that she was quite influential and successfully blurred the advertising/information boundary in her videos (Martínez and Olsson, 2019). YouTube influencer endorsements are not always successful, however. For example, a randomised controlled trial of influencers promoting healthy eating was unable to detect positive outcomes, whereas promotions of unhealthy foods lead to increases in their consumption (Coates et al., 2019).

2.5 Topics on YouTube

Many different and exotic topics are posted to YouTube. For example, it supports criminal investigations providing victims and the authorities with the opportunity to circulate video evidence including information relating to missing people or material showing crimes being committed (Kellner and Kim, 2010; Tucker, 2007). This section reports some important types of content on YouTube, using some of YouTube’s own category names, as analysed in this thesis.

2.5.1 Beauty

Beauty vlogging is an important aspect of YouTube, generating many YouTube celebrity influencers (Bishop, 2018; Banet Weiser, 2017; Nathanson, 2014). These give make-up and fashion advice (usually from bedrooms) and, when they have a large enough audience, can earn money through paid product endorsements and can also expect to receive free products to review (Bishop, 2018; Hou, 2018; Dryden, 2016). Nevertheless, in order to be successful vloggers need to carefully consider what YouTube wants in terms of video content, from a commercial and advertising point of view, to remain visible to audiences (Bishop, 2019; Jarboe, 2012).

2.5.2 Causes

YouTube provides an important forum for exposing public interest concerns, such as poor company practices, contributing to mainstream media coverage, affecting movements with the stock markets, and supporting acts of revenge (Ayres, 2009; Sykora and Panek, 2009a; Gueorguieva, 2008). It has been investigated for public awareness information about the global water crisis (Krajewski et al., 2019). No studies seem to have claimed that YouTube is effective for disseminating non-political cause-related information, however, and one has found that end users do not use it to seek sustainable development information (Kang, 2019).

2.5.3 Education

Although not one of the original purposes of YouTube, it hosts many educational videos (Belanger and Jordan, 2020; Tackett et al., 2018; Olasina, 2017; Duncan et al., 2013; Haran and Poliakoff, 2012). Today’s students have grown up with a wide range of interactive and multi-media technologies that can be integrated into teaching and learning (Belanger and Jordan, 2020; Savage and Barnett, 2017;

Duval et al., 2016). Due to the prominence of technology and the Internet in their day-to-day world our current students used to be referred to as the Net Generation (Duncan et al., 2013). The use of video is an effective motivational and educational tool (Luo, et al., 2020; Kurt, 2019; Klobas et al., 2018; Jones and Graham, 2013).

Educators are developing their understanding of the pedagogical benefits of YouTube as a key tool, innovative and fresh approach for developing teaching and learning (Savage and Barnett, 2017; Balbay and Kilis, 2017; Styati, 2016; Forbes, 2015). Two key factors that must be considered when using YouTube for educational purposes are the credibility of the video content and the time that it can take educators to find the materials they require (Jones and Graham, 2013). The length of educational YouTube videos can have an impact in whether someone will make the decision to watch, with a length of over 15 minutes being off-putting (Greenberg and Zanetis, 2012). Related to this, students, as a result of the technological world they have grown up in, have increased expectations in 'on demand' content and the concept of instant and immediate feedback which can be also be addressed, to some extent, through educational interactions with the YouTube platform (Clifton and Mann, 2011; Duffy, 2008; Skiba and Barton, 2006; Prensky, 2004).

Discussion, active participation and critical thinking are effective teaching and learning tools (Buzzetto-More, 2014) and YouTube's features of sharing and commenting can be used to develop these important skills (Kellner and Kim, 2010). YouTube also provides students with the opportunity to share educational videos and engage with and create online learning communities where all have the power to contribute, support and have their voice heard (Hilner, 2012; Tan and Pearce, 2012; Logan, 2012). Using YouTube, students can also develop their ability to participate in self-directed learning (Liu, 2010). There are also a wide range of YouTube videos that have provided people with the opportunity to educate themselves on many different topics for free (Buzzetto-More, 2014; Kellner and Kim, 2010).

Through the use of the YouTube platform there is the potential to bring aspects of educational practice more up-to-date, inspire and motivate students, maintain their attention and make the learning experiences they undertake be more memorable, interactive and relevant to real life (Luo, et al., 2020; Kurt, 2019; Klobas et al., 2018; Christensen, 2013; Duncan et al., 2013). YouTube also allows educators to rapidly and easily disseminate and distribute learning materials, content and information (Christensen, 2013). Educators are producing and posting videos of their lectures and How-to visual demonstrations to further support, deepen and enhance the learning experiences of their students (Ahn and Shin, 2016).

A significant educational benefit in using YouTube is that it can support and enable performance and musical creativity and development (Forbes, 2015; Waldron, 2012; Cayari, 2011; Rudolph, 2009). Music educators have therefore integrated the YouTube platform into their practice (Waldron, 2012; Cayari, 2011; Rudolph and Frankel, 2009). Not only does YouTube provide students with the opportunity to express their creations in a multimedia format, but also provides them with access to share them with a potentially large and global audience and to a range of instant and constructive feedback (Cayari, 2011). Students can also use YouTube as a source of inspiration for developing their work (Liberatore et al., 2019; Waldron, 2012). Audiences may tend to be geographically close to the creator rather than international, however (Saurabh and Gautam, 2019).

One investigation has tried to estimate the cognitive value of an educational video from a set of ten properties, finding that length and speaker gender had little predictive power but that more relevant properties (modality, spatial congruity) were more important (Shoufan, 2019).

2.5.4 Gaming

YouTube hosts many videos with footage of amateur gamers playing a computer game (Arthurs et al., 2018). These can showcase the skills of the user and have interest for the challenge as well as giving gameplay insights and para-social interaction through the commentary of the gamer and comments on the video. Live streaming games may add to the excitement (Göring et al., 2019). A prominent YouTuber for both gaming and para-social connections is PewDiePie (Fägersten, 2017). As mentioned above, gaming videos seem to be a male genre (Ferchaud et al., 2018).

The commentaries on gaming videos can also be used for usability testing (May, 2018).

2.5.5 Health

It is used regularly to support health and well-being (Anthony et al., 2013; Lo et al., 2010; Madathil et al., 2015). YouTube has provided the health industry with the opportunity to get important messages out to the public and through videos have provided the public with more accessible and multimedia ways of disseminating important information, for example, about the benefits, issues and risks of immunization (Lo et al., 2010; Steinberg et al., 2010; Keelan et al., 2007). YouTube has also been used to support viewers' knowledge and understanding of current medical and mental health news, awareness and issues (Lewis et al., 2011). There is also misleading health information on YouTube (Goobie et al., 2019), and it can sometimes gain an audience (Yiannakoulias et al., 2019). For example, an investigation of prostate cancer videos disturbingly found that videos with higher quality information were less viewed and received fewer likes per view (Loeb et al., 2019).

Information about YouTube video consumption patterns has also been used, along with other social media data, to gain insights into trends in the spread of diseases (Watad et al., 2019).

2.5.6 News

Important and contemporary ideas, news, issues and global events are now the subject and focus of much commentary and discussion, with YouTube providing a significant mainstream platform for this to take place (Klobas and McGill, 2019; Broderson et al., 2012; Sykora and Panek, 2009; Van Zoonen, Mihelj and Vis, 2011; Crane et al., 2010; Van Zoonen, Vis and Mihelj, 2010). In addition, some YouTube videos discussing current issues have triggered news and media coverage themselves (Van Langendonck, 2009).

An investigation of YouTube news videos that were most tweeted or most popular on four news website channels found that positive news videos dominated this popular set, with surprising and socially relevant content being important to garner views (Al-Rawi, 2019). Political news videos in target audience languages may also be used to disseminate foreign policy content online. A study of Russia Today suggests that this strategy may not always reach the target audience, however (Orttung and Nelson, 2019).

Videos in the News and Politics categories on YouTube host the most interactive discussions in YouTube, in the sense of comments with the highest proportion of replies from other commenters (Thelwall, Sud, and Vis, 2012).

2.5.7 Politics

YouTube has the power to impact on and influence national politics (Lillie, 2008), but this influence may be relatively small (Baumgartner and Morris, 2010). YouTube is used as an important part of the political process (Halpern and Gibbs, 2013; Broderson et al., 2012; Baumgartner and Morris, 2010; Agichtein et al., 2008; Gueorguieva, 2008; Hang and Yun, 2008; Im, 2010; Thorson et al., 2010; Duarte et al., 2007). YouTube has provided political operatives with an additional tool to reach more of the undecided and disinclined groups of voters (Kellner and Kim, 2010). Political operatives also research into which YouTube videos potential voters are watching and are using this to produce targeted political advertising and to disseminate their candidate's message to greater effect (Weber et al., 2013). YouTube videos and the associated discussions and comments are helping candidates to find out some of the thoughts, opinions and concerns of the voting public (Garcia et al., 2012). YouTube videos have also provided political candidates and voters greater access with each other, particularly in users being able to discuss their concerns (Kellner and Kim, 2010; Gulati, 2010). Indirect political messages may also be spread through satire in video, but these do not seem to be effective with those that disagree with them (Bowyer et al., 2017).

As mentioned above videos in the News and Politics categories on YouTube host the most interactive discussions (Thelwall et al., 2012).

2.5.8 Marketing

YouTube provides marketers and advertisers with a valuable tool and a rich source of market consumption, insights and trends (Sokolova and Kefi, 2020; Bellman et al., 2009). YouTube videos can also be used by members of the public who are unhappy with a product or service that they have received and are able to make their opinions known to a much wider audience (Kellner and Kim, 2010; Ayres, 2009). Some users have developed videos to look like consumer generated content but have covertly hidden advertising messages, techniques and product promotion into content (Freeman and Chapman, 2007). Viewers find some of the advertising techniques used by YouTube to be annoying and a high proportion skip these when they can (Dehghani et al., 2016). In 2014, 88% of US companies were using social media websites, including YouTube, for their advertising and marketing purposes (Dehghani et al., 2016). For example, the success of an e-cigarette brand has been attributed to its social media campaign, including YouTube videos (Huang et al., 2019). Nevertheless, companies are reliant on people becoming their distributors and sharing their message, and videos, with other users within their social groups and networks (Barry et al., 2014).

YouTube has been used as an online tool for the dissemination and diffusion of information, innovation and ideas (Klobas et al., 2019; Aytar et al., 2018; Kardas and Brien, 2018; Lillie, 2008). Some diffusion can take the form of trying to reach millions of YouTube users whereas other approaches are more interpersonal with people undertaking more social exchanges where ideas and innovations are developed (Dynel, 2014; Bou-Franch et al., 2012; Lillie, 2008). Ideas shared within YouTube, these might be adopted by relevant companies or distributors (Lorenzo-Dus et al., 2011; Bou-Franch et al., 2012; Lillie, 2008).

YouTube seems to be an effective venue for product advertising (Duffett et al., 2019). This creates the possibility to investigate customer reactions to advertising through their online interactions. For example, a Dove skincare product campaign has been investigated through YouTube video comments (Feng et al., 2019).

Since being bought by Google, YouTube has adjusted its focus to favouring content that is considerably more aligned with their advertisers' demands and commercial interests (Bishop, 2018). This emphasis upon using the platform for commercial purposes demonstrates a significant move away from original open philosophy of why YouTube was created. It has also has the effect of pushing, and rewarding, users to producing video content that is more in line with this commercialisation priority, so that their productions are more visible to audiences and as a result has emphasised biases (Bucher, 2017) within the platform and hidden more diverse and creative content (Bishop, 2018; Cunningham et al., 2016; Jarboe, 2012).

2.6 Uploading and Categorisation of Videos - YouTube

When a video is uploaded to YouTube the uploader must provide a title (up to 100 characters), enter a description (up to 5000 characters) and choose a thumbnail picture (either a still from the video or a picture relating to the video) (YouTube, 2020). Individuals have a range of other decisions to make about the video including its visibility, whether they will allow comments (different levels are available) and ratings (likes and dislikes), and when the video will become available for viewing. Then any age restrictions have to be chosen and the user must also highlight whether or not it is a paid promotion. The uploader can then add their own tags (up to 500 characters) relating to the video, using words that relate to the content and material of the production in order to increase its chances of being found by other YouTube users. The accuracy of tagging is dependent upon spelling and the wide variety of vocabulary and terminology that individuals might use when searching for videos (Greenaway et al., 2009). One of the difficulties with YouTube is that only the uploader is able to decide which tags to assign to a video and this cannot be added to by viewers (Greenaway et al., 2009), who may come up with additional and possibly more accurate tags which could support the visibility of the video. The accuracy of tagging therefore lies with the uploader of the video (Toderici et al., 2010; Greenaway et al., 2009) Greenaway et al (2009) suggest that tagging was an extension of the description and title, and that individuals relied on YouTube categorisation process to establish the overall category of the video. YouTube then requires the uploader to manually assign one of their broad categories (e.g. Comedy, Education, Entertainment, Gaming, Sports…) in which to place their video (Ortega-León et al., 2019). Although, the categories have developed over the years they there have not been substantial changes (YouTube, 2020) The uploader will need to carefully consider their decision as they are only allowed to choose one of the board categories (YouTube, 2020). This could lead to people wrongly categorising their videos to increase the potential audience, by choosing a more popular category. An additional problem is that categories can be understood in different ways depending upon the uploader (Ortega-León et al., 2019; Toderici et al., 2010). In order to increase the views of their content (and as a result improve their income) some uploaders provide misleading information, such as popular phrases. This 'clickbait' has become a problem for social media platforms (Zannettou et al., 2018).

2.7 Research with YouTube Data

YouTube provides the public with access to its data and most videos are accompanied with various metrics that can be used for analysis (Arthurs et al., 2018; Richier et al., 2014). The data which can be extracted from YouTube is a sophisticated and valuable research resource with the potential to demonstrate contemporary social problems and issues and can also provide valuable insights into society's social behaviours online (Anthony et al., 2013; Snelson et al., 2012; Chenail, 2011; Jang, 2011). A substantial proportion of the research into YouTube has had a humanities agenda and focused on investigating genres, topics and different types of information encompassed within the videos (Thorson, et al., 2010; Steinberg et al., 2010). It has extended to discussing and analysing the impact that YouTube can have on society because of the information that is disseminated within the videos that it hosts (Lewis et al., 2011).

Even though YouTube has been the focus for some significant and large-scale quantitative research (Ding et al., 2009; Gill et al., 2007) the majority has still tended to be of a limited scale and more qualitative in nature (Thelwall et al., 2011). YouTube as a social network website, the mechanics involved with searching for videos within the website and the use of video data for social sciences and health have all been the focus of different research projects (Lange, 2007; Keelan et al., 2007; Cunningham and Nichols, 2008; Jang, 2011). Research relating to YouTube has also focused on topics such as time series analysis, view count prediction, collective responses, political issues, audience partisanship and the news (Crane and Sornette, 2008; Szabo and Huberman, 2010; Garica et al., 2012; Weber et al., 2013; Crane et al., 2010).

Other research has specifically focused on elements of the popularity of online content and videos. Lee et al. (2010), have, with the growth of user generated content and social sharing sites such as YouTube, researched the popularity of online content. They suggest that some content becomes popular based on a range of factors and that this differs from user to user and from content to content. One of their key findings is that these multiple and varying factors make it difficult to predict the popularity of online content. Lee et al. (2010) specifically focus on the popularity of thread-based discussions and the content within these, and tries to predict a popularity metric. Nevertheless, although their findings are interesting in terms of general online content and they discuss YouTube as a platform for discussion it is not the focus of the research. In contrast, Figueiredo et al. (2014) focused on YouTube video content and its popularity. Their research investigated users' perceptions and opinions of the content of popular videos presented in pairs. Users were asked to consider video content in terms of enjoyment, their willingness to share and whether they thought it would become popular. The key finding was that when there was a consensus relating to the content of a video it almost always became popular (Figueiredo et al., 2014). This research did not focus on why users choose to watch specific video, determine video popularity based on key associated metrics or key popularity predictors. In addition, the study uses a relatively small number of respondents to establish its findings (Figueiredo et al., 2014).

Other research has analysed the sharing of YouTube videos and the impact that this can have on them becoming virally popular (Broxton et al., 2010). They used a substantial amount of YouTube videos, focused on the socialness of the videos (specifically the importance of sharing videos) and discussed patterns of virality between highly and low social videos. Although this research (Broxton et al., 2010) demonstrates the importance and impact of video dissemination through various online social platforms and ranks these accordingly it does not focus on the factors that encourage users to specifically watch certain videos. Borghol et al. (2011) focused on researching the popularity evolution of videos but focused specifically on those that had been recently uploaded. They suggested that the current popularity of a video is not a reliable predictor of the future of popularity and analysed the time taken for a popular video to peak. This study only focused on modelling and tracking popular videos that had recently been uploaded and did not investigate popularity factors or user decision making behaviours (Borghol et al., 2011). Chowdhury and Makaroff (2013) discuss the popularity growth of videos based on early views, but does not consider a wider range of factors for predicting popularity.

Figueiredo et al. (2011) have focused on the development of YouTube popularity over time. They investigated the popularity growth patterns for individual videos with consideration of users' access (specifically through referrers) and of some of the factors that attract them to the videos selected within the study. The findings focus specifically on three different sets of videos extracted from YouTube and shows that the main referrers that attract users to videos are the YouTube search and internal lists, however, although video metrics are discussed in relation to popularity no in-depth investigation or analysis of this seems to have been done. Although not the key focus of the study it also suggests that mobile devices are having an important impact on video popularity.

The 'peak day' of viral videos has been estimated through a modelling approach that exploited associated metadata (Jiang, et al., 2014). The findings suggested that viral videos have a higher level of socialness and have shorter titles, durations and life spans. They also propose that the popularity of the uploader and the upload time of videos with similar content can have an impact on the viral popularity. This study focuses specifically on viral videos and does not investigate a wider range of videos, or predicting factors or user behaviours to any significant depth.

Other YouTube research has focused on the following:

- The discriminatory elements and biases of YouTube due to advertisers demands (Bishop, 2018)
- Using random prefix sampling to estimate the number of videos within YouTube (Zhou et al., 2011)
- Vivisecting and measuring the YouTube video delivery platform (Adhikari, 2012)
- Analysing the research priorities of YouTube (Snelson et al., 2012)
- Masculinity and video blogging (Morris and Anderson, 2015),
- The use of YouTube and touch screen technologies with the motor impaired (Anthony, 2013)
- The development of YouTube video memes (Shifman, 2012)
- Modelling of view count dynamics within YouTube (Richier et al., 2014)

- YouTube has been investigated as an educational engagement tool (Buzzetto-More, 2014; Mayoral and Tello, 2010; Ashraf, 2009; Duffy, 2008)
- As a platform for supporting learners by providing them with worked examples (Christensen, 2013)
- Using YouTube (and Facebook) to develop student-centred learning and engagement opportunities (Cuvas and Kohle, 2010).
- Using YouTube as a tool to support the meeting of affective learning objectives (Snelson and Elison-Bowers, 2009a)
- User generated videos as effective learning resources (Chenail, 2011)
- How YouTube videos can support clinical education skills (Duncan, et al., 2013)
- Teaching music through the development of online communities (Waldron, 2013)
- Critical pedagogy and video media activism within education (Kellner and Kim, 2010).
- Online deliberation and discussion within YouTube (Halpern and Gibbs, 2013)
- The practice of commenting and forming discussions based on YouTube videos (Thelwall et al., 2012)
- Performing genders within YouTube (Wotanis and McMillan, 2014)
- The consumption, production and distribution of music through YouTube (Cayari, 2011).
- Industries use social networking sites (Kim et al., 2014)
- Motivations for sharing information and providing social support through social media (Oh, 2015)
- As a platform for fostering relationships and a medium for disseminating protest videos (Meek, 2012)
- Measuring and analysing online friendship networks within YouTube (Mislove et al., 2007)
- User interactions and responses within online video social networks (Benevenuto et al., 2009a) and general social networking through YouTube (Lange, 2008)
- Investigations relating to the participatory culture of YouTube (Burgess and Green, 2009)
- The propagation of video content based on social network structures (Yoganarasimhan, 2012)
- Participant frameworks (Dynel, 2014)
- Analysing features of social networking and dynamics within video-sharing services (Halvey and Keane, 2007)
- Growth behaviour of social differences of YouTube videos (Broxton et al., 2010)
- Analysing how videos that become viral are shared (Broxton et al., 2013)
- The evolution of videos that become popular and their growth over time (Figueiredo et al., 2011).
- The ranking of YouTube videos and tag space and integration (Choudhury et al., 2009)
- Video quality descriptors (Crane and Sornette, 2008)
- What viewers are watching in YouTube and what they are sharing (Dimopoulos et al., 2013), discussion relating to user sharing probabilities (Nwana et al., 2013)
- Investigation of the characteristics of short video sharing on the internet (Cheng et al, 2007).
- YouTube video recommendation system (Davidson, 2010)
- User participation, interaction and consumption on YouTube (Khan, 2017; Khan and Solomon, 2013)
- YouTube as a tool and platform for the diffusion and sharing of innovation (Lillie, 2008)
- Analysing the feedback from communities through commenting on YouTube videos (Siersdorfer et al., 2010)
- The analysis and modelling of YouTube traffic with a focus on enhancing network design (Ameigeiras et al., 2012)
- The analysis of workload characteristics based on usage patterns, file properties, video popularity and referencing (Gill et al., 2007)
- YouTube and network traffic (Gill et al., 2007)

- University campus network traffic relating to YouTube video requests (Zink et al., 2008)
- Discussing user generated video files (Mitra et al., 2011; Cheng et al., 2008; Cha et al., 2007)
- The future of the YouTube platform and members of society broadcasting themselves (Jarrett, 2008)
- Utilising user generated content website as video on demand websites (Cha et al., 2007)
- The discussed and analysis of news-based videos on YouTube (Peer and Ksiazek, 2011),
- YouTube as an alternative form of journalism (Poell and Borra, 2012)
- YouTube literature and its distribution (Snelson, 2011)
- The citing of YouTube videos within academic publications (Kousha et al., 2012)
- The issues of YouTube providing access to underage alcohol marketing and promotional content (Barry et al., 2014)
- Advertising tobacco content (Freeman and Chapman, 2007) and the influence that YouTube advertising can have on young users (Dehghani et al., 2016)
- Over popular and social trends (Gaunt, 2015)
- The negative impact of adopting shock tactics within YouTube videos (Henke, 2013)
- Violence in web-based online video entertainment (Weaver et al., 2012)
- Issues of transmitting threats through YouTube (Celis, 2013)
- Extracting information and data relating to elections (Shah, 2010)

Despite this wide range of research, there does not appear to be any research focusing on the samples of data that the YouTube API provides, analysing this and comparing it to data collected from YouTube users (UK) using questionnaires. This contrasts, for example, with the situation for web search engines, which have been systematically investigated to assess their performance over time (Bar-Ilan, 2002; Bar-Ilan and Peritz, 2004; Rousseau, 1999). YouTube views are potentially a new altmetric, in the sense of an indicator of impact of scholarly work (Priem et al., 2011), because they are derived from the social web and may shed light on an aspect of academic impact for some videos. YouTube users may therefore want to know what makes a video popular and How-to estimate the impact of their videos (Haran and Poliakoff, 2012; Sugimoto and Thelwall, 2013). The development, continual growth and changing nature of YouTube makes it almost impossible to capture its, and user, behaviour at a point in time (Gill et al., 2007). In addition, the size of YouTube and the number of users any sample sized used would not be able to be used to represent the whole population and will always limit the generalisability of research findings (Dehghani et al., 2016).

2.7.1 YouTube Application Programming Interface (API)

YouTube provides users and developers with access to video data and statistics though its Applications Programming Interface (API) (Abisheva et al., 2014; Richier et al., 2014; Zhou et al., 2010; Bent, 2010; Gill et al., 2007). This interface lets applications access data and most statistics that are normally available to anyone accessing the YouTube website (Abisheva et al., 2014; Richier et al., 2014; Zhou et al., 2010; Bent, 2010). A high proportion of the research focusing on YouTube and the information and data that it provides uses the API (Figueiredo et al., 2014; Richier et al., 2014; Chowdhury and Makaroff, 2013; Ameigeiras et al., 2012; Figueiredo et al., 2011; Thelwall et al., 2011; Chatzopoulou et al., 2010; Choudhury, Breslin and Passant, 2009; Gill et al., 2007). The API provides users with samples of YouTube data and statistics, but it is crucial to determine what type of sample is provided to ensure interpret the results of any analysis of the data.

2.8 Summary

Posting, watching, commenting, discussing, rating and sharing videos has become a popular form of entertainment. The ability to produce, post and interact with videos has become easier, particularly with mobile phones.

YouTube has become one of the most popular websites. High levels of online video content are uploaded to the website daily and it has become the world's largest source of a wide range of user-generated content. Some of the videos posted have become popular and viral, whereas others have few views. YouTube has made it easier for people to watch, upload and interact with online videos, and has provided users to a global platform to express themselves in various ways. The website has also become an area for social interaction, networking and the formation of networks and communities. YouTube has also become a tool for a variety of purposes including:

- The replacement of traditional TV – Entertainment
- Humour and comedy
- News and current affairs
- Marketing and advertising
- Sharing information and innovation dissemination
- Communication, debate, discussion and social interaction
- Expressing opinion and complaining
- Political campaigns
- Advice
- Education and personal development
- Health and well-being
- Revenge
- Criminal investigations
- Research

A wide variety of online videos are now available on demand. Viewers can choose what they want to watch and when. Online videos offer differences in length, life span and content in relation to more traditional television programming. Social networking and connections may greatly influence video watching patterns. YouTube is most popular with 18 to 34-year olds but it is unclear whether it used more by males or females.

Popularity is a social construct that relates to something that is liked, used, discussed and/or shared significantly within a social group or network. The factors that may influence whether something becomes popular include the following.

- Topic
- Emotional responses
- Prominence and attention within homophilous or hetrophilous social groups
- Social processes, such as key and influential people with social or significant standing and their opinions, information cascades, word-of-mouth (on and off line) and other diffusion processes (diffusion theory)
- Society's expectations, structures, pressures and influences
- Banning or denying access

Some videos have become popular due to viral sharing. Viral videos, unlike popular videos, rarely generate social views across long periods of time. Viral videos generally have a short shelf-life because social appetite wanes quickly. The popularity (in the sense of frequently encountered) of a video is best measured by its view count. Videos that are more popular tend to receive more comments and ratings (both positive and negative). Factors that help videos become popular or viral (i.e., frequently shared in a short period) include the following.

- Social web, email or blog sharing or recommending
- Offline or online discussions

- Entertainment, music, cute, humorous, inspirational, fails, remixes, parodies, and memetic content
- Invoking strong emotional responses
- Innovative, creative or engaging
- High quality production values
- Being banned or flagged as unsuitable
- Producer or uploader, title, title length, video duration, day and time of day uploaded, current affairs and issues, originality, accessibility, and the age and gender of typical viewers

Despite the known or hypothesised factors above, it is difficult to predict video popularity of the varied nature of popular videos as well as ongoing changes in popular culture.

3 Aims and Research Questions

This thesis addresses several broad gaps in the literature, addressed through two types of data. No previous study that has collected data from YouTube and online users and then compared this to determine video category popularity and key factors influencing decisions to watch videos. This research is based around collecting data relating to YouTube from two different, but key sources:

1) Through extracting video data and metrics from the YouTube API; and,
2) Through questionnaires given to a convenience sample of human web users.

From the collected data this research reports an in-depth analysis into the YouTube API (to understand the sample provided), video category popularity, user interaction, decision making and watching behaviours, and the differences between genders, ages and user levels.

3.1 Research Aims

The following are the research aims of this PhD thesis:

- To determine the type and scope of sample provided by the YouTube API – before any data analysis it is important to understand the sample so that conclusions are valid and any biases identified.
- To identify which YouTube categories are the most popular i.e. have the highest number of average views – this will be achieved through a process of analysing two sets of data. One extracted from the YouTube API and the other collected using a questionnaire from a convenience sample of web users.
- To identity and quantify factors or features which have an impact on the popularity of YouTube videos – because of analysing the collected data determinations will be made, based on both samples, in the popularity of YouTube videos.
- To determine to what extent users are interacting with YouTube videos and to highlight, discuss and analyse different levels of engagement across the various categories – this will also be achieved through collecting, processing and analysing data that has been extracted from the YouTube API and from a questionnaire.
- To identify UK YouTube users' watching behaviours, decisions and habits in what videos they are watching and the key factors that influence these – again this will be achieved by processing, analysing and comparing the data collected from both sources.
- To identify the metrics which have a significant relationship with viewing behaviours – from the data collected each of the various YouTube metrics will be analysed and those that have the most impact in users watching videos will be presented.
- To determine the key factors which influence users' decisions to watch YouTube videos.

A set of focused research questions have been formulated to address the above aims.

3.2 Research Questions

This study addresses the following eight research questions.

RQ1: How does the YouTube API select the sample of videos that it returns for a category search?

This question concerns a different issue than the main aims of this research (popularity, user behaviour and influencing factors) and focuses on the nature of the sample provided by YouTube via its API. This is an important question because users will see this sample if they search for videos by category. It is also important as a preparatory question for some of the other research questions because it will provide an understanding of the sample provided so that any biases can be taken into consideration. Many studies have used information from the YouTube API, but none have reported a detailed analysis of the sample provided (Figueiredo et al., 2014; Richier et al., 2014; Chowdhury and Makaroff, 2013; Ameigeiras et al., 2012; Figueiredo et al., 2011; Thelwall et al., 2011; Chatzopoulou et al., 2010; Choudhury, Breslin and Passant, 2009; Gill et al., 2007).

The following research question (RQ2) focuses on the differences between various metrics and popularity across the categories. It is addressed with the data sample extracted from YouTube through the API.

RQ2: What is the age, length, and popularity of videos in each YouTube category and how do these vary between categories?

YouTube provides users with the opportunity to access information and associated metrics relating to the content that is uploaded to its website and can be used to develop a deeper knowledge and understanding of videos. This research question addresses key aspects of the information and metrics to provide a clearer understanding of a sample of YouTube videos and their popularity. It also addresses any key similarities and differences across the main YouTube video categories.

The following three research questions (RQ3, 4 and 5) address factors that influence the popularity of YouTube videos. They are addressed with the data sample extracted from YouTube through the API.

RQ3: Which categories of YouTube video are the most popular?

To investigate the popularity of YouTube videos, it is logical to investigate whether some *types* of video are more popular. Although there are many ways to classify the type of a video, one important characteristic is its YouTube category. This is important because the viewer may find a video through a category search and because the categories have been defined by YouTube, to group videos in a way that they believe is relevant to users (Ortega-León et al., 2019). Thus, any discussion of YouTube video popularity should start with the categories in which they have been placed. The research will consider both the overall average number of views and the average number of views per day to provide a clear and more detailed picture of what viewers are watching (Frasco, 2014; O'Neill, 2011; Cutler, 2009). Nevertheless, since popularity should not be exclusively thought of in terms of views alone (Brodersen et al., 2012; Chatzopoulou, et al., 2010; Benevenuto et al., 2009a), other research questions address different aspects of popularity.

RQ4: Which categories of YouTube video do users comment on most?

Videos that generate many interactions can be influential even if they are not highly viewed because interactions may represent active engagement. Interactions are also important because they may lead to viral or other sharing. This research question thus addresses a second, and complimentary, aspect of popularity. Since more viewed videos often receive a higher proportion of comments and ratings (Wotanis and McMillan, 2014; Pinto et al., 2013; Bent, 2010), there should not be a big difference between the answers to RQ3 and RQ4 but this has not been previously tested.

RQ5: How does the length, like count, dislike count and comment count of a YouTube video relate to its popularity?

Video popularity is likely to be related to factors other than its subject (or category, as in RQ3). The properties that most affect popularity are impossible to quantify on a large scale, such as the design creativity and professionalism, topicality of the issue, popularity of the art genre displayed (e.g., international pop star or local folk group) or usefulness of the information conveyed (e.g., in DIY videos). Nevertheless, with a large volume of data it may be possible to identify general trends in simple properties, which may give useful insights into popularity on YouTube. RQ5 addresses length, like count, dislike count and comment count because these are available from the YouTube API and can therefore be obtained on a large scale. Whilst this is essentially a type of convenience sample of video properties, rather than a theoretically-informed selection, it is a practical method to gain robust properties for a large sample of videos. It is therefore a logical starting point for this type of research. Overall, RQ3, RQ4 and RQ5, will provide insights into some factors that affect the popularity of videos.

This thesis also investigates the popularity of videos from the user perspective, through a survey of UK YouTube users. Although YouTube used to display time series information alongside some videos until November 2018 (see the descriptive text for this video: https://www.youtube.com/watch?v=xQp8dPoM2XE), it does not report the age, gender, or other activities of the viewers of a video. It is possible to get some insights into who views popular videos from the content and usernames of comments left underneath each one but since about 1 in 1000 views translates into a comment, this gives little information about the typical YouTube user. Hence, surveys, interviews, ethnography or focus groups are needed to gain useful insights into the UK YouTube user perspective. Of these, surveys are most suitable for insights into video popularity because they can reach the most people and may yield quantitative information that can be compared with the individual video popularity information.

The following two research questions (RQ6 and 7) concern the popularity of video type and the factors that influence decisions to watch across the gender, age and user levels of viewers. They are addressed with the data sample collected through questionnaires.

RQ6: What are the main gender and age differences in the types of YouTube video that are the most popular?

Many videos seem likely to be popular with restricted demographics on YouTube, such as k-pop videos, gardening advice and ballroom dance instructions. Whilst this is obvious at the level of individual videos, it is not clear which types of videos would most appeal to different demographics in general. This research question addresses RQ3 from the user perspective to give more fine-grained information. As for RQ3, and for the same reasons, YouTube video types are analysed only in the YouTube categories. The user characteristics of gender, age and level of YouTube use (i.e. how often they watch YouTube videos) were selected as three salient characteristics that seem likely to reveal differences. Previous research has found variations within these areas separately, but not from a systematic comparison (Andone et al., 2016; Gaunt, 2015; Perrin, 2015; Smith, 2014; Duggan and Brener, 2013). Whilst other differences are likely to influence video choices, including social class, level of internet access, nationality, ethnicity and culture, these were not investigated for the pragmatic reason that not enough data could be obtained about them from a focused questionnaire sample (as justified below) to give statistically significant conclusions.

RQ7: Which factors influence the decision to watch a YouTube video for different genders and ages?

As discussed for RQ5, this thesis investigates a selection of factors other than the topic of a video that influence the likelihood of it becoming popular. This is also investigated in RQ7 from the user perspective to assess whether demographics affect the influencing factors. This will determine which factors could be manipulated, such as by video producers, marketers, companies and politicians, to increase the likelihood of UK YouTube users watching their videos. This question addresses visible YouTube metrics which demonstrate the preferences, decisions and opinions of others: likes, dislikes, views, and comments. This is relevant because popularity can occur virally by following, copying, or imitating (or going against) the behaviours, opinions, preferences and decisions of others or groups within their social groups or society (Eger, 2015; Oh and Syn, 2015; Boyd, 2014; Milkman and Berger, 2014).

The final research question (RQ8) concerns the combined data collected from YouTube and the questionnaires.

RQ8: What influences the decision to watch a YouTube video?

The final research question ties together the findings of RQ2 to 7 and critically evaluates them to provide conclusions about what has been deduced about YouTube watching preferences.

4 Methods

4.1 Research Design

The research method focused on collecting two different sets of data relating to the watching and accessing of videos within the YouTube website. Gathering two sets of data meant that the findings would represent different samples of video watching behaviour which could be analysed separately and then compared to find common themes and key differences. The samples collected would provide two different perspectives of user interaction: YouTube data and YouTube users. Due to the substantial user base (with YouTube being the 3rd most popular website) of YouTube, it was not possible to analyse all YouTube videos or users and so a sampling approach had to be devised.

1) **Method 1** - Submitting category-based searches within the YouTube website to produce a large sample of video information (RQ1, RQ2, RQ3, RQ4 and RQ5);
2) **Method 2** - Survey a sample of YouTube users through distributing a questionnaire to potential respondents within a convenience sample (RQ6 and RQ7).

The research design incorporated both quantitative and qualitative data collection in a multi-methodology. However, as the two approaches used were not mixed (i.e. were independent of each other and not blended) this was not a mixed methods design. The quantitative data (YouTube metadata) were analysed using quantitative data analysis methods (explored in further detailed, below). The qualitative data (questionnaire data) were analysed using quantitative data analysis methods (also explored further, below). Onwuegbuzie & Combs, (2010) define the application of quantitative data analysis to qualitative data as 'crossover' analysis. Hitchcock and Onwuegbuzie (2020, p63) expand on the definition, saying that crossover analysis comes from "…from using techniques from one tradition (e.g., quantitative) to analyse data associated with the other tradition (qualitative)…and vice versa." Therefore, for the qualitative data at least, the data analysis methodology was a comparison of percentages using 95% confidence intervals to determine statistically significant differences in the findings.

The first sample was collected from the YouTube API using the data extraction program Webometric Analyst. Searches, based on the key YouTube video categories (at the time), and were run over two periods of time, using daily and five-day searches for comparison. A data extraction program was used with the YouTube API because it would have been too time consuming and impractical to search each category and extract enough data manually. This data was initially processed, compared and analysed to gain a greater understanding of the sample provided by the YouTube API in order to determine any patterns within the way in which videos were chosen (RQ1). The two sets of YouTube data were combined and analysed, focusing on metric variations and popularity of videos between categories (RQ2), which types of video were the most popular (RQ3), user interaction with types of video (RQ4) and the relationship between key video metrics and popularity (RQ5).

One of the key metrics used to determine the popularity of videos within this thesis was the views received. As views are one of the most important indictors of popularity by those who produce and upload videos to YouTube (Dynel, 2014; Frasco, 2014; Figueiredo et al., 2011; Chatzopoulou et al., 2010; Gill, et al., 2008; Zink, et al., 2008). Views over the period the video had been uploaded to YouTube were also used to determine its popularity over time. This was a pragmatic approach that relied on data that was freely available on YouTube. In addition, other metrics were also used in conjunction with the views as these can also determine elements of popularity (Brodersen et al., 2012; Chatzopoulou, Sheng and Faloutsos, 2010; Benevenuto et al., 2009a; Benevenuto et al., 2009b; Cheng et al., 2008; Benevenuto et al., 2008). Although it did not include other factors that were known to be relevant to video popularity, such as novelty, it enabled a large enough sample to be extracted from the YouTube API for a systematic analysis. All data collected, such as, views, dislikes, likes, comments, video age, video length, user gender, user age, user level, category and key influencers was also analysed and used to address the aspects of the research questions relating to user behaviours, engagement, interaction and the factors that influence people to watch YouTube videos.

A survey was also conducted to investigate YouTube user behaviour and decision making. Questionnaires were used instead of interviews to provide a larger sample. A mixture of hard copy and electronic questionnaires were used to ensure a higher return rate. Despite research suggesting poor return rates (Cohen et al., 2017; Wood and Smith, 2016; Mukherji and Albon, 2015; Bell, 2014; Punch, 2009; Silverman, 2005; Walliman, 2001), due to the use of a convenience sample, 81% (534/660 respondents) of the questionnaires where completed and returned. The results were analysed for differences in category popularity (RQ6) and influencing factors in terms of decisions to watch specific types of videos (RQ7) across gender, age and user level.

The YouTube metadata and survey results were analysed to give an overall picture of what influences users to watch YouTube videos (RQ8).

4.1.1 YouTube Data Sample

To generate the overall sample of video data the computer program Webometric Analyst was used to extract multiple sets of metrics from YouTube's API. Webometric Analyst is a computer program designed to automatically extract information and data from a variety of web-based sources and websites such as YouTube, Flickr and Twitter (Thelwall, 2015). YouTube category-specific queries were used to generate large samples of videos from different topics so that multiple types of video could be investigated. The alternative, a set of keyword queries, would be difficult to construct systematically enough to give interpretable results because YouTube does not reveal the frequency of queries used by users, although the categories are public and visible to all users.

Before analysing the data extracted from the YouTube API, it was important to understand the sample provided by the searches and any biases in it. As the YouTube API search algorithm is unknown, its results were analysed to identify the nature of the videos that it returned from category-specific searches (RQ1), driven by the following specific questions.

- To what extent does the sample provided by the YouTube API differ over time and between YouTube categories?
- What factors and/or metrics, if any, does the YouTube API employ to determine the sample that it provides?

4.1.2 YouTube Categories

To analyse the sample provided by the API (RQ1), compare video category information (RQ2), video popularity (RQ3), and user behaviour and interaction (RQ4 and 5) data needed to be extracted from the YouTube website. A new YouTube account was set up on a computer with the history and web cache deleted to ensure that any previous watching behaviours, preferences or biases were not able to affect the searches being submitted.

At the time the searches were submitted, YouTube organised its videos into 23 categories and 22 of these were used within this research (YouTube, 2015). The Music category was not included within this research because it did not return any results from the YouTube API queries for unknown reasons. The 22 categories used within this research were: Animation; Automotive; Beauty; Best of; Causes; Comedy; Cooking; DIY; Education; Entertainment; Fashion; From TV; Gaming; Health; How-to; Lifestyle; News; Non-profit; Politics; Science; Sport; Tech.

4.1.3 API - YouTube Data Extraction Searches

To generate an appropriate sample of data, to address research questions RQ1, RQ2, RQ3, RQ4 and RQ5, video category searches were submitted using Webometric Analyst through the YouTube API. Two series of searches were run for data triangulation and comparison. The first set of category searches was submitted at five-day intervals for 95 days (starting on 29/12/15), providing 20 days of data for each category. The second set of category searches ran daily for 30 days (starting on 4/4/16).

For each category an initial search was run to extract a set of video identification numbers (video IDs). This provided approximately 500 video IDs for each of the 22 categories each day (a total of approximately 11000 video IDs). A second query was then submitted to extract the metadata for each video, such as upload date, views, dislikes, likes, comments, length, for each of the 500 video IDs within each category. For example, on the first day, a search for the category Animation was carried out using Webometric Analyst within YouTube. Approximately 500 video IDs were obtained. Using those 500 video IDs, 500 additional searches were carried out (also on the same day) to obtain the metrics for each video. This paired searching was repeated every day for each of the 22 categories for the span of both sets of searches, five-day interval and daily. This yielded 20 sets of data for each category on the five-day searches and 30 sets of data for the daily searches. The raw data was entered into two separate MS Excel spreadsheets, one for five-day searches and one for the daily searches.

4.2 Analysing the YouTube API Sample

Both sets of generated data, five-day and daily, were processed and analysed separately, and then compared to investigate the API sample for RQ1.

4.2.1 Analysis of Five-day Search Data

The first issue investigated was the extent to which the same search returned different videos at different points in time. To assess the percentage of videos that were repeated across the 20 five-day searches the data was processed in two ways. First, across all the categories consecutive searches were compared for the percentage of videos that appeared in both. For example, Search 1 was compared to Search 2, then Search 2 was compared to Search 3 and so on across all the Daily searches. Within this comparison framework, consecutive day searches were identified as a sub-group of searches to determine whether there was any change in results over time, as well as attempting to identify any whole-period changes in category results. The percentage of videos returned by each search that occurred in the next search was coloured coded in a table (higher percentages in darker green graded through yellow and orange with the lowest percentages represented in red). Then a comparison was made, showing the percentage of videos that appeared in Search 1 compared to subsequent searches for all categories. For example, Search 1 was compared to Search 2, then Search 1 was compared to Search 3 and so on across all the Daily searches. The percentage of videos returned in Search 1 that also occurred in each subsequent search was presented in a similar colour-coded table.

To establish the number of times that videos were repeated throughout the search process the percentage frequency of video appearance across the 95-day period of 20 Five-day searches was also tabulated for all categories. The table showed what percentage of videos for each category appeared between 1 and 20 times across all the Five-day searches. For example, the percentage of Animation videos that appeared 13 times across the searches. To determine any characteristics in relation to the videos that the YouTube API chose in every search these were extracted from the data. The videos that appeared within all 20 searches were processed and the mean metrics for these was tabulated. This data was then analysed to see whether there were any patterns or key similarities with these videos.

The videos that only appeared once across the searches, i.e. those without any repeats, where also isolated and the video data metrics processed and tabulated. This was then analysed and compared to the average metrics of those videos that appeared across all searches. The sample of videos extracted from the Five-day searches was also processed and analysed in the 'comments', 'likes', 'dislikes', 'days' posted, 'length' in seconds and 'view count'. The data was then tabulated for each metric and analysed accordingly for any patterns.

4.2.2 Analysis of Daily Search data

The above was repeated for the daily search data.

4.3 Method 1 – YouTube Periodic Category Searches

The periodic YouTube API category search data collected was used to investigate and determine users watching behaviours, interactions and which videos were more popular (RQ2, RQ3, RQ4 and RQ5). The separate data obtained through both the Five-day and Daily searches was amalgamated into a larger dataset and duplicate videos and associated metrics were removed leaving only unique YouTube API data for each of the categories. The remaining data was then divided based on YouTube video category and was then entered into a separate MS Excel spreadsheet for processing, analysis and comparison.

The total number of videos extracted from YouTube was recorded, with the number of individual videos (i.e. with repeated information removed) and the percentage of individual videos for each category. This was used to investigate any patterns, similarities or differences between the samples that were extracted for each of the categories. With the repeated data removed the findings were entered in tables for analysis and to determine any patterns in video popularity and any user behaviours (RQ2, RQ3, RQ4 and RQ5):

- The average number of days the videos within each of the categories had been posted to YouTube (providing the age of the videos)
- The average dislikes and dislikes per day for each of the categories – calculating the dislikes per day took into consideration the age of the video that had been posted
- The average likes and likes per view for each of the categories
- The average views and views per day for each of the categories
- The average comments and comments per view for each of the categories
- The average length of video in minutes – converted from seconds

As the length of a video was known to have an impact on whether a YouTube user will watch it the average views per day for each of the categories was compared to the corresponding length in minutes.

To determine a clearer picture of the data particularly in user interaction and engagement (RQ4) and how YouTube metrics relate to popularity (RQ5) the data for each of the categories was then sorted into bands for each of the associated metrics. The percentage of videos within each banding was calculated and then tabulated for each of the metrics for each category. The bandings for each metric were as follows:

Days – 1 to 30, 31 to 60, 61 to 180, 181 to 360, 361 to 720, 721 to 1800, 1801+

This was used to determine any categories that had less turnover in the videos that were posted to YouTube. For example, if a category had a higher percentage of videos that had only been uploaded to YouTube for a shorter period then this suggested that more videos were regularly added. In contrast, categories with a higher percentage of older videos (in the time they had been uploaded to YouTube) may have had fewer videos posted regularly.

Dislikes – 0, 1 to 10, 11 to 100, 101 to 250, 251 to 500, 501 to 1000, 1001 to 2000, 2001+

Likes - 0, 1 to 10, 11 to 100, 101 to 250, 251 to 500, 501 to 1000, 1001 to 2000, 2001+

Comments - 0, 1 to 10, 11 to 100, 101 to 250, 251 to 500, 501 to 1000, 1001 to 2000, 2001+

The percentage of dislikes, likes and comments across the bands was analysed to determine YouTube user's interaction with, and feelings towards, the videos within each of the categories. Higher percentages within the upper bands suggested that people had a greater opinion or emotional connection, which may have been positive or negative, to the videos within a category. The dislike and like percentages could also reflect opinions about the accuracy or usefulness of the video(s) in the cases where advice, support or instructions were provided.

View count – 0, 1 to 10, 11 to 100, 101 to 1000, 1001 to 10000, 10001 to 100000, 100001 to 1000000, 1000001+

Analysing the banding of views helped to show the percentage of videos within each category and show those that had a substantially higher or lower number of views. Those categories with a higher percentage of videos within the higher viewing bands were used as another indicator of video category popularity. For example, if a category had 50% of its videos within the 1000001+ banding it suggests that a substantial number of people were watching that type of video.

Length – 0, 1 to 30, 31 to 60, 61 to 300, 301 to 600, 601 to 1800, 1801 to 3600, 3601+

Investigating the different bandings of video length showed which were the most common and would have helped to establish if there was a pattern or patterns in longer or shorter videos. The length of videos may have been determined by the content relating to the different videos. This information was used in conjunction with the view count of videos to determine if there is a type of video which was watched most often and would show any preferences that people had in this metric.

4.3.1 Factors associating with popularity for the periodic searches

The amalgamated data (with repeats removed), from the Five-day and Daily searches, was then used to seek correlations between the metadata (likes, dislikes, comments and length) and the view counts (RQ5). Correlations were calculated separately for each of the YouTube video categories in case the results differed between categories.

Correlations between view counts and the other metrics (except length) would be misleading because older videos tend to have more views, likes, dislikes and comments. Thus, even if popularity and likes did not influence each other, the two would have a positive correlation. This factor could be eliminated in theory by dividing each metric (except length) by age to give per day metrics. This would also be misleading, however, because videos in some categories (e.g., music, news) can expect to get most of their views over the first few weeks on the site. To partially get round both problems, the videos in each category were split into bands of at least 50 videos by age, so that videos were only correlated in sets of similar ages. This reduces but does not eliminate age as a spurious factor because some bands still contain videos with a variety of ages. Spearman correlations were calculated between view counts and each metric within each bands, and then the correlations were averaged across the bands.

The Spearman correlation compares the ranks of two variables and determines the extent to which they are in the same rank order (Coolican, 2014). Based on the sample sizes within each banding, predefined critical values were used to determine statistical significance. For a 5% significance +/- 0.279 was used and for a 1% significance +/- 0.363 was used based on a sample size of 50 (Ramsay, 1989). This is an approximate value because each correlation is the average across multiple bands, which increases its statistical power, but individual bands include videos with a (narrow) range of ages, contributing a spurious positive tendency to the correlations.

A second Spearman correlation was used to compare the rate of attracting likes, dislikes and comments (i.e., per view) against the number of views to see if more popular videos attracted a higher rate of interaction (if so, high interaction could then be a partial cause of their popularity). Although it would be possible to correlate, for example, likes per view against views to see if the two relate, this could give misleading answers for two reasons.

First, the high proportion of videos with few views combined with the low proportion of likes per video overall would introduce too many zero likes per view ratios, skewing the overall correlation. This is an artefact of discrete data. For example, if a set of videos expect 1 like per 20 views then a set of 20 videos with one view each could be expected to have 19 likes per view ratios of 0 and one of 1, giving strange results. To overcome this limitation, a threshold minimum of 500 views was set to filter out videos with too few views to generate reliable likes per view ratios.

Second, some videos had unusually many or few likes per video. Few likes might occur because the owner had banned likes for a period. Many likes might occur due to spam, video owner encouragement, or false promises of random rewards for likes. To circumvent this, the middle 50% of videos were analysed, after ordering them by like, dislike or comment ratio.

Spearman correlations were used again for the same reason, with 95% confidence intervals (see below) given to show the variability of these point estimates of the underlying population correlation (Cumming and Calin-Jageman, 2016; Cumming and Finch, 2005).

4.4 Method 2 - Collecting User Data – Questionnaires

This section discuss the justification for using questionnaires to collect data from respondents and the process for developing the questionnaire for distribution. This data was used to address RQ6 and RQ7, and was compared to the data extracted from YouTube to address RQ8.

4.4.1 Rationale for the use of Questionnaires

Three strategies were considered in how the data could be collected from respondents (to address RQ6 and RQ7) and these were:

- Individual interviews
- Group interviews
- Questionnaires

Individual interviews could have provided a greater level of control over the questions, there would have been the opportunity to elaborate on questions and points, and any misunderstandings or difficulties with the questions could have been dealt with throughout the process (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014; Creswell, 2014; Kumar, 2014; Punch and Oancea, 2014). It would have also enabled the researcher to ensure that there was a greater level of accuracy in the responses provided, they could have picked up on body language and non-verbal ques, and could have built up more of a professional rapport with the interviewee (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014). Other advantages could have been that the researcher could have chosen respondents who are more suitable for the study and there would have been greater opportunity for discussion and the development of ideas (Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Kumar, 2014; Punch and Oancea, 2014; Clark-Carter, 2010). Nevertheless, interviews would be time consuming to conduct and record systematically (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016). Other key considerations were that interviews would have provided a much smaller and limited sample size, it would have been difficult to find representative respondents and there was also the possibility of process being influenced by interviewer bias (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014).

Adopting a group interview approach would have had the additional benefits of providing respondents with a more relaxed environment to open up in and would have given them the opportunity to share ideas and be stimulated by others' ideas and thoughts (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014). This type of interview situation could have been dominated by one or two interviewees and reducing others contributions, and could have had the potential for discussions to move away from the core topic or topics (Wood and Smith, 2016; Wellington, 2015; Creswell, 2014; Punch and Oancea, 2014; Clark-Carter, 2010). In addition, people could have felt under pressure to conform to the norms of the group rather than expressing their own personal opinions and the data would have been difficult to analyse and quantify due to the group nature of the information provided (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014).

Overall, even though individual and group interviews could provide accurate information, the opportunity to expand on answers and responses further, and choose YouTube users that the process would be too time consuming and would not have provided a large enough sample size, neither of these strategies were used within this research (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014; Creswell, 2014; Kumar, 2014; Punch and Oancea, 2014).

Questionnaires were therefore used to collect data from a sample of respondents relating to their use of YouTube, how they access videos and their watching habits, decisions and preferences. Using a questionnaire provided the opportunity to gather a wider range of responses from a much larger group of respondents in a standardised and objective manner (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Mukherji and Albon, 2015; Bell, 2014). The questionnaire made the process of collecting data relatively quick and in a short period of time, particularly using an online version, however, the development of the questionnaire did take some time to develop and refine (see below) (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014). Although research had established that the return rate on questionnaires could potentially be low, due to using a convenience sample this was not an issue with a return rate of 81% (Cohen et al., 2017; Wood and Smith, 2016; Mukherji and Albon, 2015; Bell, 2014). The use of a questionnaire provided a means to collect mostly quantitative data which was relatively easy to tabulate and process (Cohen et al., 2017; Wood and Smith, 2016). It was much easier to maintain respondents' anonymity using questionnaires, it provided them with the opportunity to complete it at their own leisure and to respond more truthfully (Wellington, 2015; Bell, 2014; Kumar, 2014).

As the questions within the questionnaire were standardised and it was not possible to explain any difficulties, misunderstandings or different interpretations from respondents it was important to pilot the questionnaire with various focus groups (see below) (Arthur et al., 2017; Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014). The use of open-ended questions was kept to a minimum to reduce the generation of large amounts of data that would need a substantial amount of time to process also respondents were provided were provided with limited space to keep their responses concise and focused (Wood and Smith, 2016; Mukherji and Albon, 2015; Wellington, 2015; Bell, 2014; Creswell, 2014; Kumar, 2014; Punch and Oancea, 2014). There was an awareness that respondents might answer more superficially if the questionnaire took too long to complete therefore the number of questions were kept to a minimum (Bell, 2014; Punch and Oancea, 2014). Although it was difficult to determine how truthful and how much thought respondents were putting into the questions, this is unlikely to be a problem due to the non-sensitive nature of the subject matter (Arthur et al., 2017; Cohen et al., 2017; Mukherji and Albon, 2015; Bell, 2014).

When handing out the questionnaires there could have been a danger of respondents telling the researcher what they think they wanted to know (Cohen et al., 2017; Mukherji and Albon, 2015), however, this was addressed through reinforcement with respondents, when writing questions and removing any bias within the questions. Although it could have been difficult using a questionnaire to find out the respondents' thoughts in depth (Mukherji and Albon, 2015; Simmons, 2008) this was not relevant to this study so was not an issue. One of the most important issues for social science questionnaires is that low response rates mean that the respondents may be unrepresentative of the sample. In particular, they may have particularly strong feelings on the topic or may have more spare time than others. To address the non-response issue (Cohen et al., 2017; Wood and Smith, 2016; Wellington, 2015; Bell, 2014; Creswell, 2014; Bryman, 2004), therefore a sampling strategy was chosen to ensure a high response rate, as described below, although at the expense of the representativeness of the sample.

4.4.2 Development of the Questionnaire

In general terms there were factors that needed careful consideration when developing the questionnaire for this research and to be able to address the relevant research questions (RQ6, RQ7 and RQ8). It was important to ensure that the questionnaire was focused on the information required (RQ6, RQ7 and RQ8), was concise, that the layout was clearly structured and overall it was well presented with a professional finish (Cohen et al., 2017; Wood and Smith, 2016; Robertson et al., 1990). Although different types of questions were considered in generating an appropriate sample relating to YouTube use, the use of mainly closed questions would be more effective at eliciting direct responses from respondents and this supports the analysis and comparison of the data (Mukherji and Albon, 2015; Blaxter et al., 2010). When writing the questions, it was ensured that they were unbiased and produced in a precise way in order not to confuse the respondents (Mukherji and Albon, 2015; Blaxter et al., 2010; Simmons, 2008; Silverman, 2005). There was an emphasis placed on keeping the questions as short, clear, focused and direct as possible (Arthur et al., 2017; Denscombe, 2010; Edwards et al., 2002). In addition, the questions were written and presented in a manner that would be accessible to respondents and provided them with the ability to respond as easily as possible (Arthur et al., 2017; Wellington, 2015; Bell, 2014). The questions were designed to effectively record respondents' behaviours, decisions, preferences and influences relating to YouTube video selection and use (Bryman, 2004; Field and Hole, 2003), so that they matched the to the research questions and overall aims of the study (Mukherji and Albon, 2015; Kumar, 2014). It was important to ensure that the questions were ordered logically (Bell, 2014; Simmons, 2008) and that the questionnaire was designed to ensure that accurate replication of the research could be carried out in the future (Bryman, 2004). As part of the development process piloting opportunities were taken with a variety of focus groups to make further developments, modifications and refinements to the questionnaire (Mukherji and Albon, 2015; Anderson, 1998). Once developed the questionnaire was produced and distributed in both paper (handed out) (Kumar, 2014; Walliman, 2001) and electronic (Mukherji and Albon, 2015) formats to increase response rates (Cohen et al., 2017; Wellington, 2015; Bell, 2014).

To pilot, test and develop the questionnaire with various focus groups an initial version needed to be produced (Cohen et al., 2017; Wellington, 2015; Bell, 2014). When constructing the initial questionnaire, the first question focused on asking respondents to show which age range they fell into, but not requiring them to share their age or date of birth. The information was important for the study as research has established the significance of age in online and YouTube use (Gaunt, 2015; Perrin, 2015; Smith, 2014; Duggan and Brener, 2013; Griffiths et al., 2004). The ages ranges used within the questionnaire were those used by OxIS (Blank et al., 2019; William et al., 2013) in their surveys of technology and online use. As with age, gender had also been confirmed as an important aspect of online and YouTube use and collecting information relating this was deemed to be necessary for this research (Oh and Syn, 2015; Duggan and Brener, 2013; Chappell, 2012; Rainie et al., 2012; Mayoral et al., 2010; Madden, 2009; Molyneaux et al., 2008). As the OxIS surveys (Blank et al., 2019; William et al., 2013) also use and discuss education levels, this could be useful information to collect and therefore was included within the questionnaire. It was also important to know if respondents were a YouTube user and how often they watch videos on the website (RQ6 and RQ7). In determining respondents searching behaviours they were asked question relating to how they found and accessed YouTube videos in general and through the website itself. It was also necessary to find out what factors influence respondents' decisions to watch videos and which of these were the most important and why (RQ7). Finally, respondents were provided with an open-ended question to give them the opportunity to make any further comments or statements about their use of YouTube. As a result

Questionnaire Version 1 (see Appendix 3) was produced and then piloted with the first focus group of 7 previous work colleagues who provided oral feedback relating to their experience of answering the questions presented to them.

The first pilot group felt that the questionnaire followed a considered structure, the language used was straightforward, it was not too long and was quick to complete. They needed greater clarity in what they were expected to fill in and where, and that some simple instructions and shaded boxes would support this. Some of the questions needed to be reworded as it was not clear what was being asked for. They thought that Question 10 should just be an open-ended question rather than repeating all the options from Question 9. Finally, two members of the piloting group felt that Pinterest needed to be removed as videos are rarely accessed through this website and that the overall presentation could be refined as it did not look finished.

Because of evaluating Questionnaire Version 1 and from the comments and feedback from the focus group developments and improvements were made to the initial questionnaire to produce a revised version. Questionnaire Version 2 (see Appendix 4) was produced by adding shading to the response boxes to make it clearer for respondents in where they needed to put their answers and responses. Further instructions were added to the end of questions to provide greater clarity in what was being asked, how they needed to respond and where. The lay out and structure of the questions was developed so that the whole questionnaire could fit onto one side of A4 paper. Pinterest was removed as an option from Question 7 (How do you usually access YouTube videos?) and an additional metric was added to Question 9 which has been missed off the first version of the questionnaire. Finally Question 10 was changed so that respondents could just simply choose which they thought was the most important factor in influencing their decision to watch a video rather than selecting from the repeated list of options from Question 9.

Questionnaire Version 2 (see Appendix 4) was then sent to my supervisor for suggestions, aspects to develop, and elements to change, remove or add. It was established that the ethical blurb would need to be added to questionnaire and this produced on an additional page (see Appendix 2) so that respondents could retain this information. In the question relating to gender it was suggested that Other and Prefer not to say needed to be added as options within this section. There was a recommendation that the question 'Do you watch YouTube videos?' (Question 5) be removed as this would be answered more effectively by Question 6 with the introduction of the options Less than once per year and Never and therefore the numbering of questions needed to be updated.

It was also recommended that the categories within Question 6 (On average how often do you watch YouTube videos?) needed to be more precise and less vague. The options of Daily, Weekly, Monthly and Annually were changed to the following (respectively):

- On most days I have watched at least one YouTube video;
- During most weeks I have watched at least one YouTube video;
- During most months I have watched at least one YouTube video;
- During the past year I have watched at least one YouTube video;
- I never watch YouTube videos.

In Question 7 (How do you usually access YouTube videos?) it was suggested that the phrase 'Select the ones you use regularly' be removed and replaced with 'In the last year' as regularly is not specific enough. As with Question 6 there was a recommendation that the options being provided to respondents needed to be more precise and that this would give the question greater clarity. Therefore YouTube, Email, Blogs, Facebook and Twitter were changed to the following (respectively):

- Through accessing the YouTube website;
- From a hyperlink sent to you in an email;
- From a hyperlink in a Blog;

- From a hyperlink or post on Facebook;
- From a hyperlink or Tweet on Twitter.

In addition, from advice a further option of 'From verbal recommendation' was added as an option. It was further suggested that the options being provided within Question 8 (How do you find videos on the YouTube website?) also needed to be more specific and contain more information to make it clearer what was being asked of respondents. Therefore Homepage, Most viewed page, Most popular page, YouTube recommendations and Search were changed to the following (respectively):

- Videos posted on the homepage;
- Videos posted on the most viewed page;
- Videos posted on the most popular page;
- Videos posted on the recommendations bar;
- Through the YouTube search facility.

In addition, it was also suggested that the option of 'Videos posted on your subscriptions page' needed to be added to Question 8.

From the supervisor feedback it was also established that Question 9 (Which of the following factors are important in deciding whether you watch a YouTube video?) was too vague given that some respondents might have watched videos for lots of different reasons. After further discussion the question was changed to 'How important do you feel the following are when deciding whether to view a YouTube video?' and the further option of Thumbnail picture was also added. In addition, a grading system was introduced for each of options so that respondents could rate their importance in influencing their decision to watch a video. The grading system was made up of the following choices: Irrelevant, Not very important, Slightly important and Very important.

It was decided that the open-ended question, Question 10 (Which of the factors you have chosen in question 9 is the most important? And why?), needed to be removed. This was replaced with a question which asked respondents to consider the last 10 videos they had watched and to rate each of the factors in how they might have influenced their decision to watch. Another grading system was introduced using the following options: Never, Rarely, Sometimes, Mostly and Always.

Through discussions, feedback and evaluating previous versions of the questionnaire an additional question, which should have been in the first version of the questionnaire, was added. This new question provided respondents with the opportunity to show what types of YouTube videos they like to watch using the general YouTube categories used within the metadata extraction from the YouTube API and supported the addressing of RQ6. As a result, Questionnaire Version 3 (see Appendix 5) was produced based on all these developments with a restructuring of the question numbers being required. Questionnaire Version 3 was then piloted with 10 colleagues for further feedback, suggestions and developments.

From the comments provided by this second focus group the title of the questionnaire was changed from 'How and why we watch YouTube videos' to 'How and why you watch YouTube Videos' to make it more specific to the respondents. The introduction phrase was shortened and made more focused, Standard qualification was changed for School qualifications and other aspects of wording were tightened up to ensure clarity of what was been asked of the respondents. It was also suggested to make some of the key words and phrases bold and underlined to emphasise them with respondents. Because of feedback and through discussion it was decided that Questions 8 and 9 were essentially asking respondents for the same information with just the wording being slightly different. These questions were replaced with a question asking them to consider the most recent videos and what out of the list of options had influenced them in their decision to watch. They were provided with a grading scale of Never, Rarely, Sometimes, Mostly and Always. An additional question was introduced which required respondents to select what factors might influence them to not watch a video.

All changes and developments were made and a new version, Questionnaire Version 4 (see Appendix 6), was produced and piloted with a new focus group of 8 colleagues. From their comments and feedback, the questionnaire was clear, well structured, easy and straight forward to complete, not too long and quick to complete. A colleague suggested that Question 9 (Which of the following would you consider when deciding NOT to watch a video on YouTube?) was unnecessary and rather confusing, and that the information was essentially covered in Question 8. Again, all these issues were addressed and as a result Questionnaire Version 5 (see Appendix 7) was developed.

After reviewing Questionnaire Version 5, the order of the questions was changed. The question relating to the type of video respondents watched was moved before the one asking what factors influence their decision to watch.

Questionnaire Version 6 (see Appendix 8) was then piloted with a group of 20 computer science undergraduate students who were asked for suggestions and criticisms about the questionnaire. They liked the questionnaire, were happy with the wording, structure and thought that all aspects were clear. Nevertheless, they made two suggestions relating to Question 6 (How have you accessed YouTube videos in the past year?) and these were to add 'Through the YouTube App (phone or tablet)' and 'From a Google search'. From these recommendations the final version of the questionnaire was produced (see Appendix 9).

Ethical approval was then sought and given for the distribution of the final questionnaire because of following and successfully completing the University of Wolverhampton (2018) ethical approval process (see Appendix 1 and Ethical Considerations Chapter).

The final questionnaire was produced in two formats, paper (hard copy) and electronic, through Google Forms, to ensure the highest return rate could be achieved. This was then distributed to people over 18 years of age and were my current students (handed out), my past students (electronically thorough social media) and through my educational and professional networks from working within two primary schools, five universities, further study and from working with a wide range of midland-based schools through teacher training courses (electronically). No children or vulnerable people were used within the research. All respondents were provided with:

- An information sheet explaining the research, its purpose and uses of the research data - enough information to be able to make an informed decision whether to participate within the research – this was written in simple, non-technical terms (avoiding jargon and abbreviations) and could be easily understood by a lay person (see Appendix 2).
- Reassurances that their data would remain protected, secure and anonymous (no names or personal identifying information would be collected).
- A consent form to show that they were willing to be involved in the research and their right to withdraw at any time - questionnaires were numbered for this purpose (see Appendix 2),
- Information explaining the date by which the responses needed to be completed by and where they needed to return them to in the case of the paper copies.
- Reassurance that the researcher did not feel that because of being involved within the research that they would come to any harm and that there was no possibility of physical or psychological distress.

The information collected from the questionnaires was entered into MS Excel spreadsheets to facilitate organisation, processing, sorting and analysis of the data (see below – Questionnaire Analysis).

4.4.3 Questionnaire Sampling

Sampling was a major consideration for this research as it is not possible to study every individual within a population, in this case all YouTube users (Mukherji and Albon, 2015; Wellington, 2015; Bell, 2014; Silverman, 2005). The sample needed to directly reflect the area being researched, be accessible, contain an adequate number of respondents to answer the research questions and provide appropriate data readily and quickly (Cohen et al., 2017; Wood and Smith, 2016; Merriam, 2009; Sliverman, 2005). The precision of the statistics drawn from the sample depends on the size of the sample and as a result reducing the sampling error (Cohen et al., 2017; Kumar, 2014; Bryman, 2004). OxIS (Blank et al., 2019) suggests that within the UK internet use ranges between 95% and 100% across the age ranges 18 to 54 (25 to 34 being the highest with 100% and 35 to 44 being slighter lower with 95%), with 80% being aged 55+. Moreover, 95% of students and 95% of those educated to Higher Education level are internet users (Blank et al., 2019).

The sample for this research was therefore made up of my current students (95% internet users), and my past higher education students (95% internet users), and other internet users that I have access to through various education and professional networks. This was a convenience sample that was a biased subset of YouTube users for many reasons. Firstly, they were all from the UK. Second, they had (nearly) all accessed higher education. Third, most had taken education-related higher teaching degrees. Fourth, they had mostly lived in the UK Midlands area. Nevertheless, based on the findings of the OxIS surveys (Blank et al., 2019; William et al., 2013) the sample were likely to be almost exclusively internet users. The primary advantage of using this convenience sample was that it generated a much higher response rate than usually possible with questionnaires (Arthur et al., 2017; Cohen et al., 2017; Bell, 2014; Creswell, 2014; Blaxter et al., 2010; Bryman, 2004), especially those that were internet-related. Thus, the sample bias was partly compensated for by a greatly reduced non-response bias.

The sample skewing towards people working within UK education, specifically teacher education, also needed to be taken into consideration when analysing the findings (Arthur et al., 2017; Cohen et al., 2017). This was a known bias, whereas all alternatives were likely to generate a much lower response rate and would have had a significant unknown non-response bias (Arthur et al., 2017; Cohen et al., 2017). Thus, the survey generated reliable information about a demographic of YouTube user but caution was exercised when generalising the findings to all YouTube users.

4.4.4 Questionnaire Analysis

The data from the questionnaires was entered into an Excel spreadsheet and then from this further spreadsheets were produced related to specific elements of the data collected (e.g. by gender, age, user level, questions). All raw data was converted into percentages, based upon the number of respondents within the sample obtained, and then compared across genders, ages and users for each of the questions. For each set of percentages relating to a specific aspect of the data confidence intervals of 95% were calculated (see below) in order to determine statistically significant differences in the data. From a frequentist statistical perspective a 95% confidence interval for a percentage reflects the belief that if the data was collected repeatedly, then the true percentage would lie within the 95% confidence interval 95% of the time.

For each of the questions answers the percentage of female responses were compared to those from males within the sample, taking into consideration the level of variation, as represented by the 95% confidence intervals. Then the percentages (and corresponding 95% confidence intervals) for each of the different age groups for females were compared to determine any key similarities and statistically significant differences in the findings. This process was repeated for male ages, female user levels (see Appendix 12) and male user levels (see Appendix 12).

Table 4.1 was used to organise which survey questions addressed the research questions RQ6 and RQ7 to further support the analysis of the questionnaire data.

Table 4.1. The survey questions which address research questions RQ6 and RQ7

Research question	Survey questions
RQ6: What are the main gender and age differences in the types of YouTube video that are the most popular?	1) What is your age? 2) How would you define your gender? 5) On average how often have you watched YouTube videos in the past year? 8) Which type(s) of video have you watched in YouTube during the past year?
RQ7: Which factors influence the decision to watch a YouTube video for different genders and ages?	1) What is your age? 2) How would you define your gender? 6) How have you accessed YouTube videos in the past year? 7) Which of the following methods have you used, if any, to find videos through the YouTube website in the past year? 9) Considering only the most recent videos that you have watched on YouTube, how often have the following influenced your decision to watch? 10) Are there any further comments you would like to make about how you find YouTube videos?
The following survey questions were used to provide further data relating to participant demographics:	
3) Which of the following formal qualifications have you completed? 4) If applicable, what was the major subject of your undergraduate degree?	

4.4.5 Confidence Intervals

Due to adopting a sampling approach there would be variability in the data collected from the questionnaires and it would not perfectly reflect the total population (UK users). Therefore confidence intervals were calculated to show the estimated variations and range of plausible values within the findings, assess the precision and accuracy of the data, and reflect a closer estimate of the true population (Cumming and Calin-Jageman, 2016; Cumming and Finch, 2005). Within this research a confidence interval of 95% was calculated and presented within each of the graphs (Cumming and Calin-Jageman, 2016; Cumming and Finch, 2005). As above, a 95% confidence interval for a percentage reflects the belief that if the data was collected repeatedly, then the true percentage would lie within the 95% confidence interval 95% of the time.

The larger the sample size (in this research females) then the narrower the confidence intervals (i.e. less variation) will be and thus give a smaller margin of error in the results and a narrower confidence interval. A smaller sample size (in this research males) will produce wider confidence intervals (i.e. greater variation) and a larger margin of error.

Through comparing the findings (percentages) 95% confidence intervals were used to determine statistically significant differences in the data collected from the questionnaires. If the 95% confidence intervals overlap, when comparing the percentages, then it was concluded that there was no statistically significant difference between those specific findings (Cumming and Calin-Jageman, 2016; Cumming and Finch, 2005). However, if there was no overlap, of the 95% confidence intervals (i.e. no variation in the percentages), then the difference was determined to be statistically significant (Cumming and Calin-Jageman, 2016; Cumming and Finch, 2005). When there is a small overlap between confidence intervals, the difference may be statistically significant but this possibility was ignored in the results for simplicity of analysis and because the reporting of multiple simultaneous tests increases the chance of at least one false positive occurring.

4.5 Comparison of YouTube and Questionnaire Findings

To address RQ8 the key findings from the YouTube data and questionnaire data were considered together. Where possible and appropriate the sample of data extracted from YouTube was compared to the findings of the questionnaire. Focusing specifically at the factors associated with video popularity and those that influence YouTube user's decisions to watch a video. It was important to consider the two scopes of popularity covered by the two samples of data collected. The YouTube metadata focused upon the average popularity of videos (determined by the average views of individual videos) and the questionnaire data relates to the popularity of categories of video. However, there is a key overlap within this data in that a category would become more popular if there are popular individual videos of that type. Where results could not be compared they were discussed in relation to previous research.

In terms of video category popularity, a scatterplot was produced where respondents watching preferences (in terms of category) were compared against the average number of views per video from the YouTube metadata to see if there were any patterns, trends or discrepancies. In addition, the percentage of respondents watching videos from a category against average daily views per video from the YouTube averages was also plotted, again to see if there were any patterns or discrepancies.

Any preferences relating to video length and age from the questionnaires was compared with the findings of the average video length from the YouTube metadata. The findings relating to the impact of view counts could not be compared as it would be impossible to determine the impact of views counts on view counts from the YouTube metadata as this would be tautological. The correlations focusing upon relationships between various metrics and total view count, and other YouTube metadata relating to possible influencing factors (e.g. likes, dislikes and comments) were also compared to the findings and influences presented within the questionnaires.

4.6 Ethical Considerations

Understanding the potential ethical impact or implications were of paramount importance in this and any other (Cohen et al., 2017; Arthur et al., 2017; Wellington, 2015; Bell, 2014; Kumar, 2014). Within this research the method used for generating a data sample was separated into two parts, Method 1 (YouTube Data – RQ1, RQ2, RQ3, RQ4 and RQ5) and Method 2 (Questionnaire – RQ6 and RQ7), and as a result the ethical considerations for each part were considered individually due to the different approaches that were adopted. As RQ8 will be addressed by using, comparing, analysing and discussing the data and findings from Methods 1 and 2 the ethical considerations will have already been adopted.

4.6.1 Method 1 - Extracting Data from YouTube

For Method 1 the accepted areas of ethical risk were matched with the requirements of the University of Wolverhampton (University of Wolverhampton, 2018). All data collected and used within this part of the research was obtained from the public domain in the web. Consent for the information to be used had already been granted by the data being available to the public. No personal data or information relating to someone, group or organisation was collected or used within this part of the research. There were no animal or human respondents being used within this part of the research and therefore there was no contact between them and the researcher. Similarly, there was no need for the use of deception regarding the aims, focus or research questions of the study in relation to social desirability effects. No resources or materials from webpages or websites (e.g. photographs, videos, documents, software, etc.) was used within this part of the research and therefore permission was not needed from any respondents, groups, organisations or web masters.

Based on the above, Method 1 within this research project fell into University of Wolverhampton 'Category 0' as this part of the research was non-hazardous, did not employ respondents, used only existing material that is publicly and legally available in the UK and overseas and did not meet the criteria for either University of Wolverhampton Category A or B (University of Wolverhampton, 2018). This part of the research was therefore deemed to have had minimal ethical impact in gaining consent or using deception and did not therefore require formal ethical approval (University of Wolverhampton, 2018).

4.6.2 Method 2 - Questionnaires

For Method 2 the accepted areas of ethical risk were again matched with the requirements of the University of Wolverhampton (University of Wolverhampton, 2018). The questionnaire was only distributed to respondents over the age of 18, and in addition no children or vulnerable people were used within the research. No interviews, observations, tests (or testing), tasks or activities involving animals, people, pairs, groups or those at high risk were undertaken. All questionnaires were anonymous and any respondent who provided information or data for the research were provided with information about the purpose and uses of the research data. The respondents were also provided with enough information, in line with the University of Wolverhampton requirements, to be able to make an informed decision whether to participate within the research (University of Wolverhampton, 2018). The questions asked within the questionnaire were appropriate and were produced to add to what is already known about popularity, interaction and user behaviour within YouTube. The design of the questionnaire was appropriate in the research questions being asked and all potential bias was considered and addressed. No alterations were made to the questions or design of the questionnaire after approval was provided by the University of Wolverhampton. There were no foreseen risks in people participating within the study and respondents were informed about their right to withdraw from the study. No respondents were discussed within the research.

Due to the anonymity, the study questionnaires were individually numbered for withdrawal purposes and respondents were informed that they needed to make a note of their questionnaire number for this reason. It was explained to respondents that the researcher did not feel that because of being involved within the study that they would come to any harm and that there was no possibility of physical or psychological distress. All respondents were provided with a user-friendly information and consent form which contained all the relevant information about the study and what their involvement would entail. The information sheet was written in simple, non-technical terms (avoiding jargon and abbreviations) and was easy to understand by a lay person (see Appendix 2). There was no need for the use of deception regarding the aims or focus of the research in relation to social desirability effects. All data collected in relation to the research project was always secured, within a locked filing cabinet, on a password protected computer or password encrypted data stick, to ensure confidentiality. Any resource implications were carefully considered and as a result there were no financial implications from the research. There were also no potential or actual conflicts of interest.

Based on the above, Method 2 within this part of the research project fell into the University of Wolverhampton 'Category A' as this part of the project involved the participation of people, rather than secondary data source, but was not deemed hazardous to the physical or psychological welfare of the participant or the investigator. Ethical approval was sought and granted by the University of Wolverhampton (see Appendix 1).

5 YouTube API Category Search Result Changes Over Time

This chapter addresses **RQ1**, investigating the nature of the videos returned by the YouTube API category search from the perspective of identifying potential biasing factors compared to a random sample. Since YouTube's survival depends in part on delivering videos that interest its users, the default perspective for analysing the results is that they are intended by YouTube to reflect user interests. These interests are localised by YouTube to those of the UK to some extent as this is where the searches were undertaken (Broderson et al., 2012).

5.1 Video Repetitions in Daily Category Searches

Comparison between consecutive search results: In all categories there is a substantial overlap between the videos returned by YouTube API category searches on consecutive days for the same category (Table 5.1). The percentage of repeats between concurrent searches across most of the categories remains relatively consistent with the 60% - 70% repeats each time (Table 5.1). The slight increase in the percentage of repeats between searches 20 to 23 could be due to a media event that changed uploading or viewing behaviours or it may be a technical issue at YouTube, such as a software update. Other temporary cross-category changes are suggested by mainly red vertical stripes at 8-9 and 29-30, and partial green vertical stripes at 18-19 and 27-29. Thus, there seem to be regular alterations in the extent of daily repetition for YouTube categories.

Repeats in the Entertainment and Education categories are slightly less common than in the others (partial horizontal red stripes in Table 5.1). For Entertainment, this may reflect a high volume of uploads or changing tastes whereas Education would contain content with longer term value, for example a demonstration of calculating using long division which would not need to be regularly updated, although there may be many uploads in the category. The Causes, How-to and Non-profit categories have a higher level of repeats between searches – these may have videos that maintain longer-term interest, a smaller pool of videos for the API to select from, or fewer new videos. The greater fluctuation in the From TV category (some high values, some low values) indicates periods of high and low stability, which might reflect traditional TV show interests or periodic batch uploading of TV content from producers or pirates.

Table 5.1. Daily Searches - The percentage of videos returned by each search that also occurred in the next search. High values are green, mid-range values are yellow and low values are red.

Search	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-10	10 -11	11-12	12-13	13-14	14-15	15-16	16-17	17-18	18-19	19-20	20-21	21-22	22-23	23-24	24-25	25-26	26-27	27-28	28-29	29-30
Animation	68.2	68.0	64.8	68.3	70.6	61.2	65.4	46.2	64.2	70.2	72.1	60.5	66.7	73.6	70.0	61.9	70.3	74.7	66.9	79.8	79.1	77.5	71.8	67.4	65.1	63.3	82.4	79.2	45.2
Automotive	66.3	59.5	70.1	65.9	52.7	54.2	73.5	47.6	51.3	71.7	53.9	57.9	76.7	72.2	64.7	76.7	75.7	74.5	67.7	92.1	74.3	63.3	70.6	68.0	70.1	54.1	67.2	61.8	50.6
Beauty	73.3	70.9	77.3	74.6	70.3	67.9	72.4	59.7	62.4	76.5	74.9	69.9	78.9	82.0	61.0	63.4	74.7	75.1	72.3	81.4	80.1	77.5	70.9	75.7	65.7	73.2	78.6	78.0	49.2
Best of	63.6	63.0	61.0	61.4	64.2	65.2	75.4	46.3	56.2	71.5	59.9	61.9	63.2	60.1	60.8	49.5	70.3	69.9	67.0	79.8	64.5	72.6	61.3	65.0	60.0	66.7	75.3	74.6	45.1
Causes	70.6	72.7	82.7	71.0	77.1	69.9	70.7	63.9	67.2	76.9	77.7	73.3	73.9	73.0	76.7	73.7	78.5	88.1	73.0	83.3	82.1	77.7	68.9	71.7	75.7	78.5	87.9	77.8	58.7
Comedy	69.7	65.8	82.5	70.8	73.3	67.4	75.1	51.4	67.8	70.3	72.4	70.3	73.3	77.3	66.7	65.9	68.6	87.5	69.9	84.7	85.8	80.1	67.9	72.6	66.7	68.2	84.4	85.0	44.9
Cooking	68.6	58.8	59.2	61.8	62.3	58.2	72.4	59.0	60.8	69.2	68.5	61.3	69.6	75.4	62.2	63.6	78.0	83.2	72.8	75.9	73.4	74.8	65.6	78.4	74.7	75.1	78.9	73.0	58.0
DIY	74.3	66.0	75.0	71.3	68.2	70.1	70.6	58.5	57.1	67.1	71.1	63.6	70.4	67.4	67.3	63.3	72.1	76.7	68.6	84.3	80.9	72.9	64.3	66.9	70.5	66.6	79.9	81.6	52.8
Education	55.8	61.3	66.3	62.9	56.6	48.1	64.0	57.8	59.3	53.8	62.9	59.6	47.3	68.5	57.5	58.2	64.4	61.8	57.5	56.2	66.5	68.3	59.6	78.7	70.9	73.4	60.2	61.4	63.8
Entertainment	63.4	53.0	66.0	68.2	45.8	45.8	63.6	56.3	51.3	56.1	56.7	51.0	55.3	69.6	57.9	54.2	71.2	78.7	41.9	73.2	57.6	74.9	66.9	72.4	63.1	65.0	82.9	62.4	45.9
Fashion	65.7	70.0	74.9	60.9	68.8	65.6	76.0	54.3	65.5	66.3	70.9	68.2	69.2	70.3	67.5	66.2	71.4	77.9	68.1	74.6	76.6	74.8	72.6	74.3	65.6	69.0	77.7	77.4	46.5
From TV	52.1	78.4	80.0	30.5	82.7	42.5	88.1	39.7	31.4	75.4	44.1	47.2	87.3	85.0	46.5	36.2	82.0	88.5	42.0	66.4	67.3	85.5	68.8	24.7	74.5	76.1	58.3	87.0	41.0
Gaming	78.9	74.3	77.4	70.6	72.8	68.7	80.0	52.0	73.0	77.9	77.9	62.0	68.2	71.3	71.6	64.3	71.5	87.7	71.8	81.6	76.5	77.8	69.3	83.7	70.1	71.2	84.3	86.0	57.4
Health	66.3	72.4	76.8	70.2	69.2	73.0	75.5	54.2	66.5	72.8	71.4	69.3	70.6	72.6	68.8	62.4	70.2	76.1	63.4	78.3	77.1	75.2	63.7	69.2	70.4	70.8	69.9	73.8	46.7
How to	77.0	76.1	80.1	79.8	74.8	71.3	75.7	58.8	75.2	73.3	79.6	71.9	78.0	79.8	70.7	64.6	79.9	81.3	74.9	89.1	82.8	85.5	79.5	83.9	76.3	71.0	91.2	92.4	54.7
Lifestyle	66.8	64.6	69.3	69.0	70.6	64.4	75.9	66.1	66.4	69.0	72.6	70.7	65.6	70.1	71.4	68.6	68.8	75.4	70.9	78.6	84.3	80.4	73.8	76.0	70.7	73.3	79.3	77.8	49.7
News	69.5	66.3	78.6	66.6	68.3	68.9	77.0	42.3	62.4	74.7	76.8	66.8	69.9	69.3	65.1	61.0	68.3	87.2	64.5	87.4	85.2	78.4	69.2	79.7	66.0	57.0	91.3	89.6	19.9
Non-profit	78.7	86.1	71.0	77.0	74.4	73.1	89.5	67.6	67.8	72.3	83.0	83.7	80.1	71.5	58.3	70.4	86.9	94.3	83.4	77.5	80.8	83.6	69.1	73.7	88.1	76.7	83.1	70.7	69.6
Politics	71.4	72.3	71.5	71.6	70.9	72.3	70.5	62.8	70.9	77.2	74.0	74.1	70.6	63.6	68.9	64.2	76.6	81.6	69.9	77.4	78.8	75.6	66.6	78.6	74.4	66.5	74.4	75.5	54.8
Science	69.4	72.4	70.5	65.3	69.5	61.0	66.0	53.4	68.6	66.7	63.9	70.8	63.9	70.4	71.7	63.3	73.3	69.3	59.2	80.1	79.6	73.0	70.6	72.3	69.8	69.8	76.5	75.4	53.7
Sports	68.8	64.7	75.8	65.4	68.2	63.2	75.4	57.4	60.3	68.1	74.8	71.3	71.0	78.1	68.6	55.7	71.1	76.0	65.6	77.4	79.1	70.1	67.8	75.3	65.6	67.2	73.5	75.0	48.0
Tech	68.4	66.4	68.6	69.0	70.0	71.6	76.5	56.5	59.0	67.3	59.7	67.9	67.9	66.4	60.7	63.7	72.4	75.5	64.7	69.0	77.7	72.5	62.9	67.2	70.5	68.0	77.7	78.1	52.2

Comparison between the first and subsequent search results: The number of repeated category search results decreases over time (a left-to-right shift from green to red in Table 5.2). The rate of decline differs between categories (different coloured horizontal stripes in Table 5.2). The News, From TV, Fashion and Comedy repeats decrease at a greater rate than the other categories (mostly red horizontal stripes in Table 5.2). This suggests either a greater number of videos uploaded to these categories providing a wider selection for the API to take its sample from, or changing user interests tracked by the YouTube video selection algorithm. This hypothesis is based on the data sample collected and therefore cannot be generalised to all search results provided by the YouTube API.

News has no repeats with the first search from the 15th search onwards. The most likely explanation for this is that the daily changing nature of mainstream media news is being followed by the YouTube News category video selection algorithm by largely ignoring older videos. This hypothesis cannot be directly tested, however.

From TV has a few anomalies in terms of days when repeats increase (uneven colour changes in its horizontal bar in Table 5.2). This could relate to the story arcs within video material, what is being watched in traditional TV (often on a weekly schedule) and/or content of the TV programmes acting as the sources of these videos.

The number of repeats is high but decreasing for Non-profit, Causes and Automotive (relatively green horizontal bars in Table 5.2). This suggests a smaller pool of videos in these topics, or a core of videos with long term value that new videos cannot compete with.

Table 5.2. Daily Searches - The percentage of videos returned by the first search that also occurred in each subsequent search. High values are green, mid-range values are yellow and low values are red.

Search	1 - 2	1 - 3	1 - 4	1 - 5	1 - 6	1 - 7	1 - 8	1 - 9	1 - 10	1 - 11	1 - 12	1 - 13	1 - 14	1 - 15	1 - 16	1 - 17	1 - 18	1 - 19	1 - 20	1 - 21	1 - 22	1 - 23	1 - 24	1 - 25	1 - 26	1 - 27	1 - 28	1 - 29	1 - 30
Animation	68.2	59.0	56.6	58.1	53.0	43.4	47.1	38.9	36.2	37.8	39.8	37.7	38.0	38.7	34.0	34.0	32.6	34.4	33.3	32.8	32.1	30.9	29.1	26.8	29.9	28.8	27.2	28.9	23.3
Automotive	66.3	64.6	65.5	60.2	66.4	57.5	57.2	56.5	53.6	53.5	50.0	48.2	48.8	45.0	50.1	51.0	46.7	45.6	48.5	47.5	49.6	44.0	43.3	46.1	38.2	40.9	48.3	40.9	44.9
Beauty	73.3	60.9	62.9	57.1	49.6	45.3	45.6	46.3	38.5	41.9	41.3	40.3	41.2	37.4	35.1	30.4	31.5	28.3	27.7	28.0	29.2	27.5	24.5	25.8	26.4	21.1	23.5	22.2	15.2
Best of	63.6	60.4	55.0	52.7	46.8	43.6	45.2	41.3	42.6	39.2	34.7	34.8	33.1	32.7	30.6	30.7	28.2	31.1	30.3	28.4	27.3	28.4	27.4	27.7	30.5	28.1	28.2	28.7	27.0
Causes	70.6	65.3	65.6	60.0	63.2	59.9	56.5	50.9	51.8	52.7	48.7	46.5	45.7	45.1	46.0	43.4	42.2	41.7	41.9	40.9	41.8	40.6	40.1	38.9	38.7	38.4	38.5	37.3	37.8
Comedy	69.7	57.6	53.1	47.4	44.2	37.7	36.9	28.2	28.9	25.5	25.7	20.9	19.5	17.9	17.2	16.8	17.2	16.4	13.8	13.9	12.9	12.4	10.1	9.8	9.0	7.3	7.8	7.5	4.2
Cooking	68.6	54.3	61.1	56.0	52.0	47.8	48.6	42.0	43.7	43.6	42.9	41.7	42.5	35.9	38.9	38.0	38.7	38.9	35.3	34.2	35.1	34.5	36.3	37.8	33.7	32.1	31.5	32.9	28.6
DIY	70.7	66.1	57.4	50.6	56.3	51.8	47.8	38.3	42.6	37.6	37.7	37.1	33.9	38.2	37.0	34.9	30.0	33.3	30.0	28.1	27.0	25.2	27.7	23.8	23.0	21.4	22.9	20.5	15.4
Education	55.8	51.6	53.7	49.4	43.4	39.2	40.6	40.4	33.8	38.6	34.6	31.7	34.4	33.5	34.5	29.4	29.7	28.5	30.9	30.1	29.5	24.5	25.1	24.3	24.4	22.8	29.4	23.9	23.4
Entertainment	63.4	52.5	50.3	46.7	38.5	38.7	43.6	37.7	31.8	31.8	33.6	30.5	30.4	28.7	31.2	29.7	29.4	25.3	26.2	24.9	28.3	28.9	26.8	26.0	23.6	23.4	27.8	27.1	24.0
Fashion	65.7	57.2	51.2	49.2	44.5	38.8	37.4	36.4	32.9	32.1	30.2	30.3	27.3	26.6	24.4	20.9	19.2	19.7	20.2	18.1	15.9	16.5	14.3	14.7	10.1	6.5	6.2	7.8	1.1
From TV	52.1	43.2	60.0	18.4	30.7	17.2	16.4	34.9	15.0	32.3	10.8	17.0	18.6	18.0	16.8	9.2	10.1	8.7	13.0	13.8	9.9	15.5	37.5	10.6	9.2	10.1	8.3	4.6	10.6
Gaming	78.9	64.9	55.7	53.8	47.6	46.1	45.6	39.3	40.6	39.1	34.7	29.2	29.0	27.2	25.7	21.4	23.4	20.4	18.6	17.5	17.5	18.8	20.1	21.9	17.9	16.3	19.2	16.3	15.5
Health	66.3	59.0	60.9	50.1	47.5	44.0	42.0	37.8	40.8	39.2	37.3	34.0	35.6	36.9	35.7	30.3	34.4	32.9	27.3	27.8	26.7	27.1	26.2	26.9	25.6	22.3	27.3	23.5	17.3
How to	77.0	62.5	58.1	52.2	48.4	40.3	39.2	34.2	31.3	33.2	29.5	29.6	28.2	26.3	23.6	23.4	21.5	22.5	22.1	23.0	21.5	20.8	21.3	21.7	19.9	19.0	19.2	20.3	17.1
Lifestyle	66.8	60.8	57.2	54.7	53.5	50.9	49.3	43.6	42.0	41.2	45.0	43.2	36.3	42.7	40.6	39.3	35.0	39.5	37.8	36.6	37.4	37.6	35.4	36.2	34.9	32.4	33.6	30.9	23.7
News	69.5	45.3	40.0	27.2	20.7	16.9	15.8	11.9	8.6	8.4	6.8	3.3	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Non-profit	78.7	76.2	71.4	68.8	71.1	68.3	68.2	55.1	66.2	63.6	60.9	61.0	59.5	60.0	55.9	59.8	59.1	59.3	60.1	48.7	58.2	57.3	48.4	54.9	58.5	50.9	51.8	54.2	49.4
Politics	71.4	65.6	64.4	60.1	58.1	52.3	52.3	51.7	52.7	50.6	47.2	43.5	42.5	44.8	46.8	41.9	42.7	39.4	38.5	36.8	38.4	41.1	35.9	36.6	36.6	38.3	36.8	36.7	30.8
Science	69.4	65.5	62.4	52.8	48.6	48.1	48.9	44.1	40.6	43.3	40.1	39.6	36.7	40.4	36.5	34.6	32.5	33.9	30.8	31.6	31.3	29.8	30.7	28.6	27.9	26.4	28.8	28.5	26.6
Sports	68.8	62.4	62.4	53.5	53.7	45.6	42.9	43.5	36.6	33.4	34.0	34.7	34.9	32.3	35.4	27.4	26.8	27.2	25.1	26.2	24.7	24.3	24.4	22.6	19.0	19.2	18.9	20.3	16.0
Tech	68.4	60.6	59.7	56.8	51.2	48.3	49.5	48.7	42.6	41.9	39.7	36.9	37.4	37.1	36.5	37.1	34.7	33.8	32.8	33.3	31.7	30.5	30.4	32.0	32.3	29.5	30.1	29.8	29.1

Videos within multiple search results: Most videos occur only once over 30 days for most categories (Table 5.3). There is a slightly greater spread of percentages of videos being repeated from 1 to 6 times across the categories Comedy, Gaming, News, Non-profit, From TV and How-to. As before, this may be due to the size of the pool of videos to choose from or the YouTube algorithm reflecting changing user interests.

News provides a higher proportion of videos repeated 3, 4, 5 and 6 times, but has no videos that are repeated more than 15 times (Table 5.3). This supports the above conjecture that the News category is affected by the news cycle covering current and topical events with a limited lifespan. **This gives the strongest evidence yet that the YouTube algorithm is affected by public interest and not just the pool of available videos.**

Table 5.3. Daily Searches - Video appearance frequency over 30 days. High values are green, mid-range values are yellow and low values are red (categories in decreasing order of repetition).

No of appearances	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Automotive	69.6	5.8	1.9	1.7	1.3	1.0	0.9	1.1	0.9	1.4	1.2	0.8	0.5	0.7	0.5	0.7	0.6	0.7	0.7	0.6	0.6	0.6	0.8	1.1	0.9	1.0	0.4	0.8	0.6	0.8
Science	69.5	2.8	2.7	2.4	2.2	2.0	1.5	1.7	1.3	0.9	1.1	1.2	1.1	0.9	1.0	0.6	0.5	0.5	0.5	0.3	0.5	0.5	0.4	0.4	0.5	0.3	0.7	0.8	0.8	0.6
Health	68.8	3.8	3.0	3.3	1.8	1.9	1.9	2.1	1.2	1.2	1.3	0.8	0.8	1.0	0.6	0.5	0.5	0.4	0.4	0.5	0.5	0.3	0.4	0.2	0.2	0.4	0.4	0.7	0.4	0.7
Lifestyle	68.5	4.9	2.3	2.4	1.9	1.7	1.3	1.4	1.3	0.8	1.0	1.1	1.1	0.8	0.8	0.4	0.7	0.6	0.4	0.5	0.4	0.5	0.4	0.5	0.6	0.6	0.7	0.8	0.7	1.0
Education	68.4	3.4	3.0	2.2	2.8	2.1	2.0	1.9	1.5	1.1	1.1	1.0	1.0	0.9	0.6	0.7	0.7	0.6	0.8	0.4	0.6	0.3	0.3	0.4	0.5	0.5	0.2	0.3	0.4	0.4
Tech	68.1	3.8	3.4	2.7	2.7	1.9	2.0	1.7	1.2	1.0	0.7	0.9	0.9	0.7	0.7	0.5	0.7	0.6	0.4	0.3	0.6	0.3	0.3	0.6	0.5	0.4	0.5	0.4	0.6	0.9
Cooking	66.8	3.4	2.6	3.5	2.6	1.6	1.8	1.7	1.4	0.9	1.2	0.8	1.2	0.7	0.7	0.5	0.7	0.5	0.5	0.5	0.9	0.7	0.5	0.3	0.1	0.6	0.8	0.9	0.6	1.0
Politics	66.7	5.8	2.3	2.8	1.7	1.8	1.6	1.3	1.0	0.9	0.9	1.1	0.7	0.9	0.6	0.7	0.6	0.7	0.5	0.5	0.7	0.5	0.6	0.5	0.6	0.8	0.6	0.8	0.8	1.1
Causes	63.1	7.9	2.5	2.2	1.9	1.6	1.5	1.3	1.6	0.7	0.7	0.8	1.0	1.0	0.9	0.4	0.7	0.7	0.6	0.9	0.8	0.3	0.5	0.5	0.5	0.4	0.5	1.0	1.4	1.9
Fashion	63.0	4.6	3.5	3.9	3.5	3.0	2.2	2.1	1.2	1.4	1.1	1.0	1.1	1.2	0.8	0.7	0.6	0.6	0.5	0.7	0.4	0.5	0.4	0.4	0.5	0.4	0.2	0.2	0.3	0.1
Sports	61.6	5.3	4.4	4.2	3.2	3.0	2.1	2.3	1.3	1.0	1.5	0.8	0.8	0.8	0.8	0.6	0.5	0.4	0.4	0.5	0.5	0.3	0.4	0.5	0.4	0.5	0.2	0.6	0.5	0.4
Entertainment	58.7	9.1	5.5	3.6	3.5	2.7	2.4	1.7	1.7	1.3	0.9	0.9	0.9	0.8	0.9	0.7	0.4	0.4	0.4	0.5	0.4	0.5	0.4	0.2	0.3	0.4	0.3	0.3	0.1	0.0
Animation	58.3	8.9	5.0	4.1	3.1	2.6	2.4	1.7	1.2	0.8	1.2	0.8	1.0	0.6	0.8	0.5	0.6	0.5	0.4	0.4	0.4	0.4	0.3	0.2	0.5	0.6	0.5	0.6	0.7	0.6
Beauty	55.7	5.2	4.6	4.9	3.3	2.8	2.3	2.4	1.5	1.2	1.6	1.8	1.3	1.3	0.9	0.8	0.5	0.7	0.8	0.7	0.6	0.6	0.4	0.6	0.5	0.6	0.6	0.7	0.9	0.5
Best of	55.6	11.1	5.1	4.3	2.9	2.8	2.3	2.2	1.6	1.1	1.3	0.9	0.8	1.0	0.7	0.6	0.3	0.5	0.5	0.4	0.3	0.3	0.4	0.3	0.6	0.6	0.7	0.4	0.4	0.1
DIY	54.3	8.6	5.9	4.0	2.9	2.9	2.4	2.3	1.6	1.6	1.5	1.1	0.9	0.9	0.9	0.7	0.5	0.8	0.4	0.6	0.6	0.4	0.5	0.6	0.4	0.3	0.4	0.3	0.6	0.9
Comedy	47.3	13.0	7.8	5.6	4.4	3.5	3.1	2.2	1.4	1.3	1.2	1.4	0.9	1.1	0.8	0.4	0.5	0.8	0.4	0.5	0.3	0.3	0.3	0.3	0.3	0.2	0.1	0.3	0.2	0.3
Gaming	37.5	16.6	8.4	5.9	4.4	3.7	2.4	2.8	2.0	2.4	1.4	1.9	1.2	1.5	0.8	0.9	0.5	0.7	0.5	0.4	0.5	0.5	0.3	0.3	0.2	0.3	0.2	0.2	0.9	0.7
News	29.6	18.3	15.0	11.9	6.3	4.6	3.8	2.8	2.1	1.6	1.4	1.2	1.0	0.3	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Non-profit	26.9	18.5	13.6	4.6	2.1	2.5	1.4	1.9	1.8	1.4	1.8	1.2	1.0	1.3	1.1	1.0	0.8	0.6	0.6	0.6	0.9	1.1	1.2	1.4	1.2	1.9	1.9	0.9	2.5	2.6
From TV	17.3	21.0	17.4	7.5	5.0	3.6	4.1	2.0	2.8	3.1	1.3	2.6	0.2	1.0	1.3	1.1	1.8	2.6	1.6	0.3	0.2	0.3	0.2	0.3	0.5	0.0	0.2	0.7	0.0	0.2
How to	16.0	10.4	8.1	7.3	4.6	3.4	2.6	1.8	1.3	1.4	1.3	1.0	0.7	0.9	0.8	0.6	0.4	0.7	0.3	0.3	0.7	0.4	0.2	0.1	0.2	0.4	0.4	0.3	0.4	0.4

5.2 Metadata and Metrics for Daily Category Search Videos

Metadata for videos returned by all searches within a category: Categories with many videos appearing in all 30 results seem to have relatively stable topics (Table 5.4). Best of has the most videos that appeared within every sample, suggesting that being a top video is a relatively stable characteristic.

There is no clear relationship between stability and the mean values of the other metrics (Table 5.4). The existence of substantial differences in metadata between categories with similar numbers of videos suggests that the cause of the differences is the category topic rather than the characteristics of its videos. Entertainment and From TV have 1 video that appeared in every search. There is a substantial difference in the metrics for these except length, which could be a coincidence (Table 5.4). There are also substantial differences in the metadata averages for Gaming and Animation with 20 videos and Health and DIY with 26 videos. Thus, **there isn't a simple cross-category relationship between metadata values and likelihood to be selected by the API**.

One clear finding from the metadata is that the frequently repeated videos are young – less than four months old in all except two cases. This is important because YouTube has decades of videos, so its algorithm is selecting almost exclusively recent videos for frequently repeated videos. With one exception (From TV: video with title, **"**Very creepy, disturbing children's cartoon, banned from TV", posted in 2007 and with over 15 million views: https://www.youtube.com/watch?v=cqi5F5MqqTQ) the results do not include classic viral YouTube videos. For example, Gangnam Style (3.4 billion views: https://www.youtube.com/watch?v=9bZkp7q19f0) and Charlie Bit My Finger (8500 million views: https://www.youtube.com/watch?v=_OBlgSz8sSM) are absent from the results.

Table 5.4. Daily Searches – Number of videos and metadata averages for videos appearing in all results sets for a category. High values are green, mid-range values are yellow and low values are red. The metadata is from the first occurrence of the each video in the dataset.

30 Appearances	Videos	Days	Title Words	Dislikes	Likes	View Count	Comments	Length
Best of	120	71	12	55	1704	320505	186	933
Causes	56	91	8	35	627	45126	196	584
Non-profit	47	156	8	0	9	695	4	565
Politics	37	73	9	39	1009	34952	126	1041
Tech	33	71	9	63	2361	113747	883	2192
Lifestyle	31	80	9	49	716	45658	203	891
Cooking	28	67	10	64	2484	133623	230	660
DIY	26	64	8	311	32079	451146	2329	372
Health	26	77	7	5	196	7802	27	1594
Automotive	23	98	8	2	139	4408	32	695
Animation	20	75	7	599	20670	1345041	1755	156
Gaming	20	67	9	150	11368	394395	1256	931
Science	19	72	7	179	6388	426794	534	1482
How to	16	81	8	503	14486	862542	1235	292
Education	15	67	7	2	30	1923	13	1650
Beauty	14	67	10	211	7650	526747	685	390
Sports	12	73	6	244	4238	717957	570	355
Comedy	8	58	12	174	9883	475210	972	1828
Fashion	5	67	7	108	6248	130373	499	783
Entertainment	1	54	9	96	21379	250505	1148	318
From TV	1	3076	8	4526	55792	14812772	59049	305
News	0	0	0	0	0	0	0	0

Metadata for videos returned by only one search for a category: There are wide variations in the average metadata properties of videos returned only once by API searches (Table 5.5). The API may therefore ignore the associated metrics when taking a sample, use more complex popularity indicators (e.g., views in the last day), or may reference the values to within-category averages. Nevertheless, videos with no repeats (Table 5.5) tend to have lower means than always repeated videos (Table 5.4) across the metrics associated with the categories. Since higher popularity metrics (Likes, Dislikes, Views, and Comments) are a logical consequence of being frequently recommended, it is not clear whether this is a cause or effect of the recommendation. Either the videos were selected for category searches partly based on their popularity, or their category recommendation quickly resulted in popularity, triggering the YouTube system to retain them as category recommendations.

With two exceptions (Non-profit, From TV), the videos occurring once are under two months old. Given that YouTube has decades of videos to choose, **the YouTube category search API algorithm primarily selects young videos**. There does not seem to be a systematic pattern in terms of title words or video length, however.

Table 5.5. Daily Searches – Number of videos and associated mean data for videos occurring in a single result set for a category. High values are green, mid-range values are yellow and low values are red.

No Repeat	Videos	Days	Title Words	Dislikes	Likes	View Count	Comments	Length
Health	2503	37	9	2	29	972	5	563
Tech	2456	37	9	2	61	1843	7	593
Science	2386	36	10	1	39	2539	6	4171
Education	2360	39	8	1	8	2682	2	1454
Fashion	2347	36	8	5	171	8736	19	260
Lifestyle	2172	36	8	1	10	363	2	405
Politics	2166	35	10	0	3	371	1	1334
Automotive	2129	38	8	0	0	55	0	129
Cooking	2127	38	8	3	51	4077	7	420
Entertainment	2093	59	8	7	121	25800	14	845
Sports	2047	37	9	6	115	11555	12	406
Animation	1955	40	11	17	285	31738	32	3341
Best of	1906	55	10	42	1106	142873	91	3050
Causes	1859	36	9	0	4	327	1	654
DIY	1728	40	9	22	826	18245	64	255
Beauty	1674	36	9	6	213	7581	28	2798
Comedy	1648	38	12	9	134	20552	12	3951
News	1222	35	9	18	339	20548	91	753
Gaming	1107	41	10	20	610	18272	236	1736
How to	583	61	9	110	3667	135812	387	735
Non-profit	489	85	8	1	73	4063	10	452
From TV	106	431	10	5	95	33768	23	293

Comments per video: A high proportion of videos in all categories have few (0 to 10) comments (Table 5.6), confirming that **the API does not wait for a video to be extensively discussed before returning it in search results**. Automotive, Education and Non-profit have a higher percentage of videos with 0 to 10 comments. Either videos in these categories receive fewer comments or there are fewer videos to choose from, so less popular videos are returned by the API. There do not appear to be substantial similarities within each of the bandings when compared across the different categories.

Table 5.6. Daily Searches – Percentage of videos by number of comments. High values are green, mid-range values are yellow and low values are red.

DS - Comments	0 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 100	101 to 500	501 to 1000	1001 to 5000	5001 to 10000	10001 to 50000	50001 to 100000	100001+
Animation	32.62	6.39	4.35	3.14	2.79	6.51	17.88	7.43	14.78	3	1.11	0	0
Automotive	89.96	4.73	1.51	0.67	0.57	1.54	0.8	0.22	0	0	0	0	0
Beauty	29.27	6.78	5.16	3.29	2.82	11.28	24.17	8.8	7.54	0.75	0.14	0	0
Best Of	24.24	4.31	2.68	2.48	2.06	9.15	27.55	10.52	14.65	1.45	0.78	0	0.12
Causes	67.82	5.65	2.33	2.95	2.04	5.52	9.41	2.3	1.79	0	0.2	0	0
Comedy	54.44	12.32	6.97	4.17	3.02	5.87	7.65	2.81	2.4	0.25	0.1	0	0
Cooking	39.07	9.36	6.38	3.98	2.55	10.38	21.82	3.42	3.02	0	0.03	0	0
DIY	18.34	5.84	4.34	3.77	2.78	12.17	28.86	9.79	11.81	1.7	0.61	0	0
Education	83.04	3.2	2.51	1.98	0.55	1.5	5.38	1.16	0.67	0	0	0	0
Entertainment	62.61	6.95	3.71	2.46	1.92	7.16	8.98	3.02	2.59	0.32	0.29	0	0
Fashion	56.19	5.78	3.6	2.21	2.86	7.55	15.27	4.6	1.87	0.08	0	0	0
From TV	75.12	1.41	1.3	1.7	0.64	0.98	8.49	1.96	6.21	0.12	1.21	0.87	0
Gaming	18.01	5.27	4.02	2.78	1.78	8.12	29.03	13.8	14.39	2.55	0.26	0.01	0
Health	79.97	5.2	1.69	1.66	1.04	3.77	4.51	1.32	0.84	0	0	0	0
How to	19.06	5.53	4.99	2.8	2.51	9.4	25.91	9.95	16.28	2.77	0.81	0	0
Lifestyle	54.14	7.81	5.62	3.39	3.34	7.97	14.77	2.11	0.84	0	0	0	0
News	38.36	6.27	4.37	3.46	2.8	9.9	24.78	4.88	4.76	0.12	0.17	0.03	0.08
Non-profit	93.73	2.78	1.27	0.31	0.07	1.32	0.28	0.22	0.01	0	0	0	0
Politics	66.93	7.74	3.47	2.47	1.16	6.22	9.22	1.72	0.82	0.13	0.13	0	0
Science	60.95	7.56	3.51	3.14	1.51	5.01	9.78	3.28	3.85	0.89	0.51	0	0
Sports	52.65	4.91	3.99	2.64	2.41	7.66	18.4	3.47	2.78	0.57	0.52	0	0
Tech	63.08	4.21	2.82	1.57	2.37	4.82	9.6	5.18	4.45	1.52	0.37	0	0

Likes per video: There is a high percentage of videos extracted by the API that have few, or no, likes, particularly Automotive, Causes, Education, From TV, Health, Non-profit, Politics and Science (Table 5.7). This suggests that **the API does not primarily select highly liked videos**.

There is a rise in the percentage of videos in the '101 to 200' and '501 to 1000' bands are due to increases in band width. The categories Animation, Beauty, Best of, DIY, Gaming and How-to have a high percentage of videos with more than 4000 likes – this could reflect the overall popularity or longevity of videos in these categories (Table 5.7). These categories (except Animation) may provide more instructional videos that would have longer-term value.

There are substantial differences in the percentages of likes in each of the columns across all the categories. The number of likes that a video receives is therefore not used in a simplistic cross-category way in the sample that the API provides (Table 5.7).

Table 5.7. Daily Searches – Percentage of videos by number of likes. High values are green, mid-range values are yellow and low values are red.

DS - Likes	0	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 3000	3001 to 4000	4001+
Animation	8.36	13.02	3.20	2.08	1.52	1.50	1.18	0.85	0.70	1.15	0.95	5.51	3.71	2.45	2.24	7.22	4.35	2.25	3.77	2.80	31.19
Automotive	41.87	36.90	5.50	2.75	2.31	2.13	1.29	1.23	0.66	0.28	0.43	2.52	1.33	0.16	0.07	0.28	0.01	0.07	0.22	0.00	0.00
Beauty	10.99	8.29	4.59	2.47	2.21	2.10	1.23	1.39	0.48	0.90	1.09	7.07	4.70	3.55	2.04	7.02	4.44	2.75	5.46	3.65	23.57
Best of	10.92	7.42	1.75	0.97	0.93	0.64	0.53	0.45	0.62	0.66	0.51	3.53	2.93	2.48	2.00	10.41	6.89	4.31	6.43	3.84	31.78
Causes	22.56	33.23	7.40	4.62	2.79	1.22	0.73	1.66	1.35	1.28	0.87	6.88	2.72	1.15	1.26	3.39	1.43	1.63	2.03	0.29	1.54
Comedy	6.11	11.95	6.35	4.32	3.78	3.54	2.71	2.49	2.31	2.17	1.71	14.42	8.81	4.81	3.33	8.22	2.53	1.94	2.09	1.53	4.88
Cooking	16.71	11.15	4.57	2.90	1.86	1.54	2.26	1.70	1.12	1.35	0.99	7.95	6.55	3.37	2.83	9.45	4.65	3.78	3.58	3.30	8.38
DIY	7.55	6.46	1.84	1.24	1.14	1.01	0.83	0.71	0.95	0.71	0.57	4.53	3.94	3.38	3.17	9.30	4.78	4.40	5.74	4.94	32.82
Education	26.06	42.48	7.27	3.83	3.08	1.65	1.04	0.95	0.39	0.43	0.26	2.89	2.17	0.81	0.71	1.63	0.95	0.48	1.76	0.71	0.65
Entertainment	22.03	21.34	4.93	3.95	3.11	2.26	2.58	1.70	1.62	1.08	0.95	4.37	3.49	3.53	2.05	5.82	3.01	1.89	2.31	1.73	6.24
Fashion	17.95	16.63	6.53	5.15	3.89	2.40	2.08	1.19	1.15	0.87	1.21	5.68	3.04	2.43	2.24	7.18	4.73	3.37	4.68	1.81	5.78
From TV	56.87	15.65	1.53	0.14	0.35	0.64	0.20	0.06	0.61	0.69	0.09	2.51	2.68	1.62	0.75	3.46	1.07	0.14	1.27	0.64	9.04
Gaming	3.79	8.34	3.50	2.23	1.72	1.38	1.44	0.91	0.90	0.81	0.96	7.12	5.67	3.35	3.05	8.20	6.71	4.72	7.80	4.24	23.14
Health	30.94	35.52	8.45	3.07	1.73	1.97	0.94	0.40	0.90	0.72	0.58	3.90	2.98	1.03	0.86	2.39	0.51	0.36	0.52	0.47	1.77
How to	2.75	7.47	2.39	1.89	1.42	0.96	0.79	0.91	0.54	0.66	0.38	5.26	5.02	3.97	3.13	10.66	6.66	4.73	4.85	4.72	30.81
Lifestyle	18.18	24.10	7.37	3.82	2.82	1.70	1.43	1.65	0.94	1.68	1.09	7.65	4.43	2.64	2.29	7.08	2.03	3.42	2.45	0.65	2.57
News	6.09	21.30	6.42	3.77	2.89	2.59	2.86	2.13	1.41	1.71	1.55	10.30	5.62	3.97	2.70	10.25	4.29	2.48	2.68	1.47	3.51
Non-profit	49.51	38.04	3.98	1.37	1.08	1.37	0.55	0.36	0.17	0.05	0.04	1.67	0.63	0.17	0.05	0.42	0.26	0.06	0.01	0.03	0.19
Politics	25.46	30.81	9.46	4.86	3.34	1.85	1.72	0.75	1.21	0.47	0.63	4.69	1.35	1.29	1.53	3.40	1.55	1.08	1.84	0.66	2.04
Science	22.76	22.59	6.19	4.19	2.86	1.49	1.39	1.13	0.92	1.23	1.28	8.44	3.25	2.08	1.16	3.17	2.47	2.34	2.29	0.67	8.11
Sports	19.32	18.18	5.76	2.33	2.18	1.20	1.38	0.90	0.74	1.17	1.07	8.97	3.89	2.30	2.19	7.59	3.64	2.10	2.38	1.80	10.90
Tech	18.69	28.93	9.18	3.53	2.29	1.68	1.61	0.84	0.64	0.53	0.81	4.01	2.75	1.30	1.65	3.73	4.23	2.39	2.00	1.04	8.17

Dislikes per video: A high proportion of videos have few dislikes across most of the categories (Table 5.8). Automotive and Non-profit seem to have a low percentage of dislikes – also for Causes, Education and Health. The videos provided within the API sample have fewer dislikes than likes (Tables 5.7 and 5.8), which may be in line with YouTube in general. As with likes, the variation in the percentages of dislikes across the categories within the bands suggests that there is not a pattern in relation to what the API selects. It seems that **the API does not take dislikes or likes into consideration in a simplistic way when selecting videos** for a category search, although they might be minor features in the algorithm.

Table 5.8. Daily Searches – Percentage of videos by number of dislikes. High values are green, mid-range values are yellow and low values are red.

DS - Dislikes	0	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 3000	3001 to 4000	4001+
Animation	21.74	17.74	6.01	4.25	2.70	2.12	1.60	1.47	1.27	1.14	1.10	8.47	5.02	4.07	2.92	6.77	4.72	1.94	2.64	1.26	1.07
Automotive	80.21	18.72	0.56	0.34	0.14	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beauty	23.63	24.63	6.71	3.56	4.47	3.24	2.63	2.51	2.40	1.67	1.96	9.84	3.73	2.94	1.19	2.72	1.00	0.53	0.47	0.05	0.12
Best of	18.03	13.62	6.68	4.53	4.47	2.64	2.84	2.20	1.77	1.62	1.66	8.99	5.11	3.70	2.69	7.60	5.13	1.67	1.21	1.37	2.48
Causes	52.19	31.22	4.99	2.56	1.17	0.38	0.79	1.65	0.77	0.36	0.10	1.96	0.93	0.37	0.06	0.21	0.18	0.12	0.00	0.00	0.00
Comedy	14.66	34.06	13.83	8.42	6.76	3.61	2.90	1.42	1.69	1.40	0.57	4.30	2.28	1.06	0.86	1.28	0.47	0.28	0.04	0.00	0.10
Cooking	32.59	31.06	8.09	5.03	3.65	1.79	1.95	1.95	1.53	0.69	0.76	4.24	1.93	0.79	0.59	1.48	0.82	0.31	0.35	0.24	0.16
DIY	17.02	20.41	8.94	5.55	3.59	3.75	2.93	2.28	2.19	1.62	1.91	10.02	4.97	3.17	2.98	5.70	1.70	0.44	0.40	0.25	0.19
Education	68.05	21.68	2.82	1.37	0.79	0.35	0.52	0.47	0.44	0.12	0.47	1.61	0.48	0.30	0.11	0.38	0.03	0.00	0.00	0.00	0.00
Entertainment	44.29	26.77	7.47	4.08	2.07	0.89	1.31	1.04	1.05	0.23	0.55	4.31	1.52	1.01	0.22	0.93	0.74	0.46	0.36	0.25	0.45
Fashion	35.79	31.48	6.02	4.65	3.61	2.30	2.38	2.37	1.71	0.95	0.68	4.32	2.10	0.66	0.28	0.46	0.23	0.01	0.00	0.00	0.00
From TV	71.71	7.88	2.66	1.39	0.66	0.61	0.55	1.44	1.50	0.23	0.03	1.96	3.35	1.21	1.33	0.49	1.39	0.00	0.00	0.00	1.41
Gaming	13.74	24.84	10.04	6.96	4.77	3.98	2.54	2.25	1.51	1.66	0.92	8.29	6.33	3.54	2.31	3.01	1.81	0.27	0.64	0.31	0.27
Health	68.61	21.68	3.82	1.42	0.77	0.68	0.39	0.55	0.17	0.03	0.11	0.74	0.22	0.47	0.28	0.06	0.00	0.00	0.01	0.00	0.00
How to	11.28	23.64	10.41	5.65	4.40	4.14	2.36	1.91	1.89	2.34	1.69	7.80	4.25	3.15	2.01	6.67	2.07	1.17	1.21	0.78	1.19
Lifestyle	44.50	35.42	5.25	3.21	2.68	1.61	0.74	0.32	0.65	0.80	0.55	2.40	0.36	0.22	0.27	0.54	0.20	0.08	0.00	0.20	0.00
News	19.01	37.80	14.24	6.73	3.67	2.45	2.10	1.19	1.41	1.08	0.50	4.75	2.16	0.88	0.32	1.06	0.20	0.06	0.32	0.02	0.06
Non-profit	87.82	11.05	0.60	0.28	0.01	0.01	0.00	0.20	0.00	0.00	0.01	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Politics	48.76	35.10	5.41	1.77	1.67	1.60	1.04	0.39	0.31	0.85	0.19	1.37	0.69	0.34	0.12	0.08	0.00	0.00	0.19	0.10	0.00
Science	45.37	29.59	6.52	2.95	2.14	1.55	1.22	0.56	0.48	0.42	0.31	1.67	1.74	1.50	0.68	1.44	0.48	0.30	0.89	0.10	0.10
Sports	38.74	26.92	6.52	4.10	2.59	1.94	1.28	1.99	1.36	1.04	1.30	5.58	1.83	1.10	0.82	1.14	0.72	0.28	0.15	0.00	0.58
Tech	48.36	25.94	5.05	2.57	1.91	1.56	1.43	1.22	0.90	0.77	1.18	3.97	1.93	1.42	0.38	1.04	0.32	0.05	0.00	0.01	0.00

Video age: Most videos within the sample provided by the API are 51 to 200 days old (Table 5.9). Apart from the From TV category, there are no videos older than 1001 days. Best of, Entertainment, From TV and Non-profit have slightly more, older videos, suggesting longevity of the content or a lack of requirement to update material. The API does not seem to provide any particularly new videos, i.e. those within the 1 to 10-day banding, with the small exception of Health and Non-profit. Thus, **the API chooses videos that are at least 11 days old**. After this, differences between categories may be due to the longevity or competition (number of videos) of typical videos.

Table 5.9. Daily Searches – Percentage of videos by age (number of days). High values are green, mid-range values are yellow and low values are red.

DS - Days	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 2500	2501 to 3000	3000+
Animation	0.00	2.15	5.74	8.38	10.27	46.94	26.53	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Automotive	0.00	2.72	5.47	5.07	5.91	19.49	40.28	21.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beauty	0.00	2.12	6.46	6.93	14.98	65.80	3.72	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Best of	0.00	2.30	5.53	7.95	8.60	25.12	28.27	19.75	2.42	0.07	0.00	0.00	0.00	0.00	0.00	0.00
Causes	0.00	2.38	6.06	6.33	7.48	32.72	44.66	0.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Comedy	0.00	2.38	7.28	11.84	19.13	59.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Cooking	0.00	2.54	5.60	6.61	11.93	45.64	27.60	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
DIY	0.00	2.67	6.05	10.98	14.34	60.64	5.35	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Education	0.00	2.65	6.50	7.02	7.99	37.25	35.92	2.68	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Entertainment	0.00	1.17	6.36	6.45	8.12	20.83	27.27	17.65	7.61	3.68	0.86	0.00	0.00	0.00	0.00	0.00
Fashion	0.00	3.14	9.19	14.84	17.06	55.26	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
From TV	0.00	1.41	3.98	3.72	4.16	1.59	15.59	16.97	3.35	8.31	18.82	6.90	6.52	6.41	0.84	1.41
Gaming	0.00	2.14	8.20	15.77	17.31	48.94	7.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Health	0.01	3.18	8.09	8.06	10.89	36.12	33.52	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
How to	0.00	1.93	8.23	12.10	13.99	42.44	21.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lifestyle	0.00	2.51	6.41	6.46	9.39	38.58	36.65	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
News	0.00	4.06	20.09	25.51	26.10	24.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Non-profit	0.01	1.48	4.46	3.51	4.13	11.60	25.37	25.08	18.13	6.24	0.00	0.00	0.00	0.00	0.00	0.00
Politics	0.00	2.70	6.70	8.47	11.20	42.41	28.52	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Science	0.00	2.51	6.52	8.18	9.81	40.71	32.27	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sports	0.00	3.03	7.74	8.77	14.36	49.71	16.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tech	0.00	2.68	7.32	9.15	11.96	45.14	23.75	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Video length: There is an even spread in the length of the videos provided by the API (Table 5.10). Almost half (45%) of Comedy videos last for more than 1 hour (3600 seconds). The available data indicates that these are full-length professional comedy performances (live or from TV) rather than shorter user generated amateur productions. Since the focus of YouTube is user generated content (Arthurs et al., 2018; Berryman and Kavka, 2018; Crick, 2016; Dehghani et al., 2016) and issues relating to copyright (Bulakh et al., 2014; Jondet, 2008; Frey, 2007) then shorter amateur productions will outnumber longer professional content. Therefore, from the data collected, this suggests that the **API is deliberately, or as a side effect of its algorithm, favouring the longer Comedy videos**.

A fifth of the From TV videos last about a minute (61 to 70 seconds). These are probably trailers for TV shows since they are too short for full episodes and the standard length suggests that they are not ad-hoc popular scenes from longer shows. In addition, previous research has discussed that TV advert ‘spots’ generally average 60 seconds (McAllister and Stoltzfus-Brown, 2020; Baranova and Trofymenko, 2018; Wells et al., 2005).

Table 5.10. Daily Searches – Percentage of videos by length (seconds). High values are green, mid-range values are yellow and low values are red.

DS - Length	0 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 120	121 to 150	151 to 180	181 to 210	211 to 240	241 to 300	301 to 360	361 to 420	421 to 540	541 to 660	661 to 900	901 to 1200	1201 to 2400	2401 to 3600	3601 to 6000	6001+
Animation	2.41	2.43	1.64	2.71	3.04	3.68	4.11	3.64	2.76	7.89	8.47	5.4	4.68	2.85	5.06	2.73	2.17	2.58	2.39	2.59	0.68	3.44	3.04	13.09	6.52
Automotive	2.17	3.68	1.08	2.42	1.9	3.79	5.65	2.79	3.11	8.78	9.09	6.35	5.99	5.3	6.72	4.2	2.5	3.78	4.23	3.05	2	4.53	2.26	3.99	0.66
Beauty	0.3	0.29	0.15	0.49	0.99	0.76	0.97	0.78	1.65	4.11	3.34	3.8	5.23	2.11	5.07	7.66	4.74	9.83	11.06	12.64	8.38	4.91	4.3	4.19	2.26
Best of	0.3	0.26	0.25	0.31	0.29	0.23	0.69	0.9	0.72	3.2	4.47	3.49	3.41	4.23	6.39	5.77	4.56	6.22	5.58	7.9	6.46	10.44	3.39	8.47	12.07
Causes	0.46	1.1	2.86	3.14	2.47	4.7	3.81	3.92	2.93	5.46	6.72	6.67	3.99	5.32	8.02	5.22	4.71	7.81	4.58	3.14	2.86	4.61	2.05	2.3	1.14
Comedy	0.1	0.47	0.44	1.39	0.73	0.58	1.02	0.65	1.21	2.09	4.44	3.9	1.12	0.84	2.28	2.16	3.01	2.6	1.57	3.12	3.67	11.65	5.15	20.86	24.96
Cooking	1.3	0.56	0.43	2.15	2.28	1.53	2.05	3.13	1.75	3.49	3.8	4.02	4.45	8.72	11.17	7.53	5.06	9.49	6.33	7.8	4.7	4.96	1.38	1.4	0.53
DIY	2.61	0.34	0.06	0.34	0.34	0.37	0.62	0.45	1.02	3.94	4.87	4.21	6.09	6.95	13.88	12.19	9.15	14.2	6.3	6.72	3.14	1.83	0.25	0.08	0.03
Education	1.62	0.8	0.55	0.95	0.77	1.77	1.72	2.23	1.56	6.54	6.6	7.31	4.99	3.74	6.89	3.37	3.08	4.97	3.96	6.17	3.31	7.3	6.6	7.46	5.71
Entertainment	1.68	2.34	1.23	2.66	2.28	1.58	2.18	2.81	2.12	5.49	6.07	5.31	5.95	5.34	9.82	5.04	3.36	5.75	4.49	4.82	2.99	5.88	6.62	2.94	1.24
Fashion	8.06	0.9	0.74	0.99	1.02	1.41	1.38	1.61	2.4	8.88	7.15	5.82	5.4	3.77	6.17	6.13	3.77	7.44	6.54	7.7	5.19	4.63	1.53	1.15	0.25
From TV	0.95	2.31	2.37	3.2	0.95	1.59	20.03	3.41	0.81	7.3	5.23	5.72	3.61	3.58	5.74	4.3	1.07	5.66	3.35	3.64	1.82	4.94	5.86	1.5	1.07
Gaming	0.45	0.4	0.18	0.42	0.35	0.35	1.15	1.04	0.33	1.82	1.34	2.92	2.61	3.43	5.67	5.52	7.91	12.03	7.85	9.86	8.42	8.73	8.5	5.18	3.52
Health	2.99	0.93	0.41	1.6	0.91	1.35	2.12	2.54	2.47	6.41	7.25	7.27	4.64	4.04	6.87	3.91	2.77	4.58	3.01	4.19	3.22	8.48	9.44	5.85	2.75
How to	0.01	0.22	0.64	1.09	0.85	1.76	1.73	1.54	2.34	6.55	7.44	6.58	5.31	6.57	10.41	7.44	6.63	8.4	5.29	5.91	3.76	4.54	2.51	2	0.47
Lifestyle	0.47	1.04	0.84	1.54	2.28	1.27	1.83	1.38	1.91	4.72	7.33	4.32	3.82	4.68	9.17	6.79	3.36	7.84	7.48	10.37	5.51	6.79	3.76	0.97	0.51
News	0.35	0.37	0.64	2.19	3.05	4.29	4.12	3.39	3.07	8.86	8.86	8.51	4.28	3.41	5.38	5.38	2.86	3.85	4.01	5.38	4.45	8.25	2.84	1.26	0.94
Non-profit	1.24	1.02	0.81	1.27	2.41	2.47	2.15	1.64	5.14	10	9.04	9.2	4.41	6.27	7.62	4.6	3.55	4.71	4.81	2.89	2.11	6.86	2.97	1.64	1.18
Politics	1.86	0.76	0.86	1.39	0.34	0.91	0.94	1.14	1.21	3.31	3.55	5.64	3.44	4.71	4.83	5.04	3.36	6.62	4.53	6.89	3.38	10.77	9.97	10.28	4.28
Science	0.85	1.11	1.11	0.84	0.58	0.67	1.14	1.35	1.52	5.34	4.34	5.33	5.46	5.17	6.75	3.83	2.62	3.77	3.09	2.84	2.5	5.97	12.52	12.99	8.34
Sports	3.42	0.85	1.21	3.36	1.73	2.25	2.53	2.72	2.64	7.41	4.94	5.83	5.12	4.82	9.97	7.59	4.78	5.58	5.81	5.93	1.97	3.81	1.52	3.5	0.71
Tech	1.39	0.86	0.79	0.97	1.26	1.93	3.2	1.84	2.02	6.46	7.93	5.99	5.89	4.14	8.66	6.34	5.14	5.31	3.74	5	2.56	5.43	3.1	6.26	3.81

Views: Since all categories include some videos with less than 11 views, **popularity is not a requirement for the API**. Except for Animation and Best of, most videos provided by the API sample have less than 1 million views (Table 5.11). The **API is also not returning a substantial number of the most popular videos**, classing videos receiving more than 1 million views as popular (O’Neill, 2011; Cutler, 2009).

There does not appear to be a pattern in the number of views that a video receives and the likelihood of the API including some bandings within a sample. The different bands of views have quite a variation in the percentages of videos and therefore suggests that the API does not focus on a view count range when selecting a sample.

Table 5.11. Daily Searches – Percentage of videos by view count. High values are green, mid-range values are yellow and low values are red.

DS - Views	0 to 10	11 to 100	101 to 1000	1001 to 10000	10001 to 100000	100001 to 1000000	1000001+
Animation	5.3	4.8	7.4	12.4	22	27.5	20.6
Automotive	18.3	17.1	39.6	20.8	4.1	0.1	0
Beauty	3.3	4.2	17.4	17.3	28.2	26.6	3.1
Best of	4.4	6.9	7	6.3	18.1	30.5	26.9
Causes	11.5	10.4	28.9	26.3	17.7	4.7	0.5
Comedy	2.6	2.2	5.1	17.1	46.2	24.9	2.1
Cooking	8.7	6.6	14.9	23.2	28.1	15	3.4
DIY	7.2	4	5.8	14.5	33.8	29	5.8
Education	11.6	18.2	34.4	22.6	7.6	4.8	0.8
Entertainment	11.9	12.1	15.5	18.3	22.6	15.7	3.9
Fashion	13.4	4.6	18.4	25.6	26.5	10.4	1.1
From TV	34.6	23.7	14.7	3.1	5	11.1	7.7
Gaming	2.4	5.8	6.1	16.7	39.4	25.2	4.4
Health	13	15.6	35.3	23.2	9.6	3	0.3
How to	0.7	4	6.7	15.3	33.7	28.2	11.3
Lifestyle	11.9	10.6	25.9	26.1	21.5	3.8	0.3
News	0.4	3.5	15.7	31.2	38.5	10.1	0.7
Non-profit	13.5	40.1	33.5	9.9	2.7	0.5	0
Politics	14	11.3	28.9	27.8	12.9	4.5	0.6
Science	13.9	7.1	22.1	23.9	19.6	9.5	4
Sports	10.1	7.7	15.9	20.2	21.7	20.7	3.7
Tech	12.8	7.7	25	23.9	19.2	9.9	1.5

5.3 Five Day Search Results

The Five Day Searches data set consists of the results of YouTube API queries for videos from 22 categories submitted every 5 days over 95 days (20 samples). These largely echo the Daily Search patterns, with some exceptions. The key differences between search methods are discussed below (for a full discussion of the Five Day Search findings see Appendix 10).

Comparing the Five Day (Table A.1 – Appendix 10) and Daily (Table 5.1) searches, there are slightly more repeats in the Five Day sample. This seems illogical as the rate of repeats decreased over time in the Daily sample, with fewer repeats after five days than after 1 (Table 5.2: percentages in the 1-2 column are higher than percentages in the 1-5 column and there is a general decreasing trend, left to right). This suggests that **there was a change in the way that the YouTube API selected the videos between submitting the Five Day (95 days, starting 29/12/15) and Daily searches (30 days, starting 4/4/16)**.

For News (see also: Table 5.2,) after 15 Daily Searches there were no further repeats with the first search. For Five Day Searches (Table A.2 – Appendix 10) repeats continued across all 20 searches (and period of 95 days). This **invalidates the previous conclusion of a moratorium (15 days) on older News videos**. Nevertheless, the News category has relatively few long-term repeats. There could have been a change in the way that API selects it sample between the periods that the Five Day and Daily Searches were submitted, or differences in the news cycle.

The average number of likes and dislikes is higher with the Five Day searches (Table A.6 – Appendix 10) than for the Daily searches (Table 5.4), which may be due to changes in the algorithm or the season of the year when the videos were viewed.

Comparing the Five Day (Table A.11 – Appendix 10) and Daily (Table 5.9) metadata, there is a substantial difference in what the API has provided in the number of days that the videos that have been posted to YouTube. The Daily Search data (Table 5.9), apart from two categories, to suggest that the API extracts few videos that have uploaded for more than 300 days. Whereas the Five Day Search data (Table A.11 – Appendix 10) shows a much wider spread of videos all the way up to the '3001+' banding. In addition there is a substantially higher percentage of days for most categories within the '501 to 1000' band (Table A.11 – Appendix 10), where this was the '51 to 100' banding for the Daily searches (Table 5.9). The Daily search data (Table 5.9), apart from 0.01% for Health and Non-profit, had 0% for all the other categories within the '1 to 10' banding, but the Five Day data (Table A.11 – Appendix 10) provides a variety of percentages within this banding. As the differences across these two search approaches for number of days uploaded is clearly so substantial the way in which the API choses the videos to extract must have altered or there must have been a change to the algorithm that YouTube employs.

When the Five Day data (Table A.12 – Appendix 10) is compared to the Daily findings (Table 5.10) there is a similar pattern in the spread of the percentage of videos within each length band across the categories. Nevertheless, the percentages are slightly lower in the shorter video length bands (approximately 0 to 90 bands) and slightly higher in the upper middle bands (241 to 900 bands).

There is a narrower spread of view count percentages across the bandings (Table A.13 – Appendix 10) compared to Table 5.11 (Daily Searches). When compared to the Daily data (Table 5.11) the Five Day Searches (Table A.13) seem to provide more videos with a higher view count and less with a view count under 1000. There is a substantial difference between both sets of data (Table 5.11 and 5.24) either suggesting that that there has been a change in the way that the API chooses the videos, that doing the searches over a longer period provides a different set of data or that it is random.

5.4 Summary of Main API Data Patterns

The following summarises the findings of the YouTube API in relation to RQ1.

Consecutive Daily Search results from a YouTube API category differ in the degree of overlap between consecutive search results over time and between categories. The number of repeated videos does not follow an obvious pattern. There is also no time synchronisation because when the number of repeated videos increases for one category it does not necessary happen for the other categories. The number of repeats for each of the categories changes daily, suggesting that there are variations either in the API sampling strategy or the underlying data (e.g., new YouTube videos). The level of unpredictability is surprising because, for example, it would be reasonable to expect a similar and high degree of overlap between search results if the API returned the most popular videos and a selection of newer videos with fast increasing popularity.

The From TV category has strong fluctuations in the videos that are repeated. This could be related to traditional broadcast TV schedules. It was not possible to determine whether this is the reason for these fluctuations, however.

Comparing subsequent searches to the initial search, the API sample changes substantially across the categories over time. The total number of changes for the Daily Searches compared to the initial search differs between categories. In some categories the percentage of repeats reduces considerably more than others. In some categories the percentage of repeats is much higher than others compared to the initial search. In addition, there is a substantial difference in the News sample compared to the other categories. This could be due to the relevance of news videos changing more rapidly than other categories (e.g. due to current affairs or global events).

Frequently repeated videos do not dominate the Daily Search results. Two categories have a slightly higher number of videos that are repeated 1 to 3 times across the 30 searches, but the search results are not dominated by a small number of frequently repeated videos. News has no videos repeated more than 15 times, underlining the uniqueness of this category.

The metric averages (comments, likes, dislikes) for videos repeated in every search differ substantially between categories. This shows that the videos that are repeated across all searches are not chosen based on being substantially high or low in relation to a metric value. Nevertheless, as the YouTube API algorithm may vary between categories, there may be category-specific minimum metric values to give a video a higher chance of being selected and repeated. Moreover, since videos repeated daily are in a minority, it seems unlikely that the YouTube API selects videos solely based on a systematic non-random process.

The metric averages (comments, likes, dislikes) for non-repeated videos differ substantially between categories. This again suggests that these videos are not selected from a universal threshold. Nevertheless, the metric averages for each of the categories are lower for non-repeated videos than for the videos that were selected by every search. Thus, either the metrics play some role in selecting videos, the videos attract higher metrics through being more frequently selected, or another factor (e.g., publication by a successful YouTube channel) makes a video more likely to be both repeatedly selected by the YouTube API and more likely to be popular.

The videos in the YouTube API search sample are usually 11-200 days old. The algorithm therefore does not focus on new videos, which might be expected. Except for two categories (which had results of 0.1% each) the API did not select any videos younger than 11 days old. It is possible that there is a ten day indexing delay between creating a video and it entering the YouTube API. It also selected few videos aged over 200 days and only for a few categories. The only category that provided a wider spread of video ages was From TV. The YouTube API might therefore take a different approach for From TV videos. It might only choose videos within an age band for some categories or this may only be a side-effect of its selection algorithm.

There are substantial average video length differences between categories. This suggests that the search results are not determined video length. Comedy had a substantially higher proportion of longer videos, however. This could be due to Comedy videos tending to be longer or the algorithm selecting longer videos for this category.

The average views of the videos differ substantially between categories. The API therefore does not appear to focus on videos with a certain view count withaverage views naturally varying between categories for all YouTube videos (Arthurs et al., 2018).

The Daily Search data reflects a single time period and the YouTube algorithm may have changed over time. Moreover, its algorithm may not also apply to other YouTube searches, such as keyword queries. Nevertheless, for the category searches reported over 30 days the YouTube API is not dominated by any of the factors or metrics collected (e.g. likes, dislikes, comments, video length, or view count). Nevertheless, since videos that were repeated within every search had higher metric averages than non-repeated videos, it is possible that the factors play some role. For example, a proportion of videos might be selected by a method employing category-specific metric thresholds and the remainder might be selected at random or by a different mechanism. This proportion may vary between categories.

5.5 Conclusions about the YouTube Category Search API (RQ1)

The following observations demonstrate conclusively that the sample returned by YouTube API category results is not a random selection of ***all*** category videos and therefore cannot be used by researchers as a random (probabilistic) sample. YouTube API category searches:

- Rarely return just created videos: few are under 11 days old.
- Rarely return videos over a year old.
- Return the same videos in subsequent searches too often to be accounted for by chance.

YouTube API category searches also do not give a random (probabilistic) sample of ***new*** videos because the results include some older videos. The fact that it is not a random sample of these two types (all videos or new videos) is an important limitation for those using it to investigate YouTube and for the next chapter of this thesis.

Since the YouTube API category searches are not random, the next logical issue is to understand the factors that influence the selection of videos. It is also not a simple algorithm based on thresholds in any of the parameters reported above (e.g., a random sample of videos with 10000 likes). Nevertheless, the parameters can be deduced to some extent, although some speculation is needed because some properties are consistent with either (a) the number of new videos to choose from in a category or (b) YouTube algorithmically trying to select videos that users might enjoy. YouTube API category searches results are influenced by:

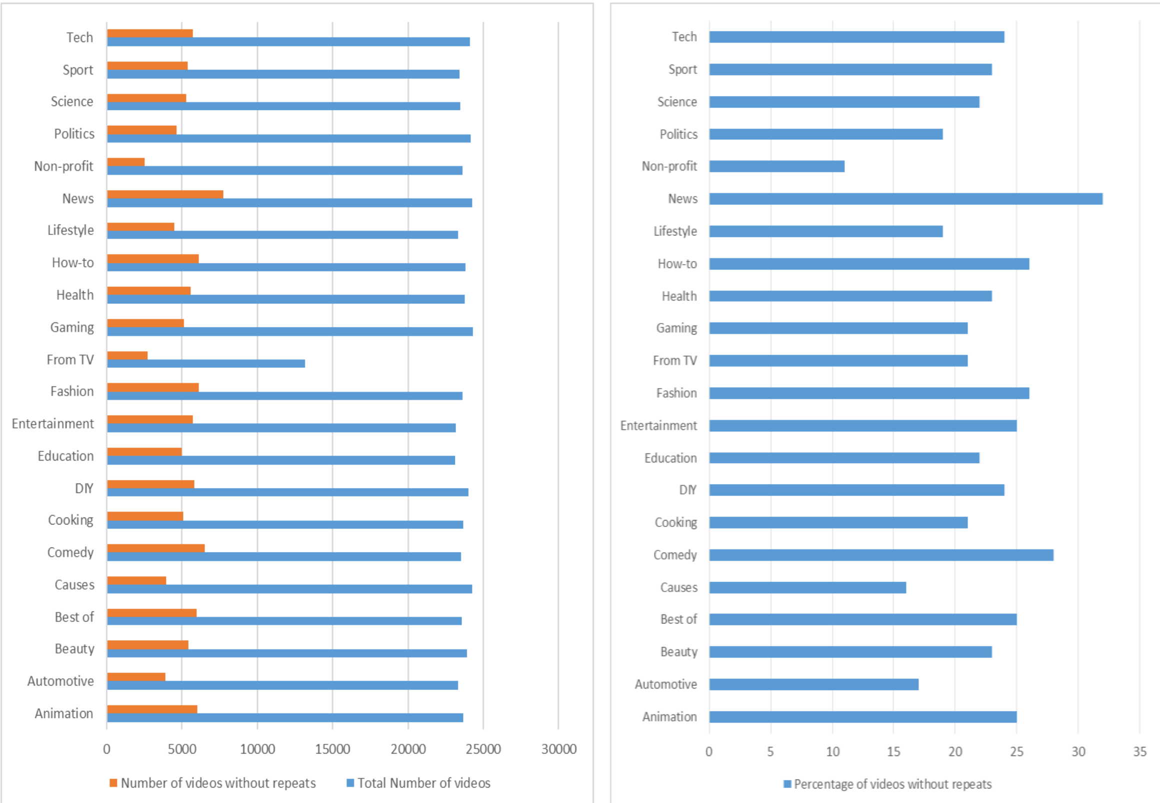
- **Video age**: Most videos are under a year old but almost all are over 10 days old.
- **Video popularity**: There are too many popular videos to be accounted for by random selection, so popular videos are more likely to be selected. Nevertheless, the results include unpopular videos (less than ten views, no likes, no comments) so popularity is not a requirement. YouTube's overall most popular videos are also largely absent, so extreme popularity is insufficient for inclusion. The algorithm either selects a proportion of videos at random (e.g., to give a mix of popular and new videos) or this is a side-effect of its algorithm.
- **The previous day's category search results**: There are far more repeats than could be due to chance. The percentage of repeats changed between the Five-day and Daily searches, suggesting an algorithmic change or a change in videos or watching habits (assuming these are inputs to the algorithm).
- **The nature of the category**: YouTube API category searches work substantially differently for the News and From TV categories, especially in terms of repeats, either due to the nature of the content uploaded or variation in the algorithm or parameters between categories. Human intervention or selection is also a possibility.

Combining the results, the algorithm is neither random nor focused only on popular videos. Instead it may have a mixed strategy, selecting videos partly based on popularity and partly at random, but always taking age, the nature of the category, and the previous day's results into account.

6 YouTube API Category Search Video Properties

Chapters 6 and 7 analyse the videos returned by the searches combined into one large set for each category, addressing **RQ2, RQ3 and RQ4**. Whilst the previous chapter demonstrates that the API sample is not random, it also shows that it contains videos with a wide variety of metric scores so it is neither dominated by new nor by highly popular videos. Although it is a convenience sample and its videos are not representative of YouTube as a whole – because of the almost complete omission of videos over 200 days old for the data collected from the Daily Searches and the oversampling of popular videos – it can illustrate the nature of videos in the categories. This is relevant because the category search results are part of the YouTube user experience and because differences between categories may give insights into wider differences within YouTube. Thus, the results provide some, but not conclusive, evidence about videos in the categories. Chapters 6 and 7 discuss the YouTube category search results at face value, and the above sampling limitations are returned to in the conclusions.

Except for From TV (13,183), there are at least 23,000 videos in each category for analysis (Figure 6.1). The Non-profit, Causes, Automotive, Lifestyle and Politics categories have the fewest non-repeated videos (Figures 6.1 and 6.2). This might be due to the way that the YouTube API category search algorithm works but it seems more likely that there are fewer of these type of videos within YouTube.



Figures 6.1, 6.2. **The total number of videos with information extracted from YouTube with and without repeats, and the percentage of videos without repeats for each category.**

6.1 Video Age (Days)

There are substantial differences between categories in the average age of videos, with From TV videos being over seven times older than News videos (Figure 6.3). From TV videos may last longer because they contain high quality professional content. As a result, YouTube in this case is probably being used as a form of video repository. YouTube is therefore meeting the needs of those who have grown up within an instant access culture and are the 'on demand' generation of video watchers.

News and Gaming have newer videos, this will be due to the nature of these topics and the need for them to be updated regularly to ensure viewers continue to access the content of these topics (Figure 6.3) (Lin et al., 2019; Arthurs et al., 2018; May, 2018; Zeiler, 2018). News and current affairs videos will be uploaded throughout each day to reflect the new issues presented by the global media (Al-Rawi, 2019). As a result, the relevance of typical news-related videos quickly diminishes and further views become less likely. Similarly, due to video games being released and updated regularly, videos relating to these will also be uploaded regularly and can become obsolete quickly (Lin et al., 2019; May, 2018; Zeiler, 2018). This may explain why News and Gaming have a higher turnover of videos than the other categories and fewer older videos.

Categories which have content that is changing or needs updating more regularly will have fewer old videos. Popular categories may also have a higher turnover and therefore additional newer videos but this is not reflected in the data since these categories are not dominated by unusually young videos (Figure 6.3). Categories with older videos will probably have content that needs less updating and these may have more repeated viewing value and appeal.

There is a substantial difference between ages of How-to and DIY videos (Figure 6.3). Since these are relatively similar categories the DIY videos may have more up-to-date content. DIY videos might relate to more contemporary methods and techniques, and ways of fixing items around the home which might need more updating as technologies and styles change and develop. How-to videos might instead relate more to skills, techniques and topics that need less updating such as drawing, sewing or knitting. Cooking also has a high average number of older videos, which could be due to people repeatedly viewing a favourite recipe or cooking technique. As a result, there is probably less need for new cooking videos.

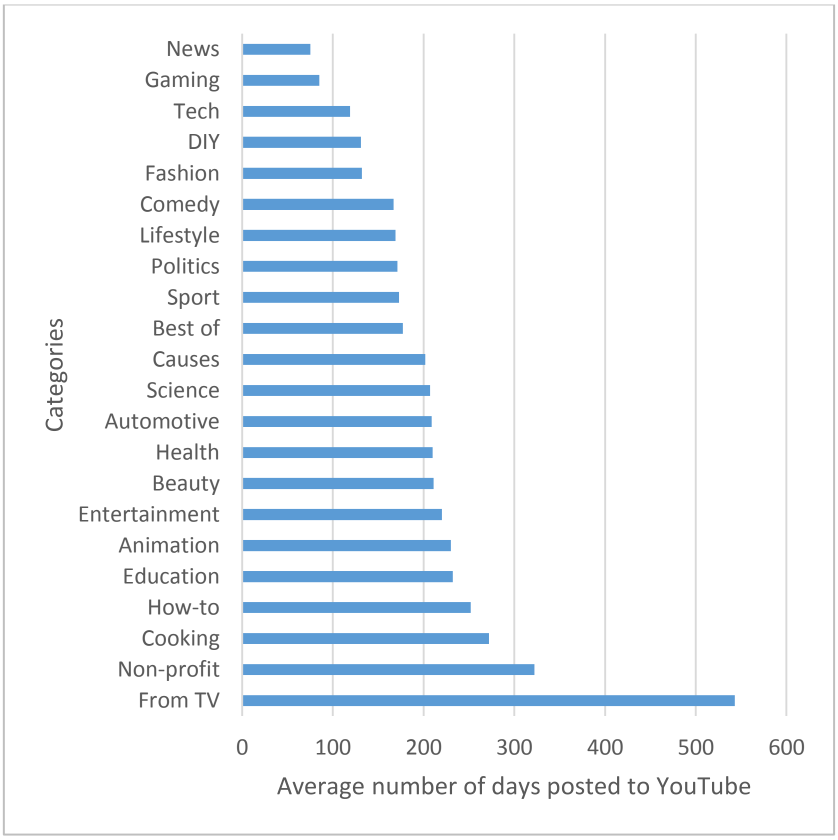


Figure 6.3. **The average video age (days posted to YouTube) by category.**

Most of the videos across the categories within the sample had been posted to YouTube for between 1 to 180 days (Figure 6.4). From TV has a greater spread of percentages across the bands of days, suggesting that less of these types of video are posted regularly and may reflect when programmes are broadcast by more traditional means. This category has a higher proportion of older videos that are probably high-quality professional content which individuals are accessing on an on demand basis (Zannettou et al., 2018).

News has the highest percentage (87%) of videos aged 1 to 60-days, which would be due to the changing nature of the news (Figure 6.4) (Al-Rawi, 2019). This could also be the case for Politics, Tech, Fashion and Gaming as 69-70% of the videos within the sample for these categories are within the 1 to 60-day banding and again could reflect the changing nature of these topics.

Approximately 60% of the Non-profit videos have been posted for more than 60 days which suggests a lack of regular uploads or turnover in the content or material within this category (Figure 6.4). In How-to, 50% of the videos within this category have been available for more than 60 days, again suggesting a long shelf-life for useful videos (Figure 6.4).

Overall, the categories that have a higher proportion of more recent videos might be those where the content needs to be more up-to-date. It also suggests that these categories are more popular and therefore there is a need for new content regularly. Videos that have been uploaded more recently are more popular. In those categories with a higher proportion of older videos, it seems that this content does not need updating regularly, that it does not need to be new or contemporary, or that there is just less activity from viewers in relation to these types of video.

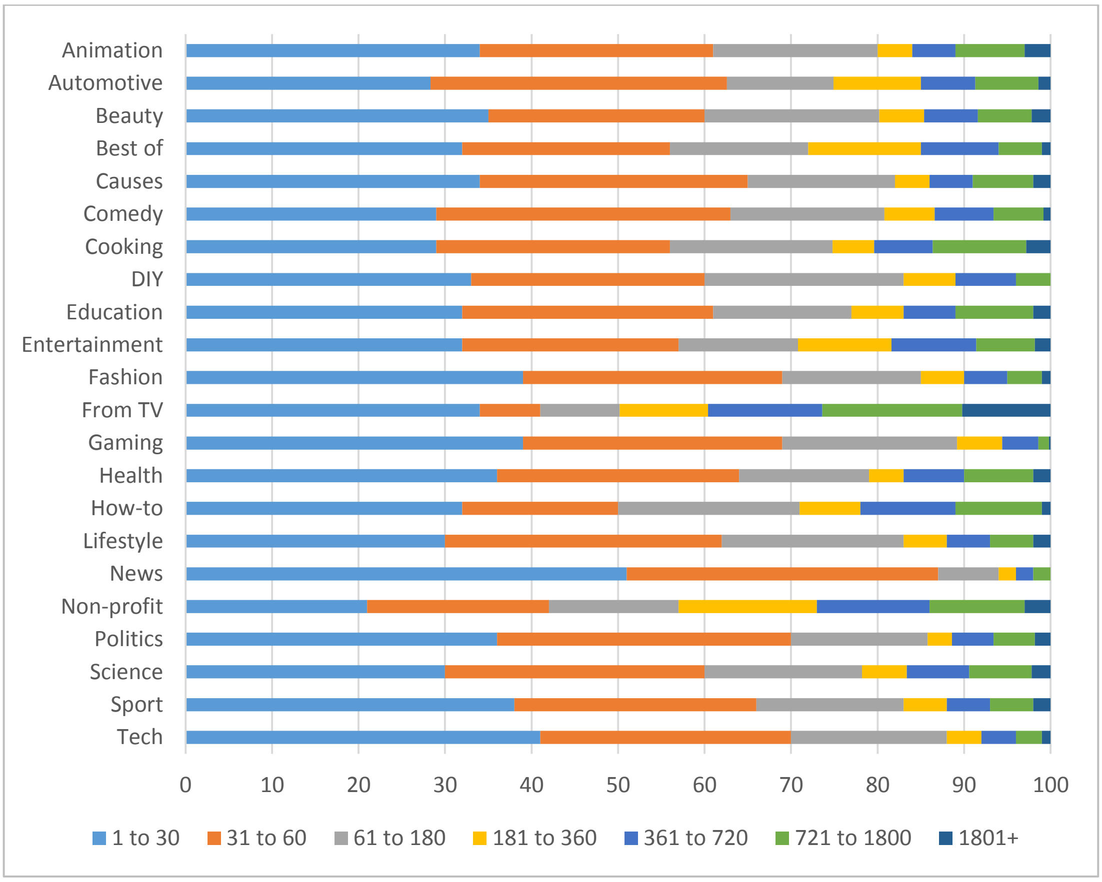


Figure 6.4. **The percentage of days within bands that the videos within the sample had been posted on YouTube.**

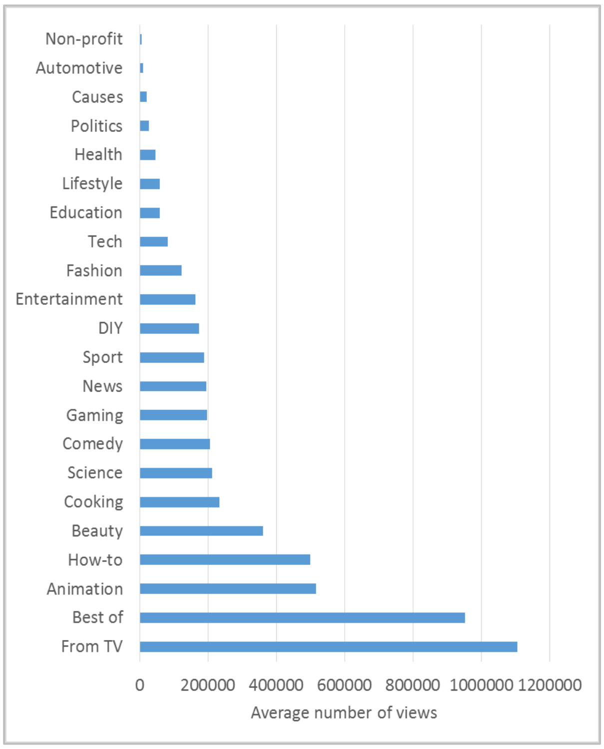
6.2 Views

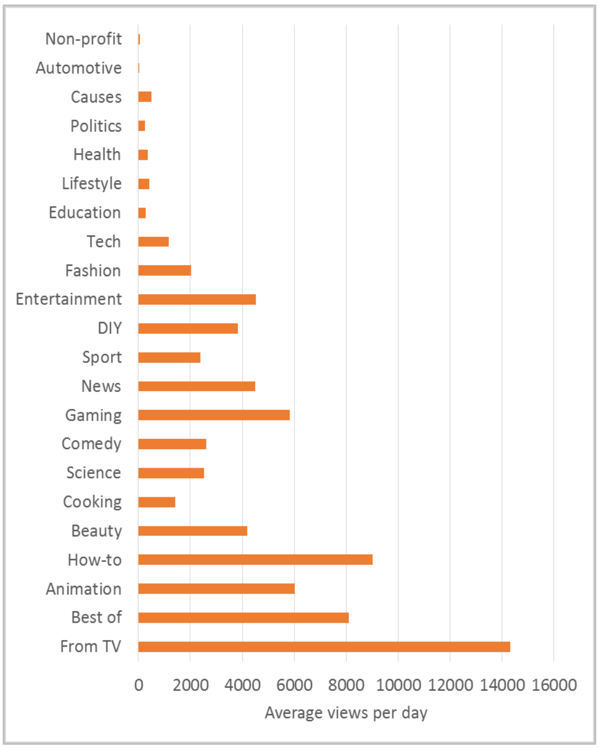
There are substantial differences in the average popularity of videos between categories, however this is measured. Both average daily and total views are reported, although neither is a perfect measure of popularity (Figures 6.5, 6.6 and 6.7). Total views favour older videos that have had longer to attract viewers than younger videos. In contrast, average daily views favour newer videos, assuming that a video attracts most of its viewers when relatively young.

From TV videos are viewed most (Figure 6.5), by people using YouTube as an on-demand forum for traditional TV content (Zannettou et al., 2018). This confirms the popularity of the From TV category. From TV has the highest proportion of older videos (Figure 6.3) and has an unfair advantage for total views because its videos have had longer to accumulate views, but its videos also attract the most views per day. Best of (unsurprisingly), Animation and How-to also have a relatively high number of average overall views and views per day (Figures 6.5 and 6.6).

Gaming, News and DIY have relatively more daily views than total views (Figure 6.6). As previously suggested, Gaming and News content tends to change regularly. These videos may be popular for a short period, becoming out of date quickly (Figure 6.3). DIY videos may be accessed daily but become out of date quickly in the advice and support they offer, especially if offering season-related advice (e.g., gardening, winter protection). It is also clear that there is a higher turnover of newer videos, again suggesting that content becomes out of date relatively quickly (Figure 6.3). The Entertainment category has a much higher average daily watch placement than overall views, suggesting that these videos are watched more when they are newer, but have less long term appeal. Although Cooking has many views it has a few daily views – this suggests that these videos are popular but are probably revisited over longer periods of time (Figures 6.5 and 6.6). This is also supported by Cooking having more older videos (Figure 6.3).

Non-profit, Automotive, Causes, Politics, Health, Lifestyle and Education have lower numbers of average overall and daily views (Figures 6.5 and 6.6). These seven categories are also the lowest in emotional connection, interaction and engagement considering the dislike and like metrics (Table A.14 – Appendix 11). Automotive and Non-profit are particularly low both in overall and daily views (Figures 6.5 and 6.6) and dislikes and likes (Table A.14 – Appendix 11).





Figures 6.5, 6.6. **The average views and views per day for each category, in the same order.**

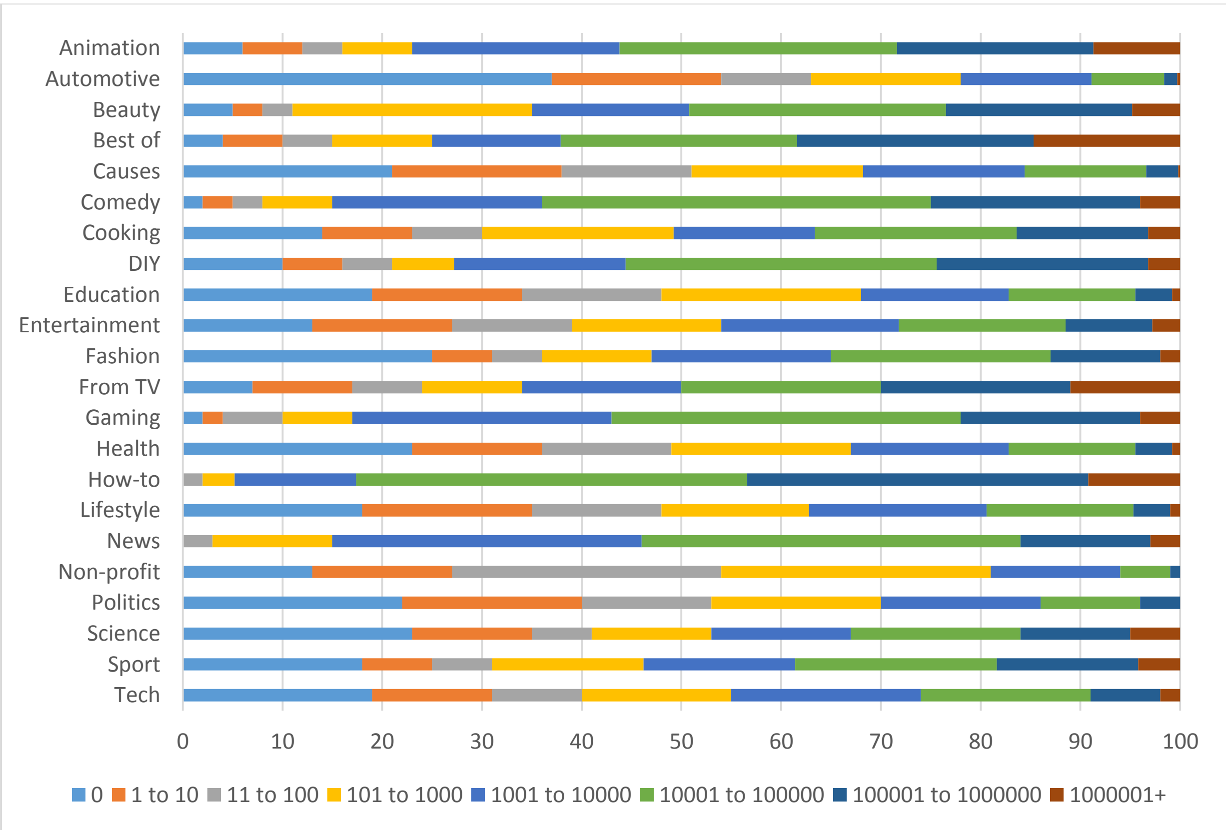


Figure 6.7. **The percentage of views within bands that the videos within the sample have received.**

6.3 Dislikes and Likes

There are substantial differences between categories in the number of likes and dislikes per video, the number per view and the ratio of likes to dislikes (Figures 6.8, 6.9 and 6.10, and Table A.14 –

Appendix 11). The number of likes and dislikes for a video is partly due to the number of viewers and partly due to the reactions of viewers and so the ratio of likes to dislikes is the most informative about the attitude of viewers towards a video. Videos receive considerably more average and likes per view than dislikes across all the categories, suggesting that viewers are more likely to respond to a video when it has a positive impact on them (Figures 6.8 and Table A.14 – Appendix 11).

From TV, Best of, How-to, Beauty and Animation receive the most dislikes (Table A.14 – Appendix 11). This suggests a greater level of connection with these videos to trigger this response. The level of dislikes and dislikes per view that some categories receive, for example Gaming, DIY and Beauty, could be a comment on poor quality of advice, support or usefulness they provide. The wording of the Best of category description in YouTube suggests a higher standard of video, either in content or production quality. People may be using the rating system to show that they disagree with the video being classed as Best of. There is a high level of average overall dislikes for From TV which could firstly relate to the production quality of the video as viewers have probably accessed these videos due to already liking, or being a fan of, the content (Table A.14 – Appendix 11). In addition, the level of dislikes for the From TV category could also relate to the accuracy of the video title, description or thumbnail picture. It is less likely that viewers will be rating the content of the video as these will be copies taken from traditional sources.

How-to, Best of, DIY, From TV, Animation, Gaming and Beauty receive the most likes (Table A.14 – Appendix 11). DIY receives a substantially high average number of likes per view showing that these videos generate a high level of response from viewers. In addition the categories Gaming and How-to also have high levels of average likes per view (Figure 6.8). This suggests that viewers have more positive interaction and connections with the videos within these three categories in relation to:

- Their enjoyment of the content
- The quality of the video production
- How useful it was
- How accurate it was in the title and thumbnail
- The level of positive emotion it generated

Automotive, and Education receive the lowest average number of overall dislikes and likes, and dislikes and likes per view (Figures 6.8 and Table A.14 – Appendix 11). This suggests that there is less interest in these categories of video or that viewers of these categories are more passive in their watching habits.

Comedy and associated themes are consistently established as being popular and based on the type of content and material in these types of videos it seems strange that this category does not generate more substantial levels of either positive or negative emotion, interaction and connection.

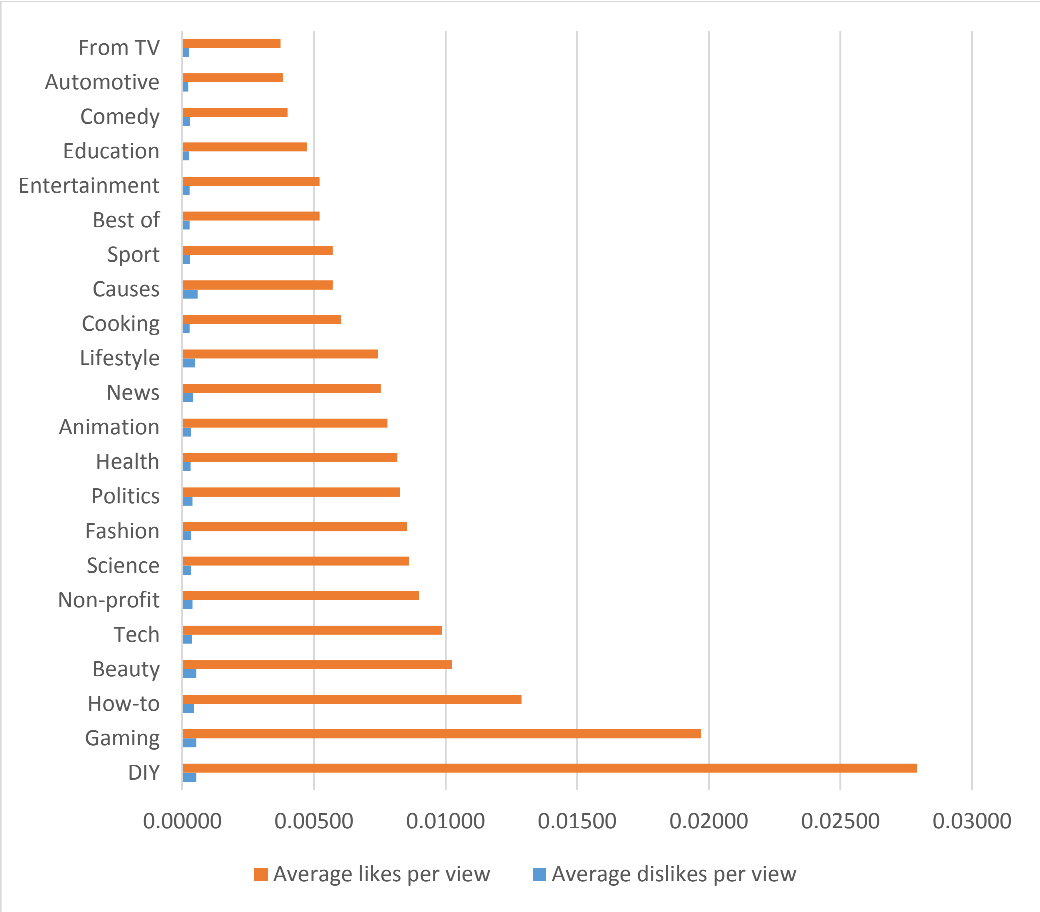


Figure 6.8. **The average number of dislikes and likes per view for each of the categories.**

All categories receive substantially more likes than dislikes (Figure 6.9). The two categories that receive the highest ratio of likes to dislike are DIY and Gaming, but these are not the most popular videos based on views or views per day (Figures 6.5 and 6.6). Tech and Health are also categories that are not particularly popular based on views but also receive a higher ratio of likes to dislike (Figure 6.9). Considering the ratio of likes to dislike has had a particular impact in terms of the positivity rating of Causes, which is now the lowest in the graph (Figure 6.9) Therefore, this data suggests that it is not necessarily those video that are watched more frequently that get the higher ratio of likes to dislike. It also shows that just because videos are watched to a greater extent does not mean that they will be rated. Categories that provide support, help and guidance get the highest like ratio, which could represent appreciation for the creator, an acknowledgement of usefulness or to help other searching for the same thing (Figure 6.9). Categories that are more subject-based receive the lower ratios of likes to dislike. This could be because some these types of video are more entertainment-based and do not necessarily lend themselves to being rated. Overall, liking and disliking videos seems to be a statement of usefulness rather than entertainment.

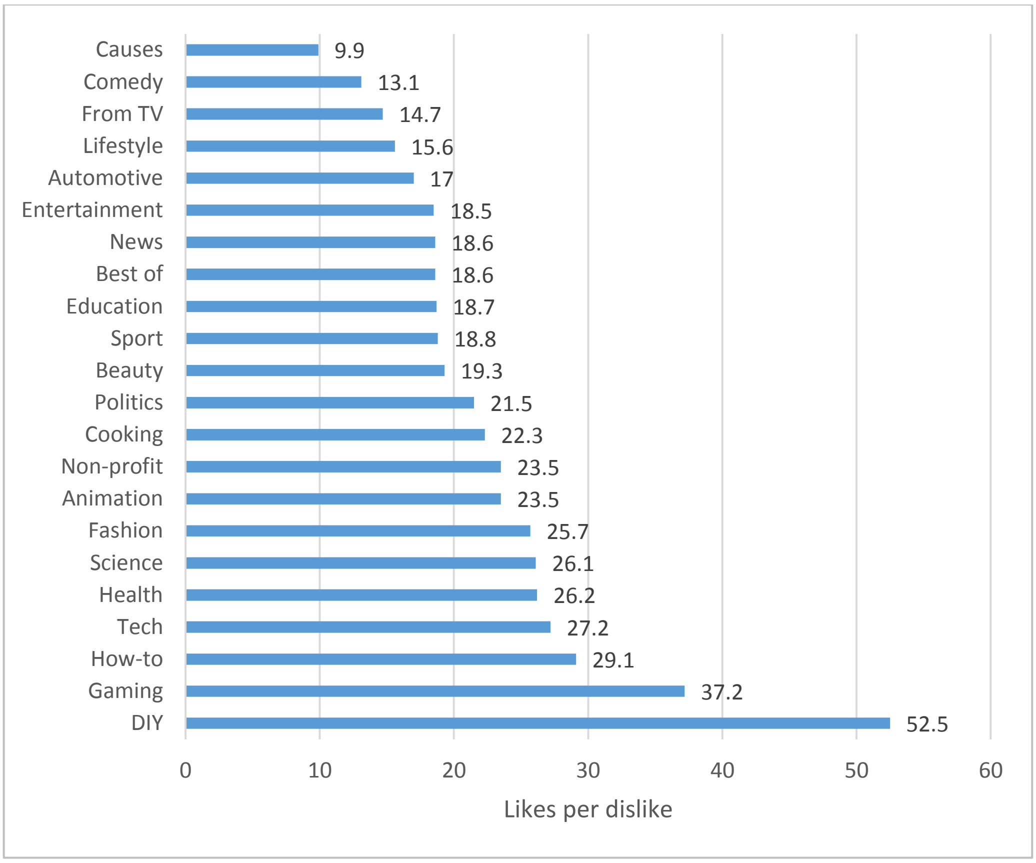


Figure 6.9. **The average likes per dislike for each of the categories.**

Automotive and Non-profit clearly receive a low percentage of dislikes, with 85% of both these categories receiving no dislikes at all and have no videos with more than 100 dislikes (Figure 6.10). This suggests that the videos:

- Content is exactly what the viewers were expecting and reflects the title and thumbnail
- Are of an expected quality in how they have been produced
- Content does not establish a connection with or invoke an emotional response that would encourage someone to click the dislike icon
- Have a low level of interaction

This could also relate to Education, Causes and Health which also have a high percentage of videos within the no dislikes banding (Figure 6.10).

The categories receive fewer dislikes with most being between 1 and 100 (Figure 6.10). The main categories that seem to have a more substantial number of dislikes, within the 11 to 100 banding, are:

- Comedy: this could be a statement relating to the content of the videos with the viewer confirming whether they found it funny or that they were offended by some of the material used;
- DIY: which could relate to the usefulness or accuracy of the information, advice, support or demonstrations provided;
- Gaming: these videos may contain advice and support relating to playing computer or video games and might be a negative statement highlighting the poor quality or inaccuracy of the information provided;
- How-to: again this might be a statement relating to the quality and accuracy of the information, advice and support provided by the video;

- News: this suggests a negative emotional reaction to the content of the news story being covered or the perspective presented in these videos.

How-to, News, Gaming, Comedy and Best of have the lowest average percentage of videos with no dislikes, with 26% or less, which suggests that these videos invoke a greater level of negative emotional connection (Figure 6.10). How-to, Best of, Animation, Gaming and DIY have the highest percentage of videos that get the most dislikes over 101. This suggests a greater level of connection with these videos, even though it is negative. The content of some videos may not encourage viewers to rate them, so it is difficult to suggest that these categories are less popular based on this metric alone.

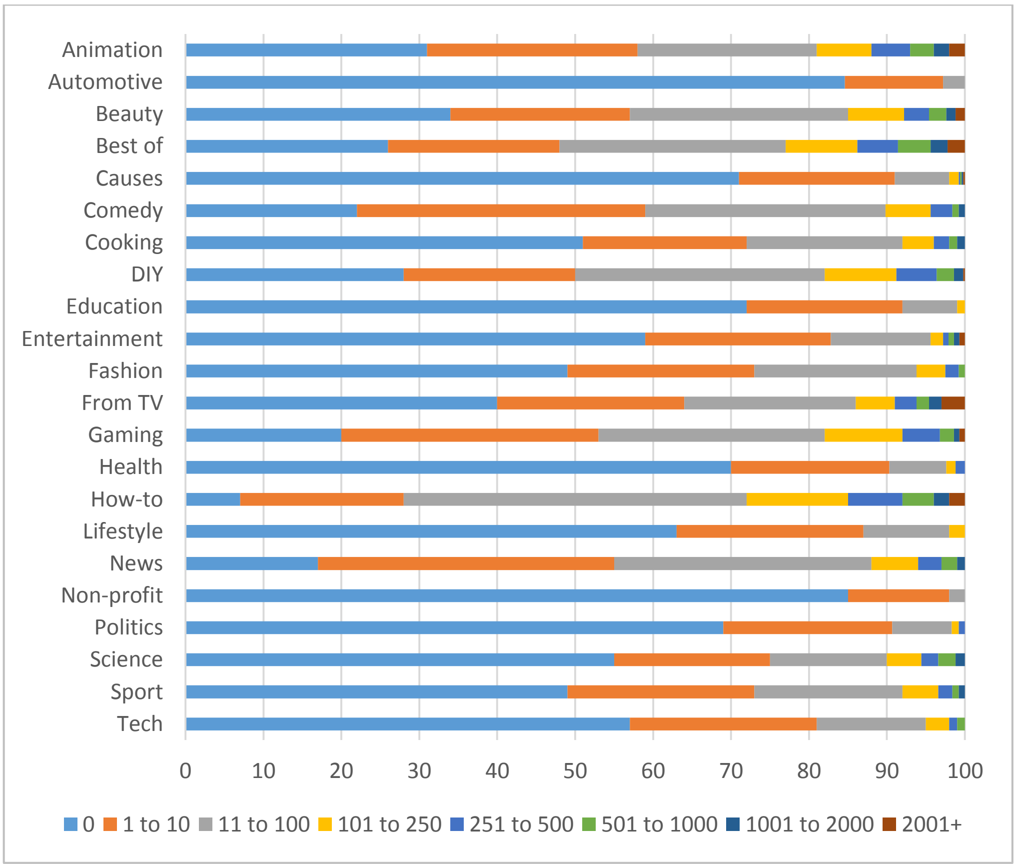


Figure 6.10. **The percentage of dislikes within bands that the videos within the sample have received.**

Automotive, Causes, Non-profit, Politics, Health and Education have a substantial percentage of videos that have received no likes suggesting a lack of positive engagement or ambivalence from viewers (Figure 6.11). The lack of likes could show categories that are less popular with viewers, however, this is difficult to determine from this data. It could just be the case that the type of content in these videos does not require a response or rating from viewers. Whereas How-to, Gaming, News and Comedy have a low percentage of videos with no likes showing a greater level of engagement and connection from people and also suggests that these categories are more popular with users (Figure 6.11).

The categories How-to, DIY, Best of, Gaming and Beauty have a highest percentage of likes within the 2000+ banding (Figure 6.11), suggesting a high level of positive engagement and connection with these type of videos. Considering How-to, DIY, Beauty and, to some extent, Gaming it could be that viewers are positively rating the advice, support and information contained within these videos and how useful it is. It might be that they want to provide recommendations to other users and as a result could have the potential to increase the popularity of these videos. The category Best of, as explained previously, suggests quality in the content and viewers may be reinforcing this by positively rating it (Figure 6.11).

If other users, through the rating system, feel that these videos and their content are of high quality then they are more likely to watch. Poor quality videos, in content and production, can have a negative impact in relation to users choosing to watch or recommend them.

How-to has a particularly high percentage of likes within the 2000+ band which suggest a substantial level of positive engagement and also suggests, in likes, that this category is particularly popular (Figure 6.11). Videos that receive more likes stand a greater chance of being selected by YouTube's various recommendation facilities and system, as a result might be given greater prominence on the website and therefore are viewed more regularly and gain popularity.

Overall, there is a higher percentage of likes than dislikes suggesting that people are more likely to rate videos when they have had a positive watching experience (Figures 6.10 and 6.11).

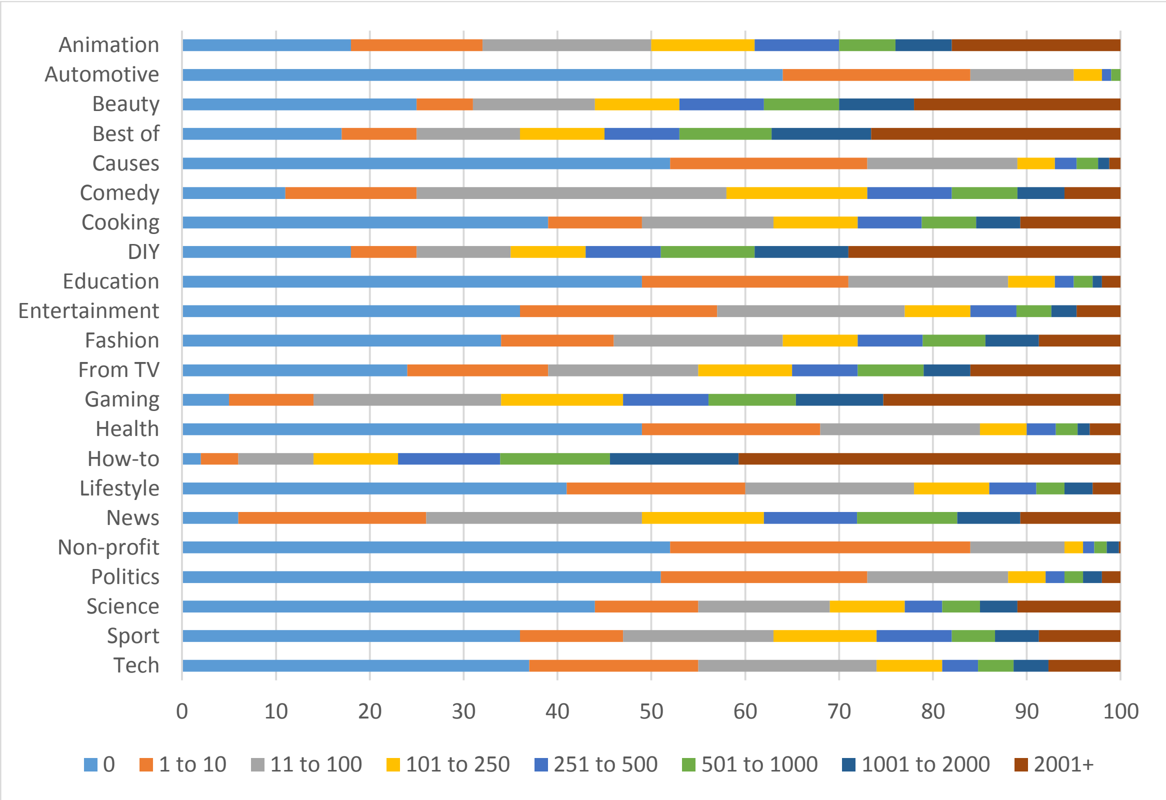


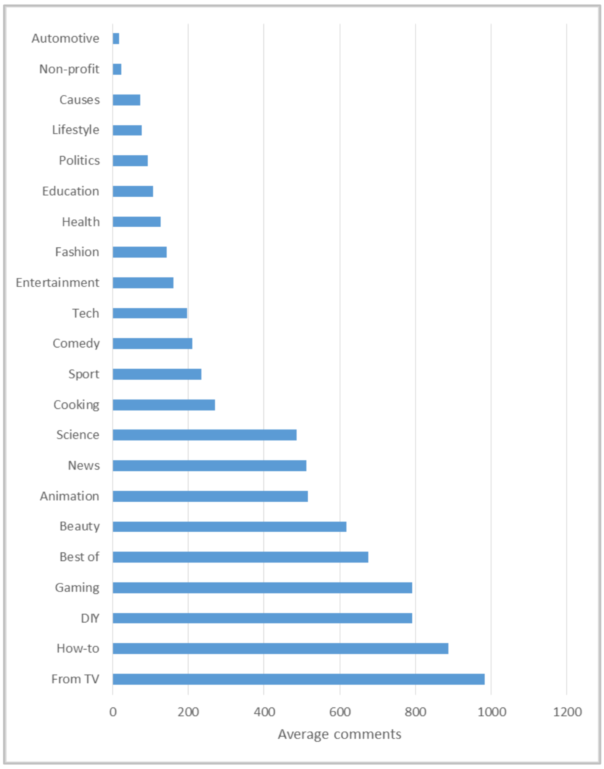
Figure 6.11. **The percentage of likes within bands that the videos within the sample have received.**

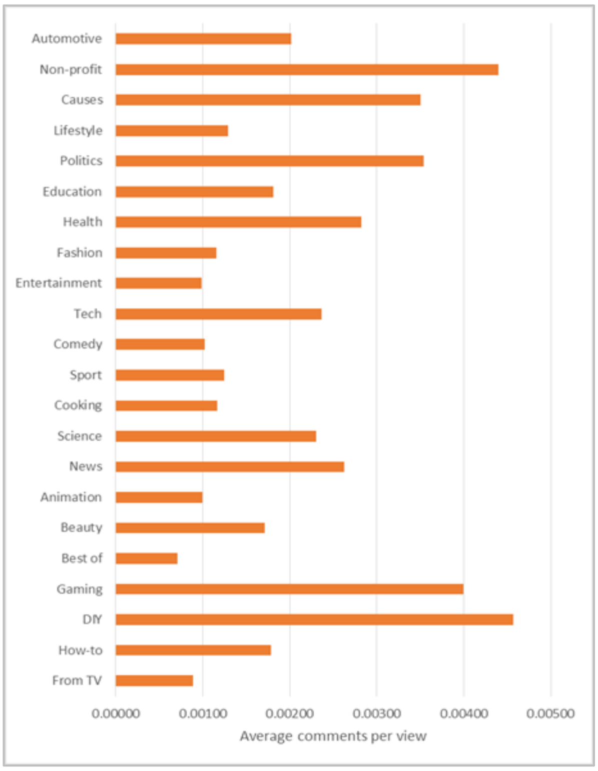
6.4 Comments

There are substantial differences between categories in the number of comments per view. Taking the time to write a comment relating to a video, or responding to comments, takes more effort than clicking the dislike or like button, and therefore demonstrates a stronger level of emotional connection, interaction or engagement. Comments can also be used as a means of interacting or connecting with other people who have responded to the video and its content, as these discussions develop it is possible that they will no longer relate to the original topic and become more social in nature. Although the comments may initially relate to what has been watched in the video they could evolve into other topics of discussion or social interaction. It is also possible that comments can be quite controversial, angry, antagonistic, personal, aggressive, anti-social or abusive in nature and can lead to arguments or inappropriate interactions that can sometimes move away from the video's content or material.

Differences in average comments that each of the categories receive partly reflects the popularity of these different types of video and also the likelihood of viewers commenting (Figure 6.12). Nevertheless, due to the differences in terms of the views and daily views of each of the categories (Figures 6.5 and 6.6) it is difficult to draw conclusions from the average comment data alone (Figure 6.12). Comments per view provide a clearer picture in terms of categories that viewers interact with (Figure 6.13). The categories with the highest amount of average comments per view are DIY, Non-profit, Gaming, Politics and Causes which suggests that viewers are more likely to interact with and discuss the content within these types of videos. DIY and Gaming both have a lot of average comments and also has a substantially high level of comments per views, showing that these types of video garner substantial interaction from viewers (Figure 6.13). Although From TV has the greatest amount of overall average comments (Figure 6.12), it has a considerably low number of average comments per view (Figure 6.13). This suggests that viewers are interacting to lesser extent with these types of video based on the number of views. As previously demonstrated From TV has the highest number of average views and views per day (Figures 6.5 and 6.6) therefore although substantial levels of viewers are watching these types of video the content does not encourage users to make substantial levels of comments and discussion. This further supports the idea that YouTube is being used more as a free on demand service to access traditional TV content rather than user produced videos. How to and Best of also have high levels of average comments, but again receive a sustainably lower number of comments per view. It could be due to the high levels of views and daily views that some of these categories receive (Figure 6.5 and 6.6) that has had a huge impact on the comments per view that they receive, however, it does suggest that viewers are less likely to interact with these types of video.

Non-profit has a low number of average comments (and average and daily views – Figures 6.5 and 6.6), but has a high number of comments per view, suggesting that although watched less than some other categories it generates more discussion from viewers. There are several further categories, Automotive, Causes, Politics, Health, Lifestyle and Education, that have substantially low numbers of average views and daily views (Figures 6.5 and 6.6) but have substantially higher levels of comments per view (Figure 6.13). This also suggests that although these types of video are not watched much by viewers they do provide content that users want to interactive with, comment on and discuss. Overall, categories that are viewed more regularly do not necessarily garner more user interactions.





Figures 6.12, 6.13. **The average number of comments and comments per view for each of the categories.**

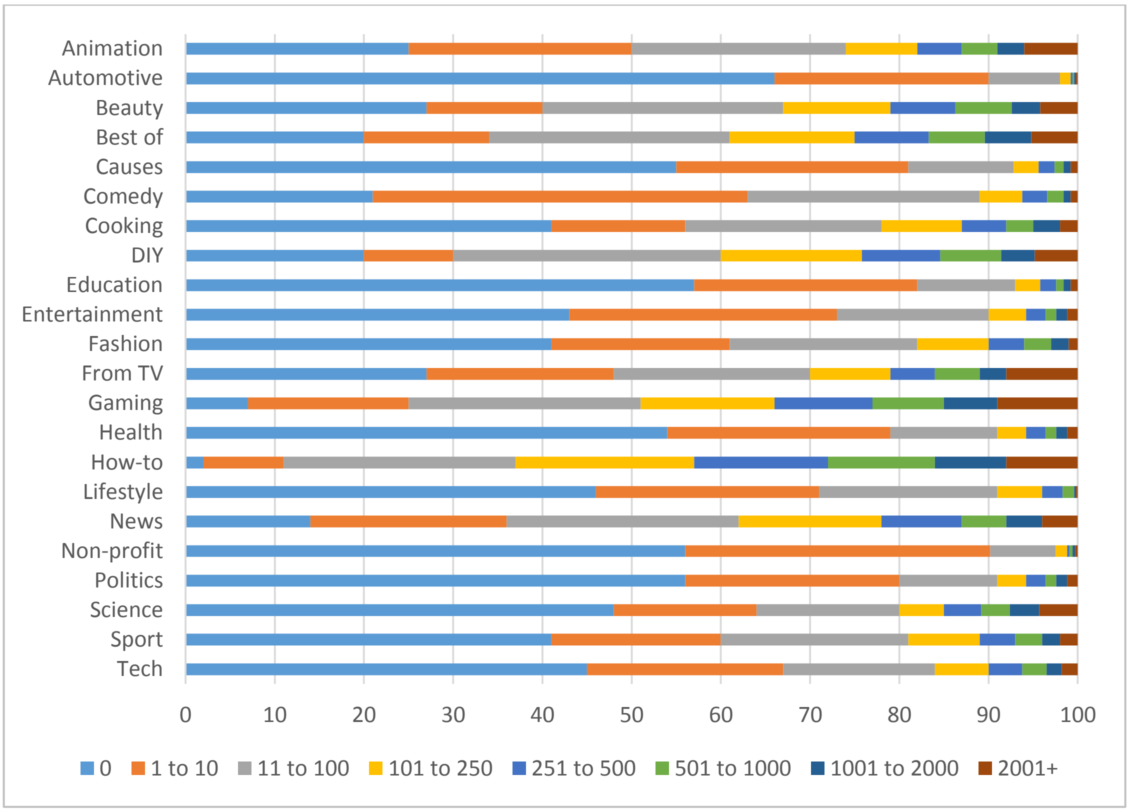


Figure 6.14. **The percentage of comments within bands that the videos within the sample have received.**

6.5 Length

There are substantial differences in the average lengths of videos across the categories with these ranging from 5 minutes to 54 minutes (Figure 6.15). The categories that have the longest average videos are Comedy (54 mins), Science (46 mins), Best of (41 mins), Animation (37 mins) and Gaming (26 mins) (Figure 6.15). The categories that have the shortest average videos are Automotive (5 mins), DIY (6 mins), Fashion (8 mins), Cooking (9 mins) and Lifestyle (9 mins). Comedy videos are on average the longest with an average length of 54 minutes which could relate to the length of stand-up comedy shows, either for the whole show or approximately 1 hour either side of the interval.

Videos that focus on providing advice, support, help, demonstrations of skills, techniques or methods, or DIY type activities might be shorter so that they are easier for someone to access, follow and digest. Videos that are accessed for more entertainment-based purposes might be longer. Nevertheless, most traditional TV programmes are generally either 25 to 30 minutes, 40 to 45 minutes or 60 minutes in length, but this does not match with the average length of From TV and Entertainment videos. Video length will be analysed further where the data has been banded to provide a clearer picture for each category (Figure 6.16). When compared to the view count data, there is not a trend for the average length of videos in a category to associate with the average number of views for that category (Figures 6.5 and 6.6).

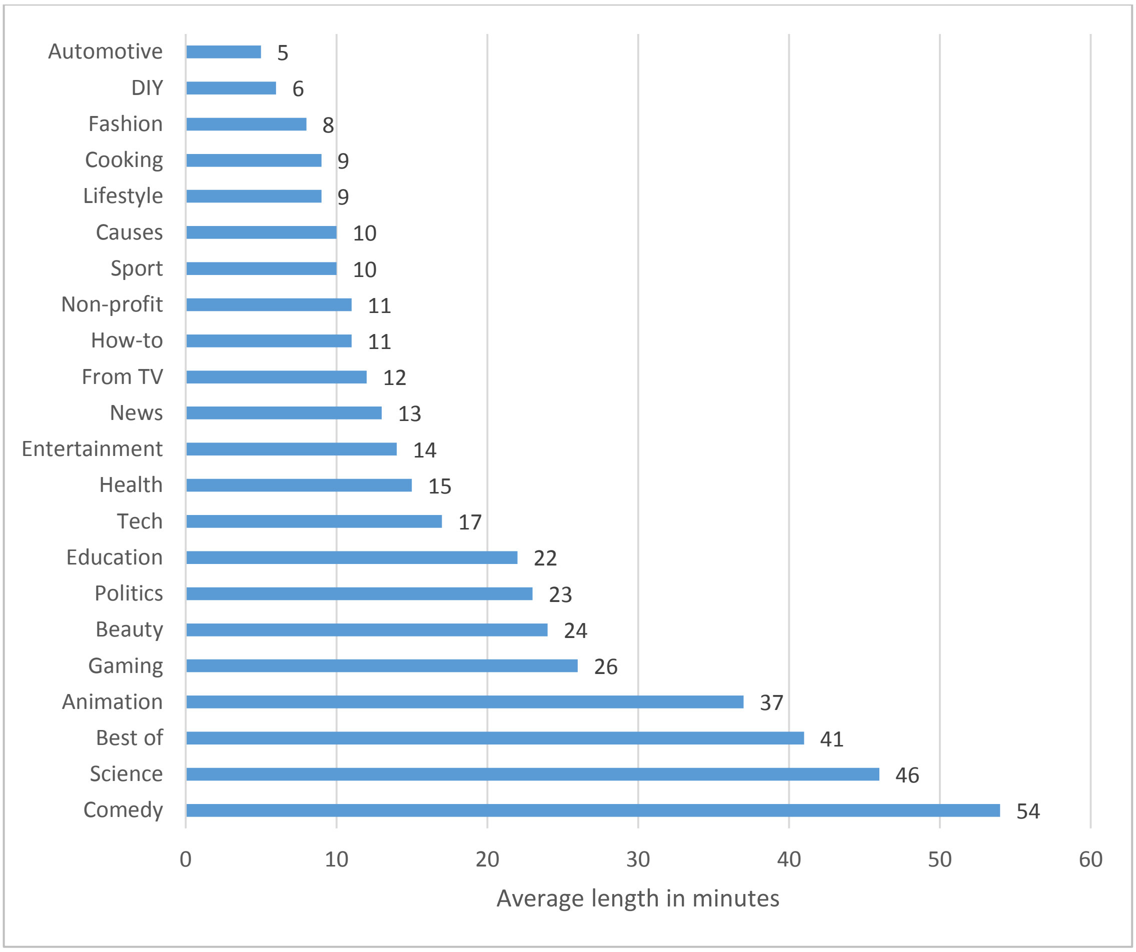


Figure 6.15. **The average length of video (minutes) for each of the categories.**

Entertainment, From TV and Sport have 1% of videos that have no length, therefore these must be just a picture, probably being used as an advertisement, rather than any playable content (Figure 6.16). Most of the videos across the categories have their highest percentage of videos in the 1 to 5 minute (61 to 300 seconds) banding which suggests the following:

- That users prefer producing videos within these lengths or that they feel this is the length of video watchers prefer;
- That this banding is the most popular length of video with users;
- Most video content happens to fit into this length of time.

The exceptions to this are the categories Comedy, Best of, Science and Gaming. With Comedy (42%), Best of (29%) and Science (27%) all the have their highest percentage of video length within the over 1-hour band and Gaming (28%) being in the 5 to 10-minute band (Figure 6.16). This suggests that those producing videos within the Comedy, Best of and Science categories are more focused on the content rather than the shorter runtimes which are generally considered to be more popular. The possible advice-based nature of Gaming content could support the finding that videos of an educational nature should be under 15 minutes to have an impact. News, Lifestyle, How-to, Fashion, Cooking and Automotive only have 2% of their videos that are an hour or longer (Figure 6.16) and this is probably due to the nature of these videos:

- News – may be done in small reports either covering one story or just the key headlines
- Lifestyle and How-to – shorter videos of information, instruction and advice
- Fashion – advertisements, reviews, vlogs (video logs) and instruction

- Cooking – length of the recipe being demonstrated
- Automotive – short advertisements, reviews or car related How-to videos

DIY does not have any videos that are an hour or more in length which could also relate to the content within these types of video (Figure 6.16). People probably do not want to watch advice or instructional videos that are too long and supports the finding that these types of video should ideally be 15 minutes or less. To support this further 77% of DIY videos are within the 1 to 10-minute bands (Figure 6.16). Any DIY videos that are longer in nature are probably broken up into easier to digest chunks and presented in several parts. Automotive, Fashion, Sport, Causes and Health have the highest percentage of shortest videos falling within the 1 second and 60 second bandings. These could be short advertisement or information videos for each of these categories.

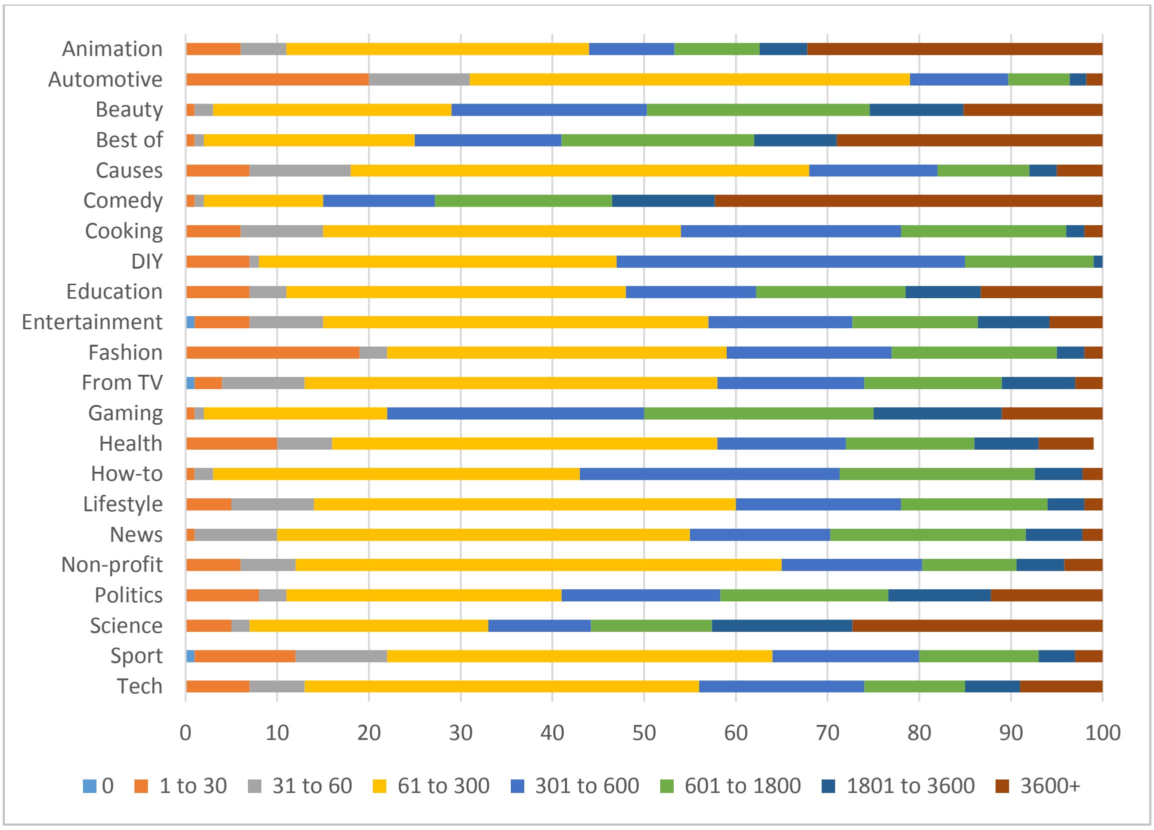


Figure 6.16. **The percentage of videos within each category that fall into the different length bandings.**

6.6 Summary

The following summarises the key findings about the YouTube category search videos extracted from their metadata in relation to RQ2, RQ3 and RQ4.

- ***There are substantial differences in the average ages of videos between categories.*** For example, From TV videos are over 7 times older than News videos. In general, categories with videos that may have longer term value tend to return older videos. This seems to be primarily due to the retention of relatively old videos by the category search because all categories return reasonably similar proportions of younger videos.
- **There are substantial differences in the average popularity (views) of videos between categories.** For example, the average popularity of From TV videos is many orders of magnitude higher than that of Non-profit videos. Whilst entertainment videos seem to be the most popular, on average, informational videos (How To, Beauty, Cooking) are moderately popular. Political videos (politics, causes) are perhaps surprisingly unpopular overall.

- **There are substantial differences between categories in the number of likes and dislikes per video, the number per view and the ratio of likes to dislikes**. For example, DIY gets 7 times as many likes per view as From TV. Relatively high levels of likes suggests more active use of the videos, rather than passive entertainment. For dislikes, the ratios are much smaller, with Causes getting twice as many dislikes per view as From TV. The ratio of likes to dislikes varies more substantially, with DIY attracting five times more likes per dislike than Causes. A high number of dislikes or a high ratio of dislikes to likes suggests a controversial topic or one with varied opinions in society. The converse suggests a useful topic (Rubenking, 2019; Feitosa and Botelho, 2017).
- **There are substantial differences between categories in the number of comments per view.** For example, DIY videos receive seven times as many comments per view as Best Of. Helpful videos seem to generate the most comments and passive entertainment the least. Controversial topics (e.g., Non-profit, Causes) seem to generate many comments, due to discussions (Hussain et al., 2018; Shapiro and Park, 2018).
- **There are substantial differences in the average lengths of videos between categories.** For example, Comedy videos average about an hour but automotive typically last for 5 minutes. It is not clear whether the algorithm selects videos differently by length between categories, or if different lengths are most popular within different categories.

From these results the category search videos differ greatly in all respects in average properties between categories. Except for the length results, the differences seem logical in terms of reflecting the different types of video in a plausible way. The length results, in contrast, suggest that the algorithm selects directly or indirectly for length differently between categories.

7 Factors associating with YouTube Video Popularity

This short chapter addresses RQ5: How does the length, like count, dislike count and comment count of a YouTube video relate to its popularity?

7.1 Factors Correlating with View Counts (age bands)

The popularity (views) of videos significantly and positively correlates with likes, dislikes, and comments but not length across all categories, even when comparing similar-aged videos (Table 7.1). The **lack of a correlation between view counts and length** is surprising, given the previous research suggesting a preferred video length and the possibility that specific lengths would be most popular in some categories (e.g., long comedy videos in Entertainment and short pop videos in music). It seems logical that the more a video is watched, the more dislikes, likes, or comments it would receive, so these positive correlations are expected.

Relatively strong positive correlations between view counts and the three interaction metrics (likes, dislikes, comments) suggests that **there are few category differences in the rate at which people interact with videos from a category**.

High correlations between likes/dislikes/comments and view counts

- **Animation, From TV** – The high correlations suggest that few animations or TV shows are niche (loved by small audiences) or overrated (widely viewed by an unappreciative audience).
- **Best of** – Since this is a multi-genre category, the uniform relationship between popularity and the interaction metrics suggest a degree of uniformity in selection. Thus, this category probably does not include many examples of videos from low correlation categories.
- **How-to** and **DIY** – Home help videos are viewed to the extent that they are useful (Rubenking, 2019; Feitosa and Botelho, 2017) and serve popular tasks. The high correlation suggests the absence of genre-crossing videos in this category, for example, such as comedy DIY videos or non-profit DIY videos (e.g., for recycling).

Relatively weak positive correlations between view counts and the interaction metrics suggests that people interact at different rates with some videos than others within the category.

Low correlations between likes/dislikes/comments and view counts

- **Non-profit** – Videos from high profile organisations (e.g., Greenpeace) might attract casual viewers that do not interact or trolls that interact disruptively (comments, dislikes). In contrast, viewers of the outputs of smaller, more personal, non-profits might feel more induced to interact with them, particularly through likes and comments.
- **Automotive** – This category may contains multiple genres of videos that attract different types and rates of interaction. For example, fast car videos might attract relatively strong interactions whereas sales videos might generate little activity.
- **Education** – The rate at which people interact with an educational video might depend on whether it is from a professional educational channel that encourages viewer interactions or amateur.

Differences in the strength of correlations between view counts and the interaction metrics suggests that the videos are liked, disliked, or comment on at different rates within the category.

Strength differences in correlations between likes/dislikes/comments and view counts

- **Comedy** – Some Comedy videos were liked and commented on at lower/higher rates than other Comedy videos, even though they are all disliked at similar rates. No plausible reason for this could be conceived.

- **Gaming** – Some Gaming videos were disliked and commented on at lower/higher rates than other Gaming videos, even though they are all liked at similar rates. Gaming videos might occasionally be controversial, for example provoking discussion if a player exhibits anti-social behaviour in a co-operative team game.
- **Tech** – Some Tech videos were disliked at lower/higher rates than other Tech videos, even though they are all liked at similar rates. Some videos may have been controversial (watched but disliked), such as product announcements for competing technologies.

Table 7.1. The average Spearman correlations between each YouTube metric and view counts, by category. Videos are split into age bands with 50+ videos per band before the calculations to reduce the influence of video age. The reported correlations are the cross-band averages (see Methods). High values are green, mid-range values are yellow and low values are red.

Category	Dislikes	Likes	Comments	Length
Animation	0.906**	0.894**	0.786**	0.085
Automotive	0.539**	0.739**	0.529**	0.198
Beauty	0.840**	0.804**	0.741**	-0.093
Best of	0.909**	0.897**	0.818**	-0.001
Causes	0.732**	0.822**	0.689**	0.025
Comedy	0.885**	0.803**	0.701**	0.059
Cooking	0.827**	0.803**	0.747**	-0.119
DIY	0.870**	0.865**	0.831**	0.137
Education	0.677**	0.774**	0.679**	-0.050
Entertainment	0.842**	0.878**	0.769**	0.062
Fashion	0.849**	0.849**	0.790**	0.054
From TV	0.876**	0.897**	0.870**	0.069
Gaming	0.822**	0.875**	0.726**	-0.107
Health	0.769**	0.845**	0.697**	-0.025
How-to	0.888**	0.883**	0.795**	-0.132
Lifestyle	0.777**	0.820**	0.733**	0.053
News	0.841**	0.822**	0.722**	-0.115
Non-profit	0.504**	0.731**	0.554**	0.202
Politics	0.768**	0.837**	0.724**	-0.153
Science	0.845**	0.867**	0.800**	-0.243
Sports	0.854**	0.860**	0.794**	-0.064
Tech	0.822**	0.862**	0.762**	-0.057
* 0.05 (5%) Significance - +/- 0.279 for n=50				
** 0.01 (1%) Significance - +/- 0.363 for n=50				

7.2 Factors Correlating Directly with View Counts

In general, the more views a video attracts, the *less* likely it is that each viewer likes, dislikes or comments on the video, as evidenced by negative correlations between view counts and like, dislike and comments per view (Figure 7.1). Thus, although videos with more viewers get substantially more interactions (Table 7.1), each individual view is less likely to lead to a like, dislike or comment for more popular videos (Figure 7.1). This is the opposite from that to be expected if viewers were encouraged by the number of comments, likes or dislikes to view a video. A possible explanation for the negative correlation tendency is that individual likes and dislikes are less influential in larger numbers. Whilst the first few likes might make a noticeable difference to the popularity of a video, a video with over a thousand likes already would not benefit much from one more (and it might not even be visible in the rounded score), which may discourage users. The position for comments might be different but with the same result because if there are more comments than shown in the video page without requesting more, then it is more difficult to check for relevant previous comments or to engage in dialog with other commenters.

There are some exceptions to the negative correlations. The category Animation represents what might be expected from virality or intuitive logic about influence. Animations with high likes per view or high comments per view are more popular, whereas animations with few dislikes per view are less popular. Thus, it is possible that Animation viewers are influenced by audience reactions.

Ignoring values where the error bars cross the x-axis, three other fields have statistically significant positive correlations. Dislikes associate with high view counts for Non-profit and Automotive. The positive correlations from Automotive might have been due to a proportion of dissatisfied viewers trying advice and leading to bad results. For example, one of the highest ratios of dislikes per view was for, “Easy Way to Remove Automotive Window Tint”, which seems likely to be something that could lead to disaster if performed incorrectly. For non-profit, people might dislike a video for covering unpleasant topics (even with good motives). One of the highest dislike ratios was for a video about children in pain, “Whole Child LA - Our Favorite Non-Profit!”.

For News, both likes and comments per view have a small but statistically significant positive correlation with total views. The magnitude of the correlations are too small to be informative, however.

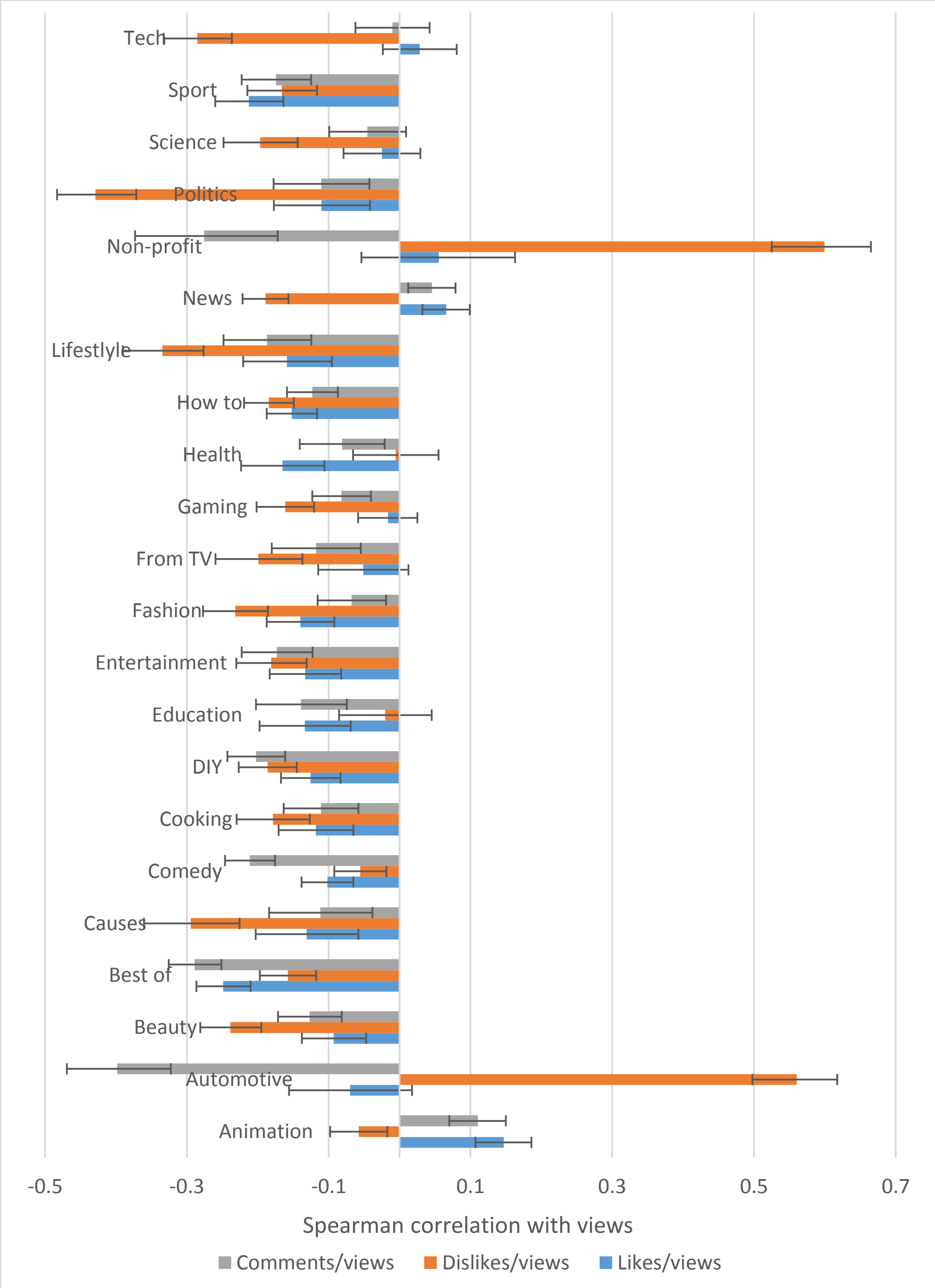


Figure 7.1. **Spearman correlations between each YouTube metric and view counts, by category with 95% confidence intervals. Qualification: at least 500 views; only the middle 50% of videos are included, ranked by metric ratio, to eliminate outliers.**

7.3 Summary

Although video length does not seem to influence popularity in any category, there is a tendency in all categories for more viewed videos to be more liked, more disliked and more commented on. This is unsurprising given that each visitor may do one of these three things and, other factors being equal, the more visitors the more likely each one is to occur. Nevertheless, the weaker relationship in some categories points to heterogeneity in video reception within them. Overall, however, the more popular a video is, the less likely each viewer is to like, dislike or comment on it, so videos seem to have decreasing interactions as they become more popular. A few categories have exceptions to this.

8 YouTube User Perspectives

A total of 660 questionnaires were given out, with 534 (81%) returned. This return rate is much higher than is expected for questionnaires (Cohen et al., 2017; Bryman, 2004). This generates a low non-response bias, helping to justify the biased initial sample strategy. The respondents comprised 444 females, 87 males, 1 other and 2 that did not disclose their gender. The three respondents who didn't disclose their gender or defined themselves as other are not included in the discussion and analysis of the data. The analysis is divided into three sub-sections: gender, age and user level (See Appendix 12) (RQ6 and RQ7).

8.1 Gender Differences

Some demographic information about the survey respondents is summarised here for background context. Most respondents are within the '18 to 24' and '25 to 34' year categories across both genders (Figure 8.1) which are most common ages of YouTube user (Gaunt, 2015; Perrin, 2015; Smith, 2014). Since 95% of both the female and male respondents are between 18 and 54, this data is representative of internet users in age but not gender (Blank et al., 2019). Despite the substantial difference in number between female and male respondents the percentages for each of the age categories are relatively consistent (Figure 8.1).

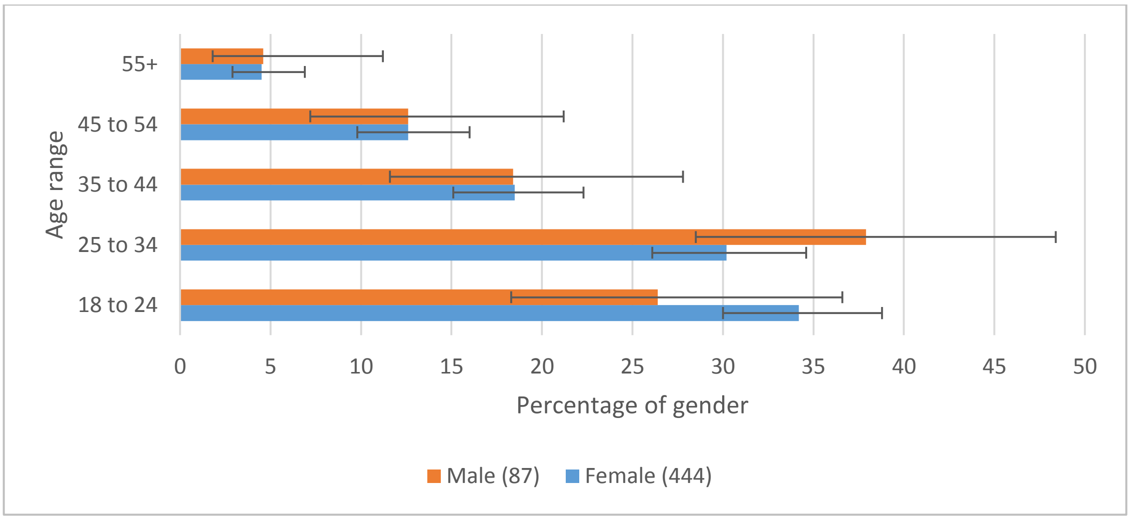


Figure 8.1. **Female and male respondents by age range with 95% confidence intervals**.

Nearly all (89%) female and male respondents are educated, or being educated, to Higher Education level (Figure 8.2). From the OxIs survey, 95% of internet users (18+) were educated (or being educated) to this level (Blank et al., 2019), but the figure may be lower now. Nevertheless, the survey sample does not seem to be far out of line for education. The gender percentages across the education levels are relatively consistent (Figure 8.2).

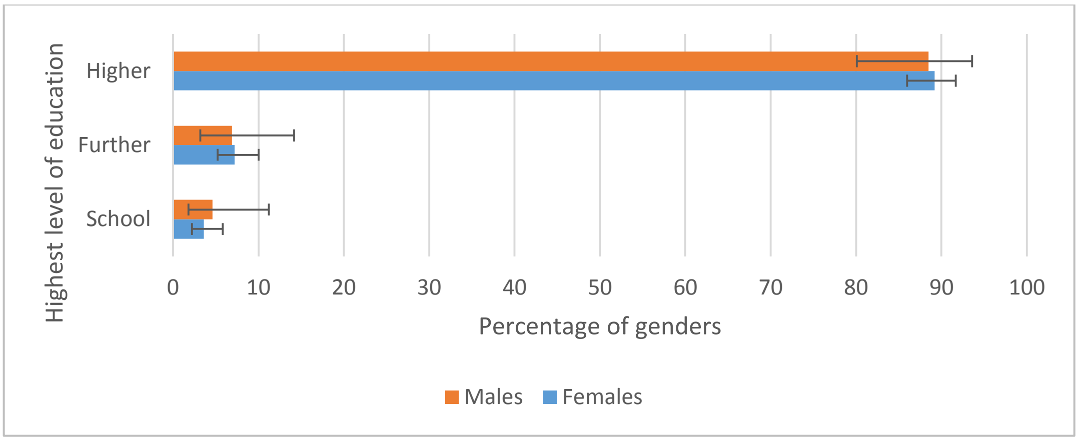


Figure 8.2. **The education levels of female and male respondents with 95% confidence intervals.**

8.1.1 YouTube Use and Method of Accessing Videos

Information about the frequency of YouTube use and methods of accessing videos is summarised here to give context to the main results. Within the sample, 85% of the females and 90% of the males were frequent YouTube users, accessing videos daily or weekly. These high percentages are unsurprising given the popularity of YouTube (Arthurs et al., 2018; Dehghani et al., 2016; Gaunt, 2015; Buzzetto-More, 2014). The statistically significantly higher percentage of male (58%) than female (36%) daily users aligns with a range of evidence that YouTube is used more frequently by males (Fisher and Ha, 2018; Mayoral et al., 2010; Madden, 2009) but contradicts other evidence that use is similar across genders (Rainie et al., 2012) and that females are more prominent users (Oh and Syn, 2015; Chappell, 2012). It is possible that males use YouTube more frequently, but both genders are similarly likely to use it. Weekly, but not daily, use is statistically significantly higher for female (49%) than male (32%) respondents. The percentages of both monthly and yearly use by the respondents are relatively low across both genders within the research: few respondents are occasional users.

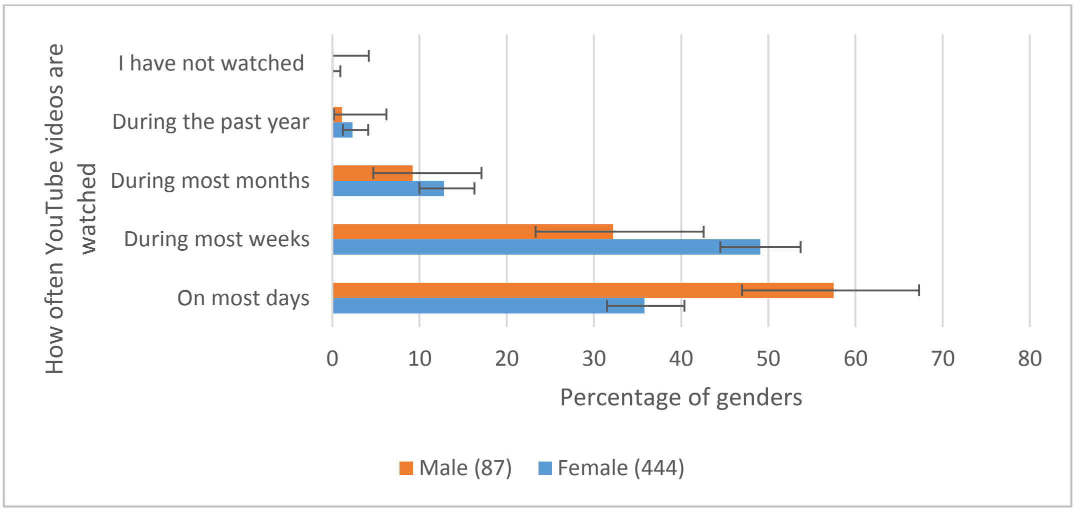


Figure 8.3. **Female and male respondents watching YouTube videos with 95% confidence intervals.**

The most popular way for both females and males to access YouTube videos (Figure 8.4) is through the website itself (Welbourne and Grant, 2016; Duncan et al., 2014; Barry et al., 2014; Buzzetto-More, 2014). Some of the videos that the respondents choose to watch could therefore be presented to them by YouTube through the homepage and its various recommendation systems (Figueiredo et al., 2014; Anthony et al., 2013; Borghol et al., 2012; Figueiredo et al., 2011; Zhou et al., 2010).

A high proportion of both genders also access videos through the YouTube app (Figure 8.4), which could be due to the greater use of mobile devices and technologies, and improvement and developments relating to roaming access to the web (Carbonell et al., 2018; Rein and Venturini, 2018; Andone et al., 2016; Gaunt, 2015; Buzzetto-More, 2014; Kim et al., 2013). Although there is a slight overlap between the upper and lower confidence intervals, this method of accessing YouTube videos seems to be more popular with male respondents (Figure 8.4). Again, as a result of using the YouTube app there is a higher possibility of the users being influenced, in the videos they watch, by those which are directly presented to them through the homepage and its various recommendation systems (Figueiredo et al., 2014; Anthony et al., 2013; Borghol et al., 2012; Figueiredo et al., 2011; Zhou et al., 2010).

Both genders seem to access a high proportion of the YouTube videos that they watch by following a hyperlink on a Facebook post (Broxton et al., 2013) and as previous research has highlighted there is a greater chance of videos being shared on Facebook (Vingilis et al., 2018). This suggests influence from social network groups and web pages visited. This is supported by evidence that videos that are shared or recommended by friends, members of social networks or a trusted source are more likely to be watched (Arthurs et al., 2018; Hayes et al., 2018; Feitosa and Botelho, 2017).

Few males and very few females access videos through hyperlinks within blogs (Figure 8.4). This could be due to the comments and discussion-based nature of blogs, with few links to videos. Fewer people may also be accessing blogs, particularly with the rise in popularity of vlogs (Hill et al., 2020; Codreanu and Combe, 2019; Berryman and Kavka, 2018).

Searching through Google is popular for accessing YouTube videos with both genders (Figure 8.4). This is supported by previous research which suggests that YouTube videos are regularly accessed through external means such as Google searches (Figueiredo et al., 2011; Szabo and Huberman, 2010; Zhou et al., 2010; Benevenuto et al., 2009; Paolillo, 2008; Cheng et al., 2008). People may use Google search to get a wider range of options than YouTube gives, particularly when searching for a recommended video. They may also search for specialised content that might not be easy to search for within YouTube.

Male respondents use more methods to access YouTube (Figure 8.4, Figure 8.5) and are at least as likely as females to use any given method. This supports the evidence that the male respondents generally spend more time browsing and searching for YouTube videos (Fisher and Ha, 2018).

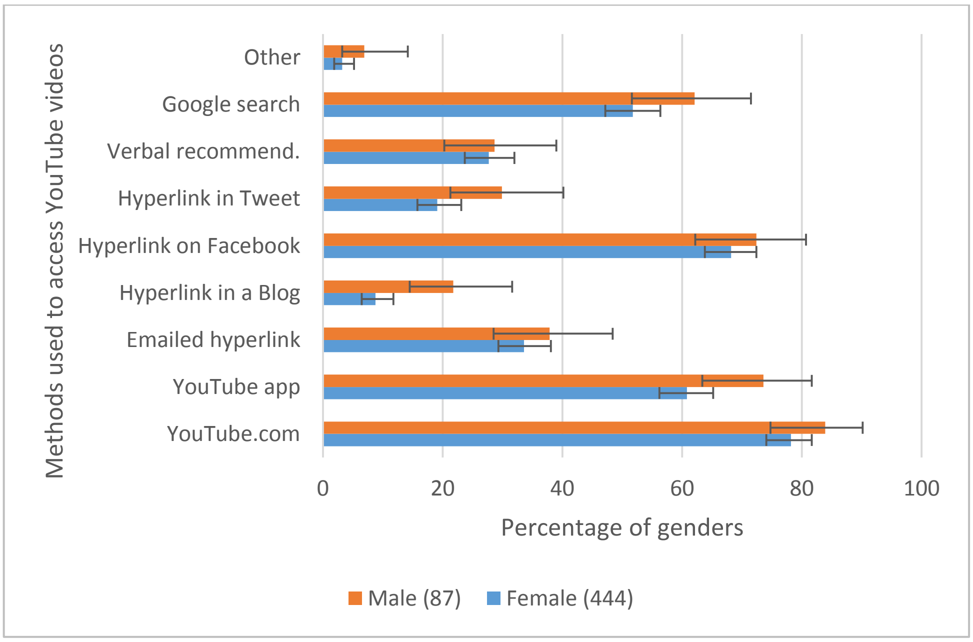
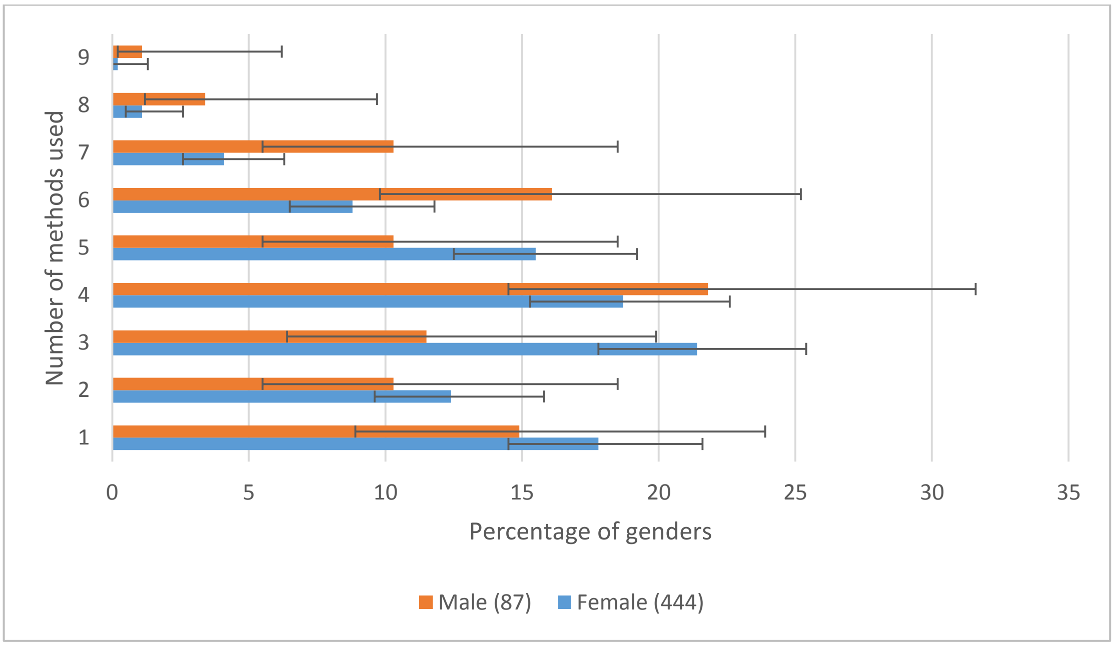


Figure 8.4. **The methods female and male respondents have used to access YouTube videos with 95% confidence intervals.**



Figures 8.5. **The number of methods used by females and males to access YouTube videos with 95% confidence intervals.**

The most popular method of accessing videos through the YouTube website for both genders (79%) is its inbuilt search facility (Figure 8.6), confirming prior (but slightly dated) findings (Pinto et al. 2013; Zhou et al., 2010; Crane and Sornette, 2008). Visiting the website or using the YouTube app suggests more control over what is viewed, compared to following hyperlink recommendations from other sources. YouTube website users may tend to access popular videos since these (and user-specific preferences) are presented by YouTube through the search facility.

The YouTube homepage is used by 32% of females and 49% of males, supporting findings that suggests that there is a greater chance of videos being watched when they are recommended by YouTube (Wilhelm et al., 2018; Figueiredo et al., 2011). Since the homepage is used more by males, they may be more susceptible to YouTube's recommendations based on what they previously watched (Figure 8.6). YouTube encourages people to access videos by presenting them with various options based on their previous viewing history and preferences and this may influence males more.

The subscription pages and the recommendation bar seem to be more important to males than females (Figure 8.6). Subscriptions are used by viewers to connect to a wider range of videos and content that they relate to and can also support the development of interactions with likeminded users (Kayumovich and Annamuradovna, 2020;Buzzetto-More, 2014; Alloway and Alloway, 2012). Thus, it seems that the male respondents are using the YouTube subscribing function to provide them with greater, wider and easier access to video content that they are more motivated to watch. Male respondents may also be more motivated by making connections and developing further social networks with other users based on their commonalities (Döring and Mohseni, 2019), and could also be a result of their greater interest in accessing YouTube (Fisher and Ha, 2018).

The most popular video and most viewed video pages seem to be of least importance to both genders (Figure 8.6). This suggests that people rarely visit YouTube to find popular videos. Popular videos may be accessed instead through a range of sources, such as social media websites, or found by being prominent in search or category browsing results. Since respondents do not search for popular videos, whereas people are more likely to follow, conform or be influenced by others in their decision-making process (Welbourne and Grant, 2016; Eger, 2015; Boyd, 2014; Shifman, 2012), influence must spread in a more subtle way.

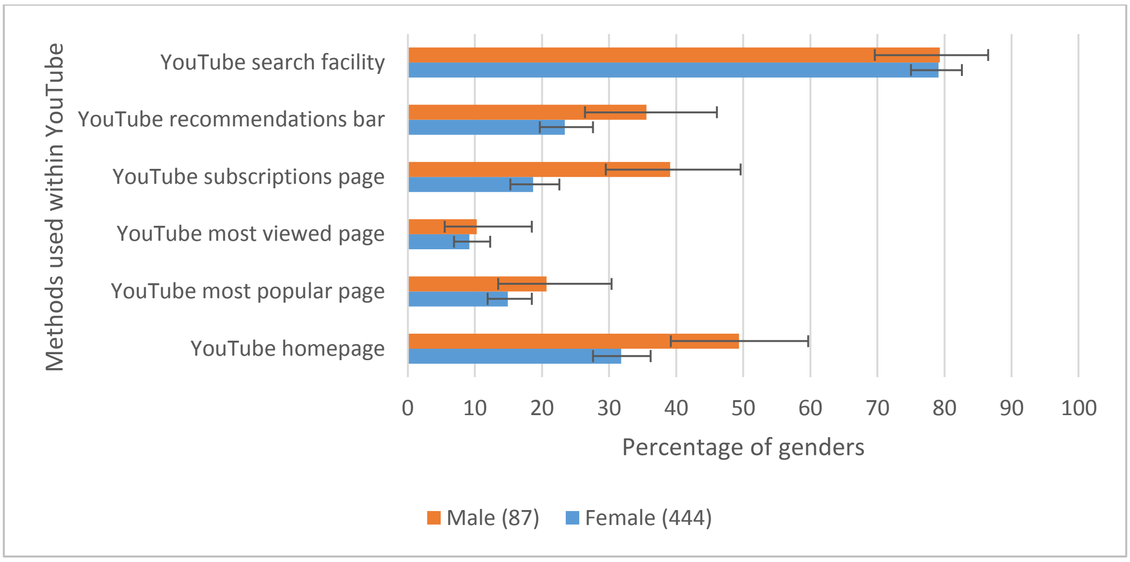


Figure 8.6. **Methods female and male respondents use to access videos in YouTube.com with 95% confidence intervals.**

8.1.2 Video Categories

The three most popular categories are the same for both genders: Comedy, Education and Entertainment (Figure 8.7). The popularity of Comedy and Entertainment on YouTube is well known (Arthurs et al., 2018; Barakat, 2014; Buzzetto-More, 2014; Guadagno, 2013; Berger and Milkman, 2013). How-to is also a common and relatively popular category for female and male respondents. This may be due to the posting of popular educational videos within YouTube and thus provides viewers with useful self-directed learning videos (Rubenking, 2019; Feitosa and Botelho, 2017; Buzzetto-More, 2014; Duncan et al., 2013; Haran and Poliakoff, 2012).

Since most respondents work within education, they are more likely than others to use YouTube for educational and work-based purposes. Animation and From TV are also relatively popular with both genders who participated within the research (Figure 8.7). Beauty, Lifestyle and Cooking are statistically significantly more popular with female than male respondents (Figure 8.7). YouTube is increasingly used for beauty and lifestyle advice from vloggers (García Rapp, 2016) and these seem to appeal more to females (Ashton and Patel, 2020; Ladhari et al., 2020). For Cooking, females may be more likely to use YouTube for free video-based recipes.

The categories that are more popular for male respondents, with no overlap in the confidence intervals, are Sport, Politics, Best of and Gaming (Figure 8.7). Content relating to sporting events can attract a large audience. Tech, although not one of the most popular categories with male, is less popular with females (Figure 8.7) (Fisher and Ha, 2018). Best of, Politics and Gaming are particularly unpopular with female respondents.

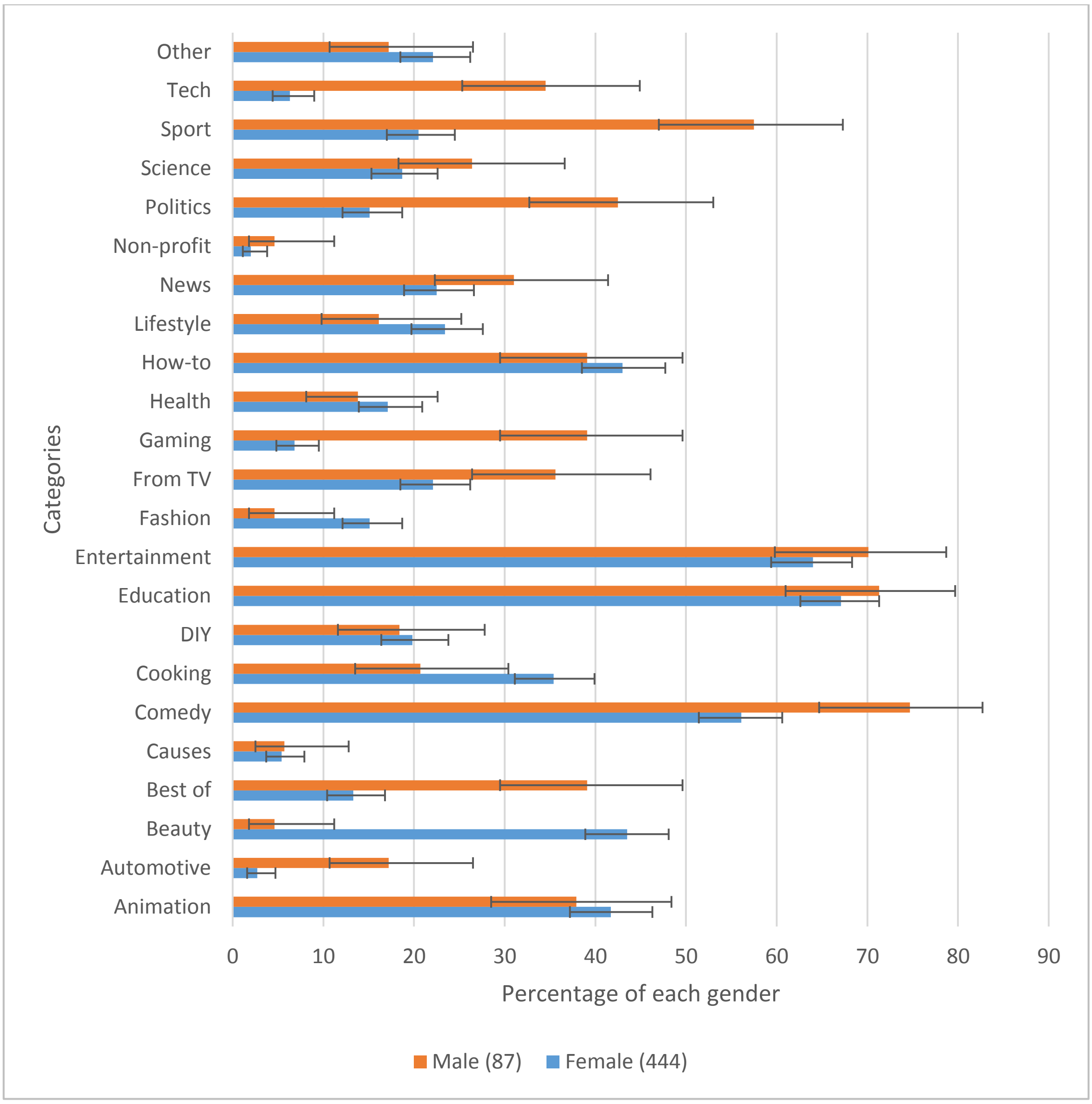


Figure 8.7. **The categories of video that female and male respondents reported watching with 95% confidence intervals.**

8.1.3 Factors Influencing the Decision to Watch a Video

The most important factor in deciding to watch a video was the title of the video, for half of both genders (Figure 8.8). The title of a video is known to be a significant factor in whether someone chooses to watch a video (Chang, 2018; Jiang et al., 2014; Konnikova, 2014; Cheng et al., 2013).

Video thumbnails are also important for both genders (Figure 8.8), for the information provided about the video (Chang, 2018). The value of a thumbnail has not been previously studied but videos need a 'hook' to encourage people to watch (Ahn and Shin, 2016; Tschopp, 2014). This 'hook' could be the thumbnail picture of the video.

A quarter of female and male respondents are influenced by the length of a video when deciding whether to watch it (Figure 8.8). Video length is known to be important when choosing to watch a video (Jiang et al., 2014; Konnikova, 2014; Tschopp, 2014; Guadagnoa et al., 2013; Berger and Milkman, 2012). Previous research suggests that shorter videos are preferred by viewers (Bentley et al., 2019; Tschopp, 2014; Guadagnoa et al., 2013; Berger and Milkman, 2012), however, and videos between 16 to 30 seconds are the most likely to be viewed (Henke, 2013).

Both genders claim that the number of comments has little impact on their decision to watch a video (Figure 8.8). They may not be concerned with others' opinions before watching a video, or may read some comments but not be interested in the number posted. Since videos with more comments are more popular, this is probably due to popular videos receiving more comments rather than commented videos becoming popular. Perhaps surprisingly, the number of likes and dislikes also seem to have little impact on someone's decision to watch, further suggesting that both genders are not directly influenced by the opinions of others. Since people are influenced by others' opinions and ideas, including for videos (Welbourne and Grant, 2016; Akdeniz, 2014; Shifman, 2012), this influence must be indirect.

A high proportion of both genders access YouTube videos through Facebook and are therefore influenced by the recommendations and the opinions of others that are part of their social group or network, but the number of views is relatively unimportant (Figure 8.8) (Dynel, 2014). However, other research has explained that videos which are more popular have a higher likelihood of being watched (Kong et al., 2018; Qiu et al., 2015). Recommendations yield more views from YouTube users, so the personal more direct connection of a recommendation seems to be more important than broad popularity (Wu et al., 2018; Jiang, et al., 2014; Frasco, 2014.

The age of a video has little influence on female and male respondents' decisions to watch it (Figure 8.8), despite new content being more appealing. Perhaps some types of content are ageless whereas others are not. Overall, respondents seem to be influenced more by video content information and direct recommendations than the opinions of unknown others.

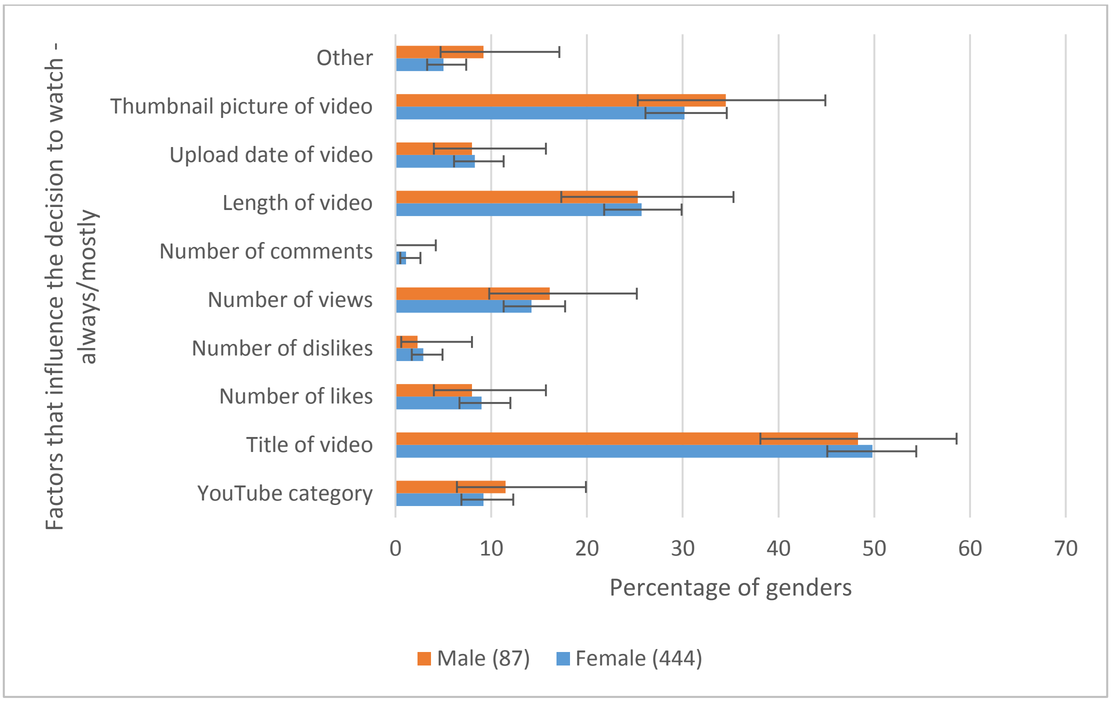


Figure 8.8. **Factors that always or mostly influence female and male respondents' decisions to watch a video with 95% confidence intervals.**

8.2 Age - Female

This section breaks down the above results by age range for the female respondents to assess whether there is evidence of differing factors influencing video choices. In most cases the sample sizes are too small to give statistically significant evidence of any except the largest differences.

8.2.1 YouTube Use

A high proportion of the respondents across the age bands are frequent, daily and weekly, users of YouTube (Figure 8.9). The higher proportions of frequent users are within the '18 to 44' age bands with substantial daily and weekly use from the '18 to 34' age bands. None of the 55+ respondents use YouTube daily and are less frequent users (Figure 8.9).

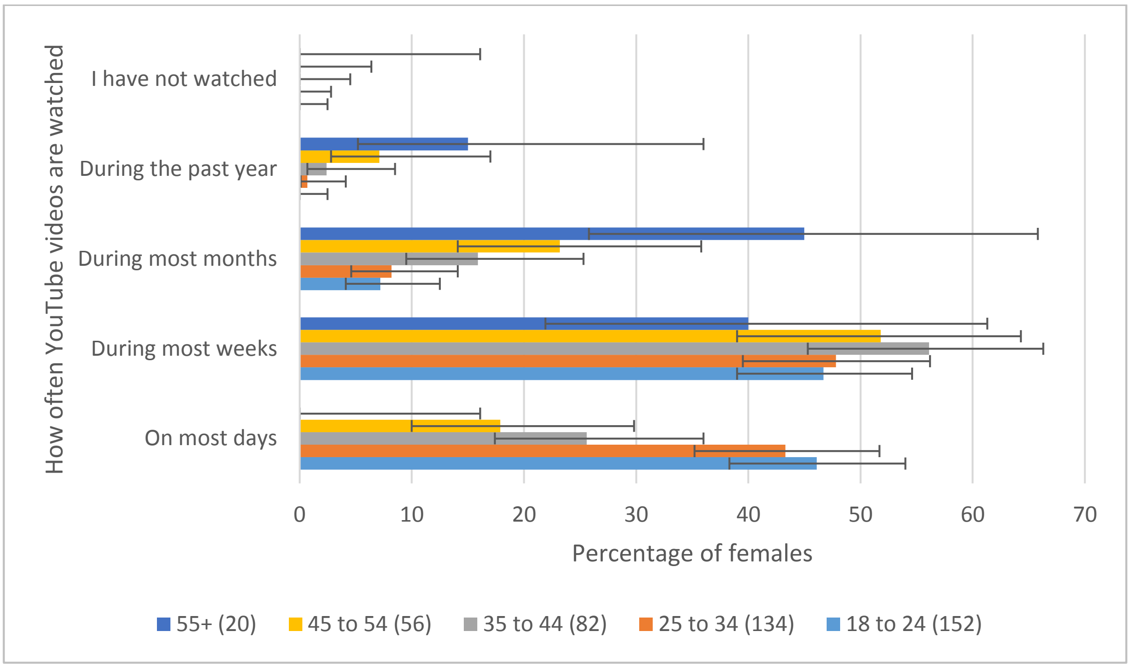


Figure 8.9. **YouTube video watching frequency by age group for female respondents with 95% confidence intervals.**

8.2.2 Accessing YouTube Videos

The most popular way of accessing videos is through the YouTube website for all age bands (Figure 8.10). The YouTube App is used most within the age bands '18 to 34', because younger people are generally more active smart phone users (Carbonell et al., 2018; Andone et al., 2016). Facebook seems to be a popular forum for accessing YouTube videos for most age bands, which could also reflect respondents use of social media websites in general (Figure 8.10) (Carbonell et al., 2018). Social media networks have had a significant impact in terms of watching online videos Mehrotra and Bhattacharya, 2017; Alloway and Alloway, 2012; Ameigeiras et al., 2012).

Hyperlinks in emails are used across age bands (Figure 8.10) and are slightly more popular with female respondents aged 35 to 54. This was a more popular method of accessing YouTube videos before the development of social media websites such as Facebook. Nevertheless, email recommendations are from a trusted source, if they know the person that sent the email. People are more likely to watch videos that have been sent from trusted sources as they act as an online form of Word-of-mouth recommendation (Kong et al., 2018; Figueiredo et al., 2014; Frasco, 2014; Jiang, et al., 2014; Anthony et al., 2013).

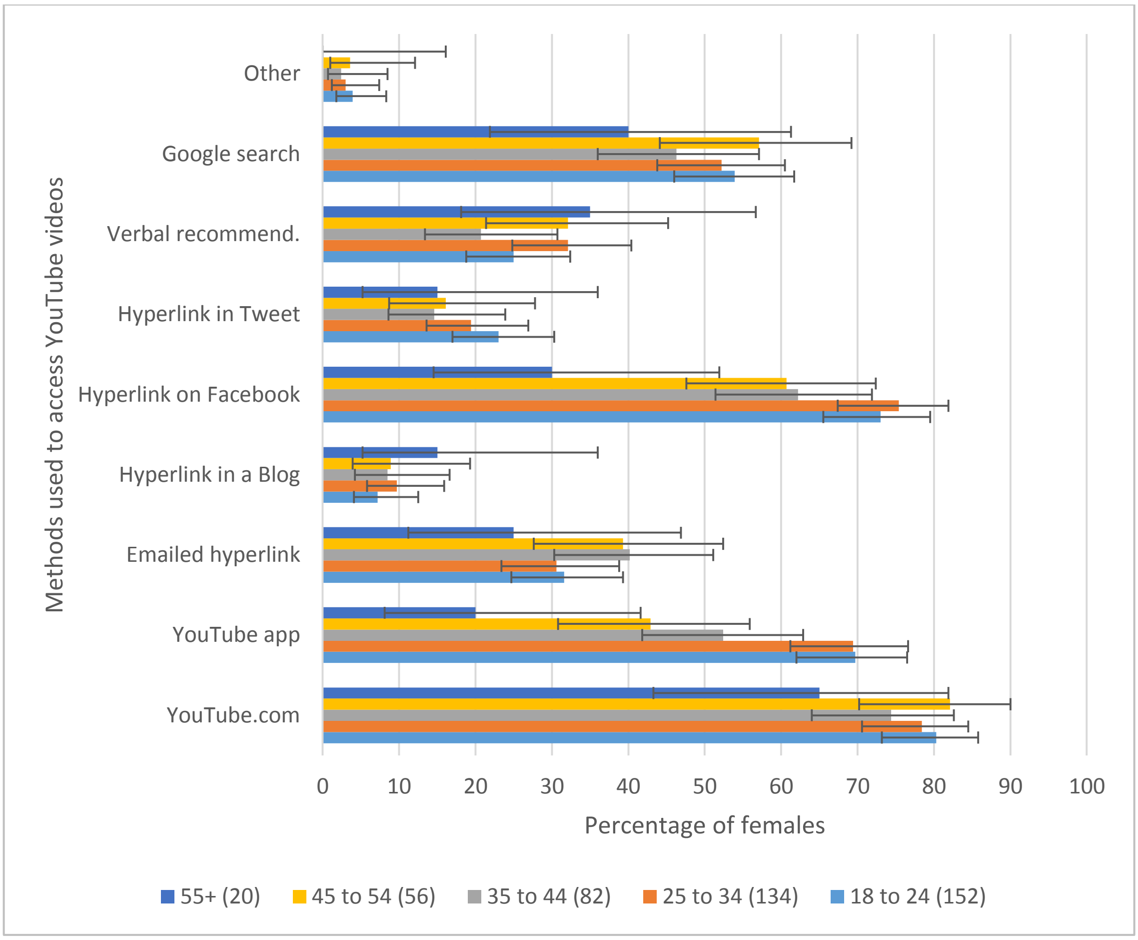


Figure 8.10. **How the different age groups of female respondents have accessed YouTube videos with 95% confidence intervals.**

Females aged 55+ seem to use few methods to access YouTube videos (Figure 8.11). This could be due to using YouTube videos for specific purposes rather than a form of regular entertainment (Fisher and Ha, 2018). Younger females aged 18 to 34 (Figure 8.11) use a wider range of methods to access YouTube videos (Gaunt, 2015; Perrin, 2015). As the bands increase in age the female respondents use a fewer methods (Figure 8.11). This may be due to younger females spending more time watching YouTube videos (Figure 8.9) (Gaunt, 2015; Perrin, 2015).

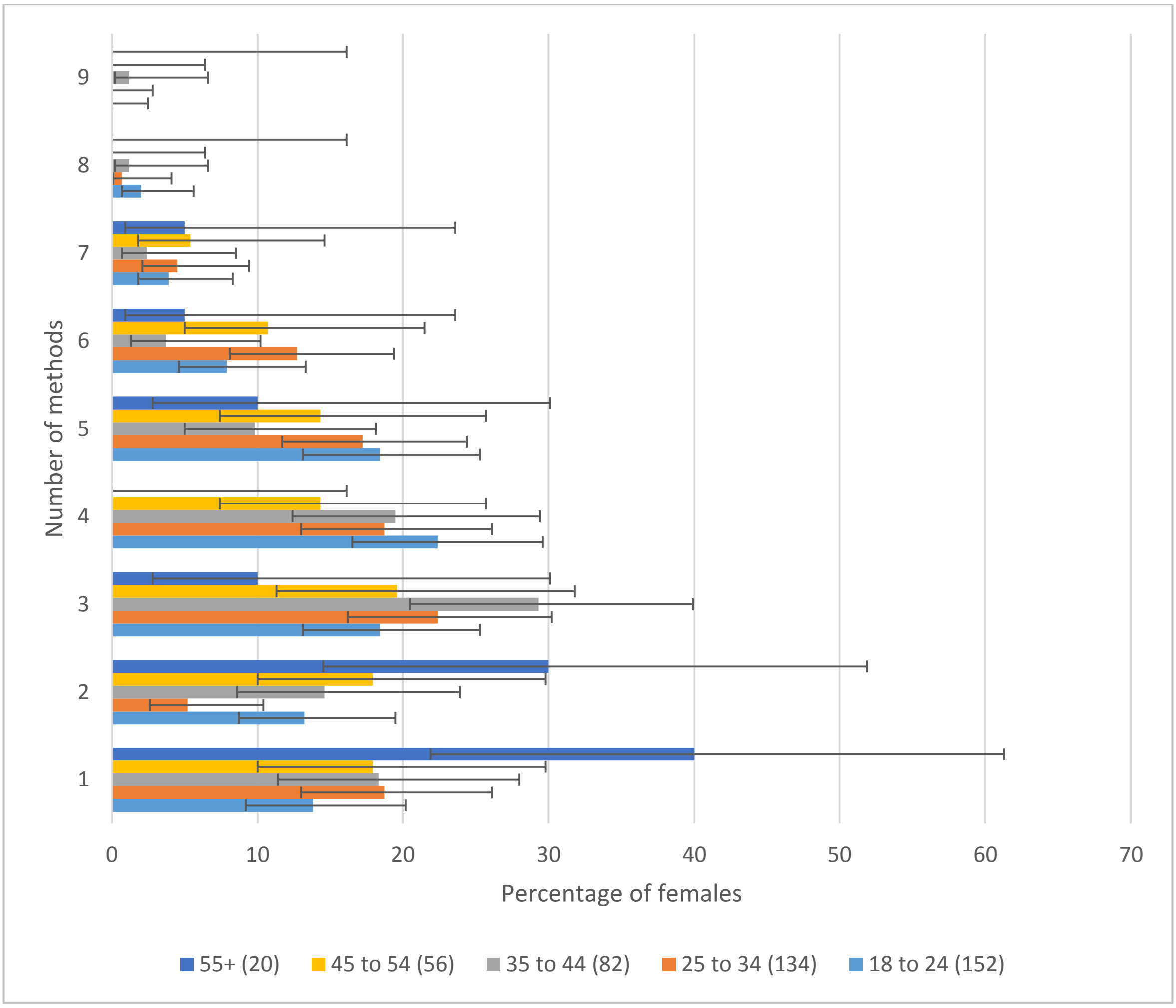


Figure 8.11. **The number of methods that different age groups of female respondents use to access YouTube videos with 95% confidence intervals.**

8.2.3 Accessing Videos within YouTube

The YouTube website search is still the most popular way that female respondents access videos across all age bands (Figure 8.12) (Pinto et al. 2013). Although the percentages are much lower, the YouTube homepage appears to be a relatively popular method for accessing videos across all age bands, but most by females aged 18 to 24 (Figure 8.12). Younger females might access YouTube for recreational and entertainment purposes rather than searching for a video. It is possible that the younger female respondent age bands interact with the suggestions on the YouTube homepage because of being more regular users as presented within the data (Figure 8.9). Female respondents 55+ seem to be selective in the methods they use within the YouTube website (Figure 8.12). They may usually have a clear purpose in what videos they want to watch rather than using YouTube as a form of entertainment like younger respondents (Dehghani et al., 2016; Duncan et al., 2014).

Female respondents across all age bands are rarely interested by the YouTube most viewed page (Figure 8.12). Although the percentages were slightly higher for some age bands, it was also the case for the YouTube most popular page (Figure 8.12). This contradicts aspects of research which suggests that individuals' choices are influenced by others and what is currently popular within society (Welbourne and Grant, 2016; Eger, 2015; Boyd, 2014).

Females aged 18 to 34 (Figure 8.12) seem to use the recommendations bar more than the other ages. Younger female respondents also use a wider range of methods within the YouTube website to access videos (Figure 8.12), which could be due to greater and more regular use (Figure 8.9).

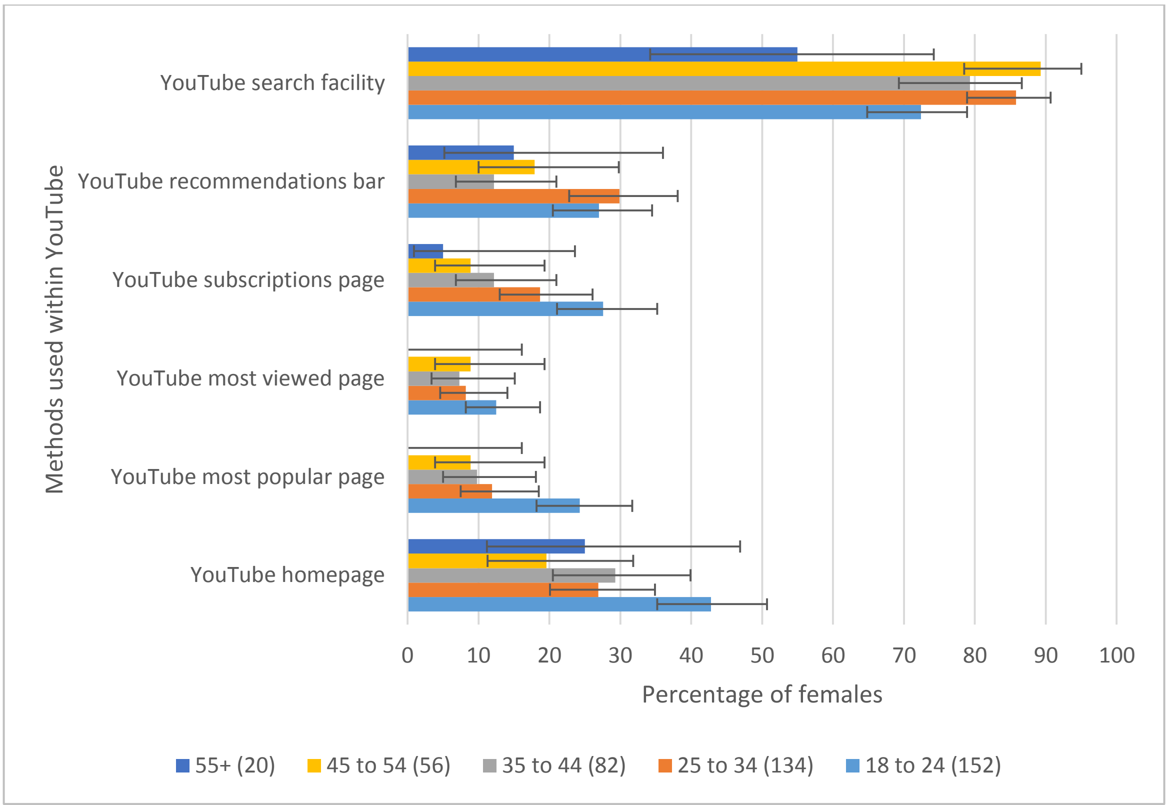


Figure 8.12. **Methods that the different age groups of female respondents use when accessing videos through the YouTube website with 95% confidence intervals.**

8.2.4 Video Categories

Education, Entertainment and Comedy are the three most popular types of video for female respondents under 55 (Figure 8.13). The three most popular video types for female respondents aged 55+ are How-to, Comedy and Education, suggesting that older females are less likely to use YouTube as a form of recreational entertainment.

The popularity of Beauty and Fashion decline with age (Figure 8.13). Videos within this category may target younger females, as they are more regular users of YouTube (Figure 8.9) (Brahmana and Vivaldo, 2018; Elven, 2018). Older females may also have more established tastes or access information relating to beauty and fashion through other means. However, other research does suggest that females overall are more interested in fashion (Fisher and Ha, 2018), beauty and lifestyle than males (Schwemmer and Ziewiecki, 2018). Health is least popular for females aged 55+ (Figure 8.13). Since health concerns increase with age, they may use more traditional methods to get health information. Politics is most popular with females aged 45 to 54 (Figure 8.13). Younger females are most likely to use YouTube to access News (Figure 8.13). Older people may prefer traditional TV or newspapers. However, research has suggested that videos which relate to events happening within the news have a greater chance of being viewed by individuals (Abisheva et al., 2014; Berger and Milkman, 2012).

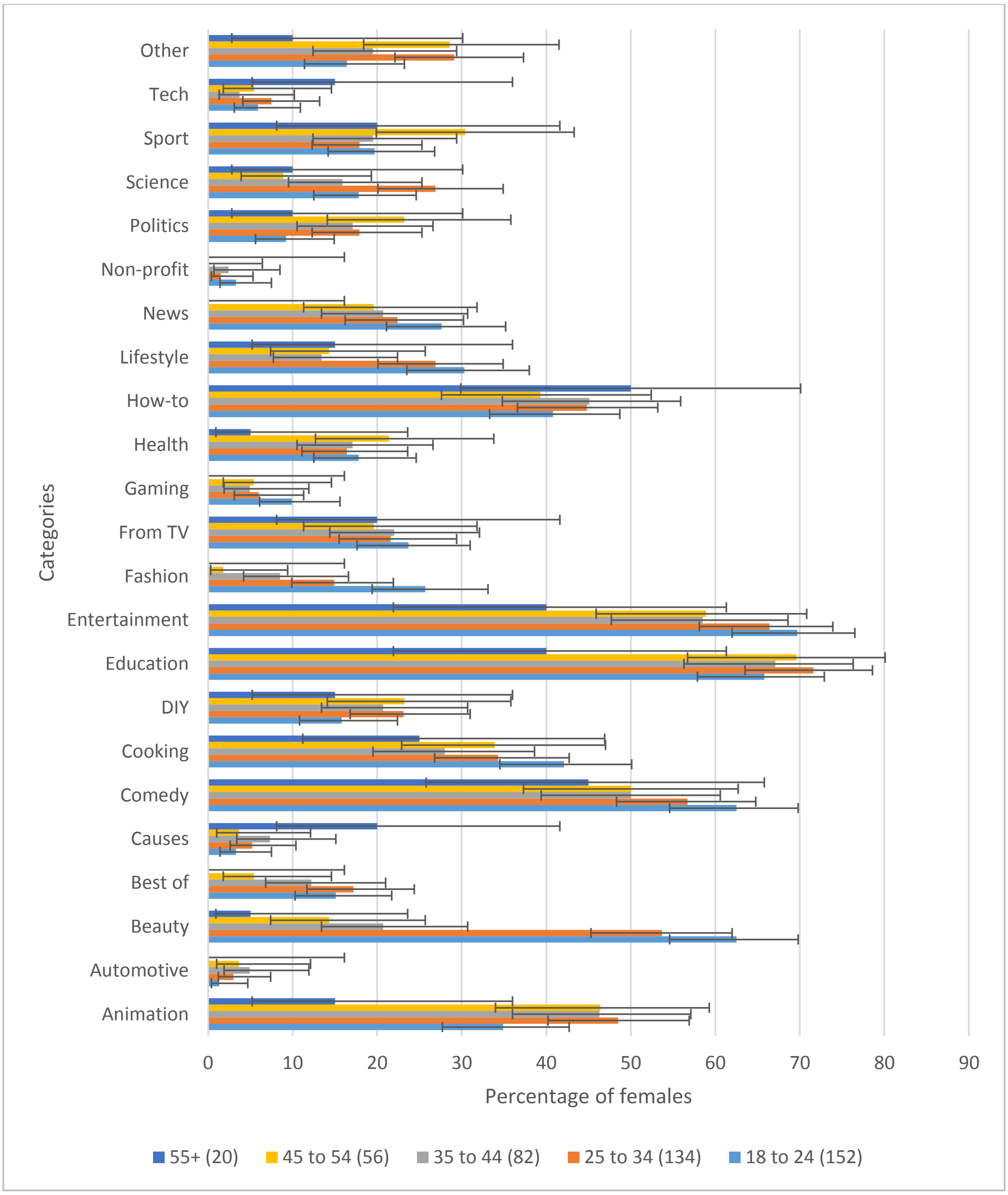


Figure 8.13. **Categories of video watched by female respondents by age with 95% confidence intervals.**

8.2.5 Influencing Factors

The most important factor in accessing YouTube videos for all female respondents across the different age bands is the title of the video (Figure 8.14) (Chang, 2018; Jiang et al., 2014; Konnikova, 2014; Cheng et al., 2013). This makes sense as the title of the video would hopefully provide the user with information relating to the content of the video. This would also be the case for the thumbnail picture as this is also an important factor for most of the female respondents in watching YouTube videos, except the '45 to 54' age band. Video length seems to most influence young females (Figure 8.14), perhaps preferring short videos (Bentley et al., 2019; Gaunt, 2015; Perrin, 2015).

Respondents choosing the Other option mainly mentioned the relevance of a video and whether it had been recommended. Since Other is highest for the '45 to 54' age band, these female respondents seem to be more influenced by their social networks. The youngest users seem to be most influenced by the number of views than the other age bands (Figure 8.14), perhaps due to increased sensitivity to trends amongst youth. This is supported by other research which suggests that individuals are generally more likely to be influenced to watch videos that are recommended by others within their social network or group as they share similar affiliations (Dehghani et al., 2016; Perrin, 2015; Oh and Syn, 2015; Kim et al., 2013).

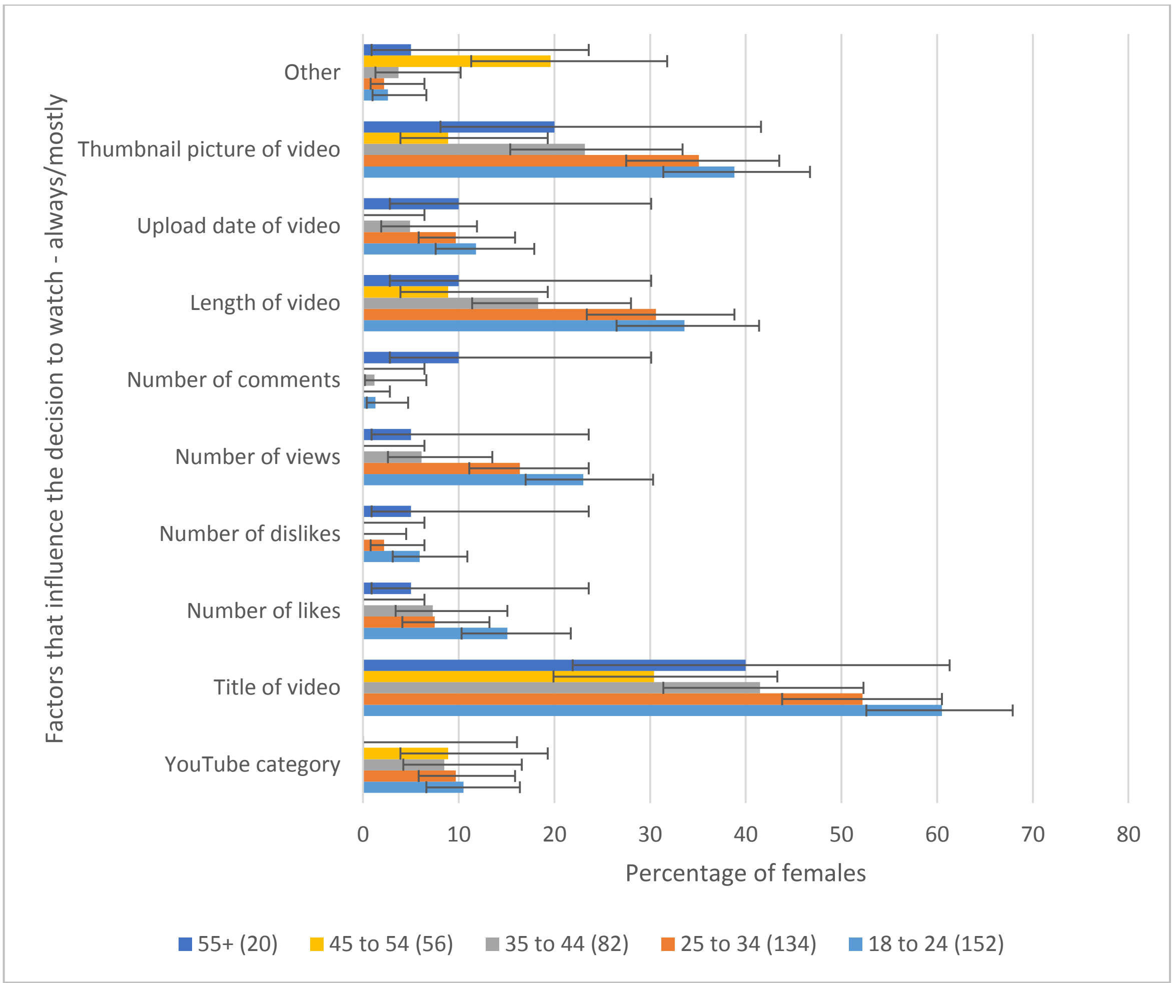


Figure 8.14. **The factors that always or mostly influence the different age groups of female respondents' decisions to watch a video with 95% confidence intervals.**

8.3 Age – Male

The following figures show the data for the various age bands for male respondents. Due to the small sample sizes, the differences are rarely statistically significant, so the discussion is speculative.

8.3.1 YouTube Use

Although there are age differences in frequency of accessing YouTube, these are not large or systematic enough to speculate about trends (Figure 8.15).

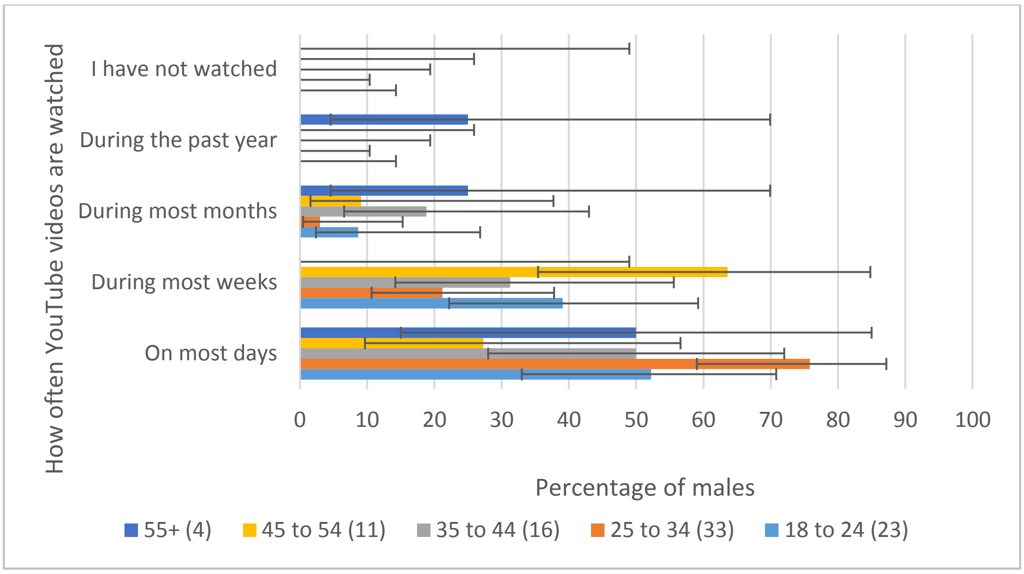


Figure 8.15. **How often the different age groups of male respondents watch YouTube videos with 95% confidence intervals.**

8.3.2 Accessing YouTube Videos

The main access method age-related trend evident in the sample is that younger males are more likely to use the YouTube app (Figure 8.16). The same trend occurs for females.

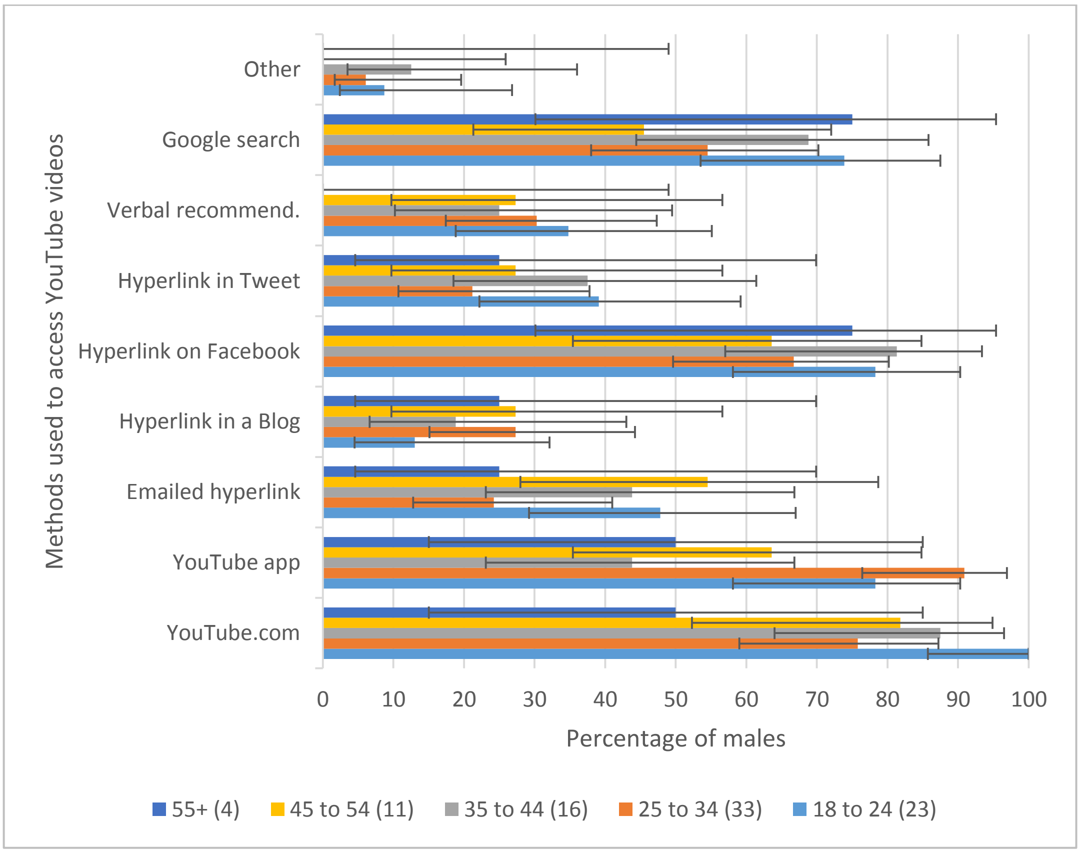


Figure 8.16. **How the different age groups of male respondents have accessed YouTube videos with 95% confidence intervals.**

Ignoring the small 55+ age group, there is a tendency for younger males to use more methods to access YouTube (Figure 8.17). This echoes the results for females.

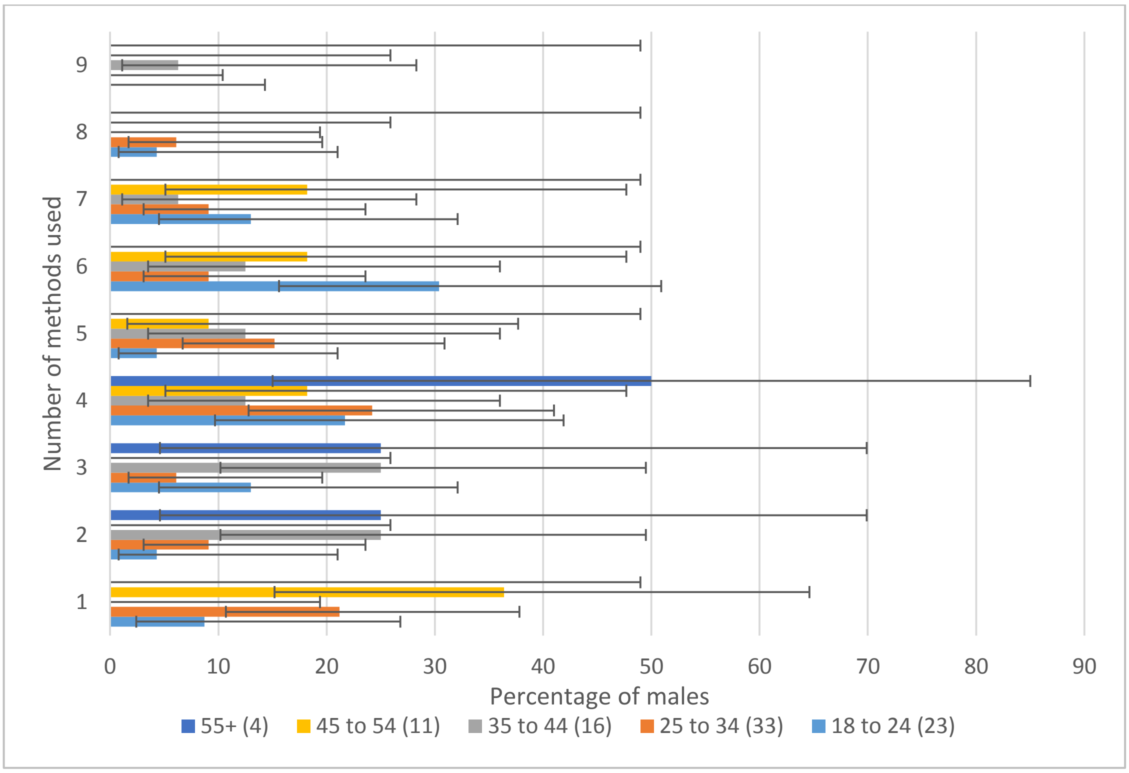


Figure 8.17. **The number of methods that different age groups of male respondents use to access YouTube videos with 95% confidence intervals.**

8.3.3 Accessing Videos within YouTube

There is a tendency for younger males to be more likely to use the YouTube home page, most popular page and most viewed page (Figure 8.18). This suggests a youth-related interest in global popularity, at least within YouTube.

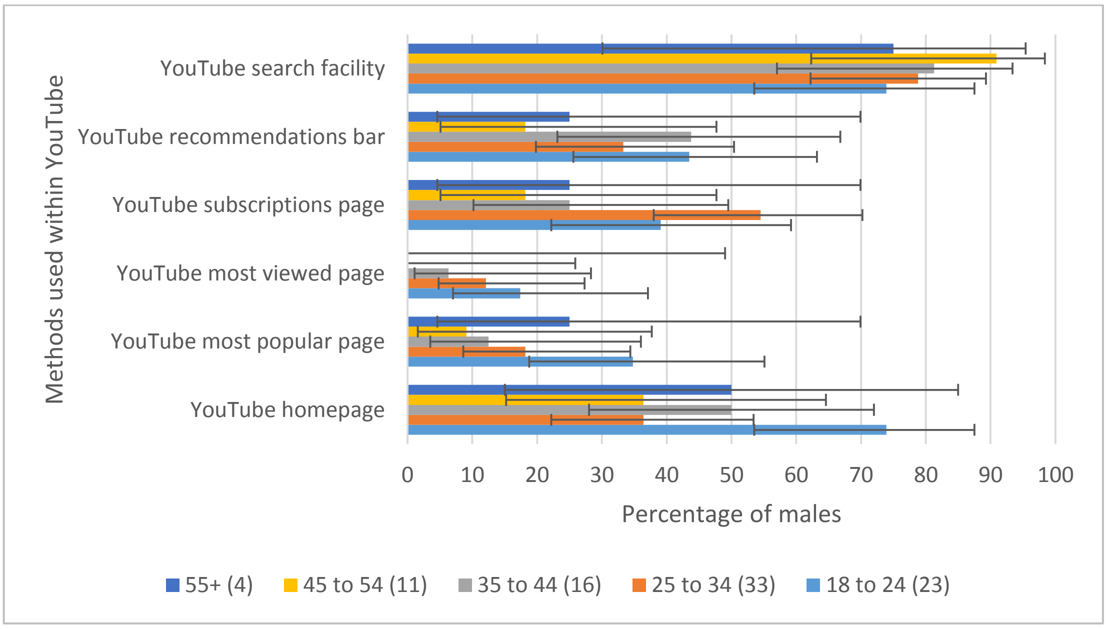


Figure 8.18. **The methods the different age groups of male respondents use when accessing videos through the YouTube website with 95% confidence intervals.**

8.3.4 Video Categories

There are non-significant age-related trends for individual categories for males, such as a tendency for younger males to be more likely to watch Sport, News, Cooking, Comedy, and Best of, and for older males to prefer From TV (Figure 8.19).

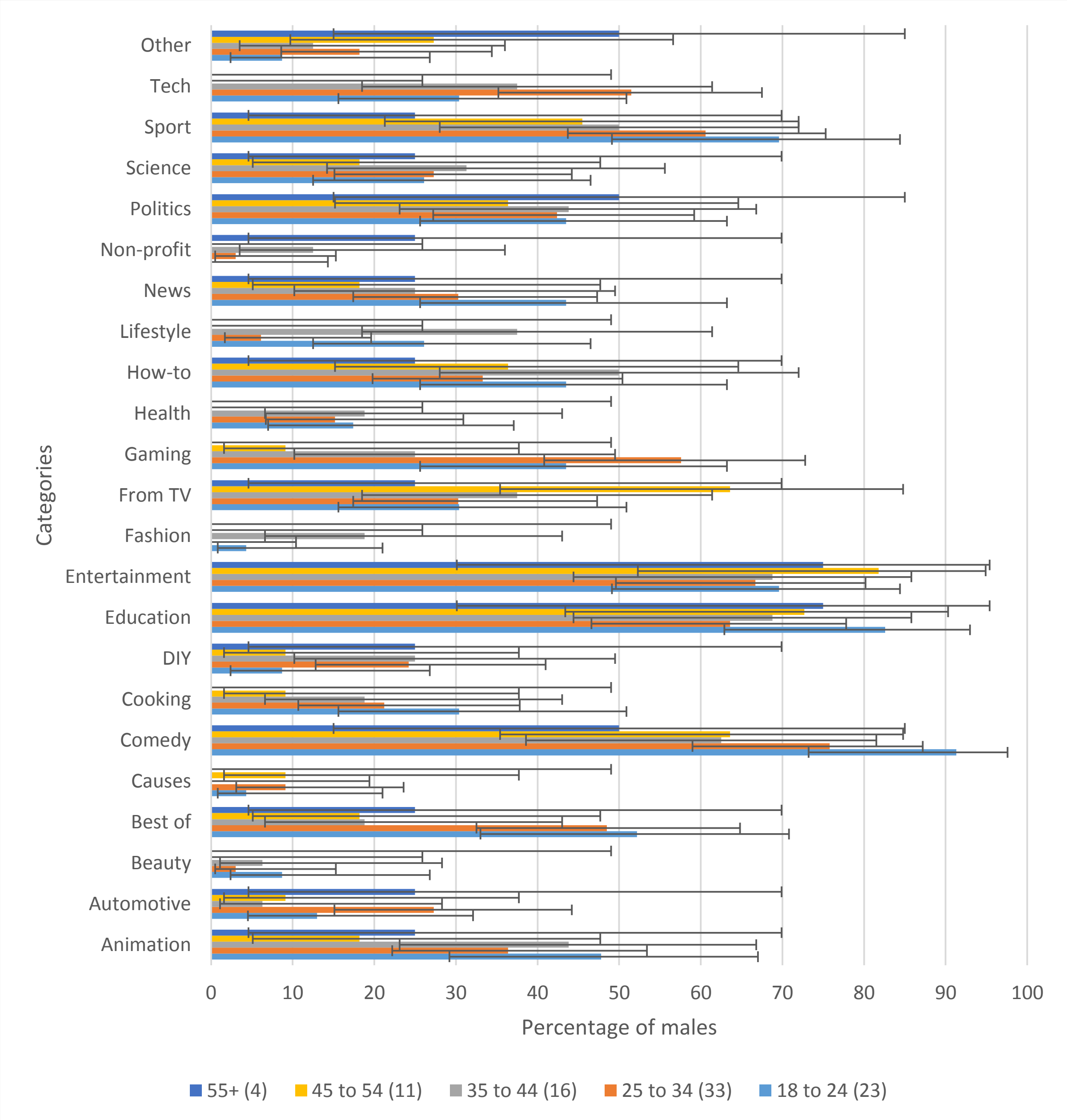


Figure 8.19. **The categories of video the different age groups of male respondents have watched with 95% confidence intervals.**

8.3.5 Influencing Factors

The title of video and the thumbnail picture of video are more influential for younger males than for older males (Figure 8.20). This echoes the situation for females.

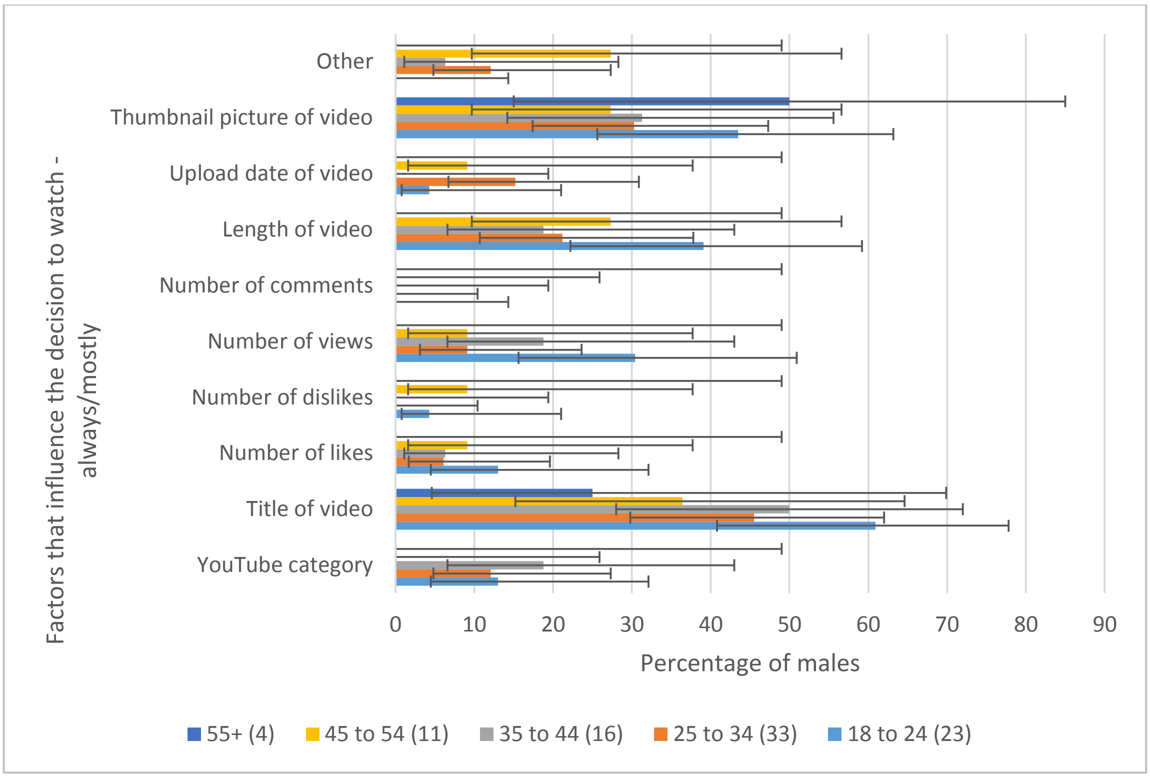


Figure 8.20. **The factors that always or mostly influence the different age groups of male respondents' decisions to watch a video with 95% confidence intervals.**

8.4 User Level

The analysis of respondent user level did not present anything new in terms of findings and therefore it can be found within Appendix 12.

8.5 Summary

The following summary considers the key findings from the questionnaire data in relation to RQ6 (gender and age differences in popular types of YouTube video) and RQ7 (gender and age differences for YouTube video watching influences). The questionnaires provided a return rate of 81% (534/660) and a sample made up with a considerably higher proportion of female respondents than male respondents. The sample was also substantially biased to people studying or working within the primary education and higher education sectors. A substantial percentage of the sample was made up of female (83%) and male (83%) respondents aged 18 to 44, with 89% of both genders being educated to higher education level. A higher percentage of male respondents (58%) use YouTube daily than female respondents (36%).

For both genders the main method for accessing YouTube videos is directly through the website. The YouTube App, hyperlinks or posts on Facebook and Google searches were also popular methods used by both genders, but male use a wider range of methods to access YouTube videos. The substantial use of hyperlinks or posts within Facebook shows that both genders can be influenced by their friends or online social networks.

Younger female respondents watch YouTube videos more regularly than older users. Younger female respondents use the YouTube App more than older respondents, and overall use a wider range of methods to access videos. Using a Google search to find YouTube videos is also relatively popular and consistent across all age bands. From the questionnaire data, younger female respondents use a slightly wider range of methods to access videos within YouTube.

In finding and accessing videos within YouTube the most popular method with all age bands is through the website's own search facility. Another relatively popular method across most female respondent age bands (less so for 25 to 34) is accessing videos through the YouTube homepage which will be suggestions made by the website based on a user's previous watching choices and preferences.

8.5.1 Gender and age differences in popular categories of YouTube video

The most popular categories of video with both genders are Comedy, Education and Entertainment. The three categories most watched by both genders are the same: Comedy, Education and Entertainment. Nevertheless, due to the sample selection bias, respondents could be accessing a higher proportion of Education videos for work-based purposes, so this may not reflect overall YouTube popularity. In addition, both genders extensively watch How-to, Animation and From TV. However, from the data collected it is not clear whether users search using key terms or by the specific YouTube categories and is therefore a limitation of the questionnaire findings.

The least popular type of video is Non-profit. The least watched type of video for both genders is Non-profit.

There are gender differences in categories of videos watched. Although there are some similarities in the types of videos that are watched it can also be seen that there are some substantial differences in the categories the genders prefer to watch. For example, Beauty and Cooking are watched more by female respondents whereas Sport and Politics are watched more by males.

Beauty and Fashion videos are more popular with younger females than with older females.

8.5.2 Gender and age differences for YouTube video watching influences

Both genders are influenced by video titles and thumbnails. There are few gender differences in factors influencing both genders to watch a video. The key influencing factor for both genders is the title of the video, as it provides insights into the content of a video before committing to watching it (Chang, 2018). The thumbnail picture of the video is also a key influencing factor for both genders for the same reason (Chang, 2018).

Both genders are little influenced by the opinions of others. This includes the number of comments and number of dislikes.

All ages of female are uninterested in most popular or most viewed video pages. The least popular methods for watching videos across all age bands are the most popular and most viewed pages. This suggests that female respondents believe that they are unlikely to watch videos just because they are popular with (unknown) others.

Younger females are more likely to be influenced by a wider range of factors. The factor that influences all ages of female respondents the most in watching content is the title of the video. From the questionnaire data younger female respondents are more likely to be influenced by a wider range of factors than older respondents. The length of the video and the thumbnail picture appear to have greater influence for younger female respondents in watching than the older female respondents.

9 Factors Affecting Decisions to Watch YouTube Videos

This chapter ties together the findings collected for RQ3-7 to address RQ8 (What influences the decision to watch a YouTube video?), relates them to prior literature, and proposes some conclusions relating to YouTube video watching preferences. The following discussions relate to the sample of data extracted from YouTube and the convenience sample respondents, and therefore are not generalizable to all UK YouTube users. Due to the limited nature of this thesis, the changing nature of popular culture and current events, and the growth, development and organic nature of YouTube there are influences that can have an impact on a user's decision-making process in watching videos that have not been investigated. These include personality type, nationality, culture, religious beliefs, socio-economic status, employment, online access, YouTube's algorithms, overlaps in categories and types of video, news, currents affairs, politics, fashions, disability and accessibility.

This thesis investigates factors associated with video popularity or that influence users' decisions to watch a video. This involves two scopes of popularity (in terms of view counts).

- **Individual videos**.
- **Categories of video**.

The two types overlap because a type of video would become more popular, on average, if there were popular individual videos of that type. The questionnaire data relates to the second type of popularity because users are not asked about individual videos. In contrast, the metadata has information primarily about individual videos and reports within category averages. For example, a high correlation between video length and view counts in the metadata would indicate that longer videos are more likely to be watched but not necessarily that viewers tend to watch long videos (because there might be many more short than long videos). This mismatch must be taken into account when comparing the survey and category API data.

9.1 Video Titles and Thumbnails

Survey users claimed that video titles and thumbnails helped them to decide whether to watch a video. This agrees with previous research into a sample of 306 videos (Chang, 2018). This is consistent with the content of a video being important in the decision to watch it.

9.2 Video Categories

There are substantial differences in the average popularity (views) of videos between categories from the YouTube API category search results (Figures 6.5, 6.6). The survey data also reported systematic differences in the categories watched (Figure 8.7). Thus, and unsurprisingly, different types of videos attract different average numbers of views. For example, the average popularity of From TV videos is many orders of magnitude higher than that of Non-profit videos. Whilst entertainment videos seem to be the most popular, on average, informational videos (How To, Beauty, Cooking) are moderately popular. Political videos (politics, causes) are perhaps surprisingly unpopular overall.

Comparing the survey and API results can give some additional insights, although the data are not directly compatible because the former covers any video in a category and the latter describes average properties of videos in a category but does not take into account the number of videos in the category (whether or not the number captured by the YouTube API sample is representative of the total number of videos available for the category).

There is broad agreement between the survey and YouTube API averages, in the sense of a very approximately linear main trend, but also five substantial discrepancies (Figure 9.1, 9.2).

- **Education**: The importance of education to YouTube is exaggerated in the survey because of the sample selection bias (primarily educators).

- **Entertainment** and **Comedy**: The mismatch seems likely to be due to interpretation by the survey users, regarding entertainment as a purpose rather than a genre, and comedy as meaning funny videos of any type rather than the Comedy category of comedians delivering stage shows.
- **Best Of**: The anomaly could be due to Best Of videos being selected for this category for tending to be popular already, and perhaps not attracting many new views from their Best Of listing. Alternatively, people may select Best Of videos only if they match another category (e.g., Comedy) and consider this as being selecting Comedy rather than Best Of.
- **From TV**: The anomaly here may be partly age-related, with younger viewers being prepared to watch the (in 2015) relatively low resolution and possibly illegally copied TV shows on YouTube.

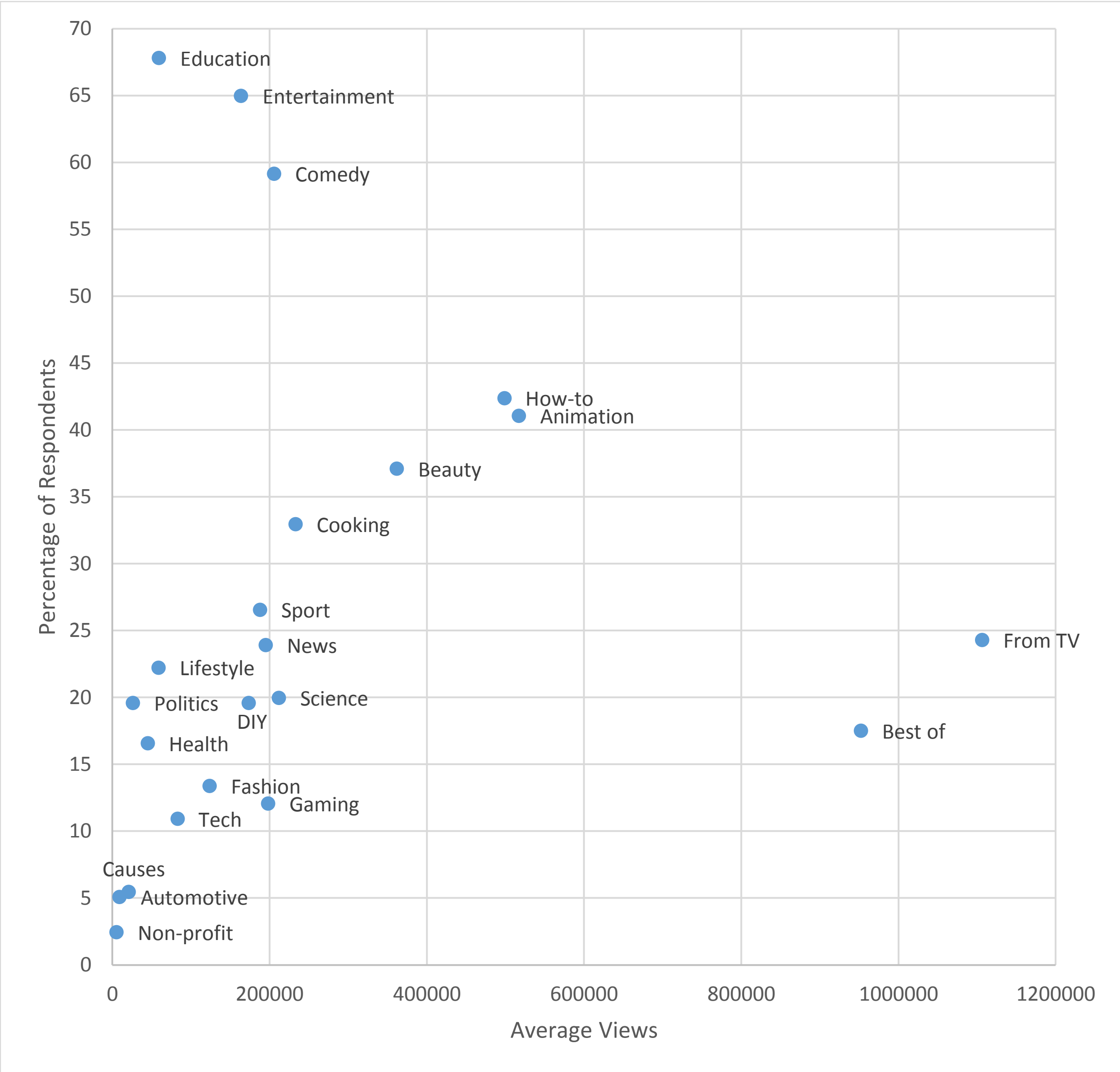


Figure 9.1. **Percentage of respondents watching videos from a category against average number of views per video from the YouTube API combined dataset.**

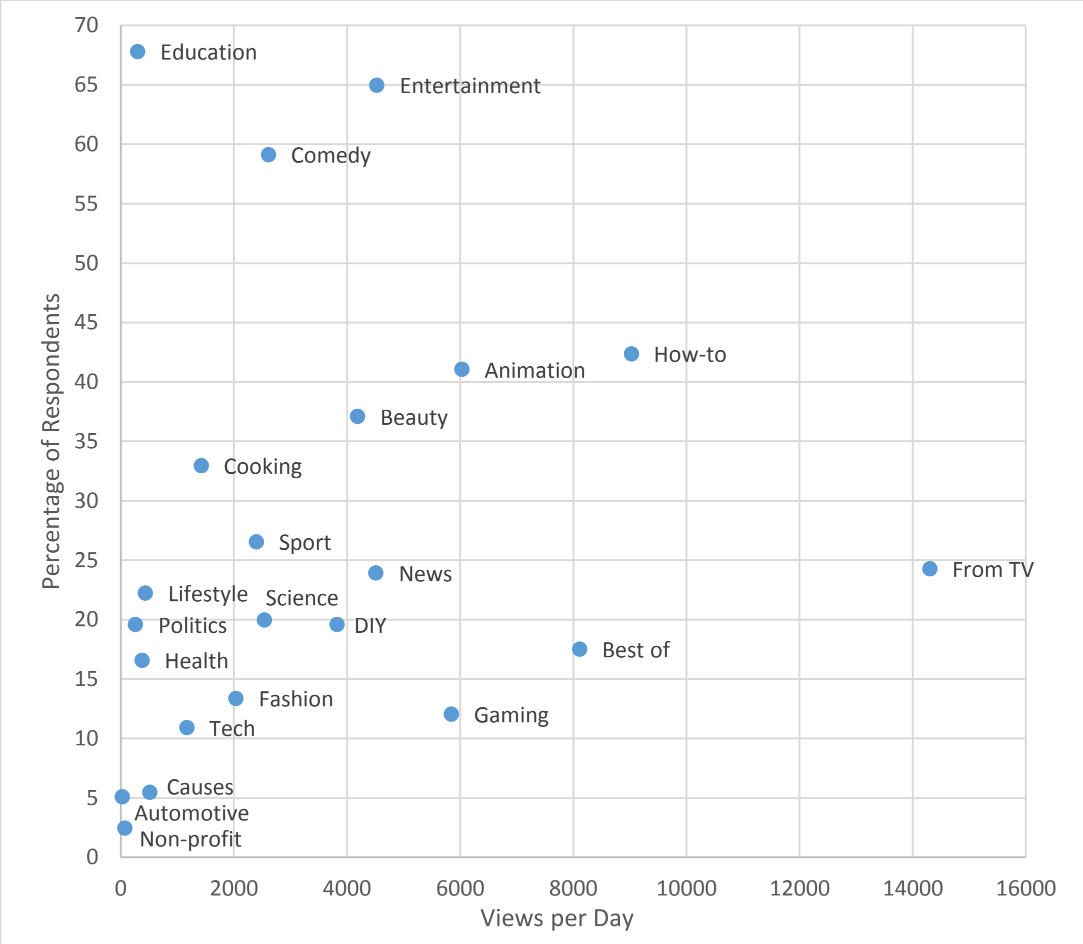


Figure 9.2. **Percentage of respondents watching videos from a category against average daily views per video from the YouTube API combined dataset.**

The results are difficult to compare with previous YouTube research, which have rarely explicitly used the YouTube categories. Nevertheless, they confirm that Education, Health and News are important on YouTube, helping to justify the many studies of these areas (see Chapter 2: e.g., Saurabh and Gautam, 2019; Madathil et al., 2015; Al-Rawi, 2019). The low rates for Causes videos are in agreement with a finding that end users don't use YouTube to seek sustainable development information (Kang, 2019), although there are many other causes. The results also give new evidence of the importance of Beauty. Although there is some YouTube Beauty research on this topic (e.g., Hou, 2018), it seems minimal compared to its importance within the site. Conversely, the relatively low numbers of views for Politics videos suggests that its importance is overplayed by the many studies about it (e.g., Bowyer et al., 2017). Since politics is often nation-specific, however, the YouTube view counts may underestimate its local importance.

Several video categories seem to have attracted little academic research focusing on them within YouTube, despite apparently being highly viewed on the site. There are multiple possible explanations for this.

- **From TV, Animation**: YouTube's importance may be primarily as a delivery platform for these, with research focusing on them in offline contexts. Previous studies have argued that watching TV-like content has been the core use of YouTube (Freeman and Chapman, 2017; Buzzetto-More, 2014; Cheng et al., 2013; Kellner and Kim, 2010; Purcell, 2010).

- **Gaming**: This aspect of gaming (help or advice videos and skill showcasing) may not be thought important to research. On YouTube, the potential to live stream games (Göring et al, 2019) may add to its attractiveness.
- **DIY, How To, Cooking**: These may cover aspects of life that are largely ignored by academic research.
- **Lifestyle, Fashion**: These may be primarily of interest on YouTube from a marketing perspective and offline research may be most interested in creativity.

9.3 Video Length

Video length varies between categories (Figure 6.25), because logically different purposes will have different durations. Video length does not have a statistically significant correlation with view counts in *any* of the categories from the YouTube metadata (Table 7.1) so there is no category in which users tend to prefer shorter or longer videos. This partially conflicts with many of the survey responses (Figure 8.8), with a quarter reporting that video length influences the decision to watch a video (Table 9.1). It is possible that viewers prefer different length videos for different purposes, even within a category, so that there is no consistent preference for a particular length in the category. Alternatively, the preferred length might be close to the average length for some or all categories so that there is not a positive overall rank correlation between length and popularity. These results conflict with prior claims that shorter videos are more popular (Tschopp, 2014; Henke, 2013).

Table 9.1. A comparison of the highest and lowest scoring responses from the questionnaire data for the female and male respondents.

	Female	Male
How often have you watched YouTube videos?	Daily **36%** Weekly **49%**	Daily **58%** Weekly **32%**
How have you accessed YouTube videos? (Highest)	YouTube **78%** Facebook **68%** App **61%** Google **52%**	YouTube **84%** App **74%** Facebook **72%** Google **62%**
How have you accessed YouTube videos? (Lowest)	Blog **9%** Twitter **19%**	Blog **22%** Verbal recommendation **29%**
Number of methods (Average)	3 Methods - **21%**	4 Methods - **22%**
Which methods through the YouTube website? (Highest)	Search facility **79%** Homepage **32%**	Search facility **79%** Homepage **49%** Subscriptions **39%**
Which methods through the YouTube website? (Lowest)	Most viewed **9%** Most popular **15%**	Most viewed **10%** Most popular **21%**
Which type of videos have you watched? (Highest)	Education **67%** Entertainment **64%** Comedy **56%** Beauty **44%** How-to **43%**	Comedy **75%** Education **71%** Entertainment **70%** Sport **58%** Politics **43%**
Which type of videos have you watched? (Lowest)	Non-profit **2%** Automotive **3%** Causes **5%** Tech **6%** Gaming **7%**	Beauty **5%** Fashion **5%** Non-profit **5%** Causes **6%** Health **14%**
Factors that influence your decision to watch - Always/Mostly (Highest)	Title **50%** Thumbnail **30%** Length **26%**	Title **48%** Thumbnail **35%** Length **25%**
Factors that influence your decision to watch - Always/Mostly (Lowest)	Comments **1%** Dislikes **3%** Upload date **8%**	Comments **0%** Dislikes **2%** Likes **8%** Upload date **8%**

9.4 Video Age

It is not possible to reliably detect the influence of age from the YouTube API data (since old videos may have attracted their views when young) although it seems likely that videos attract most of their views when they are relatively new. Similarly, news is time-dependant and pop music has short phases of maximum popularity when it is in the charts so these categories seem likely to be time-dependant. Under 10% of the survey respondents thought that video age (upload date) influenced their decision to watch a video (Figure 8.8 and Table 9.1). Overall, it seems possible that videos age has little direct influence on decisions to watch individual videos, although it must have at least an indirect influence for some video types. No previous research seems to have focused on video age as an influence on popularity, but it is implicit in many of the discussions that new content is likely to become viral than old content (e.g., Brodersen et al., 2012).

9.5 Popularity and Opinion-Related Factors

9.5.1 View counts

Only 15% of survey respondents believed that they were often influenced by video view counts in their decisions to watch (Figure 8.8). Few respondents were interested in popular video pages either (Figure 8.4 and Table 9.1). The YouTube metadata cannot provide evidence about the influence of view counts on view counts, since this would be tautological. The survey respondents might be biased towards avoiding the mainstream since they tend to be educated to degree level. Some people also actively seek to avoid being mainstream and may therefore avoid frequently viewed videos (Browne, 2013). Overall, however, the results provide weak evidence that popularity does not directly lead to further popularity through conscious user decisions.

Related to popularity, beauty vloggers sometimes attempt to make videos that follow popularity trends in the belief that their videos might be recommended to users watching popular videos (Bishop, 2019). Thus, there may be a trend for content creators to follow popularity even if viewers do not.

9.5.2 Video Dislikes and Likes

There is a substantial difference between categories in the average likes per view and a moderate difference between categories in dislikes per view (Figure 6.8). Categories with the most viewed videos tend to have the fewest likes and dislikes per view (compare Figures 6.5, 6.6 with Figure 6.8). A possible explanation is that passive videos are more popular and that the lower number of ratings is a side-effect of this.

Under 10% of the survey respondents reported being frequently influenced by the number of likes for a video and under 5% said that they were frequently influenced by dislikes (Figure 8.1 and Table 9.1). The small negative correlations between likes and dislikes per view and total views for most categories (Figure 7.1) suggests the possibility of a small negative influence: people are less likely to watch a video with many likes or dislikes. Nevertheless, a more plausible explanation is that viewers are less likely to want to rate a popular video because their individual rating would make little difference to its overall score. It is also possible that more popular videos within categories tend to be of a passive entertainment type. No evidence is available for either conclusion, however. Nevertheless, whatever the cause, neither likes nor dislikes have much influence on viewers. This agrees with previous results from a set of 306 YouTube videos (Chang, 2018).

9.5.3 Video Comments

The number of average and per view comments that videos receive again varies considerably across the different categories (Figures 6.12, 6.13. 6.14). When compared to the process of rating, through disliking and liking, the additional effort required to comment on a video and its content demonstrates a higher level of engagement from users.

Survey respondents were almost unanimous in not caring about the number of comments on a video when deciding whether to watch it (Figure 8.8). The small negative correlations between comments per view and total views for most categories (Figure 7.1), suggests that there may be a small indirect effect leading to less popular videos having more comments per video. This might be due to the greater ease of communication between commenters on videos with fewer comments because comments will be visible for longer before requiring scrolling to see. It might also be a side effect of more passive entertainment style videos in each category being more popular. Prior research has found that extensive commenting associates with controversial content (Chimel et al., 2011; Sobkowicz and Sobkowicz, 2010; Thelwall, Sud, and Vis, 2012), so controversial content may be less popular, rather than generating interest (see also: Eckler and Rodgers, 2011; Dahl et al., 2003).

9.5.4 Indirect Influences of Methods to Access Videos

The most popular method for accessing YouTube videos for all female and male respondents is through the YouTube website and App (Figure 8.4). Respondents might therefore be influenced by content on the website or App homepage. This content is selected and personalised (if the user is recognised) by YouTube's algorithms based on the user's previous watching history, with YouTube probably presenting videos that it thinks they might like, based on popularity and/or similarity to previously-watched content. Thus, **users may be indirectly influenced by the viewing habits of unknown others through the YouTube algorithm suggestions**. This power of the algorithm has been previously noted in a case study of beauty vloggers (Bishop, 2019). The results here therefore confirm the algorithmic beliefs of beauty vloggers.

Facebook is also a popular method for accessing videos across all female and male respondents and suggests that **YouTube users are influenced by the watching recommendations of their Facebook friends** (Figure 8.4). This could be due to them respecting, valuing and trusting those that they choose to be part of their online social groups and networks. Both female and male respondents listed methods that they were less likely to use to access YouTube videos with the least popular of these consistently being through hyperlinks within Blogs (Table 9.3). It seems that Blogs and Twitter are not popular methods for watching YouTube videos, perhaps be due to the more text-based nature of this social media platform.

Many users also search for videos through Google or the YouTube website. Both methods probably incorporate both popularity and personalisation to find a video that should match the need expressed by the query. Thus, unless a person searches for a previously known video, they may be indirectly influenced by popularity when selecting a video matching their query to watch.

Previous studies do not seem to have systematically analysed how the method to access YouTube influences the videos watched. Nevertheless, the findings relate to general research about influence and popularity by revealing two partially hidden mechanisms by which it can occur even for people who do not think that -or let- the opinions of others influence them.

9.5.5 Direct Influences of Methods to Access Videos

Across both genders, ages and user types the respondents are rarely accessing videos through the YouTube most viewed or most popular pages (Figure 8.4, Table 9.3). In contrast, respondents seem, through the popular use of the search facility, to know what they are looking for and therefore do not use popularity pages provided by YouTube. It also suggests that the respondents do not feel that they are interested in general YouTube popularity for its own sake. Nevertheless, due to the indirect influences discussed above, respondents could be influenced indirectly by popularity without realising it.

9.5.6 Virality

There has been extensive prior research discussing the extent to which people follow the crowd or display a herd mentality (Faafat et al., 2009; Toyokawa et al., 2019) as factors that may lead to videos becoming popular or viral. Both the survey and metadata evidence suggest that the opposite is the norm in terms of conscious behaviour. People rarely seem to watch a video because it is popular or liked although they are more likely to watch it because it is recommended by a trusted social network friend. Such friends are trusted partly because they have similar tastes or are good judges of what the recipient will like (Hayes et al., 2018; Feitosa and Botelho, 2017; Oh and Syn, 2015). Thus, the current research suggests that, on YouTube, the key determinant of virality or popularity is the value of the videos shared. In other words, a video is likely to become popular if it is valuable to the viewers, and sharing by friends (i.e., online word of mouth) may help them to find valuable videos. Thus, discussions of virality in prior research perhaps emphasise the mechanism for finding videos too much, because the key factor is video value. To give a concrete example, YouTube's most popular video, Gangnam Style, should be regarded as primarily an example of a valuable video that was able to reach a receptive audience through a variety of mechanisms, including viral sharing. It is probably not a mediocre video that became popular by accident through snowballing popularity, because the evidence suggests that this is unlikely to happen.

9.6 Viewer demographics

The decision to watch a video is influenced by the age and gender of the viewer (Fisher and Ha, 2018; Schwemmer and Ziewiecki, 2018). This is obvious in terms of video content because there are substantial gender differences in offline personal interests and information needs. Other factors also seem to have age or gender differences in the likelihood of watching YouTube videos, as revealed by the survey and previous research (Fisher and Ha, 2018; Schwemmer and Ziewiecki, 2018). This includes variations in the extent to which the factors discussed above influence watching behaviours (Table 9.2, 9.3).

Table 9.2. The **highest** scoring responses from the questionnaire data for the female and male age and user respondent groups.

	Watch daily or weekly	How have you accessed YT videos?	Number of methods (average)	Which methods through the YouTube website?	Which type of videos have you watched?	Factors that influence your decision to watch (Always/Mostly)
Female Age	18 to 24 - **93%** 25 to 34 - **91%** 35 to 44 - **82%** 45 to 54 - **70%** 55+ - **40%**	YouTube Facebook App Google	18 to 24 - 4 methods 25 to 54 - 3 methods 55+ - 1 method	Search facility Homepage	Education Entertainment Comedy How-to Animation Beauty (18 to 34) Fashion (18 to 24) Sport (45 to 54)	Title Thumbnail (18 to 44) Length (18 to 44)
Male Age	18 to 24 - **91%** 25 to 34 - **97%** 35 to 44 - **81%** 45 to 54 - **91%** 55+ - **50%**	YouTube Facebook App Google	18 to 24 - 6 methods 25 to 34 - 4 methods 35 to 44 - 2/3 methods 45 to 54 - 1 method 55+ - 4 methods	Search facility Homepage Subscriptions (25 to 34) Recommendations bar (18 to 24 and 35 to 44)	Comedy Education Entertainment Sport From TV (45 to 54) Gaming (25 to 34) Best of (18 to 24) How-to (35 to 44)	Title Thumbnail Length (18 to 24 and 45 to 54) Views (18 to 24)
Female Daily User	N/A	YouTube **81%** App **78%** Facebook **76%** Google **48%**	3 methods	Search facility **76%** Homepage **48%** Recommendations bar **35%** Subscriptions **34%**	Entertainment **76%** Education **72%** Comedy **63%** Beauty **59%** How-to **53%**	Title **54%** Thumbnail **39%** Length **30%**
Female Weekly User	N/A	YouTube **81%** Facebook **70%** App **56%** Google **53%**	3 methods	Search facility **81%**	Education **69%** Entertainment **62%** Comedy **56%** Animation **43%**	Title **49%** Thumbnail **28%**
Male Daily User	N/A	YouTube **86%** App **76%** Facebook **72%** Google **64%**	4 methods	Search facility **74%** Homepage **56%** Subscriptions **48%** Recommendations bar **46%**	Comedy **78%** Entertainment **72%** Education **64%** Sport **58%** Tech/Gaming **50%**	Title **52%** Thumbnail **32%**
Male Weekly User	N/A	YouTube **79%** App **79%** Facebook **75%** Google **61%**	6 methods	Search facility **89%** Homepage **50%**	Education **82%** Comedy **79%** Entertainment **71%** Sport **64%** Animation **57%**	Title **54%** Thumbnail **43%** Length **43%**

Table 9.3. The **lowest** scoring responses from the questionnaire data for the female and male age and user respondent groups.

	Rarest access method	Rarest YouTube website methods	Rarest types of videos watched	Factors least influencing decisions to watch
Female Age	Blog Twitter Verbal recommend. (35 to 44)	Most viewed Most popular Recommend. bar (35 to 44)	Non-profit Automotive Causes Gaming Tech Best of (35 to 54) Politics (18 to 24)	Comments Dislikes Likes Upload date
Male Age	Blog (18 to 24 and 35 to 44)	Most viewed Most popular (25+)	Causes Fashion (except 35 to 44) Automotive (except 25 to 34) Beauty Non-profit Lifestyle (25 to 34 and 44+) DIY (18 to 24 and 45 to 54)	Comments Dislikes Likes (25+) Upload date (except 25 to 34) Category (45+)
Female Daily User	Blog **13%** Twitter **23%**	Most viewed **14%**	Non-profit **4%** Automotive **5%** Causes **9%** Tech **9%** Gaming **12%**	Comments **0%** Dislikes **3%** Likes **9%**
Female Weekly User	Blog **6%** Twitter **19%**	Most viewed **8%** Most popular **11%** Subscriptions **12%**	Non-profit **1%** Automotive **2%** Causes **4%** Gaming **5%** Tech **5%**	Comments **1%** Dislikes **3%** Category **6%** Upload date **7%**
Male Daily User	Twitter **30%**	Most viewed **14%**	Beauty **4%** Causes **6%** Fashion **6%** Non-profit **6%** Health **14%**	Comments **0%** Dislikes **2%** Likes **10%**
Male Weekly User	Blog **11%**	Most viewed **7%** Most popular **18%**	Causes **4%** Fashion **4%** Non-profit **4%** Beauty **7%**	Comments **0%** Likes **4%** Dislikes **4%** Upload date **4%**

9.6.1 Age

There are some age-related differences in the ways in which YouTube is used, which may affect influences on decisions to watch videos. Ignoring males due to low numbers, younger female respondents tended to access videos in more ways and slightly more overall. There were not statistically significant differences for any methods, although the overall pattern was consistent with older users being more likely to directly search for videos rather than access them in other ways. Older females were also statistically significantly less likely to access Fashion and Beauty videos (Figure 8.13), confirming age differences in interests. Previous age-related discussions of YouTube seem to have focused on age differences in overall use rather than video watching preferences.

9.6.2 Gender

Due to the low number of male respondents, only tentative conclusions can be reached about gender differences, but in terms of categories watched, it seems likely that offline gender differences in interests have translated into different category preferences. Males also seem to use a greater range of methods to access YouTube, but it is not clear whether this would moderate the indirect influences discussed above. Prior research relating YouTube to gender has tended to discuss presenter gender (Wotanis and McMillan, 2014) rather than audience gender, but some studies have found substantial gender differences in interest depending on the topic of a video (Thelwall, 2018; Thelwall and Mas-Bleda, 2018).

9.7 Summary

As a reminder, the above discussion is based on the results of a biased survey sample (education-related adults from the UK) and a biased collection of YouTube videos (as selected by YouTube API category searches). This may affect the results but, except as otherwise discussed, does not seem likely to have had a substantial influence on the results.

Content-based factors (title, thumbnail, category) have the greatest influence on decisions to watch videos. Combined with the extensive use of Google and YouTube search facility to search for videos, this suggests that the content of a video is far more important than perceptions of others judgements of it. The main caveats to this are (a) that users are often prepared to follow suggestions from friends, who they may trust to recommend relevant content, and (b) users are probably indirectly influenced by anonymous other YouTube users through the algorithms that select videos for them to watch based on their search or previous watching history.

10 Conclusions

This thesis investigated the popularity of YouTube videos and factors that influence users to watch them. Although the findings are not representative of all UK YouTube users the results add to research about online video popularity, user behaviours and factors influencing the decision-making process. The conclusions are also relevant to people using and producing YouTube videos. As these individuals will understandably want to increase the popularity of their videos for reasons including self-promotion, earning money, marketing, information services and political election campaigns (Hou, 2019; Berryman and Kavka, 2017; Ahmed et al., 2013; Pinto et al., 2013; Kim, 2012; Ackerman and Guizzo, 2011). The primary conclusion for producers is that content is king: they should focus on the message rather than mechanisms to make videos appear to be popular, since the former is most relevant to users.

This chapter summarises the results of the individual research questions and then draws general conclusions.

10.1 YouTube API Sample

RQ1 was broken down into two further questions which are addressed here.

To what extent does the sample provided by the YouTube API differ over time and between YouTube categories?

There are substantial variations in the sample that the YouTube API selects over the two periods of time that the category searches were submitted. Although there are videos that have been repeated between subsequent searches, both daily and five-day approaches, the variations across these searches and categories suggests that the YouTube API selection process is not dominated by repeats. In addition, comparing subsequent searches to the initial search, again for daily and five-day approaches, there is substantial variation in the proportion of repeats, and these reduce at different rates across categories and searches. The reduction for some categories is more substantial than others, suggesting variations in how the YouTube API selects videos for different categories. The five-day search approach yielded a higher proportion of repeated videos when compared to the daily searches. Nevertheless, common sense suggests that having a longer period between searches would provide less repeats as the API would have a wider range of new content to take its selection from. This suggests that there may be a cyclic element to the way in which the YouTube API algorithm is applied, but this cannot be determined from watch collected within this thesis.

What factors and/or metrics, if any, does the YouTube API employ to determine the sample that it provides?

It was not possible to determine how the YouTube API algorithm is biased by all the metrics associated with videos when selecting a search sample. It was also not possible to know if these biases are altered depending on the category being searched. Nevertheless, the substantial variations in metric scores associated with each of the categories across both search methods suggest that the YouTube API is not dominated by any metric considered within this thesis (e.g. comments, likes, dislikes, length, view count) when selecting a sample of videos, with the partial exception of age. For example, none of the search results were particularly dominated by new videos (based on age) or popular videos (based on high view counts).

When analysing the associated average metrics relating to videos that are repeated across all the searches (daily or five day) due to the substantial variations across categories, these selections do not appear to be dictated or influenced by any specific level of metric. In addition, the same appears to be true when analysing the associated average metrics for the videos with no repeats.

Overall, in addressing RQ1 using category search results it is not possible to determine the algorithm that the YouTube API uses, how it works in what criteria it uses to select videos and any sample biases it may have. It is also not possible to explain how or if the YouTube API varies the way in which it applies its search algorithm to different categories. Nevertheless, it is clear that the YouTube API is not dominated by repeated videos or videos with specific levels of metrics, and that it returns a varied sample of videos, except that the video age variation is limited.

Thus, **the YouTube category API search results are biased by several factors, with the most substantial being video age (almost completely avoiding videos over a year old) but they include videos with a wide range of properties**, including length and popularity, but not age. **Research using the YouTube category API (including this thesis) should therefore not assume that it is a random sample**.

10.2 YouTube Video Metadata

RQ2: What is the age, length, and popularity of videos in each YouTube category and how do these vary between categories?

On average, **videos differ substantially between categories in their average lengths, view counts, like counts, dislike counts and comment counts**.

There are also substantial differences in the age of videos across the categories, suggesting that some content is updated more regularly but it is not clear whether this applies to the category as a whole or the video selection mechanism differing between categories. Categories that have a greater range of newer content might be addressing viewer needs, but this cannot be proven from the data collected.

RQ3: Which types of YouTube video are the most popular?

The average popularity of each category is in Figure 9.1 and Figure 9.2.

The category with the highest average overall and daily views is From TV, supporting a general trend that the **individually most watched videos tend to be high quality content professionally produced as (passive) entertainment**. **Informational videos are also individually highly viewed**, including How to, Cooking and Beauty.

RQ4: Which categories of YouTube video do users comment on most?

The average number of comments per video or per view in each category are in Figures 6.12, 6.13. and 6.14.

There is substantial variation in the number of comments per view between categories (Figure 6.13). Comments indicate a higher level of engagement with a video or a desire to communicate with its producers or other viewers.

Videos attracting the most comments per view cover participatory topics, such as Politics, Causes, DIY, How to, Gaming and Health.

RQ5: How does the length, like count, dislike count and comment count of a YouTube video relate to its popularity?

No video categories have a strong popularity preference to either short or long videos (Table 7.1).

In all categories, **videos with more views tend also to have more likes, dislikes and comments but in almost all categories, the ratio of likes, dislikes and comments per view is lower for more popular videos** (Figure 7.1).

The Animation exception to the above rule may reflect more directly quality-driven rating practices. The main two other exceptions, Automotive and Non-profit, may reflect failed self-help advice and controversial topics, respectively.

10.3 User Perspective

Due to the convenience sample used to get a low non-response bias from a high survey return (81% - 534/660), there is a sample selection bias towards people studying or working within the primary and higher education sectors. There is also a high proportion of female respondents. The survey broadly reflects the main US adult users of YouTube in terms of age and level of education.

RQ6: What are the main gender and age differences in the types of YouTube video that are the most popular?

There were too few male respondents to give clear answers to this question, although males seem to use a greater range of methods to access YouTube and may use it more regularly.

RQ7: Which factors influence the decision to watch a YouTube video for different genders and ages?

The main direct factors claimed by users to influence their decision to watch a video are its title and thumbnail (Figure 8.8). The category is also important (Figure 8.7). These are all content-related factors without substantial statistically significant age and gender differences. Most users are also prepared to accept recommendations from friends directly or via Facebook (Figure 8.10). All respondents, irrespective of gender or age, watch a substantial proportion of their YouTube videos through Facebook. These recommendations might be mediated by checks of video titles and thumbnails, however.

UK YouTube users claim to be most influenced by the content of a video when deciding whether to watch it. Most are also prepared to follow recommendations from online contacts, but few claim to be concerned by video popularity.

Users can be indirectly influenced in their watching decisions by the method that they use to find videos. The most popular method for finding and accessing YouTube videos for all respondents, gender, age and user level, is through the YouTube website and the App (with a slightly greater preference from younger respondents). For these, they will be presented with video suggestions from YouTube based on their viewing history. If, as seems likely, the YouTube selection algorithm takes popularity into account, then users may be indirectly influenced by popularity in what they watch by YouTube.

10.4 Overall Conclusions

RQ8: What influences the decision to watch a YouTube video?

Combining the survey and metadata results, the main finding is that users are rarely influenced by popularity-related information when deciding to watch a video on YouTube. Instead they are primarily concerned with the content of a video, as reflected by its category, title and thumbnail, although they are prepared to be influenced by recommendations from online interpersonal relationships.

An implication of the results is that marketing strategies to promote videos should focus on the content first, and, secondarily, seek to generate sharing through online relationships rather than focusing on attempting to generate popularity as a method to generate attention.

10.5 Further Research

This thesis has not investigated why categories or types of video are more popular with UK YouTube users, which is an important omission. In addition, the categories are broad and there may be elements within that are particularly important to YouTube users and would provide greater depth to an understanding in popularity. It would also be useful to try and determine what ways YouTube videos are used by viewers. It could be that the popularity of the categories or types of video may also be linked to the reasons why users watch online videos. Collecting and analysing the data relating to why people are choosing to access specific content could also provide a wider picture in why some videos are more popular than others. In addition, this could also provide a wider range of information relating to factors that influence user's decision making in watching videos.

Since subscribing to YouTube channels is relatively popular with younger and regular users, it would be useful to establish the influence of this practice. Developing a better understanding of which types of user subscribe to YouTube channels would also provide a more detailed picture in what types or categories of videos are more or most popular, and could be investigated in gender, age and activity level of subscriber. It would also have provided a greater depth of information to investigate the number of hours per day that respondents were spending interacting with YouTube. This research only focused on broader user levels, Yearly, Monthly, Weekly and Daily, so any further information relating to how much time people are spending watching and interacting with YouTube videos daily would have helped to identify key users and their preferences and influences. Finally, further research into users preferred video length and whether this changes based on video type, category or reason for watching would help support producers understanding of the more specific needs of viewers.

References

Abisheva, A., Garimella, V. R. K., Garcia, D. and Weber, I. (2014) *Who watches (and shares) what on YouTube? and when? Using Twitter to understand YouTube viewership.* In Proceedings of the 7th ACM international conference on web search and data mining, pp. 593-602. ACM.

Adams, B., Cha, M., Mislove, A. and Gummadi, K. (2008) *Characterizing social cascades in Flickr.* In WOSP'08, pp.13–18.

Adhikari, V., Jain, S., Chen, Y. and Zhang, Z. (2011) *How do you 'Tube'.* In Proceedings of the ACM SIGMETRICS joint international conference on Measurement and modelling of computer systems, pp. 137-138. ACM.

Adhikari, V., Jain, S., Chen, Y. and Zhang, Z. (2012) *Vivisecting YouTube: An active measurement study.* In 2012 Proceedings IEEE INFOCOM, pp.2521-2525. IEEE.

Ackerman, E. and Guizzo, E. (2011) *5 technologies that will shape the web*. IEEE Spectrum, 48(6).

Agazio, J. and Buckley, K. (2009) *An untapped resource. Using YouTube in nurse education*. Nurse Educator 34 (1), pp.23–28.

Agichtein, E., Castillo, C., Donato, D., Gionis, A. and Mishne, G. (2008) *Finding high-quality content in social media*. In Proceedings of the 2008 international conference on web search and data mining, pp.183-194. ACM.

Ahmed, Y. A., Ahmad, M. N., Ahmad, N. and Zakaria, N. H. (2019) *Social media for knowledge-sharing: A systematic literature review*. Telematics and informatics, 37, pp.72-112.

Ahmed, M., Spagna, S., Huici, F. and Niccolini, S. (2013) *A peek into the future: Predicting the evolution of popularity in user generated content*. In Proceedings of the sixth ACM international conference on web search and data mining, pp.607-616. ACM.

Ahn, D. and Shin, D. (2016) *Observers versus agents: divergent associations of video versus game use with empathy and social connectedness.* Information Technology and People, 29(3), pp.474-495.

Akdeniz, C. (2014) *Popularity: Who and What is Best Liked and Why*. California, USA: On-Demand Publishing LLC.

Amarasekara, I. and Grant, W. J. (2019) *Exploring the YouTube science communication gender gap: A sentiment analysis*. Public Understanding of Science, 28(1), pp.68-84.

Alassad, M., Agarwal, N. and Hussain, M. N. (2019) *Examining intensive groups in YouTube commenter networks*. In International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation (pp. 224-233). Springer, Cham.

Alexa. (2019) *Top 500 Sites*. [online] Available at: http://www.alexa.com/topsites Accessed: 28/12/19.

Alexander, J. and Losh, E. (2010) *A YouTube of one's own? Coming out videos as rhetorical action*. LGBT identity and online new media, pp.37-50.

Alloway, T. and Alloway, R. (2012) *The impact of engagement with social networking sites (SNSs) on cognitive skills.* Computers in Human Behavior, 28(5), pp.1748-1754.

Allsop, T., Bassett, B. and Hoskins, J. (2007) *Word-of-Mouth Research: Principles and Applications*. Journal of Advertising Research, 47(4), pp.388-411.

Almind, T. and Ingwersen, P. (1997) *Informetric analyses on the World Wide Web: Methodological approaches to 'Webometrics'.* Journal of Documentation, 53(4), pp.404-426.

Almobarraz, A. (2018) *Utilization of YouTube as an information resource to support university courses*. The Electronic Library.

Alonzo, M. and Aiken, M. (2004) *Flaming in electronic communication*. Decision Support Systems, 36(3), pp.205-213.

Al-Rawi, A. (2019) Viral news on social media. Digital Journalism, 7(1), pp.63-79.

Ameigeiras, P., Ramos-Munoz, J., Navarro-Ortiz, J. and Lopez-Soler, J. (2012) *Analysis and modelling of YouTube traffic.* Trans Emerg Telecommun Technol 23(4), pp.360–377.

Anderson, G. (1998) *Fundamentals of Educational Research (2nd Edition).* London: Taylor and Francis.

Anderson, C., Kelter, D. and John, O. (2003) *Emotional Convergence between People over Time*. Journal of Personality and Social Psychology, 84, pp.1054-1068.
Anderson, L. and Holt, C. (1996) *Information Cascades*. Journal of Economic Perspectives, Vol. 10, No. 4 (Fall 1996).
Andone, I., Błaszkiewicz, K., Eibes, M., Trendafilov, B., Montag, C. and Markowetz, A. (2016) *How age and gender affect smartphone usage*. In Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. pp.9-12. ACM.
Andress, L. (2020) *Best Time to Post on YouTube in 2020: How to Net the Biggest Audience*. How Sociable. [online] Available at: https://howsociable.com/blog/best-time-to-post-on-youtube/ Accessed: 14/5/20
Anthony, L., Kim, Y. and Findlater, L. (2013) *Analyzing user-generated YouTube videos to understand touchscreen use by people with motor impairments.* In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp.1223-1232. ACM.
Arens, E. (2020) *The best times to post on social media in 2020. SproutSocial*. [online] Available at: https://sproutsocial.com/insights/best-times-to-post-on-social-media/ Accessed: 9/5/20
Armstrong, C., Fenton, R., Lonsdale, R., Stoker, D., Thomas, R. and Urquhart, C. (2001) *A study of the use of electronic information systems by higher education students in the UK*. Program: Electronic Library and Information Systems 35 (3), pp.241–262.
Arndt, H. L. and Woore, R. (2018) *Vocabulary learning from watching YouTube videos and reading blog posts*. Language Learning & Technology, 22(3), pp.124-142.
Arthur, J., Waring, M., Cole, R. and Hedges, L. (2017) *Research Methods and Methodologies in Education (2nd Edition)*. London: Sage Publications Ltd.
Arthurs, J., Drakopoulou, S. and Gandini, A. (2018) *Researching YouTube.* Convergence, 24(1), pp.3–15.
Ashraf, B. (2009) *Teaching the Google-eyed YouTube generation.* Education+ Training, 51(5/6), pp.343-352.
Ashton, D. and Patel, K. (2020) *Girls' Bedroom Cultures*. The International Encyclopedia of Gender, Media, and Communication, pp1-6.
Ayres, C. (2009) *Revenge is best served cold on YouTube.* The Times Online [online] Available at: http://www.thetimes.co.uk/tto/law/columnists/article2051377.ece Accessed: 2/1/15.
Aytar, Y., Pfaff, T., Budden, D., Paine, T., Wang, Z. and de Freitas, N. (2018) *Playing hard exploration games by watching YouTube*. In Advances in Neural Information Processing Systems, pp.2930-2941.
Baker, M. and Gammon, D. (2008) *Vocal Memes in Natural Populations of Chickadees: Why do some memes persist and others go extinct?* Animal Behavior, 75, pp.279-289.
Baker, P. and Ward, A. (2002) *Bridging temporal and spatial 'gaps': The role of information and communication technologies in defining communities.* Information, Communication and Society, 5(2), pp.207–224.
Bakshy, E., Hofman, J., Mason, W. and Watts, D. (2011) *Everyone's an influencer: Quantifying influence on twitter.* Fourth International Conference on Web Search and Data Mining, 2011.
Bala, M. and Verma, D. (2018) *A critical review of digital marketing.* International Journal of Management, IT and Engineering, 8(10), pp.321-339.
Balbay, S. and Kilis, S. (2017) *Students' Perceptions of the use of a YouTube channel specifically designed for an Academic Speaking Skills Course.* Eurasian Journal of Applied Linguistics, 3(2), pp.235-251.
Bandura, A. (1989) *Human agency in social cognitive theory*. American psychologist, 44(9), 1175.
Bandura, A. (1997) *Self-efficacy: The exercise of control.* New York: Worth Publishers.
Banerjee, A. (1992) *A simple model of herd behavior.* Quarterly Journal of Economics, 107, pp.797–817, 1992.
Banet-Weiser, S. (2017) *'I'm beautiful the way I am': empowerment, beauty, and aesthetic labour*. In Aesthetic labour (pp. 265-282). Palgrave Macmillan, London.

Barakat, C. (2014) *How Bad Content Goes Viral*. SocialTimes.
Baranova, S. and Trofymenko, A. (2018) *Suggestive material as an aspect of stylistic implication of English advertising discourse (based on American TV commercials).* Філологічні трактати, (10,№ 3), pp.15-19.
Barber, D., Cooper, L. and Meeson, G. (2007) *Learning and Teaching with Interactive Whiteboards: Primary and Early Years*. London: Learning Matters.
Bardzell, J., Bardzell, S. and Pace, T. (2008) *Emotion, Engagement and Internet Video.* One to One Interactive. [online] Available at: http://www.scribd.com/doc/8948430/emotion-engagement-and-internet-video Accessed: 12/4/14.
Bar-Ilan, J. (2002) *Methods for measuring search engine performance over time.* Journal of the American Society for Information Science and Technology, 53(4), pp.308-319.
Bar-Ilan, J. and Peritz, B. (2004) *Evolution, continuity, and disappearance of documents on a specific topic on the web: A longitudinal study of "informetrics".* Journal of the American Society for Information Science and Technology, 55(11), pp.980-990.
Barjak, F. and Thelwall, M. (2008) *A statistical analysis of the web presences of European life sciences research teams*. Journal of the American Society for Information Science and Technology, 59(4), pp.628-643.
Barker, V. (2009) *Older adolescents' motivations for social network site use: The influence of gender, group identity, and collective self-esteem*. CyberPsychology and Behavior, 12(2), pp.209–213.
Barr, P. (2020) *The Impact of the World Wide Web on 21st Century Elections in the US.* Scholarly and Creative Works Conference, 2020.
Barry, A., Johnson, E., Rabre, A., Darville, G., Donovan, K. and Efunbumi, O. (2014) *Underage access to online alcohol marketing content: a YouTube case study.* Alcohol and alcoholism, 50(1), pp.89-94.
Bass, F. (1969) *A new product growth for model consumer durables.* Management Science, 15(5), pp.215-227, 1969.
Baumgartner, J. and Morris, J. (2010) *MyFaceTube politics: Social networking websites and political engagement of young adults*. Social Science Computer Review, 28(1), pp.24-44.
Baym, N. and Burnett, R. (2009) *Amateur experts: International fan labour in Swedish independent music*. International journal of cultural studies, 12(5), pp.433-449.
Beauchamp, G. (2012) *ICT in the Primary School: From Pedagogy to Practice*. London: Routledge.
Becker, H. (1998) *Tricks of the Trade: How-to think about your research while doing it.* Chicago and London: University of Chicago Press.
Belanger, F. and Jordan, D. H. (2020) *Cybertools for Instructors: Using the Internet and the World Wide Web to Enhance Learning in the Classroom*. Journal of Information Systems Education, 9(1), 1.
Bell, J. (2014) *Doing Your Research Project: A Guide for First-time Researchers in Education, Health and Social Science (6th Edition).* Maidenhead: Open University Press.
Bellman, S., Schweda, A. and Varan, D. (2009) *A comparison of three interactive television ad formats*. Journal of Interactive Advertising, 10(1), pp.14-34.
Benevenuto, F., Rodrigues, R., Almeida, V., Almeida, J. and Ross, K. (2008) *Identifying video spammers in online social networks.* in AIRWeb '08: Proc. of the 4th international workshop on Adversarial information retrieval on the web, 2008.
Benevenuto, F., Rodrigues, R., Almeida, V., Almeida, J. and Ross, K. (2009a) *Video interactions in online video social networks.* ACM Trans. Multimedia Comput. Commun. Appl., vol. 5, no. 4, 2009.
Benevenuto, F., Rodrigues, R., Almeida, V., Almeida, J. and Conclaves, M. (2009b) *Detecting spammers and content promoters in online video social networks*. in SIGIR '09: Proc. of the 32nd international conference on Research and development in information retrieval, 2009.

Bentley, F., Silverman, M. and Bica, M. (2019) *Exploring Online Video Watching Behaviors*. In Proceedings of the 2019 ACM International Conference on Interactive Experiences for TV and Online Video (June), pp.108-117. ACM.
Berger, J. and Milkman, K. L. (2010) *Social transmission and viral culture.* Department of Marketing, University of Pennsylvania.
Berger, J. and Milkman, K. (2012) *What Makes Online Content Viral?* Journal of Marketing Research 49:2, pp.192-205.
Berger, J. and Milkman, K. L. (2013) *Emotion and virality: what makes online content go viral?* GfK Marketing Intelligence Review, 5(1), pp.18-23.
Berk, R. (2009) *Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom.* International Journal of Technology in Teaching and Learning, 5(1), pp.1–21.
Berry, T. (2018) *Videoblogging before YouTube*. (Theory on Demand; No. 27). Amsterdam: Institute of Network Cultures.
Berryman, R. and Kavka, M. (2018) *Crying on YouTube: Vlogs, self-exposure and the productivity of negative affect*. Convergence, 24(1), pp.85-98.
Berryman, R. and Kavka, M. (2017) '*I guess a lot of people see me as a big sister or a friend': The role of intimacy in the celebrification of beauty vloggers*. Journal of Gender Studies 26(3): pp.307–320.
Bent, M. (2010) *The Pareto principle on YouTube*. In 12th Twente Student Conference on IT.
Bi, N. C., Zhang, R. and Ha, L. (2019) *Does valence of product review matter? The mediating role of self-effect and third-person effect in sharing YouTube word-of-mouth (vWOM)*. Journal of Research in Interactive Marketing, 13(1), pp.79-95.
Bikhchandani, S., Hirschleifer, D. and Welch, I. (1992) *A Theory of Fads, Fashion, Custom, and Cultural Change in Information Cascades.* Journal of Political Economy, 100 (1992), pp.992-1026.
Billig, M. (2005) *Laughter and Ridicule: Towards a Social Critique of Humor*. London: Sage Publications Ltd.
Binns, R., Veale, M., Van Kleek, M. and Shadbolt, N. (2017) *Like trainer, like bot? Inheritance of bias in algorithmic content moderation*. In International Conference on Social Informatics – September 2017. Springer, Cham. pp.405-415.
Bishop, S. (2018) *Anxiety, panic and self-optimization: Inequalities and the YouTube algorithm*. Convergence, 24(1), pp.69-84.
Bishop, S. (2019). *Managing visibility on YouTube through algorithmic gossip*. New Media & Society. 21(11-12) 2589 –2606.Blank, G., Dutton, W. and Lefkowitz, J. (2019). *Perceived Threats to Privacy Online: The Internet in Britain (Oxford Internet Survey 2019 report).* Oxford: Oxford Internet Institute
Blau, P. (1964) *Exchange and power in social life.* New York: John Wiley and Sons.
Blaxter, L., Hughes, C. and Tight, M. (2010) *How-to Research (4th Edition).* Maidenhead: Open University Press.
Brodersen, A., Scellato, S. and Wattenhofer, M. (2012) YouTube around the world: geographic popularity of videos. In Proceedings of the 21st international conference on World Wide Web, pp. 241-250.
Borghol, Y., Ardon, S., Carlsson, N., Eager, D. and Mahanti, A. (2012) *The untold story of the clones: content-agnostic factors that impact YouTube video popularity.* In Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining. pp.1186-1194.
Borghol, Y., Mitra, S., Ardon, S., Carlsson, N., Eager, D. and Mahanti, A. (2011) *Characterizing and Modelling Popularity of User-Generated Videos.* Performance Evaluation, 68, pp.1037–1055.
Bou-Franch, P., Lorenzo-Dus, N. and Blitvich, P. (2012) *Social interaction in YouTube text-based polylogues: A study of coherence*. Journal of Computer-Mediated Communication, 17(4), pp.501-521.
Boyd, D. (2014) *It's Complicated*. London: Yale University Press.

Bowyer, B. T., Kahne, J. E. and Middaugh, E. (2017) Youth comprehension of political messages in YouTube videos. New Media & Society, 19(4), pp.522-541.
Bozkurt, H., Özer, S., Şahin, S. and Sönmezgöz, E. (2018) *Internet use patterns and Internet addiction in children and adolescents with obesity*. Pediatric obesity. 13(5), pp.301-306.
Bradshaw, T. and Garrahan, M. (2008) *Rival forecast to catch YouTube. Financial Times*. [online] Available at: http://us.ft.com/ftgateway/superpage.ft Accessed: 7/4/15.
Brahmana, R. and Vivaldo, R. (2018) *Fast-fashion Social Campaign Advertisement in YouTube: Does Brand Storytelling Impacted Behaviour Intention to Use?* International Journal of Business and Society, 19(2), pp.523-535.
Branch, J., Bartholomew, P., Nygaard, C. and Paulsen, M. (2015) *Technology-Enhanced Learning in Higher Education.* Oxfordshire: Libri Publishing.
Brodersen, A., Scellato, S. and Wattenhofer, M. (2012) *YouTube around the world: geographic popularity of videos*. WWW (2012)241-250. Lyon, April 2012.
Brown, J. and Reingen, P. (1987) *Social Ties and Word of Mouth Referral Behaviour*. Journal of Consumer Research, 14, pp.350-362.
Brown, J., Broderick, A. and Lee, N. (2007) *Word of Mouth Communications Within Online Communities: Conceptualizing the Online Social Network*. Journal of Interactive Marketing, Vol. 21, No. 3, Summer 2007.
Brown, M., Bhadury, R. and Pope, N. (2010) *The Impact of Comedic Violence on Viral Advertising Effectiveness*. Journal of Advertising, 39, (1), pp.49-65.
Broxton, T., Interian, Y., Vaver, J. and Wattenhofer, M. (2013) *Catching a viral video.* In IEEE Data Mining Workshops, pp.296–304, Sydney, Australia.
Bruno, A. (2010) *Break on Through*. Billboard, Volume 122 (Issue 46), November 20, pp.10-13.
Bryman, A. (2004) *Social Research Methods (2nd Edition).* Oxford: Oxford University Press.
Bucher, T. (2017) *The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms*. Information, Communication & Society, 20(1), pp.30-44.
Bulakh, V., Dunn, C. W. and Gupta, M. (2014) *Identifying fraudulently promoted online videos*. In Proceedings of the 23rd International Conference on World Wide Web. pp.1111-1116.
Burgess, J. (2014) *'All Your Chocolate Rain Are Belong To Us?' Viral Video, You Tube and the Dynamics of Participatory Culture.* In Art in the Global Present, pp.86-96. UTSePress.
Burgess, J. and Green, J. (2008) '*Agency and Controversy in the YouTube Community*,' paper presented at Association of Internet Researchers (AoIR) IR 9.0: Rethinking Communities, Rethinking Place, 16-18 October, URL (consulted April 2009). [online] Available at: http://eprints.qut.edu.au/15383/ Accessed: 16/5/16.
Burgess, J. and Green, J. (2009) *YouTube: Online video and participatory culture.* Cambridge: Polity.
Burgess, J. and Green, J. (2013) *YouTube: Online video and participatory culture*. John Wiley and Sons.
Burgess, S., Sanderson, E. and Umaña-Aponte, M. (2011) *School ties: An analysis of homophily in an adolescent friendship network.* The Centre for Market and Public Organisation, Paper No 11/267.
Burke, S. and Snyder, S. (2008) *YouTube: An innovative learning resource for college health education courses*. International Electronic Journal of Health Education, 11, pp.39-46.
Burke, S., Snyder, S. and Rager, R.C. (2009) *An assessment of faculty usage of YouTube as a teaching resource.* The Internet Journal of Allied Health Sciences and Practice 7 (1), pp.1–8.
Burnett, D. (2013) *Criticising popular things: why is it so popular?* The Guardian [online] Available at: https://www.theguardian.com/science/brain-flapping/2013/dec/24/criticising-popular-things-why-is-it-so-popular Accessed: 6/11/14.
Buzzetto-More, N. (2012) *Social networking in undergraduate education.* Interdisciplinary Journal of Information, Knowledge, and Management, 7(1), pp.63-90.

Buzzetto-More, N. (2014) *An examination of undergraduate student's perceptions and predilections of the use of YouTube in the teaching and learning process*. Interdisciplinary Journal of E-Learning and Learning Objects, 10(1), pp.17-32.
Cai, Z., Fan, X. and Du, J. (2017) *Gender and attitudes toward technology use: A meta-analysis*. Computers & Education, 105, pp.1-13.
Carbonell, X., Chamarro, A., Oberst, U., Rodrigo, B. and Prades, M. (2018) *Problematic use of the internet and smartphones in university students: 2006–2017*. International journal of environmental research and public health. 15(3), 475.
Cardine, S. (2008) *Is education ready for YouTube?* Center for Digital Education/Converge. Available at: http://www.centerdigitaled.com/edtech/Is-Education-Ready-For-YouTube.html Accessed: 13/5/17.
Cardwell, M., Clark, L. and Meldrum, C. (1996) *Psychology for A-Level.* London: Collins Education.
Carvalho, N. and Gomes, I. (2020) *Knowledge sharing between enterprises of the same group*. In Information Diffusion Management and Knowledge Sharing: Breakthroughs in Research and Practice (pp. 403-423). IGI Global.
Cashmore, P. (2009) *YouTube: Why Do We Watch?* CNN (Dec 17, 2009).
Cayari, C. (2011) *The YouTube Effect: How YouTube Has Provided New Ways to Consume, Create, and Share Music.* International Journal of Education & the Arts, 12(6), n6.
Celis, P. (2013) *When Is a YouTube Video a True Threat*. Wash. JL Tech. & Arts, 9, p.227.
Centola, D., Gonza´lez-Avella, J., Eguı´luz, V. and Miguel, M. (2007) *Homophily, Cultural Drift, and the Co-Evolution of Cultural Groups.* Journal of Conflict Resolution, Volume 51 Number 6, December (2007) pp.905-929.
Cha, M., Kwak, H., Rodriguez, P., Ahn, Y,-Y. and Moon, S. (2007) *I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System*. in Proc. of ACM IMC, 2007.
Cha, M., Kwak, H., Rodriguez, P., Ahn, Y.-Y. and Moon, S. (2009) *Analyzing the video popularity characteristics of large-scale user generated content systems*. IEEE/ACM Transactions on Networking, 17(5), pp.1357-1370.
Chang, W. L. (2018) *Will sentiments in comments influence online video popularity?* In 2018 IEEE International Conference on Big Data (Big Data) (pp. 3644-3646). IEEE Press.
Chiang, H. S. and Hsiao, K. L. (2015) *YouTube stickiness: the needs, personal, and environmental perspective*. Internet Research. Vol. 25 Iss 1. pp.85-106
Chappell, B. (2012) *Social network analysis report*—Demographic—Geographic and search data revealed. Ignite Social Media Blog.
Chatzopoulou, G., Sheng, C. and Faloutsos, M. (2010) *A first step towards understanding popularity on YouTube.* In Proceedings of the IEEE Conference on Computer Communications (INFOCOM), 2010.
Chen, Y., Wang, Q. and Jinhong, X. (2011) *Online Social Interactions: A natural Experiment on Word of Mouth Versus Observational Learning.* Journal of Marketing Research, Vol. XLVIII, April (2011), pp.238-254.
Chen, Y-F. (2007) *Herd behaviour in Purchasing Books Online.* Philadelphia, PA: Elsevier Ltd.
Chenail, R. (2011) *YouTube as a qualitative research asset: Reviewing user generated videos as learning resources*. The Qualitative Report, 16(1), pp.229-235.
Cheng, K. (2013) *10 ways to make your video go viral.* Medium [online] Available at: https://medium.com/this-happened-to-me/10-ways-to-make-your-video-go-viral-d19d9b9465de Accessed: 9/5/20
Cheng, X., Dale, C. and Liu, J. (2008) *Statistics and social network of YouTube videos*. IWQoS ’08: Proc. of the 16th International Workshop on Quality of Service, June 2008.
Cheng, X., Dale, C. and Liu, J. (2007) *Understanding the Characteristics of Internet Short Video Sharing: YouTube as a Case Study.* Technical report, Cornell University, arXiv e-prints.

Cheng, X., Liu, J. and Dale, C. (2013) *Understanding the characteristics of internet short video sharing: A YouTube-based measurement study*. IEEE Transactions on Multimedia, 15(5), pp.1184-1194.
Chiang, H. S. and Hsiao, K. L. (2015) *YouTube stickiness: the needs, personal, and environmental perspective.* Internet Research. Vol. 25 Iss 1. pp.85-106
Chin, H., Hsieh, Y., Kao, Y. and Lee, M. (2007) *The Determinants of Email Receivers' Disseminating Behaviours on the Internet*. Journal of Advertising Research, 47(4), December, pp.524-534.
Chmiel, A., Sobkowicz, P., Sienkiewicz, J., Paltoglou, G., Buckley, K., Thelwall, M. and Hołyst, J. A. (2011) *Negative emotions boost user activity at BBC forum*. Physica A: statistical mechanics and its applications, 390(16), pp.2936-2944.
Choudhury, S., Breslin, J. and Passant, A. (2009) *Enrichment and Ranking of the YouTube Tag Space and Integration with the Linked Data Cloud.* Proceedings of the 8th International Semantic Web Conference, October 25-29, 2009, Chantilly, VA.
Chowdhury, S. and Makaroff, D. (2013) *Popularity Growth Patterns of YouTube Videos*. A Category-based Study. In WEBIST, pp.233-242.
Christensen, W. (2013) *Moving worked problems to YouTube*. The Physics Teacher, 51(8), pp.500-502.
Cialdini, R. (2009) *Influence: Science and Practice (5th Edition).* New York: Harper Collins.
Cialdini, R. and Goldstein, N. (2004) *Social Influence: Compliance and Conformity.* Annual Review of Psychology, 55(2004), pp.591-621.
Cisco Corporation. (2012) *Cisco Visual Networking Index: global mobile traffic forecast update*, 2011-2016. White Paper [online] Available at: http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11520862.pdf Accessed: 14/5/15.
Citron, D. (2009) *Law's expressive value in combating cyber gender harassment*. Michigan law review, pp.373-415.
Clark-Carter, D. (2010) *Quantitive Psychological Research.* Hove: Psychology Press.
Clifton, A. and Mann, C. (2011) *Can YouTube enhance student nurse learning?* Nurse Education Today 31, pp.311–313.
Coates, A. E., Hardman, C. A., Halford, J. C., Christiansen, P. and Boyland, E. J. (2019). Social Media Influencer Marketing and Children's Food Intake: A Randomized Trial. Pediatrics, 143(4), e20182554.
Codreanu, T. and Combe, C. (2019) *Vlogs, Video Publishing, and Informal Language Learning*. The Handbook of Informal Language Learning, pp.153-168.
Cohen, L., Manion, L. and Morrison, K. (2017) *Research Methods in Education (8th Edition)*. Oxon: Routledge.
Constant, D., Kiesler, S. and Sproull, L. (1994) *What's mine is ours, or is it? A study of attitudes about information sharing*. Information systems research, 5(4), pp.400-421.
Cooke, R. (2008) *"On-line innovation in higher education".* [online] Available at: http://fm.schmoller.net/files/online_innovation_in_he_131008.pdf Accessed: 12/11/17.
Coolican, H. (2014) *Research Methods and Statistics in Psychology*. London: Psychology Press.
Cox, D. (1963) *The Audience as Communicators*. In S.A.Greyser (Ed), Proceedings, American Marketing Association, December, pp.58-72. Chicago: American Marketing Association.
Crane, R. and Sornette, D. (2008) Viral, *Quality, and junk videos on YouTube: Separating content from noise in an information-rich environment.* In Proceedings of the AAAI Spring Symposium, 2008.
Crane, R. and Sornette, D. (2008a) *Robust dynamic classes revealed by measuring the response function of a social system*. PNAS, 105(41):15649{15653, 2008.
Crane, R., Schweitzer, F. and Sornette, D. (2010) *Power law signature of media exposure in human response waiting time distributions.* Physical ReviewE, 81(5): 056101, 2010.
Crawshaw, J. and Chambers, J. (2002) *A Concise Course in Advanced Statistics – Fourth Edition.* Cheltenham: Nelson Thornes.

Cremin, T. and Arthur, J. (2014) *Learning to teach in the Primary School*. London: Routledge.
Creswell, J. (2014) *Educational Research: Pearson New International Edition: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Pearson Educational Limited: Essex.
Crick, M. (2016) *Power, Surveillance, and Culture in YouTube's Digital Sphere*. Hershey, PA: Information Science Reference/IGI Global.
Cuevas, A. and Kohle, F. (2010) *A case study in using YouTube and Facebook as social media tools in enhancing student centred learning and engagement*. In 3rd International Conference of Education, Research and Innovation (Iceri2010), pp.3596-3601.
Cumming, G. and Calin-Jageman, R. (2016) *Introduction to the New Statistics: Estimation, Open Science, and Beyond.* Routledge Francis and Taylor Group: Ney York and London
Cumming, G. and Finch, S. (2005) *Inference by Eye: Confidence Intervals and How to Read Pictures of Data.* American Psychologist, Vol 60(2), Feb-Mar 2005, pp.170-180
Cunningham, S. and Craig, D. (2016) *Online entertainment: A new wave of media globalization?* International Journal of Communication, 10, pp.5409-5425.
Cunningham, S., Craig, D. and Silver, J. (2016) *YouTube, multichannel networks and the accelerated evolution of the new screen ecology*. Convergence, 22(4), pp.376–391.
Cunningham, S. J. and Nichols, D. M. (2008) *How people find videos*. In Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries, pp.201-210. ACM.
Cutler, M (2009) *How-to Make Your Online Video Go Viral.* Advertising Age, 80 (March), 42.
Dahl, D., Frankenberger, K. and Manchanda, R. (2003) *Does it Pay to Shock? Reactions to Shocking and Non-shocking Advertising Content Among University Students.* Journal of Advertising Research, 43, 3, September, pp.268-280.
Daly, J. (2011) *The 10 Laws of Viral Video Success*. Defuse. [Online] Available at: http://www.defuse.com/blog/the-10-laws-of-viral-video-success/ Accessed: 5/4/14.
Davidson, J., Liebald, B., Liu, J., Nandy, P., Van Vleet, T., Gargi, U. and Sampath, D. (2010) *The YouTube video recommendation system*. In Proceedings of the fourth ACM conference on Recommender systems, pp.293-296. ACM.
Dawkins, R. (1976) *The Selfish Gene.* Oxford: Oxford University Press.
De Vany, A. and Lee, C. (2008) *Information Cascades in Multi-Agent Models*. California: University of California.
Dehghani, M., Niaki, M. K., Ramezani, I. and Sali, R. (2016) *Evaluating the influence of YouTube advertising for attraction of young customers*. Computers in Human Behavior, 59, pp.165-172.
Denscombe, M. (2010) *The Good Research Guide for Small-scale Research Projects (4th Edition).* Maidenhead: Open University Press.
Dimopoulos, G., Barlet-Ros, P. and Sanjuas-Cuxart, J. (2013) *Analysis of YouTube user experience from passive measurements.* In Proceedings of the 9th International Conference on Network and Service Management (CNSM 2013), pp.260-267. IEEE.
Ding, Y., Du, Y., Hu, Y., Liu, Z., Wang, L., Ross, K. and Ghose, A. (2011) *Broadcast Yourself: Understanding YouTube Uploaders.* In ACM IMC 2011, pp.361–370, Berlin, Germany.
Ding, Y., Jacob, E., Zhang, Z., Foo, S., Yan, E., George, N. and Lijiang Guo, L. (2009) *Perspectives on social tagging*. Journal of the American Society for Information Science & Technology, 60(12), pp.2388-2401.
Döring, N. and Mohseni, M. R. (2019) *Male dominance and sexism on YouTube: results of three content analyses*. Feminist Media Studies, 19(4), pp.512-524.
Dwoskin, E., Whalen, J. and Cabato, R. (2019) *Content moderators at YouTube, Facebook and Twitter see the worst of the web — and suffer silently*. Washington Post [online] https://www.washingtonpost.com/technology/2019/07/25/social-media-companies-are-outsourcing-their-dirty-work-philippines-generation-workers-is-paying-price/ Accessed: 14/5/20
Dryden, L. (2016) *Here are the UK's top 10 most influential female YouTubers*. We the Unicorns: London. [online]. Available at: https://www.wetheunicorns.com/news/how-much-do-youtubers-make/ (Accessed: 2/12/19).

Duarte, F., Benevenuto, F., Almeida, V. and Almeida, J. (2007) *Geographical Characterization of YouTube: A Latin American View.* in Fifth Latin American Web Congress – 2007.
Duffett, R., Petroșanu, D. M., Negricea, I. C. and Edu, T. (2019) Effect of YouTube Marketing Communication on Converting Brand Liking into Preference among Millennials Regarding Brands in General and Sustainable Offers in Particular. Evidence from South Africa and Romania. Sustainability, 11(3), 604.
Duffy, P. (2008) *Engaging the YouTube Google-eyed generation: Strategies for using Web 2.0 in teaching and learning.* Electronic Journal of E-learning, 6(2), pp.119-130.
Duggan, M. and Brenner, J. (2013) *The demographics of social media users*, 2012 (Vol. 14). Washington, DC: Pew Research Center's Internet and American Life Project.
Duncan, I., Yarwood-Ross, L. and Haigh, C. (2013) *YouTube as a source of clinical skills education*. Nurse education today, 33(12), pp.1576-1580.
Duval, E., Sharples, M. and Sutherland, R. (2016) *Technology Enhanced Learning: Research Themes*. Switzerland: Springer.
Dynel, M. (2014) *Participation framework underlying YouTube interaction*. Journal of Pragmatics, 73, pp.37-52.
Easley, D. and Kleinberg, J. (2010) *Networks, Crowds and Markets: Reasoning about a Highly Connected World.* Cambridge: Cambridge University Press.
Eckler, P. and Bolls, P. (2011) *Spreading the Virus: Emotional Tone of Viral Advertising and its Effect on Forwarding Intentions and Attitudes.* Journal of Interactive Advertising, Vol. 11 (No.2), Spring. pp.1-11.
Eckler, P. and Rodgers, S. (2010) *Viral Advertising: A Conceptualization.* Paper presented at the annual meeting of the Association for Education in Journalism and Mass Communication, Denver, CO.
Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R. and Kwan, I. (2002) *Increasing response rates to postal questionnaires: systematic review*. British Medical Journal, 324, pp.1183-211.
Eger, S. (2015) *Opinion dynamics and wisdom under in-group bias and out-group discrimination*. Available at SSRN 2596546. 2015 Apr 20.
Eisentraut, R., Koch, M. and Möslein, K. (2001) *Building trust and reputation in communities and virtual enterprises*. AMCIS 2001 Proceedings, 291.
Ekeh, P. (1974) *Social exchange theory: The two traditions*. Boston: Harvard University Press.
Elven, M. (2018) *What do Millennials and Generation Z consumers want from retailers*? Fashion United: Amsterdam. [online] Available at: https://fashionunited.uk/news/business/what-do-millennials-and-generation-z-consumers-want-from-retailers/2018060730094 Accessed: 19/2/19.
Emerson, R. (1976) *Social exchange theory*. Annual review of sociology, 2(1), pp.335-362.
Entertainment Software Association. (2016) *Sales, Demographic, and Usage Data: Essential Facts about the Computer and Video Game Industry 2016*. January, 9, 2017.
Faafat, R., Chater, N. and Frith, C. (2009) *Herding in Humans.* Trends in Cognitive Sciences, 13(10), pp.420-428.
Fägersten, K. B. (2017) *The role of swearing in creating an online persona: The case of YouTuber PewDiePie*. Discourse, Context & Media, 18, pp.1-10.
Farnan, J., Paro, J., Higa, J., Edelson, J. and Arora, V. (2008) *The YouTube generation: implications for medical professionalism*. Perspectives in Biology and Medicine 51 (4), pp.517–524.
Featherstone, M., Adam, S. and Borstorff, P. (2009) *A methodology for extracting quasi-random samples from World Wide Web domains*. Journal of International Business Research, Volume 8, Number 1, pp.63 -67.
Feitosa, W. R. and Botelho, D. (2017) *Online brand content sharing on social networks – an experiment assessing the role of emotions on news and advertising online sharing*. In Becerra,

E., Chitturi, R., Henriquez Daza, M. and Londoño Roldan, J. (Vol. Eds.), Latin American advances in consumer research: Vol. 4, pp.155–158.
Feng, Y., Chen, H. and He, L. (2019) Consumer Responses to Femvertising: A Data-Mining Case of Dove's "Campaign for Real Beauty" on YouTube. Journal of Advertising, pp.1-10.
Ferchaud, A., Grzeslo, J., Orme, S. and LaGroue, J. (2018) Parasocial attributes and YouTube personalities: Exploring content trends across the most subscribed YouTube channels. Computers in Human Behavior, 80, pp.88-96.
Ferreira, N. (2020) *The Best Time to Post on Social Media in 2020*. Oberlo. [online] Available at: https://www.oberlo.com/blog/best-time-post-social-media Accessed: 9/5/20
Field, A. and Hole, G. (2003) *How-to Design and Report Experiments*. London: Sage Publications Ltd.
Fein, S. and Spencer, S. (1997) *Prejudice as a Self-Esteem Maintenance: Affirming the Self through Derogating Others.* Journal of Applied Social Pyschology, 31, pp.576-603.
Figueiredo, F., Almeida, J. and Benevenuto, F. (2014) *Does Content Determine Information Popularity in Social Media? A Case Study of YouTube Videos' Content and their Popularity.* CHI'14, April 26–May 1, 2014, Toronto, Canada.
Figueiredo, F., Almeida, J., Goncalves, M. and Benevenuto, F. (2014a) *On the Dynamics of Social Media Popularity: A YouTube Case Study*. ACM Trans. Internet Technol. 1, 1, Article 01.
Figueiredo, F., Benevenuto, F. and Almeida, J. (2011) *The Tube over time: characterizing popularity growth of YouTube videos.* WSDM'11, pages 745{754, New York, NY, USA, 2011. ACM.
Finamore, A., Mellia, M., Munafo, M., Torres, R. and Rao, S. (2011) *YouTube Everywhere: Impact of Device and Infrastructure Synergies on User Experience*. IMC'11, November 2-4, Berlin, Germany.
Fisher, A. and Ha, L. (2018) *What Do Digital Natives Watch on YouTube?* The Audience and Business of YouTube and Online Videos, 29.
Fiske, J. (1987) *Television Culture*. London: Metheun.
Frank, T. (1998) The conquest of cool: Business culture, counterculture, and the rise of hip consumerism. Chicago, IL: University of Chicago Press.
Frasco, S. (2014) *Three Ways to Increase Your Video Views.* Social Media Today.
Freeman, B. and Chapman, S. (2007) *Is "YouTube" telling or selling you something? Tobacco content on the YouTube video-sharing website*. Tobacco control, 16(3), pp.207-210.
Frey, A. (2007) *To Sue or Not to Sue: Video-Sharing Web Sites*, Copyright Infringement, and the Inevitability of Corporate Control.
Friedman, B., Khan Jr, P. and Howe, D. (2000) *Trust online*. Communications of the ACM, 43(12), pp.34-40.
Friedman, T. (2005) *The World is Flat: A Brief History of the Globalized World in the Twenty-First Century.* New York, NY: Farrar, Straus and Giroux Publishing.
Gable, P. and Harmon-Jones, E. (2010) *The Blues Broaden, but the Nasty Narrows: Attentional Consequences of Negative Affects Low and High in Motivational Intensity.* Psychological Science, 21, pp.211-215.
Galeotti, A., Ghiglino, C. and Squintani, F. (2013) *Strategic Information Transmission in Networks*. Journal of Economic Theory
Gantz, J. (2007) *A Forecast of Worldwide Information Growth Through 2010.* IDC Information and Data: MA.
Gao, W., Tian, Y., Huang, T. and Yang, Q. (2010) Vlogging: A survey of videoblogging technology on the web. ACM Computing Surveys (CSUR), 42(4), 15.
Garcia, D., Mendez, F., Serdult, U. and Schweitzer, F. (2012) *Political polarization and popularity in online participatory media.* In PLEAD, pp3-10, Nov 2012.
García Rapp, F. (2016) *The Digital Media Phenomenon of YouTube Beauty Gurus: The Case of Bubzbeauty*. International Journal of Web Based Communities. 12. pp.360-375.
Garrett-Wright, D. and Abell, C. (2011) *Using YouTube to bridge the gap between baby boomers and millennials.* Journal of Nursing Education 50 (5), pp.299–300.

Gaunt, K. (2015) *YouTube, Twerking and You: Context Collapse and the Handheld Co-Presence of Black Girls and Miley Cyrus.* Journal of Popular Music Studies, 27(3), pp.244-273.
Gerhards, C. (2019) Product placement on YouTube: An explorative study on YouTube creators' experiences with advertisers. Convergence, 25(3), pp.516-533.
Gielen, M. (2015) *Want To Know The Best Days And Times To Post YouTube Videos? Here's A Yearly Calendar*. TubeFilter [online] Available at: https://www.tubefilter.com/2015/01/12/best-days-times-to-post-youtube-videos-yearly-calendar/ Accessed: 9/5/20
Gill, P., Arlitt, M., Zongpeng Li, Z. and Mahanti, A. (2008) *Characterizing user sessions on YouTube*. MMCN 2008: Proc. of the 15th SPIE/ACM Multimedia Computing and Networking, 2008.
Gill, P., Arlitt, M., Li, Z. and A. Mahanti. (2007) *YouTube traffic characterization: A view from the edge.* In Proceedings of the ACM Internet Measurement Conference (IMC), 2007.
Gillespie, T. (2015) *Platforms intervene*. Social Media & Society, 1(1), pp.1–2.
Gilly, M., Graham, J., Wolfinbarger, M. and Yale, L. (1998) *A Dyadic Study of Interpersonal Information Search.* Journal of the Academy of marketing Science, Vol. 26, Spring, pp.83-100.
Giustini, D. (2006) *How Web 2.0 is changing medicine*. BMJ 2006; 333:1283–4.
Gladwell, M. (2000) *The Tipping Point: How Little Things Can Make a Big Difference*. New York: Little Brown.
Godwin-Jones, R. (2007) *Digital Video Update: YouTube, flash, high-definition*. [online] Available at: http://www.allbusiness.com/technology/4051526-1.html Accessed: 17/4/15.
Golub, B. and Jackson, M. (2009) *How Homophily Affects Learning and Diffusion in Networks.* Q J Econ 127(3), pp.1287–1338.
Gomm, R. (2009) *Key Concepts in Social Research Methods*. Hampshire: Palgrave Key Concepts.
Goncalves, M., Almeida, J., Santos, L., Laender, A. and Almeida, V. (2010) *On popularity in the blogosphere.* IEEE Internet Computing, 14:30-37, 2010.
Goobie, G. C., Guler, S. A., Johannson, K. A., Fisher, J. H. and Ryerson, C. J. (2019) YouTube videos as a source of misinformation on idiopathic pulmonary fibrosis. Annals of the American Thoracic Society, 16(5), pp.572-579.
Google (2019) *YouTube Partner Programme overview and eligibility*. Google [online] Available at: https://support.google.com/youtube/answer/72851?hl=en-GB Accessed: 9/5/20
Göring, S., Rao, R. R. R. and Raake, A. (2019) nofu—A lightweight no-reference pixel based video quality model for gaming content. In 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX) pp.1-6. IEEE Press.
Graham, T. (2018) *Platforms and hyper-choice on the World Wide Web*. Big Data & Society, 5(1), 2053951718765878.
Greenberg, A. and Zanetis, J. (2012) *The impact of broadcast and streaming video in education.* Report Commissioned by Cisco Systems Inc. to Wainhouse Research L.L.C. [online] Available at: http://www.cisco.com/web/strategy/docs/education/ciscovideowp.pdf Accessed: 13/10/14.
Greenberg, D. (2010) *3 Things Any Video Needs to go Viral.* [Online] Available at: http://mashable.com/2010/10/19/viral-video-science Accessed 5/4/14. Accessed: 13/10/14.
Greenaway, S., Thelwall, M. and Ding, Y. (2009) *Tagging YouTube: A classification of tagging practice on YouTube*. In Proceedings of the 12th International Conference on Scientometrics and Infometrics. pp. 660-664.
Greig, A., Taylor, J. and Mackay, T. (2007) *Doing Research with Children (2nd Edition).* London: Sage Publications Ltd.
Griffiths, M., Davies, M. and Chappell, D. (2004) *Demographic factors and playing variables in online computer gaming*. CyberPsychology & Behavior, 7(4), pp.479-487.
Grossman, L. (2006) *Best invention: YouTube*. Time Magazine.
Guadagno, R., Cialdini, R. and Evron, G. (2010) *What About Estonia? A Social Psychological Analysis of the First Internet War.* CyberPsychology, Behaviour and Social Networking, 13, pp.447-453.

Guadagno, R., Rempala, D., Shannon, M. and Okdie, B. (2013) *What makes a video go viral? An analysis of emotional contagion and internet memes.* Computers in Human Behaviour, 29(2013) 2312-2319.
Guarino, A. and Cipriani, M. (2009) *Herd Behaviour in Financial Markets: an Experiment with Financial Professionals.* Journal of the European Economic Association, 7 (2009), pp.206-233.
Gueorguieva, V. (2008) *Voters, MySpace, and YouTube: The impact of alternative communication channels on the 2006 election cycle and beyond*. Social Science Computer Review, 26(3), pp.288-300.
Gulati, G. (2010) *No laughing matter: The role of new media in the 2008 election*. The year of Obama: How Barack Obama won the White House, pp.187-204.
Gupta, A. (2015) *What is the psychological reason people enjoy watching Fail videos?* Quora. [online] Available at: https://www.quora.com/What-is-the-psycholgical-reason-people-enjoy-watching-Fail-videos Accessed: 8/5/17.
Haigh, C. (2011) *Wikipedia as an evidence source for nursing and healthcare students.* Nurse education today, 31(2), pp.135-139.
Hallinan, B. and Striphas, T. (2016) *Recommended for you: The Netflix Prize and the production of algorithmic culture*. New Media & Society, 18(1), pp.117–137.
Halpern, D. and Gibbs, J. (2013) *Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression.* Computers in Human Behavior, 29(3), pp.1159-1168.
Halvey, M. and Keane, M. (2007) *Exploring social dynamics in online media sharing*. In Proceedings of the 16th international conference on World Wide Web (pp.1273-1274). ACM.
Hang, W. and Yun, S. (2008) *How User-Generated Content (UGC) campaign changes electoral politics?* Korea Observer, 39 (Autumn), pp.369-406.
Hanson, G. and Haridakis, P. (2008) *YouTube users watching and sharing the news: A uses and gratifications approach.* Journal of Electronic Publishing, 11(3).
Hanson, W. and Putler, D. (1996) *Hits and Misses: Herd Behaviour and Online Product Popularity.* Marketing Letters, Volume 7, Issue 4, October (1996), pp.297-305.
Haran, B. and Poliakoff, M. (2012) *The periodic table of videos.* Science, 332, 1046-1047.
Haridakis, P. and Hanson, G. (2009) *Social interaction and co-viewing with YouTube: Blending mass communication reception and social connection*. Journal of Broadcasting and Electronic Media, 53(2), pp.317-335.
Hart, C. (2005) *Doing Your Masters Dissertation*. London: Sage Publications Ltd.
Hartley, J. (2004) *The 'value chain of meaning' and the new economy*. International journal of cultural studies, 7(1), pp.129-141.
Haryanto, N. and Suwito, K. A. (2020) *Gender Identity Construction of Male Beauty Vloggers on YouTube*. Journal of Talent Development and Excellence, 12(2s), pp.171-177.
Hayes, J. L., Shan, Y. and King, K. W. (2018) *The interconnected role of strength of brand and interpersonal relationships and user comment valence on brand video sharing behaviour*. International Journal of Advertising, 37(1), pp.142-164.
Haythornthwaite, C. (1999) *A Social Network Theory of Tie Strength and Media Use: A Framework for Evaluating Multi-level Impacts of New Media.* Technical Report UIUCLIS -2002/1+DKRC, Graduate School of Library and Information Science. Champaign, IL: University of Illinois at Urbana-Champaign.
Headley, M.G., and Piano Clark, V.L (2020) *Multilevel Mixed Method Research Designs: Advancing a Refined Definition*. Journal of Mixed Methods Research, 14, 2, pp.145-163
Hearn, I. (2020) *Research: The best times to post on social media in 2020*. Impactbnd [online] Available at: https://www.impactbnd.com/blog/new-research-whats-the-best-time-to-post-on-social-media-in-2 Accessed: 9/5/20
Heath, C., Bell, C. and Sternberg, E. (2001) *Emotional Selection in Memes: The Case of urban Legends.* Journal of Personality and Social Psychology (2001), 81(6) pp.1028-1041.

Heath, D. (1996) *Modest Interventions.* In Downey, G. and Dumit, J. (Eds) Cyborgs and Citadels, 67-82. School of American Research: Santa Fe.
Heffernan, V. (2009) *Uploading the avant-garde*. The New York Times, September, 3.
Henke, L. (2013) *Viral Marketing and Music Videos: When Shock Tactics Backfire*. Academy of Marketing Studies Journal, 2013.
Herzberg, F. (2008) *One more time: How do you motivate employees?* Boston: Harvard Business Review Press.
Herzberg, F., Mausner, B. and Snyderman, B. (1993) *The motivation to work.* Piscataway, NJ: Transaction Publishers.
HESA (2018) *Higher education student statistics: UK, 2016/17–Qualifications achieved*. Higher Education Statistics Agency [online] Available at: https://www.hesa.ac.uk/news/11-01-2018/sfr247-higher-education-student-statistics/qualifications#:~:text=In%20July%20to%20September%202017,%25%20(see%20figure%2016) Accessed: 1/8/20
Hills, P., Argyle, M. and Reeves, R. (2000) *Individual differences in leisure satisfactions: An investigation of four theories of leisure motivation.* Personality and Individual Differences, 28(4), pp.763-779.
Hill, S. R., Troshani, I. and Chandrasekar, D. (2020) *Signalling effects of vlogger popularity on online consumers*. Journal of Computer Information Systems, 60(1), pp.76-84.
Hilner, J. (2012) *How-to use online video in your classroom. How teachers can bring the best of YouTube and other online video services to their students*. Edutopia. [online] Available at: http://www.edutopia.org/youtube-educational-videos-classroom Accessed: 17/6/16.
Hitchcock, J.H., and Onwuegbuzie, A.J (2020) *Developing Mixed Methods Crossover Analysis Approaches.* Journal of Mixed Methods Research, 14, 1, pp.63-83.
Holahan, C. (2006) *Tech Review: AOL Video: Close But No TiVo*. BusinessWeek On-Line.
Holland, M. (2016) *How YouTube developed into a successful platform for user-generated content.* Elon journal of undergraduate research in communications, 7(1).
Hopkins, J. (2006) *Surprise! There's a third YouTube co-founder.* USA Today. [online] Available at: http://usatoday30.usatoday.com/tech/news/2006-10-11-YouTube-karim_x.htm Accessed: 3/1/15.
Horton, F. (1983) *Information literacy vs. computer literacy*. Bulletin of the American Society for Information Science, 9(4), pp.14-16.
Hou, M. (2018) Social media celebrity and the institutionalization of YouTube. Convergence, 1354856517750368.
Hsu, C. and Lin, J. (2008) *Acceptance of blog usage: The roles of technology acceptance, social influence and knowledge sharing motivation.* Information and management, 45(1), pp.65-74.
Huang, J., Duan, Z., Kwok, J., Binns, S., Vera, L. E., Kim, Y. and Emery, S. L. (2019) Vaping versus JUULing: how the extraordinary growth and marketing of JUUL transformed the US retail e-cigarette market. Tobacco control, 28(2), pp.146-151.
Huang, Y., Li, Z., Liu, G. and Dai, Y. (2011) *Cloud download: using cloud utilities to achieve high-quality content distribution for unpopular videos*. In Proceedings of the 19th ACM international conference on Multimedia, pp.213-222. ACM.
Huberman, B. and Adamic, L. (1999) *Growth dynamics of the World-Wide Web.* California, USA: Xerox Palo Alto Research Center.
Huberman, B., Pirolli, P., Pitkow, J. and Lukose, R. (1998) *Strong Regularities in World Wide Web Surfing - Report*. Science 3 – Vol. 280 pp.95-97.
Huberman, B., Romero, D. and Wu, F. (2009) *Crowdsourcing, attention and productivity*. Journal of Information Science, 35(6), pp.758-765.
Hunt, E. (2017) *LGBT community anger over YouTube restrictions which make their videos invisible*. The Guardian. [online] Available at:

https://www.theguardian.com/technology/2017/mar/20/lgbt-community-angerover-youtube-restrictions-which-make-their-videos-invisible Accessed: 14/5/20

Hussain, M. N., Tokdemir, S., Agarwal, N. and Al-Khateeb, S. (2018) *Analyzing disinformation and crowd manipulation tactics on YouTube*. In 2018 (August) IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM). pp.1092-1095. IEEE.

Hwang, A., Kessler, E. and Francesco, A. (2004) *Student networking behavior, culture, and grade performance. An empirical study and pedagogical recommendations.* Academy of Management Learning and Education, 3, pp.139–150.

Ilana, E. and Mikos, L. (2017) Young and creative: Digital technologies empowering children in everyday life. Nordicom-University of Gothenburg.

Ingram, P. and Morris, M. (2007) *Do People Mix at Mixers? Structure, Homophily, and the "Life of the Party".* Administrative Science Quarterly, 52, pp.558-585.

Jaakkola, M. (2020) *Useful Creativity: Vernacular Reviewing on the Video-Sharing Platform Vimeo*. Culture Unbound: Journal of Current Cultural Research, pp.1-20.

Jackson, M. (2008) *Social and Economic Networks.* Princeton: Princeton University Press.

Jackson, M. (2010) *An Overview of Social Networks and Economic Applications*. Handbook of Social Economics, edited by Benhabib, J., Bisin, A. and Jackson, M.

Jaffar, A. (2012) *YouTube: an emerging tool in anatomy education*. Anatomy Sciences Education 5 (3), pp.158–164.

Jang, S. H. (2011) *YouTube as an innovative resource for social science research.* In Australian Association for Research in Education Conference (AARE 2011 Conference), pp.1-16.

Jarboe, G. (2013) *YouTube Algorithm Change: 'Time Watched' Key to Higher Video Search Rankings*. Search Engine Watch: New York. [online] Available at: https://www.searchenginewatch.com/2012/10/22/youtube-algorithm-change-time-watched-key-to-higher-video-search-rankings/ Accessed: 1/12/19.

Jarrett, K. (2008) *Beyond broadcast yourself: the future of YouTube*. Media International Australia, 126(1), pp.132-144.

Jarvenpaa, S. and Staples, D. (2000) *The use of collaborative electronic media for information sharing: an exploratory study of determinants.* The Journal of Strategic Information Systems, 9(2-3), pp.129-154.

Jenkins, H. (2006) *Convergence Culture: Where Old and New Media Collide*. New York: New York University Press.

Jenkins, H., Li, X., Krauskopf, A. and Grean, J. (2009) *If It Doesn't Spread, It's Dead (Part Seven): Aesthetic and Structural Strategies*. [online] Available at: http://henryjenkins.org/blog/2009/02/if_it_doesnt_spread_its_dead_p_6.html Accessed: 12/12/17.

Jiang, L., Miao, Y., Yang, Y., Lan, Z. and Hauptmann, A. (2014) *Viral Video Style: A Closer Look at Viral Videos on YouTube.* In ACM International Conference on Multimedia Retrieval (ICMR). Glasgow, United Kingdom.

Johnson, B. and Christensen, L. (2012) *Educational Research (4th Edition)*. Los Angeles, CA: Sage Publication Ltd.

Johnson, M. R. and Woodcock, J. (2019) *The Impacts of Live Streaming and Twitch.tv on the Video Game Industry*. Media, Culture and Society, 41(5): pp.670-688.

Joinson, A. (2008) *Looking at, looking up or keeping up with people? Motives and use of Facebook*. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, pp.1027-1036. ACM.

Jondet, N. (2008) *The Silver Lining in Dailymotion's Copyright Cloud*. Juriscom.net. April, 19.

Jones, G. and Schieffelin, B. (2009) *Talking text and talking back: "My BFF Jill" from boob tube to YouTube.* Journal of Computer-Mediated Communication, 14(4), pp.1050-1079.

Jones, N. and Graham, C. (2013) *Practices and tools in online course delivery.* Learning management systems and instructional design: Metrics, standards, and applications, pp.288-302.

Juhasz, A. (2009) *Learning the five lessons of YouTube: After trying to teach there, I don't believe the hype*. Cinema Journal, 48(2), pp.145-150.

Kabooha, R. and Elyas, T. (2018) *The Effects of YouTube in Multimedia Instruction for Vocabulary Learning: Perceptions of EFL Students and Teachers*. English Language Teaching, 11(2), pp.72-81.

Kadir, S. A., Lokman, A. M. and Muhammad, M. (2018) *Identification of positive and negative emotion towards political agenda videos posted on YouTube*. In International Conference on Kansei Engineering & Emotion Research (2018, March). pp.758-767. Springer, Singapore.

Kang, S. (2019) Communicating sustainable development in the digital age: The relationship between citizens' storytelling and engagement intention. Sustainable Development, 27(3), pp.337-348.

Kardas, M. and O'Brien, E. (2018) *Easier seen than done: Merely watching others perform can foster an illusion of skill acquisition*. Psychological science, 2018, Feb 16.

Karpf, D. (2010) *Macaca Moments Revisited: Electoral Panopticon or Netroots Mobilization?* Journal of Information Technology and Politics, 7, pp.143-162.

Katona, Z. and Sarvary, M (2008) *Network Formation and the Structure of the Commercial World Wide Web.* Marketing Science Vol. 27, No. 5, September-October 2008, pp.764-778.

Kattimani, S., Manogar, M. and Naik, R. (2010) *Webometrics.* Sixth International Conference on Webometrics, Informetrics and Scientometrics & Eleventh COLLNET Meeting, October 19 – 22, 2010, University of Mysore.

Katz, E. and Lazarsfeld, P. (1955) *Personal influence; the part played by people in the flow of mass communications*. Free Press, Glencoe, I11, 1955.

Kayumovich, K. O. and Annamuradovna, F. S. (2020) *The main convenience of internet marketing from traditional marketing*. Academy, 1, 52.

Keelan, J., Pavri-Garcia, V., Tomlinson, G. and Wilson, K. (2007) *YouTube as a source of information on immunization: a content analysis.* jama, 298(21), pp.2482-2484.

Keller, E. and Berry, J. (2003) *The influential: One American in Ten Tells the Other Nine How-to Vote, Where to Eat, and What to Buy.* New York, NY: Free Press.

Kellner, D. and Kim, G. (2010) *YouTube, critical pedagogy, and media activism*. The Review of Education, Pedagogy, and Cultural Studies, 32(1), pp.3-36.

Kessler, F. and Schäfer, M. (2009) *Navigating YouTube: Constituting a hybrid information management system*. The YouTube Reader, 274.

Khan, M. (2017) *Social media engagement: What motivates user participation and consumption on YouTube?* Computers in Human Behavior, 66, pp.236-247.

Khan, M. and Solomon, J. (2013) *Advocacy, Entertainment and News—An Analysis of User Participation on YouTube.* Available at SSRN 2579596.

Kim, J. (2012) *The institutionalization of YouTube: From user-generated content to professionally generated content*. Media, Culture & Society, 34, pp.53–67.

Kim, D., Kim, J. H. and Nam, Y. (2014) *How does industry use social networking sites? An analysis of corporate dialogic uses of Facebook, Twitter, YouTube, and LinkedIn by industry type*. Quality and Quantity, 48(5), pp.2605-2614.

Kitchin, R. (2016) *Thinking critically about and researching algorithms*. Information, Communication & Society 20(1), pp.14–29.

Klobas, J. E., McGill, T. J., Moghavvemi, S. and Paramanathan, T. (2018) *Compulsive YouTube usage: A comparison of use motivation and personality effects.* Computers in Human Behavior, 87, pp.129-139.

Klobas, J. E., McGill, T. J., Moghavvemi, S. and Paramanathan, T. (2019) *Problematic and extensive YouTube use: first hand reports.* Online Information Review, 43(2), pp.265-282.

Knobel, M. and Lankshear, C. (2007) *Online Memes, Affinities and Cultural Production.* In: A New Literacies Sampler (pp.199-227). New York: Peter Lang Publishing.

Kohut, H. (1984) *How does analysis cure?* New York: International Universities Press.

Kolowich, L. (2019) *The Best Time to Post on Instagram, Facebook, Twitter, LinkedIn, & Pinterest.* Hudspot [online] Available at: https://blog.hubspot.com/marketing/best-times-post-pin-tweet-social-media-infographic Accessed: 9/5/20

Kong, Q., Rizoiu, M. A., Wu, S. and Xie, L. (2018) *Will This Video Go Viral: Explaining and Predicting the Popularity of YouTube Videos*. In Companion Proceedings of the The Web Conference, April, 2018, pp.175-178.

Konnikova, M. (2014) *The Six Things That Make Stories Go Viral Will Amaze, And Maybe Infuriate, You.* New York: The New Yorker.

Kousha, K. and Thelwall, M. (2012) *Motivations for Citing YouTube Videos in the Academic Publications: A Contextual Analysis.* In 17th International Conference on Science and Technology Indicators (STI), pp.5-8.

Kousha, K., Thelwall, M. and Abdoli, M. (2012) *The role of online videos in research communication: A content analysis of YouTube videos cited in academic publications*. Journal of the American Society for Information Science and Technology, 63(9), pp.1710–1727.

Kozinets, R. (1999) *E-tribalised Marketing? The Strategic Implications of Virtual Communities of Consumption.* European Management Journal, 17, pp.252-264.

Kozinets, R. V. and Cerone, S. (2014) *Between the Suit and the Selfie: Executives' Lessons on the Social "Micro-Celebrity".* GfK Marketing Intelligence Review, 6(2), 21.

Krajewski, J. M., Schumacher, A. C. and Dalrymple, K. E. (2019) Just turn on the faucet: A content analysis of PSAs about the global water crisis on YouTube. Environmental Communication, 13(2), pp.255-275.

Kruitbosch, G. and Nack, F. (2008) *Broadcast yourself on YouTube: really?* In Proceedings of the 3rd ACM international workshop on Human-centered computing, pp.7-10. ACM.

Kuipers, G. (2002) *Media culture and internet disaster jokes: Bin Laden and the attack on the World Trade Center.* European Journal of Cultural Studies, 5(4), pp.450-470.

Kumar, R. (2014) *Research Methodology: A Step-by-step Guide for Beginners (4th Edition).* London: Sage Publications Ltd.

Kurt, S. (2019) *Moving toward a universally accessible web: Web accessibility and education*. Assistive Technology, 31(4), pp.199-208.

Kuznetsov, S. (2006) *Motivations of contributors to Wikipedia*. ACM SIGCAS computers and society, 36(2), 1.

Ladhari, R., Massa, E. and Skandrani, H. (2020) *YouTube vloggers' popularity and influence: The roles of homophily, emotional attachment, and expertise.* Journal of Retailing and Consumer Services, 54, 102027.

Lange, P. (2007) *Publicy private and privately public: Social networking on YouTube.* Comput-Mediat Comm, 13(1), pp.361-380.

Lange, P. (2007a) *Commenting on comments: Investigating responses to antagonism on YouTube*. In Society for Applied anthropology conference, Vol. 31, pp.163-190.

Lange, P. (2007b) *The vulnerable video blogger: Promoting social change through intimacy.* The Scholar and Feminist Online, 5(2).

Lange, P. (2008) *Living in YouTubia: Bordering on civility.* Proceedings of the Southwestern Anthropological Association Conference, pp.98-106.

Lange, P. (2009) *Videos of affinity on YouTube*. In: Snickars P and Vonderau P (eds) The Youtube Reader. Stockholm: Wallflower Press, pp.228–247.

Lau, G. and Ng, S. (2001) *Individual and Situational Factors Influencing Negative Word-of-Mouth Behaviour.* Canadian Journal of Administrative Sciences, 18, pp.63-178.

Lauzen, M. (2013) *The Celluloid Ceiling: Behind-the-Scenes Employment of Women on the Top 250 Films of 2012*. Center for the Study of Women in Television and Film, San Diego State University. San Diego.

Lawrence, S. and Giles, C. (1998) *Searching the World Wide Web - Report*. Science 3 – Vol. 280 pp.98-100.

Lazarsfeld, P. and Merton, R. (1954) *Friendship as a social process: A substantive and methodological analysis.* Freedom and control in modern society 18: 66.
Lee, R. and Robbins, S. (1995) *Measuring belongingness: The social connectedness and social assurance scales*. Journal of Counseling Psychology, 42, pp.232–241.
Lee, R. M., Draper, M. and Lee, S. (2001) *Social connectedness, dysfunctional interpersonal behaviors, and psychological distress testing a mediator model.* Journal of Counseling Psychology, 48, pp.310–318.
Lee, J., Moon, S. and Salamatian, K. (2010) *An approach to model and predict the popularity of online contents with explanatory factors.* In 2010 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Vol. 1, pp.623-630.
Lerman, K. and Hogg, T. (2010) *Using a model of social dynamics to predict popularity of news*. In Proceedings of the World Wide Web Conference (WWW), 2010.
Leskovec, J., Backstrom, L. and Kleinberg, J. (2009) *Meme-tracking and the Dynamics of the News Cycle*. In KDD'09.
Leskovec, J., Singh, A. and Kleinberg, J. (2006) *Patterns of Influence in a Recommendation Network*. Pacific-Asia conference on Advances in Knowledge Discovery and Data Mining, April 09-12, 2006, Singapore.
Lessig, L. (2008) *Remix: Making Art and Commerce Thrive in the Hybrid Economy*. New York: Penguin.
Lewis, S., Heath, N., St Denis, J. and Noble, R. (2011) *The scope of non-suicidal self-injury on YouTube*. Pediatrics, 127(3), e552-e557.
Li, C. and Bernoff, J. (2008) *Groundswell: Wining in a World Transformed by Social Technologies*. Boston, MA: Harvard Business Press.
Liao, H., Liu, S. and Pi, S. (2011) *Modelling motivations for blogging: An expectancy theory analysis.* Social Behavior and Personality: an international journal, 39(2), pp.251-264.
Liberatore, M. W., Marr, D. W., Herring, A. M. and Way, J. D. (2019). *Student-created homework problems based on YouTube videos*. Chemical Engineering Education, 47(2).
Lillie, S. (2008) *Diffusion of innovation in the age of YouTube*. American journal of preventive medicine, 34(3), 267.
Lin, D., Bezemer, C. P. and Hassan, A. E. (2019) *Identifying gameplay videos that exhibit bugs in computer games*. Empirical Software Engineering, 24(6), pp.4006-4033.
Lin, K. and Lu, H. (2011) *Why people use social networking sites: An empirical study integrating network externalities and motivation theory.* Computers in human behavior, 27(3), pp.1152-1161.
Lindgreen, A. and Vanhamme, J. (2005) *Viral Marketing: The Use of Surprise*. Advances in Electronic Marketing, 2005, Idea Group Publishing, Hershey, PA, pp.122–138.
Lindstrom, M. (2009) *Essential Elements of Viral Video Success*. Advertising Age, September 19.
Lister, M. (2009) *New media: A critical introduction*. London: Routledge.
Lo, A., Esser, M. and Gordon, K. (2010) *YouTube: a gauge of public perception and awareness surrounding epilepsy*. Epilepsy and Behavior, 17(4), pp.541-545.
Loeb, S., Sengupta, S., Butaney, M., Macaluso Jr, J. N., Czarniecki, S. W., Robbins, R. and Langford, A. (2019) Dissemination of misinformative and biased information about prostate cancer on YouTube. European urology, 75(4), pp.564-567.
Logan, R. (2012) *Using YouTube in perioperative nursing education*. AORN, 95(4), pp.474-481.
Lorenzo-Dus, N., Blitvich, P. and Bou-Franch, P. (2011) *On-line polylogues and impoliteness: The case of postings sent in response to the Obama Reggaeton YouTube video*. Journal of Pragmatics, 43(10), pp.2578-2593.
Luo, J., Boland, R. and Chan, C. H. (2020) *How to use technology in educational innovation*. In Roberts Academic Medicine Handbook. pp.141-147. Springer, Cham.
Lovelock, M. (2019) 'My coming out story': Lesbian, gay and bisexual youth identities on YouTube. International Journal of Cultural Studies, 22(1), pp.70-85.

Madathil, K. C., Rivera-Rodriguez, A. J., Greenstein, J. S. and Gramopadhye, A. K. (2015) Healthcare information on YouTube: a systematic review. Health Informatics Journal, 21(3), pp.173-194.
Madden, M. (2009) *The Audience for Online Video-Sharing Sites Shoots Up*. Washington DC: Pew Internet and American Life Project.
Maier, G., Schneider, F. and Feldmann, A. (2010) *A First Look at Mobile Hand-held Device Traffic*. In PAM 2010, pp.161–170, Zurich, Switzerland.
Manning, J. and Shogan, C. (2012) *Women in the United States Congress*. Congressional Research Service.
Manyika, J. and Roxburgh, C. (2011) *The great transformer: The impact of the Internet on economic growth and prosperity*. London: McKinsey Global Institute.
Martínez, C. and Olsson, T. (2019) Making sense of YouTubers: How Swedish children construct and negotiate the YouTuber Misslisibell as a girl celebrity. Journal of Children and Media, 13(1), pp.36-52.
Maslow, A. (2015) *A Theory of Human Motivation*. Classics of Organization Theory, 50, 142.
May, J. (2018) YouTube Gamers and Think-Aloud Protocols: Introducing Usability Testing. IEEE Transactions on Professional Communication, 62(1), pp.94-103.
Mayoral, p., Tello, A. and Gonzalez, J. (2010) *YouTube Based Learning.* [online] Available at: https://www.fig.net/pub/fig2010/papers/ts07g%5Cts07g_mayoralvaldivia_tellomoreno_et_al_4098.pdf Accessed: 15/11/17.
McAllister, M. P. and Stoltzfus-Brown, L. (2020) *Television Advertising: Texts, Political Economy, and Ideology*. A Companion to Television, pp.165-182.
Meehan, E. (2006) *Gendering the commodity audience: critical media research, feminism, and political economy.* Media and Cultural Studies 311, pp.209–222.
Meek, D. (2012) *YouTube and social movements: A phenomenological analysis of participation, events and cyberplace.* Antipode, 44(4), pp.1429-1448.
Mehrotra, R. and Bhattacharya, P. (2017) *Characterizing and predicting supply-side engagement on video sharing platforms using a hawkes process model.* In Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval. pp.159-166.
McCorkindale, T. (2010) *Can you see the writing on my wall? A content analysis of the Fortune 50's Facebook social networking sites*. Public Relations Journal, 4(3), pp.1-13.
McPherson, M., Smith-Lovin, L. and Cook, J. (2001) *Birds of a Feather: Homophily in Social Networks*. Annual Review of Sociology (2001), 27, pp.415–44.
Menou, M. (1995) *The impact of information—I. Toward a research agenda for its definition and measurement.* Information Processing and Management, 31(4), pp.455-477.
Meriam-Webster.com (2019) Popular definition. [online] Available at: https://www.merriam-webster.com/dictionary/popular Accessed: 16/10/19.
Merriam, S. (2009) *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: Jossey-Bass.
Milgram, P., Bickman, L. and Berkowitz, L. (1969) *Note on the drawing power of crowds of different size.* Journal of Personality and Social Psychology, 13(2):79-82, October 1969.
Milkman, K. L. and Berger, J. (2014) *The science of sharing and the sharing of science*. Proceedings of the National Academy of Sciences, 111(Supplement 4), pp.13642-13649.
Miller, M. (2009) *Integrating online multimedia into college course and classroom: With application to the social sciences.* MERLOT Journal of Online Learning and Teaching, 5(2), pp.395–423.
Milliken, M., Gibson, K. and O'Donnell, S. (2008) *User-generated video and the online public sphere: Will YouTube facilitate digital freedom of expression in Atlantic Canada?* American Communication Journal, 10(3). [online] Available at: http://meeting.knet.ca/mp19/file.php/16/Susan/2008-AJC-Milliken-final.pdf. Accessed: 12/12/17.
Milliken, M., Gibson, K., O'Donnell, S. and Singer, J. (2008a) *User-generated online video and the Atlantic Canadian public sphere: A YouTube study*. In Proceedings of the International

Communication Association Annual Conference. [online] Available at: http://nparc.cisti-icist.nrc-cnrc.gc.ca/eng/view/accepted/?id=b7a56a1a-f962-4294-b1c3-4c529c1749a6 Accessed: 11/12/17.
Mislove, A., Marcon, M., Gummadi, K., Druschel, P. and Bhattacharjee, B. (2007) *Measurement and analysis of online social networks.* IMC '07: Proc. of the 7th ACM SIGCOMM conference on Internet measurement, 2007.
Mitra, S., Agrawal, M., Yadav, A., Carlsson, N., Eager, D. and Mahanti, A. (2011) *Characterizing web-based video sharing workloads.* ACM Transactions on the Web (TWEB), 5(2), p8.
Mohan, S. and Punathambekar, A. (2019) Localizing YouTube: Language, cultural regions, and digital platforms. International Journal of Cultural Studies, 22(3), pp.317-333.
Molyneaux, H., O'Donnell, S., Gibson, K. and Singer, J. (2008) *Exploring the gender divide on YouTube: An analysis of the creation and reception of vlogs.* American Communication Journal, 10(2), pp.1-14.
Moor, P., Heuvelman, A. and Verleur, R. (2010) *Flaming on YouTube.* Computers in Human Behavior, 26(6), pp.1536-1546.
Morris, M. and Anderson, E. (2015) *'Charlie is so cool like': Authenticity, popularity and inclusive masculinity on YouTube.* Sociology, 49(6), pp.1200-1217.
Mukherji, P. and Albon, D. (2015) *Research Methods in Early Childhood (2nd Edition).* London: Sage Publications Ltd.
Munnukka, J., Maity, D., Reinikainen, H. and Luoma-aho, V. (2019). "Thanks for watching". The effectiveness of YouTube vlog endorsements. Computers in Human Behavior, 93, pp.226-234.
Murthy, D. and Sharma, S. (2019) Visualizing YouTube's comment space: Online hostility as a networked phenomena. New Media & Society, 21(1), pp.191-213.
Nahon, K., Hemsley, J., Walker, S. and Hussain, M. (2011) *Fifteen Minutes of Fame: The Power of Blogs in the Lifecycle of Viral Political Information.* Policy Internet, 3, 1-30. [Online] Available at: http://dx.doi.org/10.2202/11944-2866.1108 Accessed: 1/6/14.
Nam, K., Ackerman, M. and Adamic, L. (2009) *Questions in, knowledge in? A study of Naver's question answering community.* In Proceedings of the SIGCHI conference on human factors in computing systems, pp.779-788. ACM.
Nardi, P. (2013) *Doing Survey Research: A Guide to Quantitative Methods*. London: Pearson Education.
Nathanson, E. (2014) *Dressed for Economic Distress: Blogging and the" New" Pleasures of Fashion*. Duke University Press.
National Center on Education, and the Economy (US). (2008) *Tough choices or tough times: The report of the new commission on the skills of the American workforce.* John Wiley and Sons.
Nelson-Field, K., Riebe, E. and Newstead, K. (2011a) *Make 'em laugh, make 'em cry: Understanding which emotions drive video sharing on Facebook.* Institute for Marketing Science.
Nelson-Field, K., Riebe, E. and Newstead, K. (2011b) *To shock or inspire: Understanding which emotions drive video sharing on Facebook*. In ANZMAC. Presented at the Annual Australia and New Zealand Marketing Academy Conference Perth, WA.
Nicholas, D., Huntington, P., Jamali, H.R., Rowlands, I. and Fieldhouse, M. (2009) *Student digital information-seeking behaviour in context*. Journal of Documentation 65 (1), pp.106–132.
Nielsen. (2012) *State of the media: The social media report*. [online] Available at: http://www.nielsen.com/content/dam/corporate/us/en/reportsdownloads/2012-Reports/The-Social-Media-Report-2012.pdf Accessed: 2/6/14.
Nikolinakou, A. and King, K. W. (2018) *Viral video ads: Examining motivation triggers to sharing*. Journal of Current Issues and Research in Advertising, 39, pp.120–139.
Nouri, M. (2018) The power of influence: traditional celebrity vs social media influencer.
Nisbett, R. and Wilson, T. (1977) *Telling more than we can know: Verbal reports on mental processes.* Psychological Review, 84(3), pp.231-259.

Norton, H. and Citron, D. (2010) *Government speech 2.0*. Denver University Law Review, 87, pp.899–901.
Nov, O. (2007) *What motivates Wikipedians?* Communications of the ACM, 50(11), pp.60-64.
Nwana, A., Avestimehr, S. and Chen, T. (2013) *A latent social approach to YouTube popularity prediction*. In 2013 IEEE Global Communications Conference (GLOBECOM), pp.3138-3144. IEEE.
Obednikovska, S., Sotiroski, K. and Mateska, E. (2019) *Web Site–Basic Internet Marketing Strategy Tool of Digital Companies*. Annals of the „Constantin Brâncuşi" University of Târgu Jiu, Economy Series, (3/2019), pp.82-91.
Oblinger, D. (2003) *Boomers, Gen-Xers, and Millennials: Understanding the "New Students".* EDUCAUSE Review, Vol. 38, No. 4, pp.36-40.
Oh, S. (2012) *The characteristics and motivations of health answerers for sharing information, knowledge, and experiences in online environments.* Journal of the Association for Information Science and Technology, 63(3), pp.543-557.
Oh, S. and Syn, S. (2015) *Motivations for sharing information and social support in social media: A comparative analysis of Facebook, Twitter, Delicious, YouTube, and Flickr.* Journal of the Association for Information Science and Technology, 66(10), pp.2045-2060.
Olasina, G. (2017) *An evaluation of educational values of YouTube videos for academic writing*. The African Journal of Information Systems, 9(4), 2.
Olsen, S. (2005) *The 'millennials' usher in a new era*. [online] Available at: http://news.com.com/2009-1025_3-5944666.html Accessed: 2/6/14.
O'Malley, G. (2011) *Media Buyers Predict Upswing in Online Video Ad Spend*. [Online] MediaPost. Available at: http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=149695&nid=126334 Accessed: 5/4/14.
Onwuegbuzie, A. J., & Combs, J. P. (2010) *Emergent data analysis techniques in mixed methods research: A synthesis*. In A. Tashakkori & C. Teddlie (Eds.) Sage handbook of mixed methods in social and behavioral research (2nd ed., pp. 397-430). Thousand Oaks, CA: Sage.
O'Neil, C. (2016) *Weapons of math destruction: How big data increases inequality and threatens democracy*. Broadway Books.
O'Neill, M. (2010) *7 Factors That Drive People to Share Content with their Social Networks*. Social Times: MediaBistro.
O'Neil, M. (2010a) *The Psychology Of FAIL Videos*. ADWEEK [online] Available at: https://www.adweek.com/digital/fail-videos/ Accessed: 6/3/16.
O'Neill, M. (2011) *What Makes A Video "Viral"?* Social Times: MediaBistro.
Oppenheim, C., Morris, A., McKnight, C. and Lowley, S. (2000) *The Evaluation of WWW Search Engines*. The Association for Information Management, Journal of Documentation, Vol. 56, No. 2, March 2000.
Ortega-León, C., Marín-Reyes, P. A., Lorenzo-Navarro, J., Castrillón-Santana, M. and Sánchez-Nielsen, E. (2019) *Video Categorisation Mimicking Text Mining*. In International Work-Conference on Artificial Neural Networks. pp.292-301. Springer, Cham.
Orttung, R. W. and Nelson, E. (2019) Russia Today's strategy and effectiveness on YouTube. Post-Soviet Affairs, 35(2), pp.77-92.
Paolillo, J. (2008) *Structure and network in the YouTube core*. HICSS'08: Proc. of the 41st Annual Hawaii International Conference on System Sciences, 2008.
Parbey, C. (2019) *Marketing How-To: The Best Time to Post on YouTube*. Edgy [online] Available at: https://edgy.app/best-time-to-post-on-youtube-live-youtube-subscriber-count Accessed: 9/5/20
Pech, R. (2003) *Memes and cognitive hardwiring: Why are some memes more successful than others?* European Journal of Innovation Management, 6(3), pp.173-181.
Peer, L. and Ksiazek, T. B. (2011) *YouTube and the challenge to journalism: new standards for news videos online*. Journalism Studies, 12(1), pp.45-63.

Perrin, A. (2015) *Social media usage.* Chicago: Pew Research Center.
Perta, A. (2020) *Stop Guessing: THIS is The Best Time to Post on YouTube! (and Get More Views).* Misfit Hustler [online] Available at: https://misfithustler.com/youtube/best-time/ Accessed: 14/5/20
Peters, K. and Kashima, Y. (2007) *From Social Talk to Social Action: Shaping the Social Triad with Emotion Sharing*. Journal of Personality and Social Psychology, 93(5), pp.780-97.
Phelps, J., Lewis, R., Mobilio, L., Perry, D. and Raman, N. (2004) *Viral Marketing or Electronic Word-of-Mouth Advertising: Consumer Responses and Motivations to Pass Along Email.* Journal of Advertising Research, 44 (4), pp.333-348.
Pintado, I. (2009) *Connection Generation: How Connection Determines Our Place in Society and Buisness*. BookSurge Publishing.
Pinto, H., Almeida, J. and Goncalves, M. (2013) *Using Early View Patterns to Predict the Popularity of YouTube Videos*. In WSDM, pp.365–374, 2013.
Plesser, A. (2007) *First Video Sharing Site Paved the Way for YouTube — ShareYourWorld.com Was There First to Launch Ten Years Back*. Beet.TV [online] Available at: https://www.beet.tv/2007/07/first-video-sha.html Accessed: 9/5/20
Poell, T. and Borra, E. (2012) *Twitter, YouTube, and Flickr as platforms of alternative journalism: The social media account of the 2010 Toronto G20 protests.* Journalism, 13(6), pp.695-713.
Pollard, A. (2014) *Reflective teaching in schools*. London: Bloomsbury.
Poposki, E. M. and Oswald, F. L. (2010) *The multitasking preference inventory: Toward an improved measure of individual differences in polychronicity.* Human Performance, 23, pp.247–264.
Porter, L. and Golan, G. (2006) *From Subservient Chickens to Brawny Men: A Comparison Of Viral Advertising To Television Advertising*. Journal of Interactive Advertising, 6(2).
Powers, K. (2019) *The Best Times to Post on Social Media*. AMA [online] Available at: https://www.ama.org/marketing-news/the-best-times-to-post-on-social-media/ Accessed: 9/5/20
Prensky, M. (2004) *Digital Game-based Learning*. New York: McGraw-Hill.
Priem, J., Taraborelli, D., Groth, P. and Neylon, C. (2011) *Altmetrics: a manifesto.* [online] Available at: http://altmetrics.org/manifesto Accessed: 7/11/15.
Prigg, C. (2012) *100 more social media statistics for 2012*. [online] Available at: http://thesocialskinny.com/100-more-social-media-statistics-for-2012/ Accessed: 5/1/16.
Punch, K. (2009) *Introduction to Research Methods in Education*. London: Sage Publications Ltd.
Punch, K. and Oancea, A. (2014) *Introduction to Research Methods in Education*. London: Sage Publications Ltd.
Purcariu, M. (2018) How-to Videos and the Construction of Authenticity in the Amateur Culture. Studia Universitatis Babes-Bolyai-Ephemerides, 63(2), pp.91-108.
Purcell, K. (2010) *The State of Online Video in 2015*. Washington DC: Pew Internet and American Life Project.
Qiu, L., Tang, Q. and Whinston, A. B. (2015) *Two formulas for success in social media: Learning and network effects.* Journal of Management Information Systems, 32(4), pp.78-108.
Qualman, E. (2009) *Socialnomics: How Social Media Transforms the Way We Live and Do Business*. Wiley, John and Sons Incorporated.
Quan-Haase, A. and Young, A. (2010) *Uses and gratifications of social media: A comparison of Facebook and instant messaging.* Bulletin of Science, Technology and Society, 30(5), pp.350-361.
Race, P. (2014) *The Lecturer's Toolkit*. London: Routledge.
Rafaeli, S., Ariel, Y. and Hayat, T. (2005) *Wikipedia community: Users' motivations and knowledge building*. In cyberculture 3rd global conference, Prague, Czech Republic, August.
Rainie, L., Brenner, J. and Purcell, K. (2012) *Photos and videos as social currency online*. Pew Internet and American Life Project.

Ramsay, P. (1989) *Critical values for Spearman's rank order correlation.* Journal of Educational Statistics, 14, pp.245–254.
Rein, K. and Venturini, T. (2018) *Ploughing digital landscapes: How Facebook influences the evolution of live video streaming*. New Media & Society, 20(9), pp.3359-3380.
Reka, R., Hawoong Jeong, H. and Barabasi, A. (1999) *The diameter of the world wide Web.* Nature 401, pp. 130-131.
Rheingold, H. (2000) *The virtual community: Homesteading on the electronic frontier*. MIT press.
Rieder, B., Matamoros-Fernández, A. and Coromina, Ò. (2018) *From ranking algorithms to 'ranking cultures' Investigating the modulation of visibility in YouTube search results.* Convergence, 24(1), pp.50-68.
Richier, C., Altman, E., Elazouzi, R., Altman, T., Linares, G. and Portilla, Y. (2014) *Modelling View-count Dynamics in YouTube.* Technical Report. [Online]. Available at: https://arxiv.org/pdf/953787 Accessed: 5/6/15.
Richter, A. and Koch, M. (2008) *Functions of social networking services*. In Proceedings of the 8th International Conference on the Design of Cooperative Systems (COOP '08). Carry-le-rouet, France, Institut d'Etudes Politiques d'Aix-en-Provence, pp.87–98.
Rigby, J., Brumby, D., Gould, S. and Cox, A. (2018) *I Can Watch What I Want: A Diary Study of On-Demand and Cross-Device Viewing.* In Proceedings of the 2018 (June) ACM International Conference on Interactive Experiences for TV and Online Video. pp.69-80. ACM.
Rihl, A. and Wegener, C. (2017) YouTube celebrities and parasocial interaction: Using feedback channels in mediatized relationships. Convergence, 1354856517736976.
Roberson, M. and Sundstrom, E. (1990) Questionnaire design, return rates, and response favorableness in an employee attitude questionnaire. Journal of Applied Psychology, 75(3), 354.
Robertson, M. (2007) *Youtube Traffic Patterns and Uploading Recommendations.* Tubular Insights [online] Available at: https://tubularinsights.com/youtube-traffic-analysis-seo/ Accessed: 14/5/20
Rodriguez, H. A., Young, M. T., Jackson, H. T., Oelschlager, B. K. and Wright, A. S. (2018) *Viewer discretion advised: is YouTube a friend or foe in surgical education?* Surgical endoscopy, 32(4), pp.1724-1728.
Rogers, E. and Bhowmik, D. (1970) *Homophily-heterophily: Relational concepts for communication research.* Public Opinion Quarterly 34(4): pp.523–538.
Rogers, M. (2003) *Diffusion of innovations (5th edition).* New York, NY: Free Press.
Rosenthal, S. (2018) *Motivations to seek science videos on YouTube: free-choice learning in a connected society.* International Journal of Science Education, Part B, Vol. 8 No. 1, pp.22-39.
Rousseau, R. (1999) *Daily time series of common single word searches in AltaVista and NorthernLight.* Cybermetrics, 2/3. [online] Available at: http://cybermetrics.cindoc.csic.es/articles/v2i1p2.html Accessed: 7/4/15.
Rubenking, B. (2019) *Emotion, attitudes, norms and sources: Exploring sharing intent of disgusting online videos*. Computers in Human Behavior, 96, pp.63-71.
Rudolph, T. and Frankel, J. (2009) *YouTube in Music Education*. New York: Hal Leonard.
Ryan, R. and Deci, E. (2000) *Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being.* American psychologist, 55(1), 68.
Saini, M. K., Gadde, R., Yan, S. and Ooi, W. T. (2012) *Movimash: online mobile video mashup.* In Proceedings of the 20th ACM international conference on Multimedia. pp.139-148.
Salganik, M., Dodds, P. and Watts, D. (2006) *Experimental Study of Inequity and Unpredictability in an Artificial Culture Market.* Science, Vol. 311, February 2006, pp.854-856.
Saurabh, S. and Gautam, S. (2019) Modelling and statistical analysis of YouTube's educational videos: A channel Owner's perspective. Computers & Education, 128(1), pp.145-158.
Saul, R. (2010) *KevJumba and the Adolescence of YouTube*. Educational Studies, 46(5), pp.457-477.
Savage, M. and Barnett, A. (2017) *Technology-Enhanced Learning in the Early Years Foundation Stage*. St Albans: Critical Publishing.

Schiffman, L. and Kanuk, L. (1995) *Consumer Behaviour*. Upper Saddle River, NJ: Prentice Hall.

Schuman, D. L., Lawrence, K. A. and Pope, N. (2019) Broadcasting war trauma: An exploratory netnography of veterans' YouTube vlogs. Qualitative Health Research, 29(3), pp.357-370.

Schwemmer, C. and Ziewiecki, S. (2018) *Social media sell-out: The increasing role of product promotion on YouTube*. Social Media+ Society, 4(3), 2056305118786720.

Scott, B. and Judge, T. (2009) *The Popularity Contest at Work: Who wins, why and what do they receive?* Journal of Applied Psychology, 94 (1), pp.20-33.

Shah, C. (2010) *Supporting research data collection from YouTube with 'TubeKit'*. Journal of Information Technology & Politics, 7(2-3), pp.226-240.

Shapiro, M. A. and Park, H. W. (2018) *Climate change and YouTube: Deliberation potential in post-video discussions*. Environmental Communication, 12(1), pp.115-131.

Shifman, L. (2012) *An anatomy of a YouTube meme.* New media & society, 14(2), pp.187-203.

Shoufan, A. (2019) Estimating the cognitive value of YouTube's educational videos: A learning analytics approach. Computers in Human Behavior, 92, pp.450-458.

Shultz, D. (2015) *Why Do We Love #Fail Videos?* Nautilus. [online] Available at: http://nautil.us/blog/why-do-we-love-fail-videos Accessed: 1/6/16.

Siefert, C., Kothuri, R., Jacobs, D., Levine, B., Plummer, J. and Marci, C. (2009) *Winning the Super 'Buzz' Bowl: How Biometrically-Based Emotional Engagement Correlates With Online Views and Comments For Super Bowl Advertisements*. Journal of Advertising Research, 49, No. 3 (September), pp.293-303.

Siersdorfer, S., Chelaru, S., Nejdl, W. and Pedro, J. S. (2010) *How Useful are Your Comments? Analyzing and Predicting YouTube Comments and Comment Ratings*. In World-Wide Web 2010, pages 891–900, Raleigh, NC.

Simmons, R. (2008) *'Questionnaires'*, in Gilbert, N. (ed.) Researching Social Life (3rd Edition). London: Sage Publications Ltd.

Skiba, D. (2007) *Nursing education 2.0: YouTube*. Nursing Education Perspectives 28 (2), pp.100–102.

Skiba, D. and Barton, A. (2006) *Adapting your teaching to accommodate the net-generation of learners.* Online Journal of Issues in Nursing, 11(2).

Sliverman, D. (2005) *Doing Qualitative Research.* London: Sage Publications Ltd.

Smith, C. (2014) *By the Numbers: 32 Amazing YouTube Statistics.* Digital Marketing Ramblings. The Latest Digital Marketing Stats, Tips, Trends and Technology [online] Available at: http://expandedramblings.com/index.php/youtube-statistics Accessed: 14/5/16.

Snelson, C. (2009) *Web-Based Video for e-Learning: Tapping into the YouTubeTM Phenomenon*. Collective Intelligence and E-Learning 2.0: Implications of Web-Based Communities and Networking: Implications of Web-Based Communities and Networking, 147.

Snelson, C. and Elison-Bowers, P. (2009a) *Using YouTube videos to engage the affective domain in e-learning.* In Méndez-Vilas A, Martín AS, González JAM, González JM. (Eds.), Research, reflections and innovations in integrating ICT in education, 3. Badajoz: FORMATEX.

Snelson, C. (2011) *YouTube across the disciplines: a review of the literature*. MERLOT Journal of Online Learning and Teaching 7 (1), pp.159–169.

Snelson, C., Rice, K. and Wyzard, C. (2012) *Research priorities for YouTube and video-sharing technologies: A Delphi study.* British Journal of Educational Technology, 43(1), pp.119-129.

Snyder, S. and Burke, S. (2008) *Using YouTube in the classroom: a how-to guide*. Int J Instruct Technol Distance Learn, 5(4), pp.45-52.

Soares, R., Combopiano, J., Regis, A., Shur, Y. and Wong, R. (2012) *2011 Catalyst Census: Fortune 500 Women Board Directors*. Catalyst. December.

Sobkowicz, P. and Sobkowicz, A. (2010) *Dynamics of hate based Internet user networks*. The European Physical Journal B-Condensed Matter and Complex Systems, 73(4), pp.633-643.

Sofia, A. M. and Widodo, S. (2020) *Personal Branding of Prosumer in the Highest Internet User Country.* Journal of Talent Development and Excellence, 12(2s), pp.98-104.

Sokolova, K. and Kefi, H. (2020) *Instagram and YouTube bloggers promote it, why should I buy? How credibility and parasocial interaction influence purchase intentions*. Journal of Retailing and Consumer Services, 53.
Southgate, D., Westoby, N. and Page, G. (2010) *Creative Determinants of Viral Video Viewing.* International Journal of Advertising, 29(3), pp.2-14.
Stauff, M. (2009) *Sports on YouTube*. In Vonderau, P., Snickars, P. and Burgess, J. (eds) The YouTube Reader. pp.236– 251. Stockholm: National Library of Sweden.
Steinberg, P., Wason, S., Stern, J., Deters, L., Kowal, B. and Seigne, J. (2010) *YouTube as Source of Prostate Cancer Information*. Urology, 75(3), pp.619-622.
Strangelove, M. (2010) *Watching YouTube: Extraordinary videos by ordinary people*. Toronto: University of Toronto Press.
Styati, E. W. (2016) *Effect of YouTube Videos and Pictures on EFL Students' Writing Performance*. Dinamika Ilmu, 16(2), pp.307-317.
Sugimoto, C. and Thelwall, M. (2013) *Scholars on soap boxes: Science communication and dissemination via TED videos*. Journal of the American Society for Information Science and Technology, 64(4), pp.663-674.
Sun, E., Rosenn, I., Marlow, C. and Lento, T. (2009) *Gesundheit! Modelling contagion through Facebook news feed.* In ISWSM'10. Technorati (2011).
Sunstein, C. (2005) *Group Judgments: Statistical Means, Deliberation and Information Markets*. NY University Law Review. 80 (2005) pp.962-1049.
Sykora, M. and Panek, M. (2009) *Media sharing websites and the US financial markets*. IADIS International Conference WWW/Internet [online] Available at: https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/6423 Accessed: 10/7/16.
Sykora, M. and Panek, M. (2009a) *Financial news content publishing on YouTube.com*. Soft Computing Applications, 2009. SOFA '09. 3rd International Workshop. pp.99-104.
Sysomos (2009) *A look inside online video engagement*. [online] Available at: http://www.sysomos.com/reports/video/ Accessed: 17/5/16.
Szabó,G. and Huberman, B.(2010) *Predicting the popularity of online content*. Commun. ACM 53(8): 80-88 (2010).
Tackett, S., Slinn, K., Marshall, T., Gaglani, S., Waldman, V. and Desai, R. (2018) *Medical education videos for the world: an analysis of viewing patterns for a YouTube channel.* Academic medicine, 93(8), pp.1150-1156.
Takacs, K., Flache, A. and Maes, M. (2014) *Is there negative social influence? Disentangling effects of dissimilarity and disliking on opinion shifts.* In: CoRR abs/1406.0900, 2014.
Tan, E. and Pearce, N. (2012) *Open education videos in the classroom: Exploring the opportunities and barriers to the use of YouTube in teaching introductory sociology.* Research in Learning Technology, 19, pp.128-137.
Tapscott, D. (1997) *Growing Up Digital: The Rise of the Net Generation*. New York: McGraw-Hill.
Tellis, G. J., MacInnis, D. J., Tirunillai, S. and Zhang, Y. (2019) *What drives virality (sharing) of online digital content? The critical role of information, emotion, and brand prominence*. Journal of Marketing, 83(4), pp.1-20.
Thackeray, R., Neiger, B., Hanson, C. and McKenzie, J. (2008) *Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media*. Health promotion practice, 9(4), pp.338-343.
Thelwall, M. and Mas-Bleda, A. (2018) YouTube science channel video presenters and comments: Female friendly or vestiges of sexism? Aslib Journal of Information Management, 70(1), pp.28-46.
Thelwall, M., Sud, P. and Vis, F. (2012) *Commenting on YouTube videos: From Guatemalan rock to El Big Bang.* Journal of the American Society for Information Science and Technology, 63(3), pp.616–629.

Thelwall, M., Vaughan, L. and Bjorneborn, L. (2004) *Annual review of information science and technology.* New Jersery: Information Today Inc.
Thelwall, M. and Wilkinson, D. (2008) *A generic lexical URL segmentation framework for counting links, colinks or URLs*. London: Library & Information Science Research.
Thelwall, M. (2008a) *Bibliometrics to Webometrics*. Journal of Information Science, 34(4), pp.605-621.
Thelwall, M. (2008b) *Extracting accurate and complete results from search engines: Case study Windows Live.* Journal of the American Society for Information Science and Technology, 59(1), pp.38-50.
Thelwall, M. (2009) *Introduction to Webometrics: Quantitative Web Research for the Social Sciences*. California: Morgan & Claypool.
Thelwall, M. (2015) *Webometric Analyst 2.0*. Statistical Cybermetrics Research Group, University of Wolverhampton, UK. [online] Available at: http://lexiurl.wlv.ac.uk/ Accessed: 9/5/16.
Thelwall, M. (2018) Can museums find male or female audiences online with YouTube? Aslib Journal of Information Management, 70(5), pp.481-497.
Thorson, K., Ekdale, B., Borah, P., Namkoong, K. and Shah, C. (2010) *YouTube and Proposition 8: A case study in video activism.* Information, Communication & Society, 13(3), pp.325-349.
Tian, S., Yu, A., Vogel, D. and Kwok, R. (2011) *The impact of online social networking on learning: a social integration perspective*. International Journal of Networking and Virtual Organisations, 8, pp.264–280.
Toderici, G., Aradhye, H., Pasca, M., Sbaiz, L. and Yagnik, J. (2010) *Finding meaning on youtube: Tag recommendation and category discovery*. In 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 3447-3454. IEEE.
Tolson, A. (2010) *A new authenticity? Communicative practices on YouTube*. Critical Discourse Studies, 7(4), pp.277-289.
Törhönen, M., Sjöblom, M., Hassan, L. and Hamari, J. (2019) *Fame and fortune, or just fun? A study on why people create content on video platforms*. Internet Research.
Toyokawa, W., Whalen, A. and Laland, K. N. (2019) Social learning strategies regulate the wisdom and madness of interactive crowds. Nature Human Behaviour, 3(2), 183.
Trier, J. (2007) *"Cool" engagements with YouTube: Part 2*. Journal of Adolescent and Adult Literacy, 50(7), pp.598-603.
Tsai, J. (2009) *Video is More Than Viral.* CRM Magazine, 13, pp.36-42.
Tschopp, K. (2014) *4 Keys to Creating Successful Video Content for the Web*. Social Times: MediaBistro.
Tucker, C. and Wycoff, T. (2016) *Blended Learning in Action: A Practical Guide Toward Sustainable Change*. London: Corwin.
Tucker, E. (2007) *Police departments deputize YouTube*. USA Today.
Turvey, K., Potter, J., Burton, J., Jonathan Allen, J. and Sharp, J. (2016) *Primary Computing and Digital Technologies: Knowledge, Understanding and Practice*. London: Learning Matters.
University of Wolverhampton (2018) Ethics Guidance. University of Wolverhampton. [online] Available at: https://www.wlv.ac.uk/research/research-policies-procedures--guidelines/ethics-guidance/ Accessed: 12/1/18.
USA Today (2007) *YouTube serves up 100 million videos a day online*. [online] Available at: http://www.usatoday.com/tech/news/2006-07-16-YouTube-views_x.htm Accessed: 08/11/17.
Van Dijck, J. (2009) *Users like you? Theorizing agency in user-generated content*. Media, culture and society, 31(1), pp.41-58.
Van Langendonck, G. (2009) *Iconic Iran video was posted in the Netherlands*. NRC Handelsblad [online] Available at: http://vorige.nrc.nl/article2280315.ece Accessed 3/1/15.
Van Zoonen, L., Mihelj, S. and Vis, F. (2011) *YouTube interactions between agonism, antagonism and dialogue: Video responses to the anti-Islam film Fitna*. New Media & Society, 13 (8), pp.1283-1300.

Van Zoonen, L., Vis, F. and Mihelj, S. (2010) *Performing citizenship on YouTube: Activism, satire and online debate around the anti-Islam video Fitna*. Critical Discourse Studies, 7(4), pp.249-262.
Vasquez, A. (2019) *Best Time To Post On Social Media In 2020 [Part 1*]. Gray Group Intl [online] Available at: https://www.graygroupintl.com/blog/best-time-to-post-on-social-media-in-2020-part-1 Accessed: 9/5/20
Vaughan, L. (2006) *Visualizing linguistic and cultural differences using Web co-link data*. Journal of the American Society for Information Science and Technology, 57, pp.1178-1193.
Verhaeghe, A., Schillewaert, N., Van Belleghem, S., Vergult, C. and Claus, D. (2007) *A New Approach For Measuring 'Buzz' Word Of Mouth And Word Of Mouse.* Paper presented at the ESOMAR World Research Conference 2007.
Vingilis, E., Yildirim-Yenier, Z., Vingilis-Jaremko, L., Seeley, J., Wickens, C. M., Grushka, D. H. and Fleiter, J. (2018) *Young male drivers' perceptions of and experiences with YouTube videos of risky driving behaviours*. Accident Analysis & Prevention, 120, pp.46-54.
Walczyk, T. (2008) *Google Video–Just Another Video Sharing Site?.* Journal of Library Administration, 47:1-2, pp.175-181.
Waldfogel, J. (2009) *Lost on the web: Does web distribution stimulate or depress television viewing?* Information Economics and Policy, 21(2), pp.158-168.
Waldron, J. (2013) *YouTube, fanvids, forums, vlogs and blogs: Informal music learning in a convergent on-and offline music community.* International Journal of Music Education, 31(1), pp.91-105.
Walliman, N. (2001) *Your Research Project: A Step-by-step Guide for the First-time Researcher*. London: Sage Publications Ltd.
Wallsten, K. (2011) *Many Sources, One Message: Political Blogs Links to Online Videos during the 2008 Campaign*. Journal of Political Marketing, 10, pp.88-114.
Wasko, M. and Faraj, S. (2005) *Why should I share? Examining social capital and knowledge contribution in electronic networks of practice*. MIS quarterly, pp.35-57.
Watad, A., Watad, S., Mahroum, N., Sharif, K., Amital, H., Bragazzi, N. L. and Adawi, M. (2019) Forecasting the West Nile virus in the United States: an extensive novel data streams–based time series analysis and structural equation modeling of related digital searching behavior. JMIR public health and surveillance, 5(1), e9176.
Watts, D. (2002) *A simple model of information cascades on random networks*. Proceedings of the National Academy of Science, USA, 99:5766-5771, 2002.
Weaver, A., Zelenkauskaite, A. and Samson, L. (2012) *The (non) violent world of YouTube: Content trends in web video*. Journal of Communication, 62(6), pp.1065-1083.
Weber, I., Garimella, V. and Borra, E. (2013) *Inferring audience partisanship for YouTube videos*. In WWW, pp.43-44.
Wei, R. and Lo, V. (2006) *Staying connected while on the move: Cell phone use and social connectedness.* New Media and Society, 8, pp.53–72.
Weimann, G. (1994) *The Influentials: People Who Influence People*. Albany, NY: State University of New York Press.
Weiner, C. (2011) *Getting The Most From Your Social Video*. [Online] MediaPost. Available at: http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=145175 Accessed: 5/4/14.
Welbourne, D. and Grant, W. (2016) *Science communication on YouTube: Factors that affect channel and video popularity.* Public Understanding of Science, 25(6), pp.706-718.
Wellington, J. (2015) *Educational Research*. Bloomsbury: London.
Wellman, B. and Berkowitz, S. (1998) *Social Structures: A Network Approach*. Cambridge, MA: Cambridge University Press.
Wells, W., Moriarty, S. and Burnett, J. (2005) *Advertising: Principles and Practice (7th edition).* New Jersey : Prentice Hall

Wesch, M. (2008) *"Context Collapse."* Digital Ethnography. [online] Available at: http://mediatedcultures.net/youtube/contextcollapse Accessed: 12/11/14.
Wesch, M. (2009) *YouTube and you: Experiences of self-awareness in the context collapse of the recording webcam*. Explorations in media ecology, 8(2), pp.19-34.
West, T. (2011) *Going Viral: Factors That Lead Videos to Become Internet Phenomena.* The Elon Journal of Undergraduate Research in Communications. Vol. 2, No. 1, 76–84, Elon University, 2011.
Wilhelm, M., Ramanathan, A., Bonomo, A., Jain, S., Chi, E. H. and Gillenwater, J. (2018) *Practical diversified recommendations on youtube with determinantal point processes.* In Proceedings of the 27th ACM International Conference on Information and Knowledge Management, October. pp.2165-2173.
William H. Dutton, W., Blank, G. and Groselj, D. (2013) *Cultures of the Internet: The Internet in Britain (Oxford Internet Survey 2013 Report).* Oxford: Oxford Internet Institute
Winkler, R. (2016) *YouTube, 1 billion viewers, no profit.* The Wall Street Journal: English Edition. [online] Available at: https://www.wsj.com/articles/viewers-dont-add-up-to-profit-for-youtube-1424897967 Accessed: 1/12/19.
Wodjao, T. B. (2020) *A double-hurdle model of computer and internet use in American households. Departement of Economics*, Western Michigan University. Fabrizio Carlevaro, Yves Croissant, Stéphane Hoareau, 49.
Wojnicki, A. and Godes, D. (2008) *Word-of-Mouth as Self-Enhancement*. HBS Marketing Research Paper No. 06-01.
Wood, P. and Smith, J. (2016) *Educational Research: Taking the Plunge*. Wales: Independent Thinking Press.
Worthington, I. (ed) (2008) *A Companion to Greek Rhetoric*. Oxford: Blackwell Publishing.
Wotanis, L. and McMillan, L. (2014) *Performing gender on YouTube: How Jenna Marbles negotiates a hostile online environment.* Feminist Media Studies, 14(6), pp.912-928.
Wright, D. and Hinson, M. (2009) *Examining how public relations practitioners actually are using social media*. Public Relations Journal, 3(3), pp.1-33.
Wu, F. and Huberman, B. A. (2007) *Novelty and collective attention*. Proceedings of the National Academy of Sciences, 104(45), 17599-17601.Younie, S. and Bradshaw, P. (2017) *Debates in Computing and ICT Education*. London: Routledge.
Wu, S., Rizoiu, M. A. and Xie, L. (2018) *Beyond views: Measuring and predicting engagement in online videos*. In Twelfth International AAAI Conference on Web and Social Media, 2018, June 15.
Wu, Q., Sang, Y., Zhang, S. and Huang, Y. (2018) *Danmaku vs. forum comments: Understanding user participation and knowledge sharing in online videos*. In Proceedings of the 2018 ACM Conference on Supporting Groupwork. pp.209-218.
Xu, W. W., Park, J. Y., Kim, J. Y. and Park, H. W. (2016) *Networked cultural diffusion and creation on YouTube: An analysis of YouTube memes.* Journal of Broadcasting & Electronic Media, 60(1), 104e122.
Yang, H. C. and Wang, Y. (2015) *Social sharing of online videos: Examining American consumers' video sharing attitudes, intent, and behavior.* Psychology & Marketing, 32(9), pp.907-919.
Yiannakoulias, N., Slavik, C. E. and Chase, M. (2019) Expressions of pro-and anti-vaccine sentiment on YouTube. Vaccine, 37(15), pp.2057-2064.
Yoganarasimhan, H. (2012) *Impact of social network structure on content propagation: A study using YouTube data.* Quantitative Marketing and Economics, 10(1), pp.111-150.
YouTube (2014) *YouTube Statistics*. [online] Available at: http://www.YouTube.com/t/press_statistics Accessed: 23/9/14.
YouTube (2020) *Upload Video*. YouTube [online] Available at: https://studio.youtube.com/channel/UCwr4BJB22ZTPln15WtU18VA/videos/upload Accessed: 21/5/20

YouTube (2015) *YouTube Help*. Google [online] Available at:https://support.google.com/youtube/topic/ Accessed: 15/7/15

Yu, H., Zheng, D., Zhao, B, Y. and Zheng, W. (2006) *Understanding user behavior in large-scale video-on-demand systems.* in Proc. ACM Eurosys, 2006.

Zannettou, S., Chatzis, S., Papadamou, K., and Sirivianos, M. (2018) *The good, the bad and the bait: Detecting and characterizing clickbait on YouTube.* In 2018 IEEE Security and Privacy Workshops (SPW). pp. 63-69. IEEE.

Zeiler, X. (2018) *Researching Let's Play Gaming Videos as Gamevironments*. In DHN, pp.334-340.

Zhou, R., Khemmarat, S. and Gao, L. (2010) *The Impact of YouTube Recommendation System on Video Views*. MC'10, November 1–3, 2010, Melbourne, Australia.

Zhou, J., Li, Y., Adhikari, V. and Zhang, Z. (2011) *Counting YouTube videos via random prefix sampling*. In Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference, November, pp.371-380.

Zink, M., Suh, K., Gu, Y. and Kurose, J. (2008) *Watch global, cache local: YouTube network traffic at a campus network-measurements and implications.* MMCN 2008: Proc. of the 15th SPIE/ACM Multimedia Computing and Networking, 2008.

Zink, M., Suh, K., Gu, Y. and Kurose, J. (2009) *Characteristics of YouTube network traffic at a campus network–measurements, models, and implications.* Computer networks, 53(4), pp.501-514.

Appendix

Appendix 1

Single response: Ethical Approval Form (Faculty of Science and Engineering)

Survey input field	Respondent's answer
Name:	

1. Please enter your surname and first name below. (SURNAME, FIRST NAME)
Foster, David

2. Please enter your University email address (e.g. M.Name@wlv.ac.uk)
[redacted]

3. Please enter the name of your Director of Studies, Principal Investigator or, for Principal Investigators, your line manager.
Professor Michael Thelwall

4. Please enter date by which a decision is required below. (Note that decisions can take up to 4 working weeks from date of submission)
4th April 2016

5. Which subject area is your research / project located?
1. Architecture and Built Environment 2. Biology, Chemistry and Forensic Science 3. Engineering 4. Life Sciences 5. Mathematics and Computer Science 6. **other**

6. Please select your School
1. School of Architecture and Built Environment 2. School of Biomedical Science and Physiology 3. School of Biology, Chemistry and Forensic Science 4. School of Engineering 5. School of Mathematics and Computer Science 6. School of Pharmacy 7. **Other (please specify below)**
RIILP

7. Does your research fit into any of the following security-sensitive categories? (For definition of security sensitive categories see RPU webpages (www.wlv.ac.uk/rpu) follow links to Ethical Guidance).
1. commissioned by the military 2. commissioned under an EU security call 3. involve the acquisition of security clearances 4. concerns terrorist or extreme groups 5. **not applicable**

8. Does your research involve the storage on a computer of any records, statements or other documents that can be interpreted as promoting or endorsing terrorist acts?
1. YES 2. **NO**

9. Might your research involve the electronic transmission (eg as an email attachment) of any records or statements that can be interpreted as promoting or endorsing terrorist acts?
1. YES 2. **NO**

10. Do you agree to store electronically on a secure University file store any records or statements that can be interpreted as promoting or endorsing terrorist acts. Do you also agree to scan and upload any paper documents with the same sort of content. Access to this file store will be protected by a password unique to you. Please confirm you understand and agree to these conditions?
1. YES I understand and agree to the conditions 2. NO (please explain below) 3. I do not understand the conditions

11. You agree NOT to transmit electronically to any third party documents in the University secure document store?
1. YES I agree 2. NO I don't agree

12. Will your research involve visits to websites that might be associated with extreme, or terrorist, organisations? (for definition of extreme or terrorist organisations see RPU webpages (www.wlv.ac.uk/rpu) and follow links to Ethical Guidance.
1. YES (Please outline which Websites and why you consider this necessary) 2**. NO**

13. You are advised that visits to websites that might be associated with extreme or terrorist organisations may be subject to surveillance by the police. Accessing those sites from university IP addresses might lead to police enquiries. Do you understand this risk?
1. **YES I understand** 2. NO I don't understand

14. What is the title of your project?
Identifying and modelling factors associated with the popularity of YouTube videos.

15. Briefly outline your project, stating the rationale, aims, research question / hypothesis, and expected outcomes. Max 300 words.

The thesis is situated partly within the literature of informetrics and partly within the emerging interdisciplinary literature of quantitative analyses of social media content. This emerging literature draws on information science, computer science and physics in addition to gaining theoretical insights from media studies, psychology and sociology.
The research will review and discuss the concept of 'popularity' and how 'something' becomes popular, and addresses the following theories and concepts:

- **Popularity as a social concept**
- **Herding behaviours**
- **Information cascades**
- **Word-of-mouth**
- **Diffusion theory**
- **Online social transmission**
- **Social networks**
- **Online video popularity and virality**
- **YouTube - its growth as a popular entertainment, educational and instructional resource**
- **YouTube videos and their popularity**
- **Decisions and behaviours of individuals in YouTube video use**
- **Analysis and comparison of YouTube data with human responses.**

<u>Aims of research</u>

To determine the type and scope of sample provided by the YouTube API

To identity factors or features which have an impact on the popularity of YouTube videos.
To quantify factors or features which have an impact on the popularity of YouTube videos.
To determine the metrics which have a significant impact in individuals viewing videos.
To determine video metrics that could be used to predict the popularity of YouTube videos.

Research questions

Which YouTube video properties associate with their popularity in the number of views?
How accurately can the popularity of a YouTube video be predicted using the factors identified above?
How do the answers to the above questions vary by the type (category) of video?

16. How will your research be conducted?
Describe the methods so that it can be easily understood by the ethics committee.
Please ensure you clearly explain any acronyms and subject specific terminology. Max 300 words

Questionnaires will be given to a range of adults (hard and electronic form).
All questionnaires will be anonymous.
No children or venerable individuals will be used within the research.
The researcher will ensure that, as far as is possible, that people who provide information/data for the research are provided with information about the purpose and uses of the research data. They will be provided with enough information to be able to make an informed decision whether to participate within the research.
The researcher cannot foresee any risks with participating within the study.
Participants will be informed about their right to withdraw from the study (due to the anonymity of the study questionnaires will be individually numbered for withdrawal purposes – it will be highlighted that participants will need to make a note of their questionnaire number for these purposes).
It will be explained to participants that the researcher does not feel that as a result of being involved within the study that they will come to no harm and that there will be no possibility of physical or psychological distress.
All participants will be provided with a user friendly information and consent form which will contain all relevant information about the study and what their involvement will entail.
The information sheet will be written in simple, non-technical terms (avoiding jargon and abbreviations) and be easily understood by a lay person
All data collected in relation to the research project will be secured at all times to ensure confidentiality.

17. Is ethical approval required by an external agency? (e.g. NHS, company, other university, etc)

1. **NO**
2. YES - but ethical approval has not yet been obtained
3. YES - see contact details below of person who can verify that ethical approval has been obtained)

18. What in your view are the ethical considerations involved in this project? (e.g. confidentiality, consent, risk, physical or psychological harm, etc.) Please explain in full sentences. Do not simply list the issues. (Maximum 100) words)

Risk: low-risk project.

19. Have participants been/will participants be, fully informed of the risks and benefits of participating and of their right to refuse participation or withdraw from the research at any time?

1. **YES (Outline your procedures for informing participants in the space below.**
2. NO (Use the space below to explain why)
3. Not applicable - There are no participants in this study

20. Are participants in your study going to be recruited from a potentially vulnerable group? (See RPU website (www.wlv.ac.uk/rpu) and follow link to Ethical Guidance pages for definition of vulnerable groups)
1. YES (Describe below which groups and what measures you will take to respect their rights and safeguard them) 2. **NO**

21. How will you ensure that the identity of your participants is protected (See RPU website (www.wlv.ac.uk/rpu) and follow link to Ethical Guidance pages for guidance on anonymity)
All questionnaires will be anonymous – participants will not be asked to provide their name. **However, questionnaires will be numbered to enable individuals to withdraw their information (within the information/consent form it will be highlighted that individuals will need to make a note of their questionnaire number for withdrawal purposes).**

22. How will you ensure that data remains confidential (See RPU website (www.wlv.ac.uk/rpu) and follow link to Ethical Guidance pages for definition of confidentiality)
Questionnaires will be anonymous. **No specific individuals will be highlighted or discussed within the research.** **Due to the nature of the project none of the information collected will be personally compromising for the participants and anonymity will be maintained throughout with no individuals being identified.**

23. How will you store your data during and after the project? (See RPU website (www.wlv.ac.uk/rpu) and follow link to Ethical Guidance pages for definition of and guidance on data protection and storage).
No specific personal information relating to participants will be collected. **All data collected will be kept in a locked filing cabinet and/or on a password protected computer.**

Appendix 2

Consent to Participate in Research

Research title: Identifying and modelling factors associated with the popularity of YouTube videos.

Introduction

My name is David Foster and I am a PhD research student at University of Wolverhampton. Would you be willing to complete a questionnaire to help with my research? You are being invited to participate in this research because you are an internet user.

Purpose

The purpose of my study is to determine and model the factors that have an impact on the popularity of YouTube videos. As part of this, I would like to know internet users' opinions about the factors that influence their decisions to watch YouTube videos. If you complete the questionnaire then the information that you provide will be analysed and compared to data extracted from YouTube to determine any patterns and/or correlations in video popularity.

Procedures

If you agree to be in this research, then please read the information contained in this form and sign and return the consent form. Please make a note of your questionnaire number. You can request that I delete your answers at any stage in the future by sending me this number. The questionnaire should take no more than 10 minutes to complete. Your responses will remain confidential and your identity will not be recorded with your answers.

Please record your questionnaire number: ____

Questions

If you have any questions or concerns about this study, please contact David Foster at [redacted] or [redacted]

___

CONSENT FORM

Research title: Identifying and modelling factors associated with the popularity of YouTube videos.

Please retain a copy of this information sheet for your records.

If you wish to participate in this study, please sign and date below.

Your Name *(please print)*

____ _______________

Your Signature Date

This questionnaire has received ethical approval from the University of Wolverhampton.

Appendix 3

Version 1 of the questionnaire:

How and why we watch YouTube videos

When answering the following questions please enter your responses in the shaded areas.

1) What is your age? (please put an 'x' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender?	Female		Male	

3) Which of the following formal qualifications have you completed?	
Standard qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable what was the major subject of your <u>undergraduate degree</u>?

5) Do you watch videos on YouTube?	Yes		No	

6) How often do you generally watch YouTube videos?	
Daily	
Weekly	
Monthly	
Annually	

7) How do you usually access YouTube videos? (choose ones you use regularly)			
YouTube		Email	
Blogs		Facebook	
Twitter		Pinterest	
Other			
Please specify			

8) If you use the YouTube website how do you find videos?			
Homepage		Most viewed page	
Most popular page		YouTube recommendations	
Search		Other	
Please specify			

9) Which of the following factors are important in deciding whether you watch a YouTube video? (please select all that apply)				
Category	☐		Title	☐
Number of likes	☐		Number of dislikes	☐
View count	☐		Number of comments	☐
Length of video	☐		Other	☐
Please specify				

10) Which of the factors you have chosen is the most important? And why?				
Category	☐		Title	☐
Number of likes	☐		Number of dislikes	☐
View count	☐		Number of comments	☐
Length of video	☐		Other	☐
Why				

11) Are there any further comments that you would like to make about your use of YouTube?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 4

Version 2 of the questionnaire:

How and why we watch YouTube videos

When answering the following questions please enter your responses in the shaded areas.

1) What is your age? (please put an 'x' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender?	Female		Male	

3) Which of the following formal qualifications have you completed? (please select all that apply)	
Standard qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable what was the major subject of your undergraduate degree?

5) Do you watch YouTube videos?	Yes		No	

6) On average how often do you watch YouTube videos? (please put an 'x' in the relevant box)			
Daily		Weekly	
Monthly		Annually	

7) How do you usually access YouTube videos? (select ones you use regularly)			
YouTube		Email	
Blogs		Facebook	
Twitter		Other	
If other please specify			

8) How do you find videos on the YouTube website? (select all that apply)			
Homepage videos		Most viewed page	
Most popular page		YouTube recommendations	
Search		Other	
If other please specify			

9) Which of the following factors are important in deciding whether you watch a YouTube video? (please select all that apply)			
YouTube Category		Title	
Number of likes		Number of dislikes	
View count		Number of comments	
Length of video		Date uploaded	
Other			
Please specify			

10) Which of the factors you have chosen in question 9 is the most important? And why?	
Factor	
Why	

11) Are there any further comments you would like to make about how you find YouTube videos?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 5

Version 3 of the questionnaire:

How and why we watch YouTube videos

When answering the following questions please enter your responses in the shaded areas.

1) What is your age? (please put an '**x**' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender? (please put an '**x**' in the relevant box)							
Female		Male		Other		Prefer not to say	

3) Which of the following formal qualifications have you completed? (please put an '**x**' in all relevant boxes)	
Standard qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable what was the major subject of your <u>undergraduate degree</u>?

5) On average how often have you watched YouTube videos in the past year? (please put an '**x**' in the relevant box)	
On most days I have watched at least one YouTube video	
During most weeks I have watched at least one YouTube video	
During most months I have watched at least one YouTube video	
During the past year I have watched at least one YouTube video	
I never watch YouTube videos	

6) How do you usually access YouTube videos? (please select, using an '**x**', the ones you have used in the last year)	
Through accessing the YouTube website	
From a hyperlink sent to you in an email	
From a hyperlink in a Blog	
From a hyperlink or post on Facebook	
From a hyperlink or Tweet on Twitter	
From a verbal recommendation	
Other	
If other please specify	

7) If you are using the YouTube website how do you find videos? (please select, using an '**x**', all that apply)	
Videos posted on the homepage	
Videos posted on the most popular page	
Videos posted on the most viewed page	
Videos posted on your subscriptions page	
Videos posted on the recommendations bar	

Through the YouTube search facility	
Other	
If other please specify	

8) How important do you feel the following are when deciding whether to view a YouTube video?
(please use an 'x' to rate the importance of each)

	Irrelevant	Not very Important	Slightly Important	Very Important
Video category (e.g. Animation)				
Title of video				
Number of likes				
Number of dislikes				
Number of views				
Number of comments				
Length of video				
Upload date of video				
Thumbnail picture of the video				

9) Considering the last 10 videos that you have watched on YouTube, how often have the following influenced your decision to watch?
(please use an '**x**' to rate each)

	Never	Rarely	Sometimes	Mostly	Always
Video category (e.g. Animation)					
Title of video					
Number of likes					
Number of dislikes					
Number of views					
Number of comments					
Length of video					
Upload date of video					
Thumbnail picture of the video					
Other					
If other please specify					

10) What type(s) of video do you typically watch in YouTube?
(please put an '**x**' in all relevant boxes)

Animation		Automotive		Beauty	
Best of		Causes		Comedy	
Cooking		DIY		Education	
Entertainment		Fashion		From TV	
Gaming		Health		How-to	
Lifestyle		News		Non-profit	
Politics		Science		Sport	
Tech		Other			
If other please specify					

11) Are there any further comments you would like to make about how you find YouTube videos?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 6

Version 4 of the questionnaire:

How and why you watch YouTube videos

Please answer the questions in the shaded areas.

1) What is your age? (please put an '**x**' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender? (please put an '**x**' in the relevant box)							
Female		Male		Other		Prefer not to say	

3) Which of the following formal qualifications have you completed? (please put an '**x**' in all relevant boxes)	
School qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable, what was the major subject of your <u>undergraduate degree</u>?

5) On average how often have you watched YouTube videos in the past year? (please put an '**x**' in the box for the highest (most frequent) option that applies)	
On most **days** I have watched at least one YouTube video	
During most **weeks** I have watched at least one YouTube video	
During most **months** I have watched at least one YouTube video	
During **the past year** I have watched at least one YouTube video	
I have **not watched** YouTube videos in the past year	

6) How have you accessed YouTube videos in the past year? (please select, using an '**x**', **<u>all</u>** the methods that you have used in the last year)	
Through searching or browsing the **YouTube website**	
From a hyperlink sent to you in an **email**	
From a hyperlink in a **Blog**	
From a hyperlink or post on **Facebook**	
From a hyperlink or Tweet on **Twitter**	
From a **verbal** recommendation	
Other	
If other please specify	

7) Which of the following methods have you used, if any, to find videos <u>through the YouTube website</u> in the past year? (please select, using an '**x**', **<u>all</u>** that apply)	
Videos posted on the **YouTube homepage**	
Videos posted on the **YouTube most popular** page	
Videos posted on the **YouTube most viewed** page	

Videos posted on **your YouTube subscriptions** page	
Videos posted on the **YouTube website recommendations bar**	
Through the **YouTube search** facility	
Other	
If other please specify	

8) Considering <u>only the most recent videos that you have watched on YouTube</u>, how often have the following influenced your decision to watch?
(please use an '**x**' to rate each)

	Never	**Rarely**	**Sometimes**	**Mostly**	**Always**
The official YouTube category of the video (e.g. Animation/Best of)					
Title of video					
Number of likes					
Number of dislikes					
Number of views					
Number of comments					
Length of video					
Upload date of video					
Thumbnail picture of the video					
Other					
If other please specify					

9) Which of the following would you consider when deciding NOT to watch a video on YouTube?
(please use an '**x**' to rate each)

	Would not consider	**Might consider**	**Would consider**
The official YouTube category of the video (e.g. Animation/Best of)			
Title of video			
Number of likes			
Number of dislikes			
Number of views			
Number of comments			
Length of video			
Upload date of video			
Thumbnail picture of the video			
Other			
If other please specify			

10) Which type(s) of video have you watched in YouTube during the past year?
(please put an '**x**' in all relevant boxes)

Animation		Automotive		Beauty	
Best of		Causes		Comedy	
Cooking		DIY		Education	
Entertainment		Fashion		From TV	
Gaming		Health		How-to	
Lifestyle		News		Non-profit	
Politics		Science		Sport	

Tech		Other	
If other please specify			
11) Are there any further comments you would like to make about how you find YouTube videos?			

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 7

Version 5 of the questionnaire:

How and why you watch YouTube videos

Please answer the questions in the shaded areas.

1) What is your age? (please put an '**x**' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender? (please put an '**x**' in the relevant box)							
Female		Male		Other		Prefer not to say	

3) Which of the following formal qualifications have you completed? (please put an '**x**' in all relevant boxes)	
School qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable, what was the major subject of your <u>undergraduate degree</u>?

5) On average how often have you watched YouTube videos in the past year? (please put an '**x**' in the box for the highest (most frequent) option that applies)	
On most **days** I have watched at least one YouTube video	
During most **weeks** I have watched at least one YouTube video	
During most **months** I have watched at least one YouTube video	
During **the past year** I have watched at least one YouTube video	
I have **not watched** YouTube videos in the past year	

6) How have you accessed YouTube videos in the past year? (please select, using an '**x**', **<u>all</u>** the methods that you have used in the last year)	
Through searching or browsing the **YouTube website**	
From a hyperlink sent to you in an **email**	
From a hyperlink in a **Blog**	
From a hyperlink or post on **Facebook**	
From a hyperlink or Tweet on **Twitter**	
From a **verbal** recommendation	
Other	
If other please specify	

7) Which of the following methods have you used, if any, to find videos <u>through the YouTube website</u> in the past year? (please select, using an '**x**', **<u>all</u>** that apply)	
Videos posted on the **YouTube homepage**	
Videos posted on the **YouTube most popular** page	
Videos posted on the **YouTube most viewed** page	
Videos posted on **your YouTube subscriptions** page	

Videos posted on the **YouTube website recommendations bar**	
Through the **YouTube search** facility	
Other	
If other please specify	

8) Considering <u>only the most recent videos that you have watched on YouTube</u>, how often have the following influenced your decision to watch?
(please use an '**x**' to rate each)

	Never	Rarely	Sometimes	Mostly	Always
The official YouTube category of the video (e.g. Animation/Best of)					
Title of video					
Number of likes					
Number of dislikes					
Number of views					
Number of comments					
Length of video					
Upload date of video					
Thumbnail picture of the video					
Other					
If other please specify					

9) Which type(s) of video have you watched in YouTube during the past year?
(please put an '**x**' in all relevant boxes)

Animation		Automotive		Beauty	
Best of		Causes		Comedy	
Cooking		DIY		Education	
Entertainment		Fashion		From TV	
Gaming		Health		How-to	
Lifestyle		News		Non-profit	
Politics		Science		Sport	
Tech		Other			
If other please specify					

11) Are there any further comments you would like to make about how you find YouTube videos?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 8

Version 6 of the questionnaire:

How and why you watch YouTube videos

Please answer the questions in the shaded areas.

1) What is your age? (please put an '**x**' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender? (please put an '**x**' in the relevant box)							
Female		Male		Other		Prefer not to say	

3) Which of the following formal qualifications have you completed? (please put an '**x**' in all relevant boxes)	
School qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable, what was the major subject of your <u>undergraduate degree</u>?

5) On average how often have you watched YouTube videos in the past year? (please put an '**x**' in the box for the highest (most frequent) option that applies)	
On most **days** I have watched at least one YouTube video	
During most **weeks** I have watched at least one YouTube video	
During most **months** I have watched at least one YouTube video	
During **the past year** I have watched at least one YouTube video	
I have **not watched** YouTube videos in the past year	

6) How have you accessed YouTube videos in the past year? (please select, using an '**x**', **<u>all</u>** the methods that you have used in the last year)	
Through searching or browsing the **YouTube website**	
From a hyperlink sent to you in an **email**	
From a hyperlink in a **Blog**	
From a hyperlink or post on **Facebook**	
From a hyperlink or Tweet on **Twitter**	
From a **verbal** recommendation	
Other	
If other please specify	

7) Which of the following methods have you used, if any, to find videos <u>through the YouTube website</u> in the past year? (please select, using an '**x**', **<u>all</u>** that apply)	
Videos posted on the **YouTube homepage**	
Videos posted on the **YouTube most popular** page	
Videos posted on the **YouTube most viewed** page	
Videos posted on **your YouTube subscriptions** page	

Videos posted on the **YouTube website recommendations bar**	
Through the **YouTube search** facility	
Other	
If other please specify	

8) Which type(s) of video have you watched in YouTube during the past year?
(please put an '**x**' in all relevant boxes)

Animation		Automotive		Beauty	
Best of		Causes		Comedy	
Cooking		DIY		Education	
Entertainment		Fashion		From TV	
Gaming		Health		How-to	
Lifestyle		News		Non-profit	
Politics		Science		Sport	
Tech		Other			

If other please specify	

9) Considering <u>only the most recent videos that you have watched on YouTube</u>, how often have the following influenced your decision to watch?
(please use an '**x**' to rate each)

	Never	**Rarely**	**Sometimes**	**Mostly**	**Always**
The official YouTube category of the video (e.g. Animation/Best of)					
Title of video					
Number of likes					
Number of dislikes					
Number of views					
Number of comments					
Length of video					
Upload date of video					
Thumbnail picture of the video					
Other					
If other please specify					

10) Are there any further comments you would like to make about how you find YouTube videos?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 9

Version 7 and final version of the questionnaire:

<u>How and why you watch YouTube videos</u>

Please answer the questions in the shaded areas.

1) What is your age? (please put an '**x**' in the relevant box)									
18 to 24		25 to 34		35 to 44		45 to 54		55 +	

2) How would you define your gender? (please put an '**x**' in the relevant box)							
Female		Male		Other		Prefer not to say	

3) Which of the following formal qualifications have you completed? (please put an '**x**' in all relevant boxes)	
School qualifications (e.g. GCSE, O-Level)	
Further Education (e.g. A-Level, BTEC, GNVQ, Fd)	
Higher Education (Undergraduate degree, Masters, Doctorate)	

4) If applicable, what was the major subject of your <u>undergraduate degree</u>?

5) On average how often have you watched YouTube videos in the past year? (please put an '**x**' in the box for the highest (most frequent) option that applies)	
On most **days** I have watched at least one YouTube video	
During most **weeks** I have watched at least one YouTube video	
During most **months** I have watched at least one YouTube video	
During **the past year** I have watched at least one YouTube video	
I have **not watched** YouTube videos in the past year	

6) How have you accessed YouTube videos in the past year? (please select, using an '**x**', **<u>all</u>** the methods that you have used in the last year)	
Through searching or browsing the **YouTube website**	
Through the **YouTube app** (phone or tablet)	
From a hyperlink sent to you in an **email**	
From a hyperlink in a **Blog**	
From a hyperlink or post on **Facebook**	
From a hyperlink or Tweet on **Twitter**	
From a **verbal** recommendation	
From a **Google** search	
Other	
If other please specify	

7) Which of the following methods have you used, if any, to find videos <u>through the YouTube website</u> in the past year? (please select, using an '**x**', **<u>all</u>** that apply)	
Videos posted on the **YouTube homepage**	
Videos posted on the **YouTube most popular** page	
Videos posted on the **YouTube most viewed** page	

Videos posted on **your YouTube subscriptions** page	
Videos posted on the **YouTube website recommendations bar**	
Through the **YouTube search** facility	
Other	
If other please specify	

8) Which type(s) of video have you watched in YouTube during the past year?
(please put an '**x**' in all relevant boxes)

Animation		Automotive		Beauty	
Best of		Causes		Comedy	
Cooking		DIY		Education	
Entertainment		Fashion		From TV	
Gaming		Health		How-to	
Lifestyle		News		Non-profit	
Politics		Science		Sport	
Tech		Other			
If other please specify					

9) Considering <u>only the most recent videos that you have watched on YouTube</u>, how often have the following influenced your decision to watch?
(please use an '**x**' to rate each)

	Never	**Rarely**	**Sometimes**	**Mostly**	**Always**
The official YouTube category of the video (e.g. Animation/Best of)					
Title of video					
Number of likes					
Number of dislikes					
Number of views					
Number of comments					
Length of video					
Upload date of video					
Thumbnail picture of the video					
Other					
If other please specify					

10) Are there any further comments you would like to make about how you find YouTube videos?

Thank you for completing this questionnaire – your time is very much appreciated.

Appendix 10

Further discussion relating to YouTube API data:

The following tables show the accumulated, processed and organised data from the 'Five Day Searches' and have been colour graded to show the differences between the values within each of the tables – the more prominent the green the higher the values, the more prominent the red the lower the values and yellow representing the mid-range values.

There are fluctuations in the percentages of repeats between searches (Table A.1) and there seems to be more substantial repeats at points ('2 and 3 to 6 and 7' and '18 and 19 to 19 and 20'). This suggests influxes of significant amounts of new videos being uploaded to YouTube at these points or could just be random.

Within the categories Automotive, Causes and Non-profit there seem to be a high level of repeats maintained suggesting a lower turnover or less videos that are uploaded to or within these categories (Table A.1). The API has fewer overall videos to select from when providing information for a search. Animation, Comedy, How-to and News seem to have lower repeats across the searches suggesting either a higher turnover more videos for the API to select from.

Comparing the data (Table A.1) with previous findings (Table 5.1) it seems that submitting the searches over a greater period between searches that at points there is a slightly greater chance of getting repeats in the API samples that are extracted. This seems illogical as new videos are being uploaded to YouTube at a significant rate and daily (Welbourne and Grant, 2016; Gaunt, 2015; Barry et al., 2014; Thelwall et al., 2012; Burgess and Green, 2009; Freeman and Chapman, 2007; Gill et al., 2007) and would provide the API with a greater selection from which to extract a sample. It could be that there was a change in the way that the YouTube API selected the videos between when the daily searches and Five Day Searches were submitted. These differences across the Daily and Five Day Searches could also just reflect the possible random nature of how the API selects videos for each sample. Nevertheless, further comparison and analysis of the data (Tables 5.1 and A.1) shows that there are less fluctuations and more consistency in the percentages of repeats for the Daily Searches suggesting that the repeats are more consistent than those submitted 5 days apart.

Table A.1. Five Day Searches - The percentage of videos in each search that also occurred in the next search. High values are green, mid-range values are yellow and low values are red.

Search	1 - 2	2 - 3	3 - 4	4 - 5	5 - 6	6 - 7	7 - 8	8 - 9	9 - 10	10 - 11	11 - 12	12 - 13	13 - 14	14 - 15	15 - 16	16 - 17	17 - 18	18 - 19	19 - 20
Animation	58.94	82.38	70.56	72.58	80.89	77.13	48.27	63.17	60.73	57.69	47.63	39.56	60.57	69.53	55.53	61.77	55.04	88.01	77.35
Automotive	85.08	95.38	91.58	91.52	90.74	91.92	79.12	86.35	86.12	86.50	80.82	75.55	84.20	90.98	87.32	85.34	81.78	95.94	93.33
Beauty	59.04	80.96	73.39	61.09	80.16	75.81	54.56	57.78	65.25	62.91	58.27	51.41	63.29	69.03	67.40	53.23	56.79	87.47	79.68
Best of	49.40	74.15	71.57	58.47	74.44	73.99	46.45	57.37	59.19	60.66	52.62	41.94	56.59	64.17	62.98	59.07	55.35	84.39	77.06
Causes	84.88	93.60	92.37	85.54	87.63	88.58	82.16	86.06	84.94	86.94	79.72	77.87	86.60	89.78	86.57	85.14	83.57	95.34	90.40
Comedy	49.49	75.30	66.26	54.58	71.34	70.21	42.86	51.04	51.63	51.43	47.40	41.62	54.67	60.48	56.65	51.41	47.78	79.51	73.68
Cooking	65.92	86.62	75.76	73.58	80.08	79.32	54.96	69.07	65.45	65.36	58.70	50.40	70.04	75.30	69.22	66.53	62.10	90.06	83.13
DIY	54.56	82.69	71.52	70.54	80.16	79.55	54.05	66.81	58.79	63.75	56.05	48.47	62.53	74.55	66.26	64.63	57.14	91.31	81.14
Education	71.40	89.21	83.84	77.15	79.55	79.55	64.17	72.44	68.08	71.69	68.15	63.80	72.34	81.41	80.69	76.22	64.99	83.43	80.53
Entertainment	52.31	78.71	77.06	62.35	77.41	72.61	42.91	58.32	56.91	58.70	62.37	49.60	60.72	66.47	65.73	65.91	57.55	83.70	79.76
Fashion	63.40	85.22	79.80	72.49	80.86	77.60	53.35	64.50	62.96	63.09	53.12	45.56	59.12	72.87	67.54	64.79	59.39	89.84	82.36
From TV	50.31	77.54	70.68	64.86	78.09	70.49	47.59	57.11	61.52	67.28	66.53	49.29	66.60	68.74	70.28	66.53	65.16	83.37	79.96
Gaming	65.05	83.43	83.23	70.30	79.92	80.85	56.94	68.15	65.11	67.29	62.02	53.05	62.22	69.04	67.89	66.06	62.02	87.90	82.19
Health	62.68	85.45	80.44	75.35	76.63	80.25	63.89	69.83	66.74	64.02	57.43	53.12	69.92	73.89	69.22	66.60	63.43	88.55	82.39
How to	47.24	79.11	66.13	58.91	71.25	69.15	38.19	46.75	48.19	45.42	43.81	27.24	42.74	51.51	49.19	51.01	39.34	84.43	75.56
Lifestyle	76.21	88.45	80.85	78.09	85.51	86.21	69.72	75.10	77.08	76.42	71.77	65.32	77.15	80.61	77.96	78.63	72.95	89.70	88.26
News	47.69	73.81	69.73	59.54	71.89	65.47	39.14	47.44	49.80	54.13	49.18	32.20	46.79	57.92	53.33	53.02	41.12	81.96	71.43
Non-profit	89.36	95.77	95.37	93.00	96.26	92.79	80.71	87.24	89.18	90.17	89.00	86.37	88.58	93.19	92.00	91.40	88.73	96.77	93.57
Politics	75.05	89.14	85.48	79.11	87.14	86.88	67.35	71.96	74.85	78.02	74.45	66.73	78.11	83.00	82.22	81.74	75.41	93.37	89.66
Science	66.73	85.92	78.30	76.63	84.65	82.65	62.55	70.90	68.83	67.41	60.04	50.30	65.12	77.94	73.75	68.90	67.14	93.25	87.58
Sports	62.55	86.91	80.49	66.60	81.84	79.72	55.38	62.40	64.68	64.10	64.42	56.43	64.59	70.64	65.99	67.54	64.11	91.35	84.54
Tech	60.24	86.29	77.44	69.57	79.75	77.98	55.89	64.43	69.18	67.35	66.26	53.78	68.98	71.17	69.59	65.79	62.68	85.71	78.14

The percentage of repeats decreases over the period of the 20 searches (over a 95 day period) (Table A.2) and is similar to the daily search data (Table 5.2). When searches are taken over a longer period the percentage of repeated videos reduces. Also, comparing the data (Tables 5.2 and A.2), the rate at which the repeated videos reduces is different for each of the categories. Although Automotive, Causes and Non-profit have a reduction in the percentage of their repeats these remain higher than the other categories. Which was also the case for these categories within previous data (Table 5.2).

The categories How-to, News and Comedy have a quicker and more substantial reduction in the percentages of their repeats and this is the case for both the daily and Five Days searches (Tables 5.2 and A.2). This further suggests that these categories have a higher turnover of new material, have an overall higher number of videos or both.

For the category News, unlike the other data collected (Table 5.2) where after 15 searches there were no further repeats with the first search, this shows (Table A.2) that there continue to be repeats across the 20 searches (and period of 95 days). It is difficult, just from this data collected, to be able to accurately determine the reason for this substantial difference between the two sets of data, particularly considering the much longer overall time frame during which the Five Day Searches were submitted. There could have been a change in the way that API selects it sample between the periods that the Daily and Five Day Searches were submitted.

Table A.2. Five Day Searches - The percentage of videos in the first search that also occurred in each subsequent search**.** High values are green, mid-range values are yellow and low values are red.

Search	1 - 2	1 - 3	1 - 4	1 - 5	1 - 6	1 - 7	1 - 8	1 - 9	1 - 10	1 - 11	1 - 12	1 - 13	1 - 14	1 - 15	1 - 16	1 - 17	1 - 18	1 - 19	1 - 20
Animation	58.94	60.25	55.44	53.23	52.31	49.69	44.40	42.18	40.08	36.03	29.90	31.53	29.47	27.40	26.56	27.57	23.79	21.54	22.44
Automotive	85.08	84.54	83.57	82.02	80.08	79.80	74.70	72.51	70.62	69.94	66.12	68.14	66.40	65.53	65.19	63.45	62.55	62.07	62.22
Beauty	59.04	57.52	54.44	50.60	50.31	45.36	43.81	41.01	39.39	40.16	35.89	36.29	34.48	33.00	33.20	32.26	27.57	28.95	26.96
Best of	49.40	47.90	46.57	43.15	44.58	42.74	39.55	40.61	37.17	34.43	29.64	31.85	29.61	29.76	27.97	28.23	26.34	27.31	28.17
Causes	84.88	84.60	84.54	81.12	80.53	80.76	78.76	78.18	75.90	76.94	73.09	74.65	72.60	73.15	72.55	70.88	70.54	70.45	69.60
Comedy	49.49	46.36	42.28	42.57	42.04	41.25	36.33	34.79	33.27	31.43	30.40	27.27	28.05	26.41	22.58	24.10	22.58	21.50	21.86
Cooking	65.92	65.39	60.61	57.11	56.02	55.06	51.03	50.00	49.29	42.68	37.85	38.20	37.45	37.15	34.61	36.49	31.85	32.86	32.13
DIY	54.56	55.19	53.74	50.90	50.40	48.99	44.74	42.59	38.99	36.46	31.85	31.49	29.86	30.30	28.46	26.83	22.13	22.22	22.92
Education	71.40	72.91	69.49	67.94	66.60	65.99	59.11	61.80	56.16	55.19	53.43	52.76	52.10	51.52	51.22	47.76	43.46	45.66	47.26
Entertainment	52.31	50.60	48.09	44.94	44.98	43.15	37.04	37.58	36.79	35.83	33.60	33.00	31.06	32.73	31.46	32.44	33.27	32.80	30.57
Fashion	63.40	63.77	61.01	54.82	52.88	50.51	42.39	42.80	39.47	37.11	35.41	27.42	31.06	30.97	31.05	30.18	27.88	28.05	29.06
From TV	50.31	50.76	49.40	46.18	46.38	44.16	44.23	47.42	45.49	43.62	40.47	40.94	39.00	38.68	41.16	38.48	36.68	36.67	37.19
Gaming	65.05	63.43	61.21	56.16	54.97	54.84	50.10	48.39	45.23	48.54	44.24	44.92	40.40	39.51	39.43	37.95	33.13	33.47	34.41
Health	62.68	63.03	62.70	61.01	57.11	58.23	52.94	51.86	50.41	54.07	44.18	46.68	44.72	44.13	42.86	42.05	43.03	43.57	41.70
How to	47.24	46.45	40.73	38.26	35.93	33.67	30.39	23.98	23.99	20.98	18.66	16.26	16.33	17.71	15.73	16.13	14.34	14.75	14.26
Lifestyle	76.21	73.81	70.56	72.01	69.18	67.28	66.26	65.84	63.08	60.77	58.87	58.47	55.31	54.95	54.31	53.23	51.23	50.71	49.60
News	47.69	43.92	39.88	38.80	37.34	34.74	29.51	29.70	27.71	26.24	23.57	17.60	19.48	18.24	17.37	17.74	15.08	15.03	14.49
Non-profit	89.36	89.13	88.33	86.60	86.07	86.03	80.29	79.01	78.56	78.03	78.60	74.95	73.55	72.95	71.80	70.40	70.22	70.56	70.08
Politics	75.05	72.34	70.36	65.52	65.35	67.31	61.22	64.95	62.27	61.90	59.15	59.92	57.03	56.28	56.57	54.36	53.46	54.22	53.75
Science	66.73	66.53	61.87	60.98	59.96	59.59	54.94	51.02	49.80	46.56	42.17	43.64	43.15	43.72	42.29	41.67	39.80	39.67	39.31
Sports	62.55	61.96	60.57	58.55	55.71	54.36	50.71	49.80	48.25	47.67	46.01	41.97	41.85	43.74	39.88	39.72	40.12	37.83	37.35
Tech	60.24	60.28	59.35	55.78	53.78	52.67	47.76	46.34	45.51	43.47	42.63	44.79	43.27	41.72	41.22	41.05	40.97	40.82	39.47

There are generally a high proportion of videos, across the categories, which were not repeated throughout the 20 samples extracted from YouTube (Table A.3). Nevertheless, the categories Non-profit, Causes and Automotive have a much higher percentage of videos, than the other categories, that have appeared within all 20 searches. This suggests that these categories possibly have a lower overall number of videos for the API to choose from when extracting a sample. Alternatively, it might suggest that the YouTube API changes the way in which it searches based on the category. Unfortunately, as this is a sample of YouTube any generalisations cannot be applied to all YouTube API searches. When the data (Table A.3) is compared to the data extracted for the Daily Searches, (Table 5.3), the overall percentage of videos without repeat is much less and does not follow the same patterns in the categories. For example, News has 55% of its Five Day Searches videos not repeated (Table A.3), but only 30% of Daily Searches videos not repeated (Table 5.3). This also occurs for some other categories and shows that there is a clear difference in the types of sample provided by the two different search approaches. Again, this could be explained by either a change in the way that the API chooses it sample (each time or for each category) or it could be a result of the API selecting videos at random that have caused these noticeable differences. Overall, both sets of data (Tables 5.3 and A.3) show that there are generally more videos that are not repeated across the searches and most of the ones that are repeated are only repeated two times.

Table A.3. Five Day Searches - The percentage of video appearances during the 95 day period**.** High values are green, mid-range values are yellow and low values are red.

No of appearances	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
News	54.9	16.5	10.0	4.4	2.8	2.1	1.5	1.3	0.8	1.1	0.8	0.9	0.6	0.3	0.1	0.3	0.4	0.2	0.6	0.4
How to	53.3	16.4	11.0	5.3	3.2	2.2	1.7	1.3	0.9	0.8	0.8	0.5	0.5	0.2	0.5	0.2	0.2	0.3	0.1	0.5
Best of	48.4	16.0	8.9	4.8	3.3	2.9	2.4	2.0	1.4	1.3	1.5	0.9	0.9	0.6	0.7	0.6	0.5	0.5	0.9	1.6
DIY	48.2	15.2	9.5	4.4	3.3	2.4	2.5	2.3	1.1	1.4	0.9	1.2	1.4	0.6	0.8	0.6	0.6	0.7	0.8	2.2
Animation	46.4	15.4	9.6	5.2	3.7	2.9	3.0	2.7	1.6	1.4	1.1	0.9	1.0	0.5	0.8	0.5	0.4	0.9	0.7	1.3
Comedy	45.8	16.3	11.3	7.2	3.8	3.0	2.2	1.9	1.5	1.2	1.1	0.5	0.8	0.6	0.4	0.3	0.7	0.4	0.3	0.7
Tech	44.6	14.8	7.6	5.3	2.5	3.0	3.0	2.1	1.8	1.1	1.4	1.8	1.4	0.7	1.0	1.0	1.0	0.8	1.6	3.5
Fashion	44.1	13.4	8.3	6.3	3.8	3.3	3.5	3.3	2.2	1.8	1.7	1.2	0.9	0.9	0.7	0.8	0.6	0.4	0.6	2.2
Entertainment	44.0	15.1	9.8	5.0	3.4	3.2	2.8	2.2	2.3	1.8	1.6	1.2	1.2	1.1	0.8	0.7	0.7	0.6	1.1	1.4
Beauty	42.4	14.7	9.8	5.9	3.4	3.6	2.8	2.5	2.1	1.6	1.3	1.3	1.4	1.2	0.8	1.5	0.7	0.5	0.8	1.7
Gaming	42.4	13.9	9.6	5.1	4.2	3.2	3.3	2.3	1.9	1.2	1.2	1.2	1.4	0.8	0.8	0.8	1.1	0.6	1.1	4.0
Health	41.4	12.4	7.5	4.3	4.3	3.2	4.0	2.1	2.3	2.3	1.8	1.6	1.9	1.3	1.5	0.9	0.8	0.9	1.9	3.5
From TV	41.1	15.9	9.1	5.5	4.2	3.0	2.6	1.9	1.8	1.4	1.4	1.6	1.1	0.8	1.0	0.9	1.1	1.5	1.6	2.4
Sports	40.8	13.7	10.0	5.5	3.6	3.1	3.0	2.7	2.0	1.8	1.4	1.5	1.2	0.9	1.1	1.0	0.7	1.0	1.6	3.5
Science	36.3	13.0	8.9	6.4	4.9	3.4	4.1	3.2	1.8	2.0	1.5	1.8	1.3	1.0	1.4	1.0	1.3	1.1	1.9	3.9
Education	35.8	10.1	6.5	5.9	3.3	3.5	4.0	2.7	2.5	2.6	3.1	2.1	2.0	1.9	1.9	1.0	1.6	2.3	2.0	5.1
Politics	35.3	10.4	6.0	4.0	3.1	2.7	3.2	2.6	2.7	2.4	2.3	2.4	2.0	2.1	1.8	2.4	2.0	1.6	2.8	8.2
Cooking	34.6	12.6	10.4	6.8	5.7	3.8	3.6	3.1	2.9	2.1	1.9	1.5	1.7	1.3	1.1	0.8	1.0	0.7	1.1	2.9
Lifestyle	28.2	11.3	7.6	5.6	4.3	4.3	3.7	2.7	2.7	2.2	1.9	1.9	2.7	1.7	2.2	1.8	1.8	1.8	3.2	8.3
Causes	27.3	7.9	6.7	4.2	3.3	2.8	2.2	2.5	2.2	1.9	1.7	2.1	2.2	1.3	1.9	2.1	1.4	1.8	4.3	20.0
Non-profit	15.5	6.2	5.5	3.6	3.9	3.4	3.9	2.4	2.1	2.2	2.7	2.8	3.0	1.4	2.3	2.2	2.1	2.4	7.9	24.4
Automotive	13.2	7.8	6.6	4.9	3.3	3.0	2.9	3.4	4.1	4.1	3.5	3.2	4.1	3.2	3.1	2.5	3.1	2.4	6.0	16.0

Comparing both sets of data relating to videos with no repeats (Tables 5.3 and A.3) shows that there are substantial differences between the Daily and Five Day Searches for most of the categories. However, this would be expected due to the different number of searches (30 for the Daily and 20 for the Five Day). Taking this into consideration there is still substantial variation in terms of what the API has returned. It is therefore impossible to determine the exact cause of the difference and it could be due to the different amount of searches, it could be merely random or there may have been a change in the way in which the API chooses the videos.

Table A.4. Compares the percentage of videos across the Five Day Searches and Daily Searches that have no repeats. High values are green, mid-range values are yellow and low values are red.

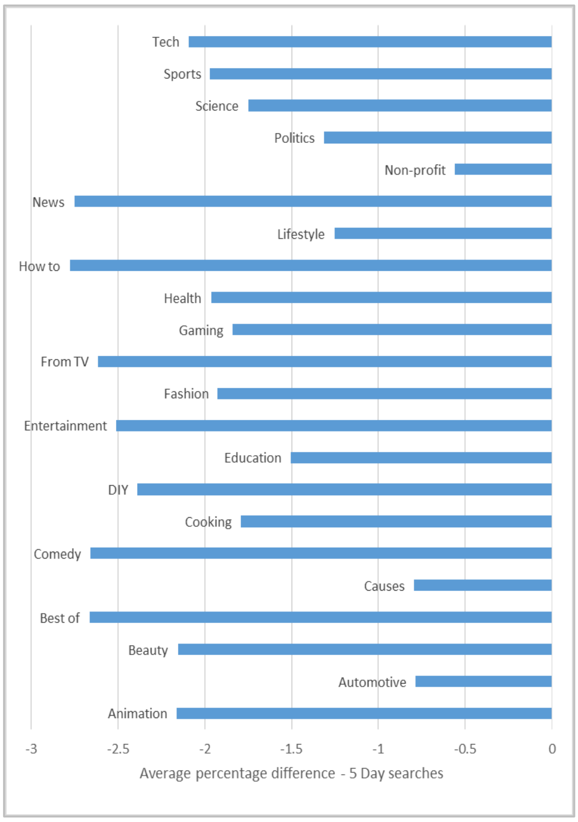
Appearance	5DS - No Repeat	DS - No Repeat
News	54.9	29.6
How to	53.3	16.0
Best of	48.4	55.6
DIY	48.2	54.3
Animation	46.4	58.3
Comedy	45.8	47.3
Tech	44.6	68.1
Fashion	44.1	63.0
Entertainment	44.0	58.7
Beauty	42.4	55.7
Gaming	42.4	37.5
Health	41.4	68.8
From TV	41.1	17.3
Sports	40.8	61.6
Science	36.3	69.5
Education	35.8	68.4
Politics	35.3	66.7
Cooking	34.6	66.8
Lifestyle	28.2	68.5
Causes	27.3	63.1
Non-profit	15.5	26.9
Automotive	13.2	69.6

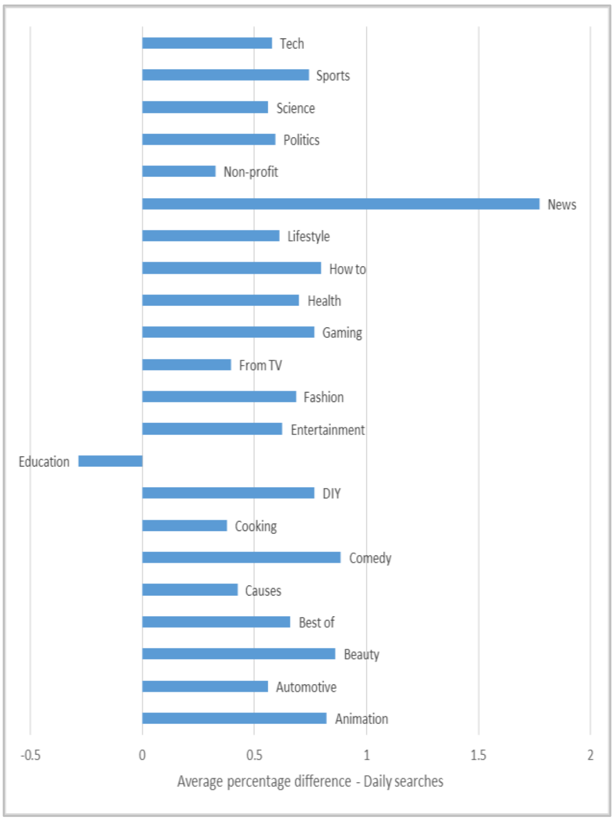
Again there are substantial differences between the percentages of videos that appeared all searches (Table A.5). However, as with the comparison of the videos with no repeat (Table A.4) there are too many variables in order to determine any specific conclusions about the API and how it selects videos. The only key finding within this data (Table A.5) is that the percentage are substantially lower than those for videos without repeat (Table A.4).

Table A.5. Compares the percentage of videos across the Five Day Searches and Daily Searches that have appeared within every search. High values are green, mid-range values are yellow and low values are red.

Appearance	5DS - All	DS - All
Non-profit	24.4	2.6
Causes	20.0	1.9
Automotive	16.0	0.8
Lifestyle	8.3	1
Politics	8.2	1.1
Education	5.1	0.4
Gaming	4.0	0.7
Science	3.9	0.6
Health	3.5	0.7
Sports	3.5	0.4
Tech	3.5	0.9
Cooking	2.9	1
From TV	2.4	0.2
Fashion	2.2	0.1
DIY	2.2	0.9
Beauty	1.7	0.5
Best of	1.6	0.1
Entertainment	1.4	0
Animation	1.3	0.6
Comedy	0.7	0.3
How-to	0.5	0.4
News	0.4	0

Due to the difference in the searches, i.e. the number of searches for each it is more appropriate to compare the average percentage difference between searches for both search methods (Figures A.1 and A.2). the average percentage difference between the subsequent Five Day searches reduces across the categories (Figure A.1). Whereas, for the Daily searches the number of videos that are not repeated on average goes up between subsequent searches, except for education (Figure A.2). This data (Figures A.1 and A.2) would generally suggest that there is a greatly likelihood of having less repeats if searching daily. However, based on other data collected it could also suggest that there has been a change in the way in which the API selects videos.





Figures A.1, A.2 . **The average percentage difference of non-repeated videos between concurrent Five Day and Daily searches, in the same order.**

The data (Table A.6) shows that the four categories that have the highest number of videos that have appeared within every search (across the 20 searches) have substantially lower number for the metrics relating to 'Dislikes', 'Likes', 'View Count' and 'Comments', however, this is not consistent throughout the other categories. Due to the significant variation in the means for each of the metrics for the categories there is no obvious pattern that suggests that the API is focusing on any metric. This further supports the idea that the API may just take a random sample from the videos within a category and that this is not based on or related to any of the associated metrics.

As with the Daily search data (Table 5.4) the mean average for the number of dislikes (Table A.6) is also significantly less than the average of likes across all categories for the videos appearing within all the searches. This suggests that people are more likely to respond positively to videos than negatively. It can also be noted that the metrics for each of the categories appear to be higher with the Five Day searches (Table A.6) than those from the Daily searches (Table 5.4).

Table A.6. Five Day Searches – Number of videos and associated mean data for videos appearing in all searches. High values are green, mid-range values are yellow and low values are red.

	Videos	Days	Title Words	Dislikes	Likes	View Count	Comments	Length
Causes	226	878	7	84	597	124841	290	674
Non-profit	219	873	8	7	204	31980	55	847
Automotive	156	755	8	10	233	52796	50	728
Politics	121	757	8	106	2528	275877	614	1899
Lifestyle	119	622	7	673	5752	1506784	1070	1008
Gaming	89	299	8	957	45943	2900792	8645	1423
Education	83	951	8	126	5088	467329	1281	1481
Tech	77	589	7	292	8537	795042	1867	2039
Science	75	599	8	597	13326	1986541	2136	1516
Sports	75	788	8	611	9559	2770814	2192	716
Health	71	846	8	147	3816	479974	943	1565
Cooking	59	604	8	820	11658	4653021	1760	689
DIY	59	101	9	492	37948	967735	3901	482
Fashion	55	473	9	456	8010	1484656	1370	835
From TV	53	1009	8	894	14572	3886478	4105	782
Best of	44	501	9	2312	37756	12477644	4553	4516
Beauty	41	507	8	12923	79381	14852211	17853	1013
Entertainment	36	615	9	903	28701	5420723	5247	918
Animation	35	332	9	2755	57178	8848304	4594	516
Comedy	22	613	10	735	13435	3281621	1464	2493
How to	17	524	10	752	19737	3272278	2389	902
News	15	550	9	1027	11016	3944642	3952	3981

There is substantial variation across the categories in the number of videos that had no repeats from News with '1994' and Automotive with '129' (Table A.7). It is clear from each of the metrics columns that there is substantial variation in the means associated with each category, therefore there is no clear pattern in what the API might focus on in how it selects videos.

The categories which have lowest number of videos with no repeats, e.g. Automotive, Non-profit, Causes, Lifestyle, Politics and Education have some of the lowest means in their associated metrics (Table A.7). However, if this was a factor taken into consideration by the API then it would be consistent across all the categories.

From comparing the data (Tables A.6 and A.7), the mean averages for the metrics are lower and the number of videos with no repeat are higher, therefore it seems that the API is more likely to extract videos that have slightly higher metrics on multiple occasions, but does not base all its selections on video metrics. It seems again, as with the Daily Searched (Table 5.5), that videos with lower metrics might have a lower profile in the API and are possibly extracted less frequently.

Table A.7. Five Day Searches – Number of videos and associated mean data for videos with no repeat. High values are green, mid-range values are yellow and low values are red.

No Repeat	Videos	Days	Title Words	Dislikes	Likes	View Count	Comments	Length
News	1994	77	10	87	1661	212823	505	742
How to	1982	343	9	197	5692	492588	891	688
Comedy	1397	218	11	59	813	211099	277	2600
Best of	1364	222	9	213	4453	776425	582	2057
DIY	1289	226	8	64	2848	144060	326	444
Animation	1276	297	10	127	3249	472701	530	2301
Entertainment	1102	293	10	29	580	96665	107	806
Fashion	1082	133	9	52	1111	141048	141	548
Beauty	1024	266	9	83	2834	190222	336	786
Tech	994	132	9	23	666	61422	161	939
Gaming	947	62	11	50	1814	71145	381	1466
From TV	907	370	10	151	1645	617418	581	752
Sports	876	172	10	35	706	101417	189	659
Health	830	219	9	7	214	19383	236	1032
Science	698	357	9	148	3212	463028	1475	1491
Cooking	695	444	8	76	1464	364591	274	546
Education	586	274	9	8	138	17622	72	945
Politics	520	180	9	6	90	17026	143	1084
Lifestyle	403	225	8	19	325	19591	59	588
Causes	309	230	8	5	73	8302	260	364
Non-profit	139	346	8	1	37	3264	59	675
Automotive	129	429	8	3	62	16598	131	455

The percentages relating to each of the bands (Table A.8) appear to be more spread out than the data representing the Daily searches (Table 5.6) going more into the '5001 to 10000' and 10001 to 50000' bands, however, there is still a higher percentage of videos within the '0 to 10' banding. Once again, the Automotive, Education and Non-profit categories have the highest percentage of videos falling within the '0 to 10' comments band suggesting that these categories receive a slightly lower proportion of video comments (Table A.8). Overall there does not appear to be any substantial similarities within each of the bandings when compared across the different categories. There does seem to be slight peaks in the '0 to 10' and '101 to 500' bandings similar to other findings (Table 5.6), but nothing to suggest that the API is choosing videos with similar numbers of comments.

Table A.8. Five Day Searches – Percentage of videos per grouping of comments. High values are green, mid-range values are yellow and low values are red.

5 DS Comments	0 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 100	101 to 500	501 to 1000	1001 to 5000	5001 to 10000	10001 to 50000	50001 to 100000	100001+
Animation	15.09	6.55	5.36	3.92	2.65	10.18	21.26	8.32	16.51	4.23	5.69	0.26	0
Automotive	57.82	10.81	4.63	3.83	2.26	9.04	8.35	1.2	1.37	0.25	0.39	0.04	0
Beauty	6.5	4.35	3.87	4.44	2.6	10	27.99	13.57	18.76	4.22	3.19	0.15	0.36
Best of	4.64	3.36	3.11	2.73	2.37	8.76	30.04	13.68	23.6	3.58	3.78	0.26	0.09
Causes	37.98	10.52	6.52	4	3.65	9.74	16.42	5.39	4.92	0.15	0.69	0.01	0.01
Comedy	27.06	10.7	6.61	5.79	4.65	11.3	21.49	5	5.63	0.93	0.77	0.01	0.07
Cooking	9.87	6.91	4.32	3.57	2.56	9.6	31.25	10.55	16.57	3.07	1.71	0.01	0
DIY	3.7	3.04	2.66	3.06	2.24	8.23	30.91	13.19	26.25	3.19	2.57	0.7	0.22
Education	41.23	10.75	5.49	3.39	2.85	9.62	14.56	4.47	5.57	0.95	0.92	0.15	0.04
Entertainment	40.02	9.07	6.11	3.86	2.93	7.83	15.66	4.63	6.74	1.27	1.68	0.2	0
Fashion	20.68	8.47	5.6	5.69	4.2	12.37	23.94	6.8	9.66	1.47	1.12	0	0
From TV	19.68	5.67	3.37	2.59	2.19	8.86	21.52	9.22	15.61	4.74	5.76	0.62	0.15
Gaming	3.42	2.45	1.59	1.15	1.09	4.91	19.41	11.12	35.46	11.67	7.26	0.46	0
Health	38.58	9.87	5.14	3.13	2.62	7.87	14.59	5.76	9.42	1.57	1.45	0.02	0
How to	3.94	3.79	3.46	2.66	2.49	9.7	33.34	13.5	20.26	3.35	3.2	0.16	0.15
Lifestyle	27.31	8.61	7.1	6.43	5.24	13.49	21.38	4.92	4.6	0.16	0.57	0.21	0
News	13.76	3.88	2.58	2.26	1.97	7.89	26.35	8.89	24.59	4.27	3.1	0.17	0.3
Non-profit	79.27	6.13	1.89	1.71	1.15	2.59	3.58	0.86	2.16	0.45	0.21	0.01	0
Politics	33.06	9.78	5.04	4.29	3.11	8.65	19.95	7.76	7.47	0.62	0.25	0	0
Science	14.36	7.31	4.54	3	2.81	8.5	21.18	10.72	19.88	3.76	3.84	0.07	0.02
Sports	13.49	6.34	5.15	4.08	2.4	11.12	27.58	9.53	15.57	2.76	1.67	0.31	0
Tech	23.37	9.24	4.55	4.11	2.82	8.27	20.34	8.71	14.48	2.18	1.93	0.01	0

There is a lower percentage of videos receiving less than 100 likes, except for the categories Automotive and Non-profit (Table A.9), particularly when compared to the Daily search data (Table 5.7). A higher proportion of categories have videos with a higher percentage of likes over 4000 than the Daily search data (Table 5.7). Searching over a longer period and at Five Day intervals seems to have provided a higher proportion of videos with more likes (Table A.9). The categories Best of, DIY and Gaming have the highest proportion of videos receiving more than 4000 likes, with Animation, Beauty and How-to also demonstrating a higher percentage of videos receiving more than 4000 likes (Table A.9). These are all the categories that also received a high percentage of videos with more than 4000 likes within the Daily search data (Table 5.7). However, as with the Daily searches (Table 5.7), there does not seem to by any patterns into which videos the API extracts and if this is related to the percentage of likes (Table A.9). There is a chance that the time scales and different periods during which the searches were submitted and a possible change in the algorithm has had an impact on the data that the APIs has provided.

Table A.9. Five Day Searches – Percentage of videos per grouping of likes. High values are green, mid-range values are yellow and low values are red.

5 DS - Likes	0	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 3000	3001 to 4000	4001+
Animation	1.69	2.04	1.67	1.36	1.33	1.23	1.40	1.25	1.08	1.12	1.20	8.65	6.46	4.74	3.04	9.58	5.98	3.49	4.27	3.10	35.34
Automotive	7.67	30.30	11.57	5.50	4.10	4.05	3.95	3.11	1.36	1.66	1.23	9.84	4.68	1.93	2.48	3.23	1.69	0.40	0.82	0.00	0.42
Beauty	1.68	0.48	0.52	0.72	0.75	0.76	0.76	0.75	0.75	0.82	0.66	4.57	5.55	4.22	3.43	10.67	6.84	4.70	7.13	4.42	39.84
Best of	0.61	0.34	0.34	0.23	0.44	0.36	0.26	0.53	0.31	0.44	0.48	3.88	3.36	2.91	2.56	9.71	7.56	4.85	7.71	5.97	47.14
Causes	3.33	10.16	10.98	7.81	6.33	4.13	3.76	2.75	3.88	2.39	1.81	11.53	5.33	4.07	2.81	7.62	2.29	1.93	2.26	1.57	3.24
Comedy	0.95	2.16	2.36	2.79	2.29	2.14	2.65	2.32	2.02	1.54	1.83	11.48	7.40	5.88	4.38	13.90	7.12	5.65	5.68	2.27	13.20
Cooking	3.31	1.58	1.14	2.22	1.69	1.37	1.70	1.68	1.94	1.01	0.80	9.33	4.69	4.67	3.55	10.25	6.00	4.49	5.21	5.28	28.10
DIY	0.99	0.17	0.30	0.39	0.66	0.59	0.53	0.44	0.43	0.43	0.40	3.81	2.99	2.03	1.70	7.34	4.79	4.34	6.34	4.61	56.73
Education	6.27	9.28	9.47	7.56	6.67	5.62	3.89	3.39	1.81	2.04	2.27	12.99	4.94	3.30	2.68	7.43	2.54	1.14	0.86	1.07	4.80
Entertainment	4.59	15.04	7.25	4.89	3.91	3.20	1.71	1.44	1.39	1.11	1.61	8.41	6.17	3.70	3.10	8.51	3.40	2.87	3.98	1.87	11.86
Fashion	2.03	2.80	2.46	2.07	2.00	1.99	2.03	1.80	1.76	1.60	1.61	10.43	7.49	5.45	4.49	12.20	6.09	4.71	6.49	2.79	17.69
From TV	6.50	7.31	1.91	2.09	1.46	1.65	1.39	1.74	1.42	1.05	0.69	8.40	4.93	2.79	2.72	9.38	4.63	2.69	4.57	2.39	30.31
Gaming	0.38	0.19	0.47	0.55	0.54	0.65	0.76	0.57	0.49	0.47	0.55	3.82	3.26	2.63	2.22	7.72	4.70	3.79	4.52	4.94	56.80
Health	6.11	9.38	9.62	6.95	5.59	4.50	3.22	1.82	1.40	1.88	1.25	9.87	6.43	4.18	3.09	6.42	2.39	1.70	2.36	2.81	9.03
How to	0.93	0.18	0.22	0.48	0.45	0.35	0.44	0.45	0.35	0.42	0.54	5.64	4.09	4.12	3.48	11.21	8.14	5.87	7.26	5.87	39.52
Lifestyle	3.39	4.96	4.79	4.74	3.56	3.25	2.74	2.71	2.55	1.92	1.37	13.15	9.43	4.75	3.81	9.00	5.94	3.65	3.85	2.57	7.88
News	1.68	5.31	3.48	2.13	1.65	1.50	1.26	1.27	0.97	0.72	0.91	8.22	4.54	3.59	3.19	13.84	6.22	4.66	6.93	4.96	22.98
Non-profit	17.26	48.93	9.65	5.54	4.70	1.65	1.53	0.45	1.23	1.04	0.58	2.97	1.53	0.40	0.64	1.07	0.36	0.00	0.02	0.00	0.47
Politics	5.61	9.88	7.43	7.46	4.02	3.56	1.94	1.92	1.47	2.00	2.02	11.50	5.27	3.79	3.50	9.71	3.73	3.59	2.32	1.94	7.31
Science	2.42	1.92	2.38	1.97	2.67	2.20	1.49	2.02	1.32	1.88	1.23	8.68	6.11	4.64	2.58	9.05	5.85	4.38	5.25	4.00	27.94
Sports	1.42	1.94	1.83	1.97	1.59	1.27	1.27	1.33	1.16	1.30	1.25	9.26	6.46	5.23	4.56	12.67	8.30	5.59	6.56	4.62	20.44
Tech	2.30	4.36	5.71	5.02	4.83	3.40	3.05	1.88	1.79	1.65	1.45	8.94	4.24	3.74	2.15	7.65	3.36	4.05	5.35	3.91	21.18

Within the data (Table A.10) there is a slightly greater spread of dislikes across the categories when this data is compared to the Daily search data (Table 5.8). Non-profit and Automotive both have substantially low percentages of dislikes, and Education and Causes are also low, which is similar to the Daily search findings (Table 5.8). There continues to be substantial variation across bandings and as a result there are no clear patterns in the different percentage bands relating to dislikes (Table A.10). The data is clearly demonstrating differences between the different categories but does not seem to be to be providing any answers in whether the metrics have any influence over the API in the videos it extracts.

Table A.10. Five Day Searches – Percentage of videos per grouping of dislikes. High values are green, mid-range values are yellow and low values are red.

5DS - Dislikes	0	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 3000	3001 to 4000	4001+
Animation	5.60	21.79	7.83	5.61	3.88	2.59	2.49	2.17	1.30	1.50	1.43	8.06	5.82	4.32	3.66	8.61	3.94	2.59	2.08	1.89	2.84
Automotive	43.86	40.54	6.36	3.21	1.37	1.45	0.95	0.45	0.24	0.13	0.15	1.05	0.00	0.23	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Beauty	3.47	15.05	10.22	7.32	5.41	4.44	2.93	2.89	2.68	2.54	1.92	11.69	7.01	3.64	3.02	6.56	3.47	1.54	1.00	1.09	2.09
Best of	2.11	10.69	6.36	5.68	4.15	2.78	3.05	2.43	2.18	1.97	1.78	11.46	7.89	4.30	3.93	10.85	6.15	2.63	3.31	2.03	4.26
Causes	18.77	44.48	11.67	6.03	3.55	3.02	1.92	1.31	0.64	0.96	0.48	3.48	1.04	0.92	0.33	0.67	0.33	0.03	0.17	0.00	0.20
Comedy	2.56	18.65	11.82	7.79	5.56	4.32	3.76	3.70	2.81	1.84	1.97	13.23	5.66	4.11	2.30	4.97	1.52	1.25	1.13	0.15	0.91
Cooking	7.30	20.65	10.58	6.21	5.94	4.08	3.12	3.22	2.57	3.09	2.07	9.92	5.29	2.73	1.73	4.82	2.19	1.54	1.32	0.62	1.00
DIY	2.41	11.70	7.15	4.98	3.78	3.98	2.75	2.58	2.47	2.37	2.16	15.23	9.99	6.68	5.57	8.71	3.35	1.99	1.39	0.48	0.28
Education	24.80	48.86	8.29	4.27	1.66	1.69	1.29	0.96	0.80	0.58	0.75	1.78	1.03	0.78	0.61	0.53	0.42	0.13	0.04	0.19	0.53
Entertainment	21.68	33.31	8.48	4.57	3.26	2.75	2.00	1.91	1.22	1.24	1.45	5.47	2.16	0.93	1.18	2.84	1.90	0.73	1.27	0.51	1.16
Fashion	8.39	27.03	12.43	7.34	5.84	4.19	3.43	2.01	1.42	1.88	1.17	8.56	4.42	2.52	1.94	3.08	1.90	0.88	0.51	0.17	0.89
From TV	14.10	20.14	8.08	5.41	3.80	2.98	2.76	2.18	1.83	1.37	1.37	9.15	4.36	2.87	2.23	4.86	2.51	1.35	3.07	1.35	4.25
Gaming	2.35	11.57	5.83	4.18	3.22	2.26	2.66	1.63	1.85	2.09	1.81	15.89	12.30	7.90	4.66	9.42	3.31	2.17	2.66	0.91	1.32
Health	25.56	39.77	7.68	4.02	2.10	1.34	1.58	1.33	1.42	0.77	0.92	5.27	2.25	1.52	0.86	2.01	0.77	0.15	0.48	0.00	0.19
How to	2.09	14.70	10.84	6.62	5.34	4.06	3.29	2.94	2.59	2.31	1.86	11.87	7.03	3.56	2.65	6.95	3.87	1.89	1.65	0.87	3.01
Lifestyle	14.61	38.86	14.23	6.44	4.71	3.54	1.63	2.06	1.50	0.85	0.60	5.87	1.22	0.47	0.65	1.50	0.63	0.03	0.01	0.00	0.61
News	4.25	21.97	10.69	5.94	3.65	2.72	2.12	2.15	2.47	1.92	1.86	11.05	5.18	5.11	3.36	7.60	3.85	1.33	1.02	0.35	1.42
Non-profit	66.17	28.38	2.18	1.13	1.08	0.07	0.03	0.80	0.00	0.00	0.00	0.36	0.32	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00
Politics	19.83	39.70	10.66	5.39	3.60	2.44	2.07	1.78	0.99	0.98	0.70	5.53	2.17	1.21	0.49	2.15	0.12	0.00	0.00	0.00	0.20
Science	6.92	24.25	11.64	5.67	4.83	2.68	2.80	2.29	1.97	1.16	0.94	9.59	4.60	3.29	2.23	7.56	3.20	2.03	0.97	0.50	0.90
Sports	4.81	22.73	10.48	5.70	5.03	4.31	3.11	2.67	2.43	2.38	1.46	11.68	5.07	3.16	2.18	6.60	2.29	1.39	0.78	0.33	1.39
Tech	11.48	34.54	8.90	5.98	3.88	3.42	2.79	2.21	1.44	1.48	0.81	7.86	3.70	2.42	2.48	3.76	1.10	0.64	0.76	0.00	0.34

Comparing the data (Table A.11) with the Daily search findings (Table 5.9), there is a substantial difference in what the API has provided in the number of days that the videos that have been posted to YouTube. The Daily Search data (Table 5.9) seems, apart from two categories, to suggest that the API extracts very few videos that have uploaded for more than 300 days. Whereas the Five Day Search data (Table A.11) shows a much wider spread of videos all the way up to the '3001+' banding. In addition the data (Table A.11) shows that there is a substantially higher percentage of days for most categories within the '501 to 1000' band, where this was the '51 to 100' banding for the Daily searches (Table 5.9). The Daily search data (Table 5.9), apart from 0.01% for Health and Non-profit, had 0% for all the other categories within the '1 to 10' banding, but the Five Day data (Table A.11) provides a variety of percentages within this banding. As the differences across these two search approaches for number of days uploaded is clearly so substantial the way in which the API choses the videos to extract must have altered or there must have been a change to the algorithm that YouTube employs.

Table A.11. Five Day Searches – Percentage of videos per grouping of days. High values are green, mid-range values are yellow and low values are red.

5DS - Days	1 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 100	101 to 200	201 to 300	301 to 400	401 to 500	501 to 1000	1001 to 1500	1501 to 2000	2001 to 2500	2501 to 3000	3001+
Animation	12.54	3.58	3.10	2.82	2.34	7.29	8.44	7.70	6.73	4.53	17.50	10.71	6.02	4.03	2.23	0.43
Automotive	0.93	0.72	0.68	0.68	0.78	4.79	7.92	7.17	7.80	8.98	30.68	17.02	7.55	3.39	0.62	0.30
Beauty	12.01	2.74	2.14	1.90	1.82	8.02	10.63	9.93	7.22	5.73	19.90	8.34	4.52	2.74	1.75	0.61
Best of	13.20	1.89	1.86	1.88	1.88	7.65	10.91	9.54	9.37	8.09	25.50	5.93	1.56	0.36	0.14	0.24
Causes	4.34	1.08	0.75	0.65	0.58	2.94	8.37	6.75	6.06	8.12	26.72	17.57	8.83	4.34	2.77	0.12
Comedy	9.31	2.83	3.43	3.42	3.58	11.02	11.44	9.97	7.04	6.08	18.10	9.03	3.02	1.03	0.48	0.21
Cooking	5.70	1.33	1.23	1.22	0.99	5.90	11.47	7.76	7.61	6.79	21.78	15.81	7.36	2.81	2.08	0.15
DIY	13.09	4.37	4.17	3.89	4.12	14.87	24.04	10.11	5.97	4.37	8.70	1.70	0.50	0.09	0.01	0.00
Education	5.97	0.70	0.62	0.47	0.36	2.48	3.60	4.37	6.73	6.70	34.49	19.59	6.75	4.66	2.16	0.35
Entertainment	11.31	3.38	2.43	2.24	1.73	8.30	10.37	8.05	7.61	6.51	19.32	11.92	3.84	1.93	0.96	0.08
Fashion	12.96	3.38	3.68	2.66	1.90	7.66	13.39	7.87	9.33	7.33	18.18	6.67	3.26	1.04	0.49	0.20
From TV	10.23	2.80	1.36	1.51	1.34	5.18	9.21	6.49	6.51	5.84	17.91	11.69	6.85	6.27	4.94	1.87
Gaming	16.15	5.37	4.13	3.76	2.57	12.32	15.00	8.39	8.01	6.30	16.01	1.65	0.30	0.04	0.00	0.00
Health	8.19	0.94	0.70	0.47	0.46	2.83	6.69	6.12	7.78	7.29	29.81	15.21	7.77	3.91	1.79	0.03
How to	16.63	4.68	2.97	2.68	1.71	7.71	10.87	8.28	6.45	5.62	19.02	8.90	2.67	1.31	0.44	0.06
Lifestyle	3.74	1.01	1.00	1.05	1.06	7.66	15.24	11.97	9.08	7.49	22.92	8.93	3.37	3.40	1.98	0.09
News	28.60	10.19	5.07	3.47	2.70	9.66	14.04	6.01	4.07	3.40	8.43	2.18	0.90	0.62	0.35	0.32
Non-profit	1.00	0.45	0.59	0.48	0.55	3.32	6.11	7.33	9.29	8.35	28.11	17.26	12.00	4.35	0.76	0.06
Politics	7.21	1.95	1.62	1.52	1.50	6.61	9.38	6.27	7.79	7.05	24.59	11.61	5.47	4.54	2.62	0.27
Science	6.06	1.34	1.29	1.12	1.13	5.42	9.43	9.64	8.41	7.32	29.62	12.01	4.67	1.70	0.83	0.00
Sports	11.21	3.13	2.29	1.76	1.71	7.01	7.55	8.51	7.40	6.96	19.18	9.53	8.30	3.36	1.86	0.24
Tech	13.45	4.18	3.77	3.54	2.73	11.39	14.65	5.88	6.78	5.58	18.55	5.47	1.85	0.93	1.05	0.20

When the Five Day data (Table A.12) is compared to the Daily findings (Table 5.10) there is a similar pattern in the spread of the percentage of videos within each length band across the categories. The percentages are slightly lower in the shorter video length bands (approximately 0 to 90 bands) and slightly higher in the upper middle bands (241 to 900 bands). Comedy and Best of have a slightly higher level of percentages from 3601 onwards (Table A.12), and this is similar for Comedy within the Daily search data (Table 5.10). Again, there are couple of 'spikes' for From TV (Table A.12) similar to those in the Daily data (Table 5.10), but not quite as substantial. Overall there are no discernible patterns to suggest that the API is choosing videos based on their length.

Table A.12. Five Day Searches – Percentage of videos per grouping of length (seconds). High values are green, mid-range values are yellow and low values are red.

5Ds - Length	0 to 10	11 to 20	21 to 30	31 to 40	41 to 50	51 to 60	61 to 70	71 to 80	81 to 90	91 to 120	121 to 150	151 to 180	181 to 210	211 to 240	241 to 300	301 to 360	361 to 420	421 to 540	541 to 660	661 to 900	901 to 1200	1201 to 2400	2401 to 3600	3601 to 6000	6000+
Animation	0.21	0.72	0.55	0.74	1.09	0.82	2.26	1.62	2.05	7.55	7.55	7.86	8.14	5.92	8.34	6.76	5.75	6.14	3.63	4.49	1.57	3.25	1.79	6.82	4.37
Automotive	0.00	0.08	0.05	0.34	0.21	0.48	0.88	1.44	1.22	4.67	6.27	6.55	7.49	5.00	10.34	7.82	7.58	10.07	7.05	7.84	3.94	5.65	2.25	2.19	0.56
Beauty	0.00	0.06	0.02	0.21	0.27	0.41	0.28	1.39	0.79	2.58	4.55	2.88	5.05	5.76	9.82	7.05	4.47	10.67	9.36	11.96	7.09	5.06	5.34	2.56	2.34
Best of	0.00	0.05	0.03	0.03	0.06	0.19	0.30	0.15	0.07	0.76	1.28	1.57	1.78	2.77	5.05	3.41	3.00	6.86	6.33	8.91	5.56	9.61	8.21	16.60	17.42
Causes	0.03	0.11	0.69	0.81	0.45	0.96	1.08	1.83	1.32	4.95	7.27	8.82	5.98	5.92	11.73	8.42	5.27	8.13	6.29	7.67	3.09	3.02	2.48	2.69	0.99
Comedy	0.04	0.10	0.18	0.12	0.04	0.08	0.12	0.12	0.28	0.85	1.77	1.32	1.60	1.37	3.01	3.34	4.03	7.32	8.01	9.67	6.04	11.76	10.90	16.73	11.19
Cooking	0.04	0.11	0.11	0.17	0.59	0.35	0.53	0.56	0.66	2.17	1.77	2.78	3.85	5.54	12.74	11.30	7.29	15.45	8.75	12.04	4.48	5.37	2.30	0.77	0.25
DIY	0.05	0.04	0.00	0.04	0.02	0.00	0.16	0.10	0.24	0.96	2.13	3.03	3.66	5.91	11.23	11.03	13.17	21.83	12.27	8.20	3.17	2.30	0.45	0.01	0.00
Education	0.02	0.01	0.02	0.15	0.11	0.15	0.55	0.62	0.79	1.67	3.76	4.53	4.69	4.56	8.81	6.17	4.31	7.57	7.42	8.76	10.19	7.81	7.57	6.47	3.28
Entertainment	0.52	0.83	0.43	1.43	1.48	0.92	1.52	1.60	1.53	4.13	4.84	5.32	5.31	6.20	8.69	6.21	4.48	6.96	4.52	5.62	3.93	10.20	8.15	3.26	1.93
Fashion	0.02	0.03	0.03	0.39	0.17	0.52	0.65	0.55	0.85	3.22	4.12	4.43	3.56	5.03	9.94	7.08	5.71	10.02	10.39	15.64	6.72	5.50	4.14	1.13	0.17
From TV	0.26	0.28	0.50	4.36	0.99	2.30	4.33	2.06	1.53	4.74	4.96	4.72	5.20	3.67	8.31	5.58	3.96	6.56	5.27	7.23	4.11	9.01	5.54	2.92	1.59
Gaming	0.01	0.00	0.00	0.13	0.05	0.13	0.14	0.29	0.29	0.28	0.43	1.03	2.05	2.35	7.41	7.50	11.09	12.77	10.47	13.54	6.53	7.29	6.34	7.67	2.19
Health	0.06	0.03	0.05	0.51	0.24	0.17	0.59	0.51	1.15	3.32	3.79	4.69	5.47	4.56	8.83	6.08	5.47	8.35	7.78	6.23	4.99	9.57	8.59	6.96	2.00
How to	0.00	0.14	0.38	0.41	0.31	0.49	0.65	0.87	1.72	2.99	3.97	3.45	4.03	4.99	8.88	8.16	7.14	9.47	8.95	11.23	8.11	7.59	3.34	2.03	0.68
Lifestyle	0.19	0.25	0.11	0.41	0.19	0.05	0.90	0.90	0.77	3.50	4.80	3.64	6.75	5.88	15.33	8.60	5.97	7.92	5.72	8.05	4.29	4.64	5.24	4.21	1.70
News	0.88	0.06	0.36	0.65	1.18	1.23	2.26	1.57	1.34	3.48	4.27	5.83	4.58	3.68	7.32	8.04	5.45	7.68	8.20	8.79	6.64	9.32	4.24	2.08	0.88
Non-profit	0.04	0.00	0.13	0.53	0.94	1.05	0.92	1.97	2.31	5.57	8.38	6.96	5.56	5.26	9.28	6.36	3.59	9.33	7.59	5.72	2.34	5.41	5.00	4.15	1.60
Politics	0.03	0.11	0.00	0.12	0.25	0.25	0.41	0.65	0.70	1.74	2.90	3.50	5.50	4.78	4.74	4.48	3.78	8.48	8.61	7.70	5.05	10.60	8.58	14.05	2.97
Science	0.03	0.02	0.00	0.12	0.10	0.11	0.33	0.33	0.53	2.41	3.91	4.43	5.16	5.05	7.72	6.26	3.36	6.85	7.46	8.08	4.14	6.13	13.63	8.41	5.42
Sports	0.03	0.18	0.23	0.38	0.28	0.60	0.94	0.79	1.84	3.42	4.77	4.68	7.23	6.51	10.93	7.21	6.50	11.81	7.93	10.63	2.71	4.12	3.88	1.35	1.02
Tech	0.01	0.01	0.08	0.04	0.39	0.27	0.80	0.90	0.57	3.38	4.12	4.94	4.97	6.16	8.13	6.01	6.02	9.50	5.84	5.72	2.97	5.25	4.00	7.40	12.52

The data (Table A.13) shows that there is less of a spread of percentages across the bandings when compared to Table 5.11 (Daily Searches). Also there are a higher percentage of videos within the '10001 to 100000' and '100001 to 1000000' view count bands (Table A.13). When compared to the Daily data (Table 5.11) the Five Day Searches (Table A.13) seem to provide more videos with a higher view count and less with a view count under 1000.

Table A.13. Five Day Searches – Percentage of videos per grouping of views. High values are green, mid-range values are yellow and low values are red.

5DS - View count	0 to 10	11 to 100	101 to 1000	1001 to 10000	10001 to 100000	100001 to 1000000	1000000+
Animation	0.1	0.1	0.4	7.7	26.6	33.7	31.5
Automotive	0.4	0.4	14.8	42.9	32.6	8.5	0.4
Beauty	0.2	0.1	0.9	5.6	25.8	45.2	22.3
Best of	0.1	0.0	0.5	4.2	15.1	33.3	46.7
Causes	1.0	0.3	2.6	29.8	46.2	18.3	1.8
Comedy	0.0	0.1	0.5	5.7	28.5	47.0	18.2
Cooking	0.8	0.4	1.1	8.0	33.5	39.2	17.0
DIY	0.1	0.2	0.8	6.7	23.7	50.6	17.7
Education	2.2	1.0	3.7	30.4	47.0	13.3	2.5
Entertainment	2.7	1.4	8.2	23.5	28.5	24.0	11.7
Fashion	0.8	0.8	2.1	11.3	39.5	34.6	10.7
From TV	2.5	2.5	4.1	11.1	19.9	29.9	29.8
Gaming	0.0	0.0	0.3	6.4	23.8	42.0	27.4
Health	2.2	1.2	5.1	28.4	40.7	17.5	4.9
How to	0.2	0.0	0.2	4.5	31.8	41.7	21.6
Lifestyle	0.8	0.4	2.2	24.3	44.7	23.4	4.2
News	0.0	0.1	1.3	15.0	32.3	34.3	16.9
Non-profit	0.7	4.9	41.9	38.7	12.2	1.5	0.2
Politics	1.5	0.9	5.2	31.5	39.6	18.7	2.7
Science	0.4	0.4	1.3	9.6	34.3	34.5	19.5
Sports	0.6	0.4	1.3	7.8	26.7	40.9	22.2
Tech	1.2	0.8	3.1	22.8	35.7	26.3	10.1

Appendix 11

Table A.14. The average number of dislikes, likes, dislikes per view and likes per view for each of the categories. High values are green, mid-range values are yellow and low values are red.

Categories	Average Dislikes	Average Dislikes per View	Average Likes	Average Likes per View	Likes per Dislike
Animation	171	0.00033	4026	0.0078	23.5
Automotive	2	0.00022	34	0.0038	17
Beauty	192	0.00053	3703	0.0102	19.3
Best of	267	0.00028	4969	0.0052	18.6
Causes	12	0.00058	119	0.0057	9.9
Comedy	63	0.00031	823	0.0040	13.1
Cooking	63	0.00027	1405	0.0060	22.3
DIY	92	0.00053	4834	0.0279	52.5
Education	15	0.00025	280	0.0047	18.7
Entertainment	46	0.00028	853	0.0052	18.5
Fashion	41	0.00033	1054	0.0085	25.7
From TV	281	0.00025	4128	0.0037	14.7
Gaming	105	0.00053	3902	0.0197	37.2
Health	14	0.00031	367	0.0082	26.2
How-to	221	0.00044	6431	0.0129	29.1
Lifestyle	28	0.00048	436	0.0074	15.6
News	79	0.00041	1468	0.0075	18.6
Non-profit	2	0.00038	47	0.0090	23.5
Politics	10	0.00038	215	0.0083	21.5
Science	70	0.00033	1826	0.0086	26.1
Sport	57	0.00030	1073	0.0057	18.8
Tech	30	0.00036	817	0.0099	27.2

Appendix 12

Questionnaire data analysed by usage frequency

Usage frequency – Female

The following figures show the collected data for the various levels of user for female respondents. Respondents accessing videos daily or weekly will be described here as frequent users. It has already been demonstrated that 159 female respondents classed themselves as daily users and 218 as weekly users (Figure 8.3), therefore, 85% (377/444) of the female sample are frequent YouTube users. The information provided by those female respondents who only access YouTube videos yearly will still be discussed and some conclusions may be established, but they will not necessarily be an accurate representation of YouTube users.

Education

A high proportion of respondents 396/444 (89%) across all types of users are educated to Higher Education level (Figure A.3). OxIS (Blank et al., 2019) shows that a high proportion of users with higher education (95%), further education (92%), and basic qualifications (70%) are internet users, and a third (36%) of users with no qualifications use the internet. Given that 42% of UK people aged 21-64 have Higher Education qualifications (HESA, 2018), this confirms that the survey respondent sample heavily over-represents people with Higher Education qualifications. The data (Figure A.3) also shows that 338/444 (76%) of the female respondents who are frequent YouTube users are also educated to Higher Education level further supporting the credibility of this sample (Blank et al., 2019).

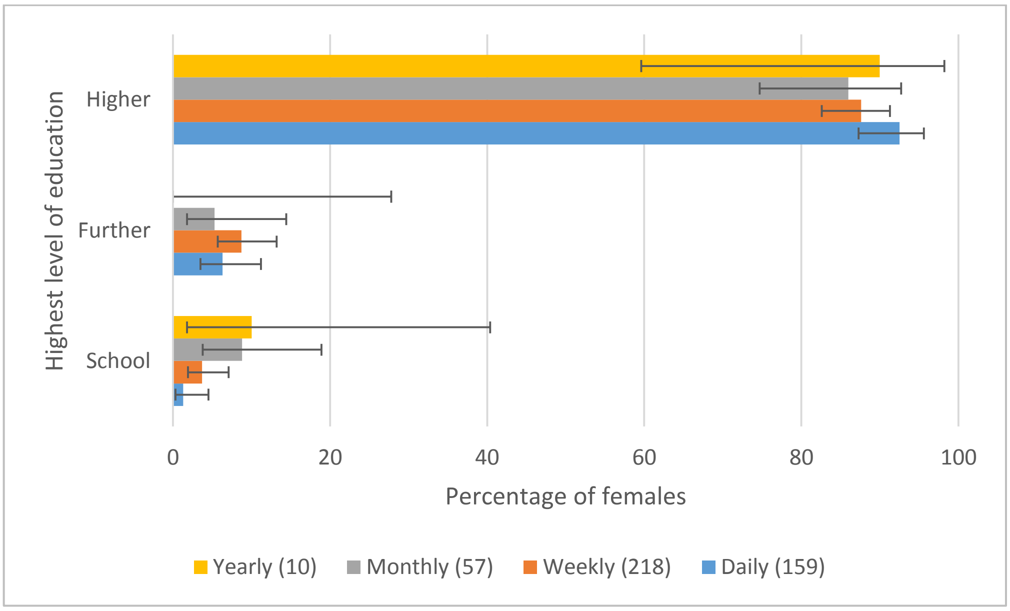


Figure A.3. **The education levels of different female user respondents with 95% confidence intervals.**

Accessing YouTube Videos

The three most popular methods for accessing YouTube videos for the female respondents who are frequent users are using the YouTube website, the YouTube App and from following a link or post on Facebook (Figure A.4). Although popular with daily female respondent users, 78%, the use of the YouTube App substantially reduces through each of the levels of YouTube usage with only 20% of yearly users accessing it (Figure A.4). It is possible that frequent users of YouTube, and other online resources and platforms, would spend more time using mobile devices and technologies, such as smart phones or tablets, and as a result use more Apps to access these resources. Google searches are also a popular method for accessing YouTube videos and is most popular method with yearly users (Figure A.4). Yearly users might not be as confident using YouTube and may have more experience and confidence in using Google. Yearly users also use hyperlinks within emails more than the other users and this could be that they trust the person who has sent it to them and, like with Google, it could be that they have more experience and confidence with email. People access online videos using a range of methods, including Google, email and social media. However, despite social media being a popular platform for accessing videos we can see that Twitter is not a popular method that female respondents, across all user levels, use to access YouTube videos and could be due to the word-based nature of this social media forum (Figure A.4). This could also be true for following hyperlinks within blogs which is also substantially low across all female respondent users.

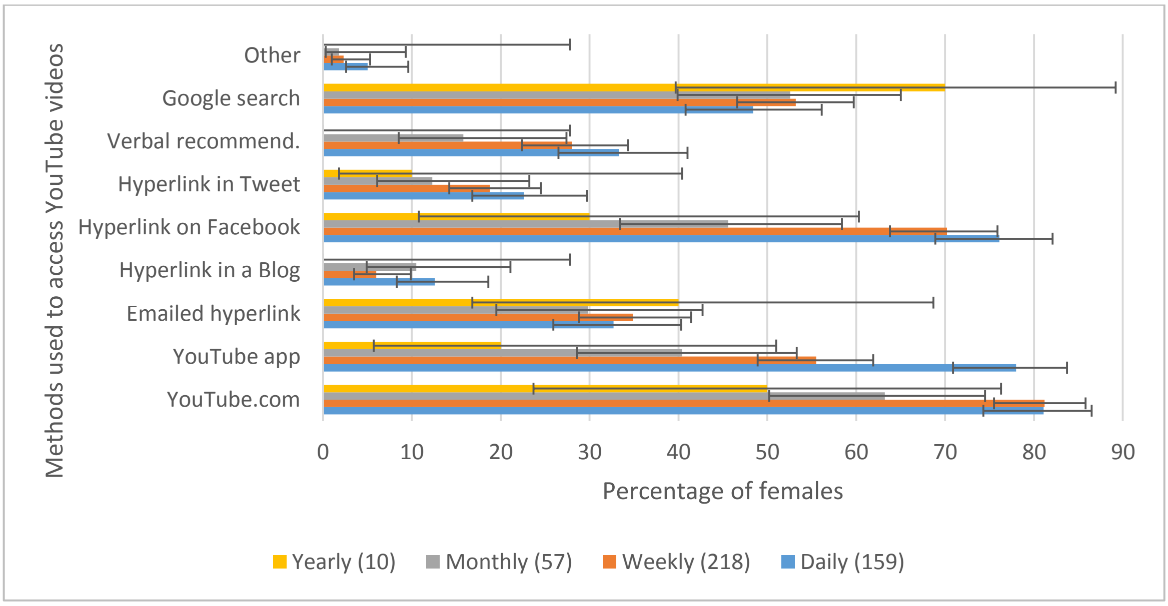


Figure A.4. **How the different female user respondents have accessed YouTube videos with 95% confidence intervals.**

The most frequent female users, daily and weekly, within the sample use about 3 to 5 methods to access their YouTube videos and have a wider spread of approaches (Figure A.5). As frequent users they probably spend more time online and have built up a wider understanding of different ways if accessing videos. Less Frequent viewers use slightly less methods, 1 to 3, to access the YouTube videos that they want to watch (Figure A.5). Overall the data (Figure A.5) shows that the more frequent that the female respondents use YouTube the greater the spread of methods that they use to find and access videos. Yearly users use a small number of methods to find and access their YouTube videos (Figure A.5).

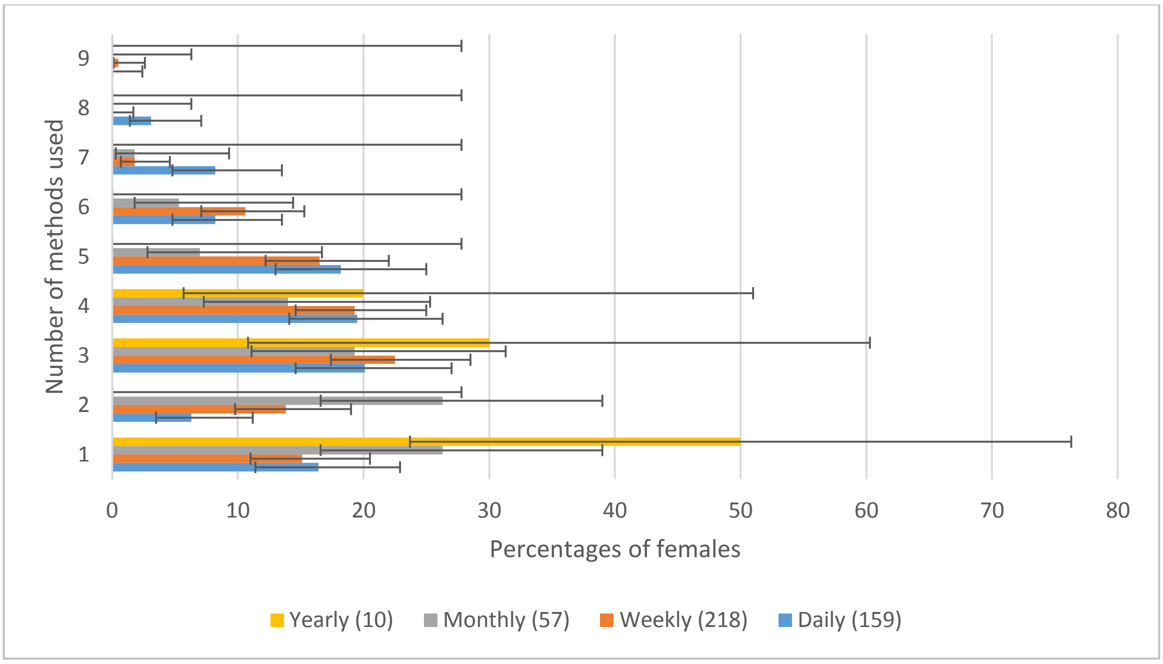


Figure A.5. **The number of methods that the different female user respondents use to access YouTube videos with 95% confidence intervals.**

Accessing Video within YouTube

The data (Figure A.6) shows that the most popular method used by all levels of female respondent users to access videos within YouTube is the website's own search facility which, as discussed previously, has some issues in the selection of videos provided by this feature. Although the search facility is still a popular method of accessing videos for weekly users they still use other methods to some extent and are also likely to watch videos that have been recommended or suggested by YouTube on the homepage (Figure A.6). Female respondents that are daily users use a much wider range of methods for watching videos within the YouTube website (Figure A.6). It is possible that daily use of YouTube might have provided them with a greater knowledge, understanding and confidence of all the different ways of findings and accessing videos. Daily users access YouTube videos from both an entertainment and specific purpose point of view (Figure A.6). As it seems through their use of multiple methods within YouTube that they are more prone to 'impulse watching' rather than just searching for and accessing videos. There is a significant drop in the popularity of using the homepage to access videos from daily through each of the user levels (Figure A.6). This suggests that frequent users are more prepared to watch videos that are suggested by YouTube based on their viewing history and further confirms that daily users might be using YouTube as a general form of entertainment. Lack of use of the most viewed and most popular pages further supports the idea that the respondents are less influenced by what others are watching when accessing YouTube videos. Even though people are influenced by the decisions of others.

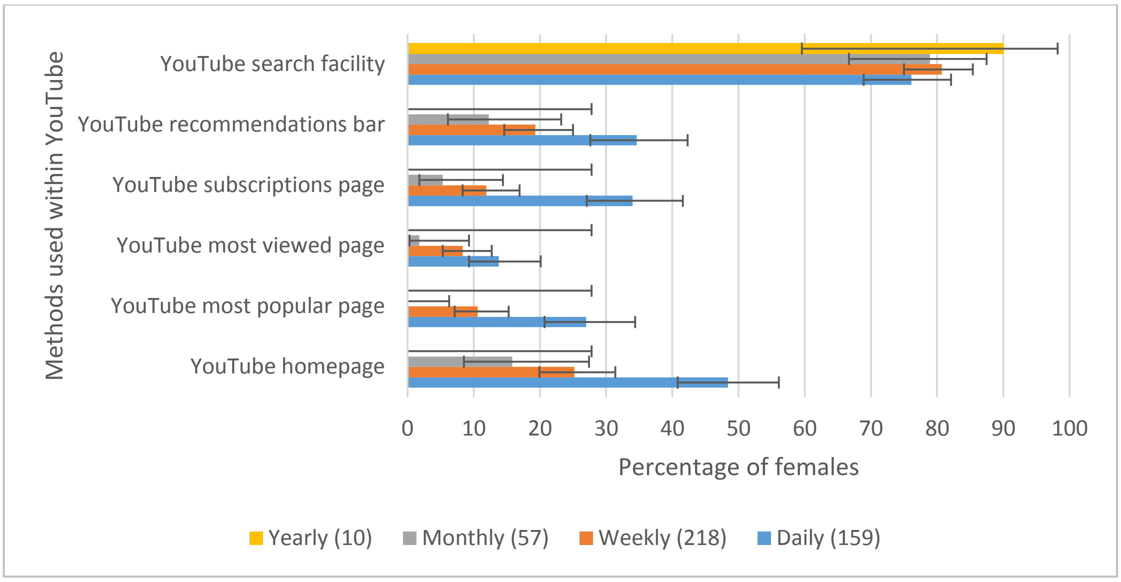


Figure A.6. **The methods the different female user respondents use when accessing videos through the YouTube website with 95% confidence intervals.**

Video Categories

The top six most popular categories for female respondents who are frequent YouTube users are Entertainment, Education, Comedy, Beauty, How-to and Animation, with most of these categories also being popular with monthly and yearly users as well (Figure A.7). Education will be popular due to the convenience sample used within this research and most respondents either working or have worked within education. The category of Beauty is less popular with monthly users, and Beauty and Animation are unpopular with yearly female users (Figure A.7). Cooking is more popular than Beauty for female respondents who use YouTube less frequently. In addition, cooking is quite popular with all female respondents across all user levels (Figure A.7). There is a decrease in popularity across most categories as we progress down through the user levels from daily to yearly (Figure A.7). This suggests that less frequent users have more specific categories that they access and usually have a reason for accessing videos, rather than browsing videos or being affected by recommendations and suggestions from YouTube (Figure A.6). Overall daily users seem to access a wider range of types of YouTube videos and again it is possible that this is due to them using this website as a form of mainstream entertainment as well as accessing it to find a video. This also seems to be true for weekly users but to a slightly lesser extent.

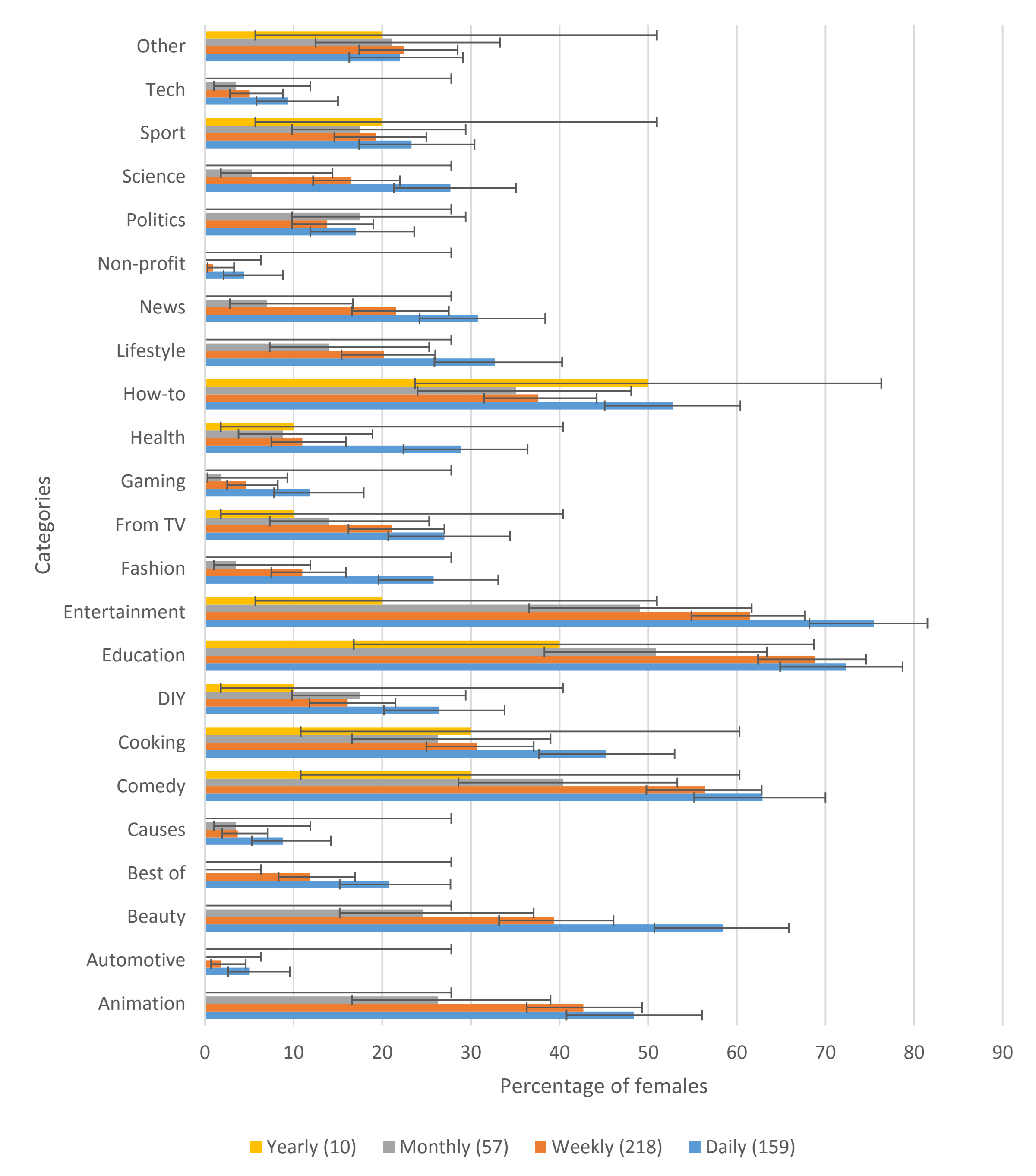


Figure A.7. **The categories of video the different female user respondents have watched with 95% confidence intervals.**

Influencing Factors

The title of the video is the most important factor in influencing the decision of the female respondents across all user levels to watch a video (Figure A.8). Thus, female respondents want information relating the video before they are willing to watch the content. The thumbnail picture of the video is also relatively important in the decision to watch a video for the female respondents who are daily, weekly and monthly users, but has no relevance for yearly users (Figure A.8). The length of the video is also a factor that seems to be relatively important to female respondents across most user levels (Figure A.8). The title, thumbnail and length are all metrics that provide information about the videos and therefore this seems more important to users. The title and length of a video can be key considerations in people choosing to watch online videos.

The number of comments and number of dislikes seem to have little influence in terms of the female respondents' decision to watch a video irrespective of user level, further supporting the idea that people do not seem to be persuaded by the opinions of others (Figure A.8). There is also a low level of influence from the number of likes that a video receives further supporting the notion that the respondents are not persuaded what others think of videos. However, people are influenced by the opinions and recommendations of others in the online videos that they choose to watch. It also seems that female respondents, particularly weekly, monthly and yearly users, are less influenced by the number of views a video has received (Figure A.8). This contradicts research that has suggest that people are more likely to watch videos that have high levels of views. The female respondents who are daily users are more influenced by a wider range of factors in whether they watch a video than other levels of user (Figure A.8). YouTube category, number of views and, to a certain extent, the age of a video (upload date) seem to be slightly more important and influential to daily users and this could be because of greater use and familiarity with the website. The video people are more likely to watch are those are new or have been uploaded more recently.

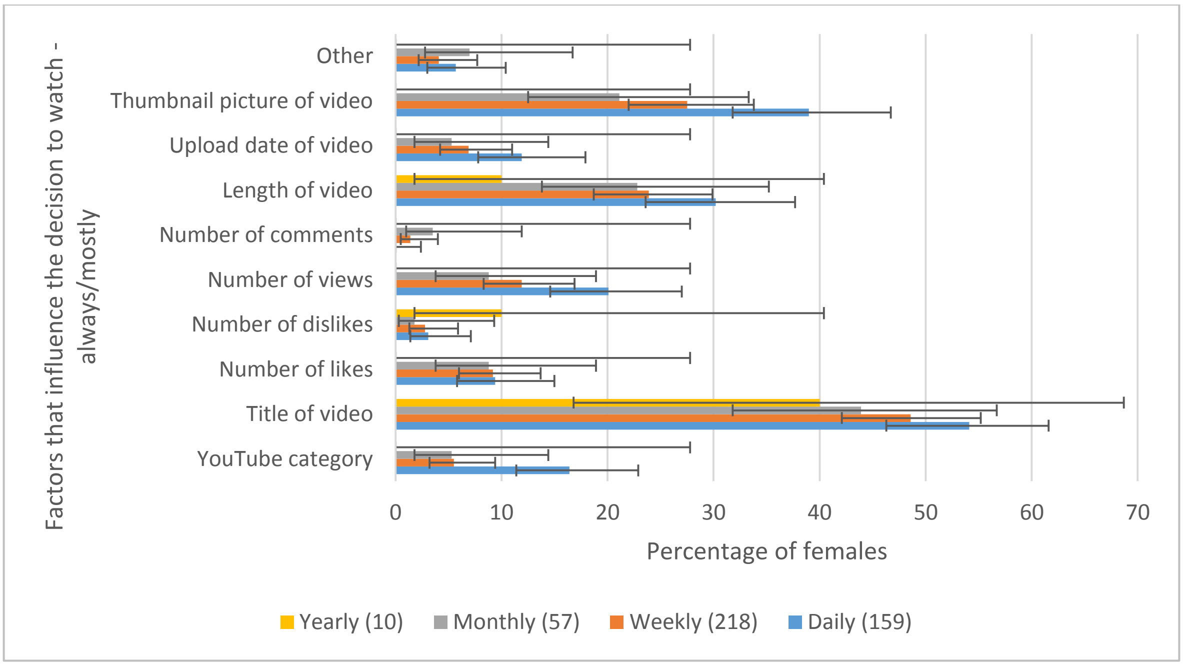


Figure A.8. **The factors that always or mostly influence the different female user respondents' decisions to watch a video with 95% confidence intervals.**

Usage frequency – Male

The following figures show the collected data for the various levels of user for male respondents. Respondents accessing videos daily or weekly will again be described as frequent users.

Education

Fifty male respondents classed themselves as daily users and 28 as weekly users, therefore, 90% (78/87) of the male sample are frequent YouTube users and clearly representative (Figure 8.3). The information provided by those male respondents who only access YouTube videos yearly will still be discussed and some conclusions may be established, but they will not be an accurate representation of this category of YouTube user.

There is a high proportion of male respondents 79% (69/87) who are frequent users and educated to higher education level. Overall 89% of all male respondent users are educated to Higher Education level (Figure A.9) supporting the credibility of the sample as internet users based upon the OxIS (Blank et al., 2019) and HESA (2018) findings.

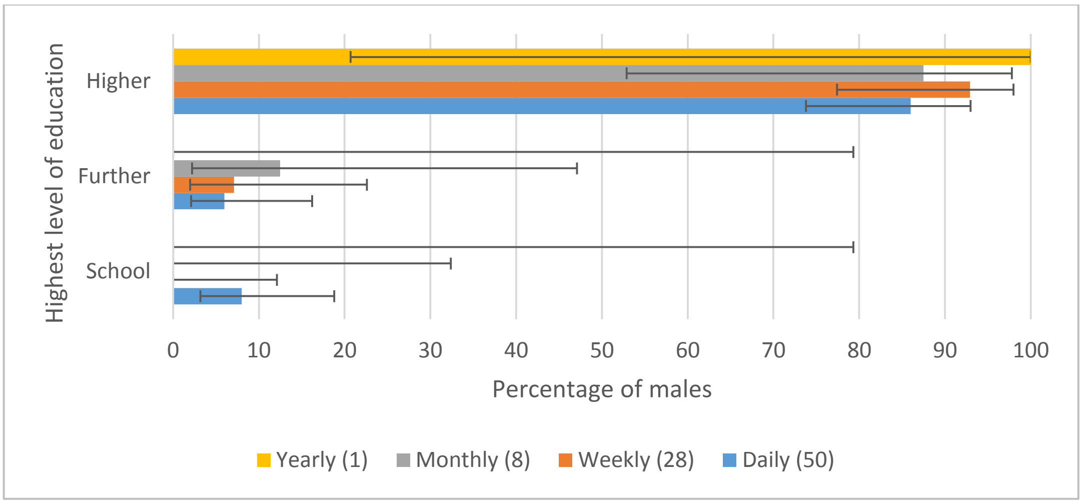


Figure A.9. **The education levels of different male user respondents with 95% confidence intervals.**

Accessing YouTube Videos

Using the YouTube website, the YouTube App, a hyperlink on Facebook and a Google search are the most popular methods for male respondents who are daily and weekly users to access YouTube videos (Figure A.10). Male respondents watching YouTube videos daily or weekly are using more methods to access their videos with 30 percentage or more across all the specified methods (Figure A.10). However, monthly users seem to use multiple methods to access YouTube videos as well. All monthly users access their videos through the YouTube website, suggesting that this is predominantly their first port of call when looking for videos (Figure A.10). Accessing videos through Facebook is also a popular method used by monthly male users. It seems that as less frequent users these respondents have more specific and focused ways of accessing YouTube videos (Figure A.10). It might also be that these respondents spend less time online and therefore only access limited or favourite websites. This could also be the reason that 50% of monthly users choose to use a Google search to find their videos (Figure A.10). Hyperlinks within emails seem to be a relatively popular way of accessing videos for all male respondents across the different levels of use (Figure A.10). Verbal recommendations and Twitter also seem to have some use with frequent male respondent users further demonstrating that these two group use a wider spread of methods to access YouTube videos. A third of daily users are accessing their YouTube videos through a hyperlink within a blog and this is substantially higher than any of the other user levels (Figure A.10). Overall, when these findings are compared to the female respondent user data (Figure A.4) there are many similar results. However, the key differences between male and female use are:

- Daily user male respondents have a much higher use of blogs than females
- Frequent male users are using Twitter more regularly than frequent female users
- Weekly user male respondents use the YouTube App substantially more than weekly female users
- Monthly male respondent users use YouTube and Facebook more than monthly female users.

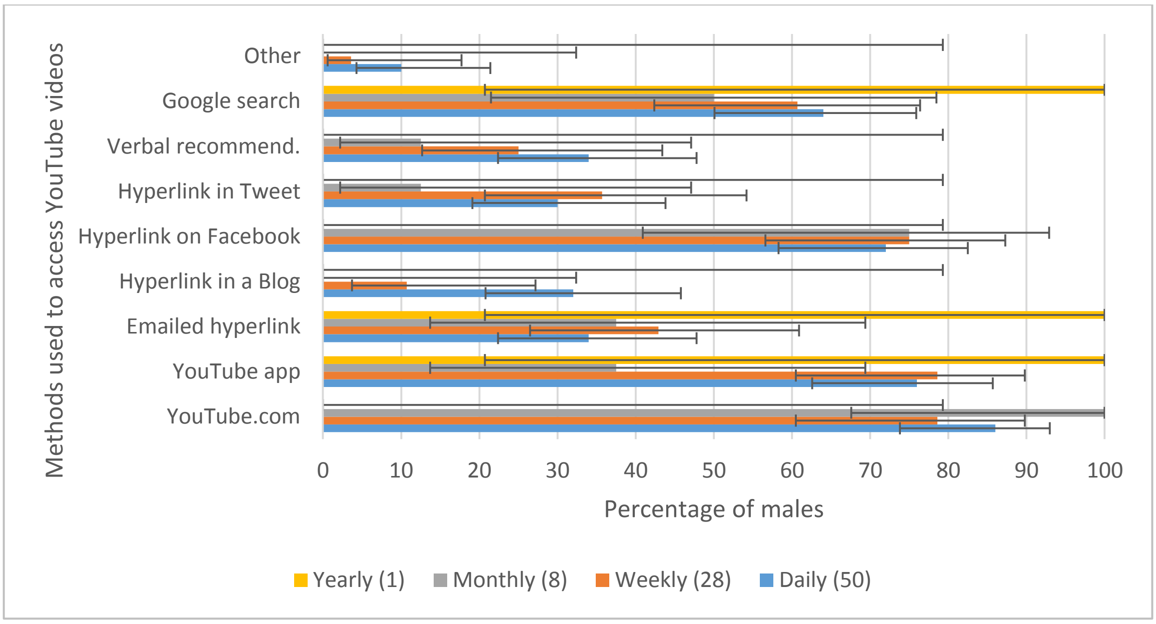


Figure A.10. **The different male user respondents have accessed YouTube videos with 95% confidence intervals.**

There is a substantial spread and difference between the number of methods each of the type of male respondent users use to access YouTube videos (Figure A.11). When compared to female respondents who are frequent users (Figure A.4) it seems that there is a wider spread in the methods used by frequent male respondent users (Figure A.11).

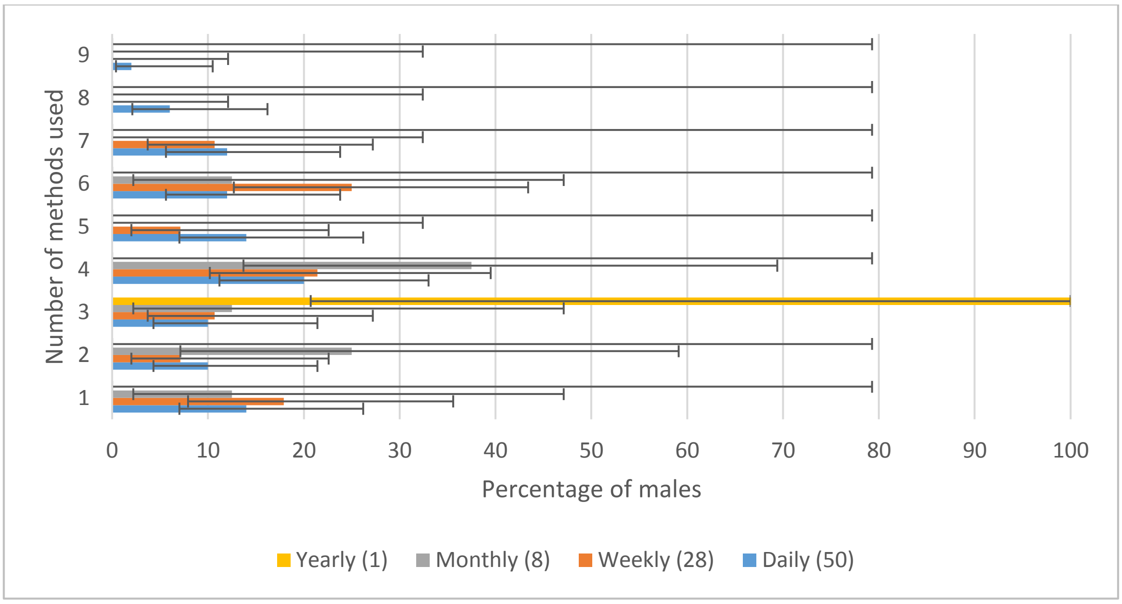


Figure A.11. **The number of methods that the different male user respondents use to access YouTube videos with 95% confidence intervals**.

Accessing Videos within YouTube

The most popular method used by daily, weekly and monthly male users for accessing videos within the YouTube website is through the site's search facility (Figure A.12). Apart from the YouTube search facility all other methods to access videos within the YouTube website appear to be more popular with daily male respondent users and then reduce through the different levels if user (Figure A.12). The least popular methods for all male users are the most viewed and most popular pages, once again, as discussed previously, suggesting that people are less influenced by what others are watching (Figure A.12). The homepage seems to be more popular with male respondents who are frequent users (Figure A.12). Frequent users may be more likely to browse YouTube as a form of entertainment and be more influenced by recommendations which are based on their interests. This also relates to the subscriptions page and recommendations bar which are also relatively popular with male respondents who are daily users (Figure A.12). Overall from comparing the respondents' data (Figures A.6 and A.12) there is similarity in the popularity of methods that they use to access videos within YouTube, however, the majority of these tend to be slightly higher for the male respondents.

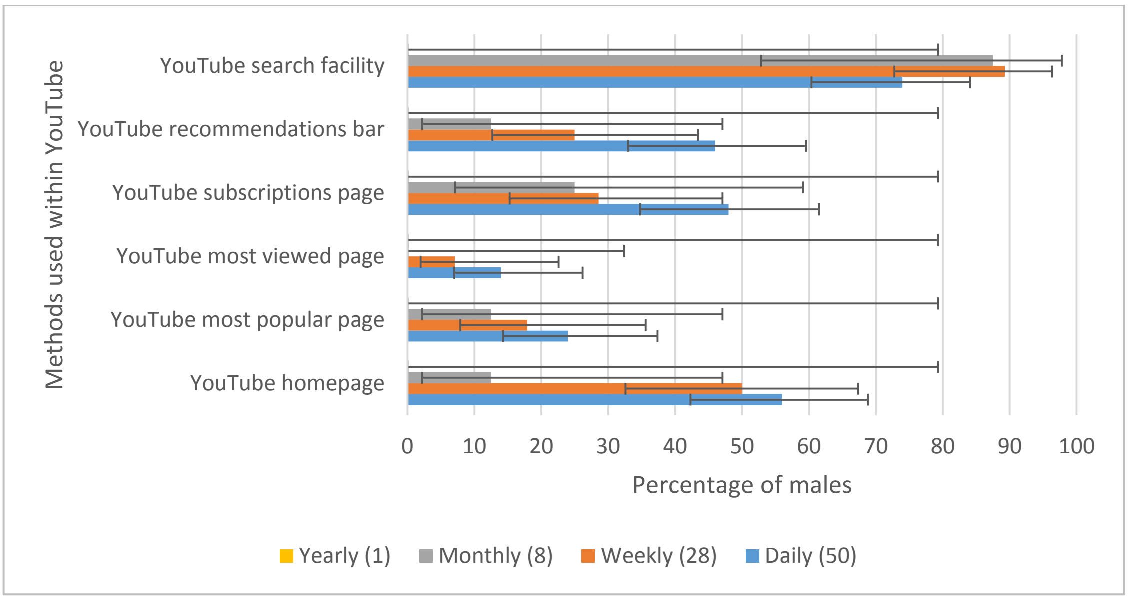


Figure A.12. **The methods the different male user respondents use when accessing videos through the YouTube website with 95% confidence intervals.**

Types of Video

The most popular categories of video that male respondents who are frequent users are Comedy, Entertainment, Education and Sport (Figure A.13). As discussed previously Education is popular with male respondent users due to the makeup of the convenience sample used within this research. Gaming, How-to, Science and Tech are particularly more popular with the male respondents who are daily users, and Animation is particularly more popular with male respondents who are weekly users (Figure A.13). However, from this data it is difficult to determine the reason for these findings and further research is needed to establish a reason for this. Beauty, Fashion, Non-profit and Causes are the least popular with all male respondents irrespective of their level of YouTube use (Figure A.13). Overall frequent male users, daily and weekly, seem to watch videos from a much wider range of categories than other male users. It is also clear by comparing the data (Figures A.7 and A.13) that although there are some similarities between the popularity of some categories, there are also substantial differences between the preferences of male and female respondents. Gender can have an impact in what type of videos are watched.

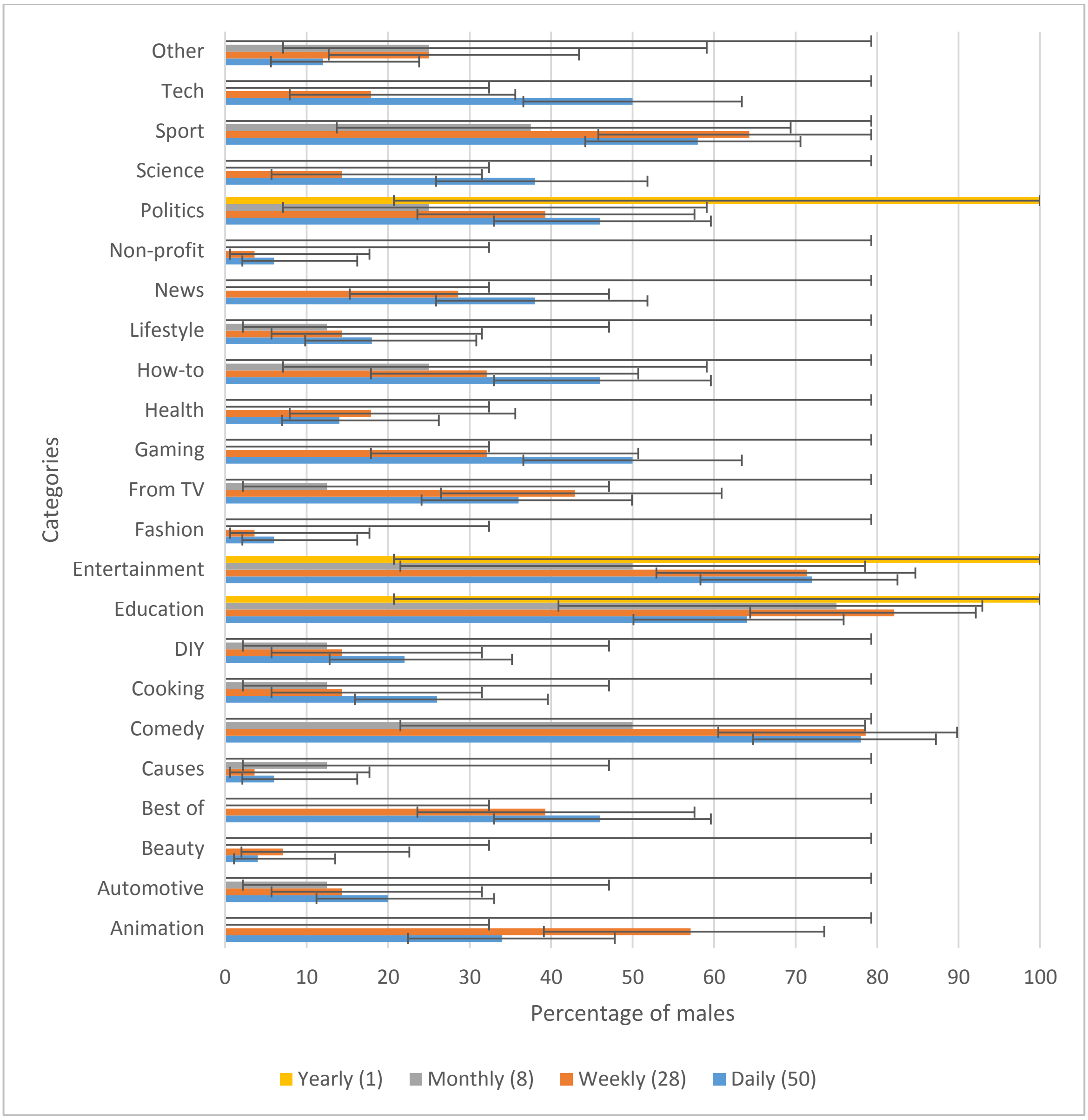


Figure A.13. **The categories of video the different male user respondents have watched with 95% confidence intervals.**

Influencing Factors

The most influential factors that support male respondents' decision to watch videos is the title of video and the thumbnail picture of video (Figure A.14). The length of the video seems to be more influential for weekly users rather than daily users, this suggests that daily users are prepared to watch videos of relevance with less concern about the length of the video (Figure A.14). Video title and length can be important factors in people choosing to watch online videos. Number of views, likes, dislikes and comments seem to have little influence on male respondents deciding to watch videos (Figure A.14). The number of views also has a relatively low level of influence over male respondents' decision whether to watch a specific video (Figure A.14).

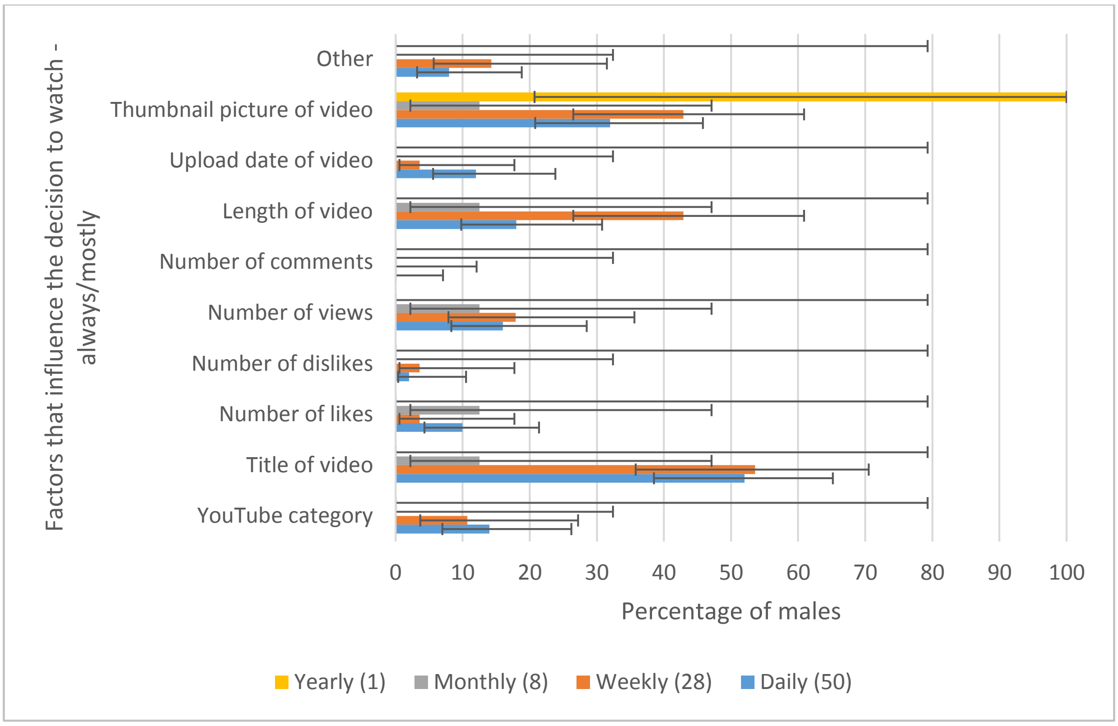


Figure A.14. **The factors that always or mostly influence the different male user respondents' decisions to watch a video with 95% confidence intervals.**