
Science communication in Podcasts: Factors that affect episode and channel popularity

Introduction

The essence of podcasting is the creation of audio or video content for an audience that wants to listen to what they want, when they want, where they want, and how they want (Burns, 2007). Science communication has traditionally been dominated by professional communicators employed directly or indirectly by the mainstream media (Valenti, 1999). With the emergence of Web 2.0, platforms such as blogs, wikis, social media and video sharing websites have redefined the mediascape (Minol et al., 2007; Brossard, 2013). Web 2.0 provides an alternative to traditional content distribution by reducing the barriers for content creators to reach an audience (Juhasz, 2009). Many Web 2.0 platforms are constructed on a participatory culture, a 'function that is most noticeably absent from most mainstream media' (Burgess and Green, 2009: 29). The audience is no longer a passive consumer, in the age of Web 2.0 the audience is now an active participant. Science communication is now conducted not only by professional communicators but also by scientists, interest groups, professional organisations and passionate amateurs across numerous Web 2.0 platforms (Nisbet & Scheufele, 2009; Lo et al., 2010; Claussen et al., 2013).

Podcasting is a particularly significant example of the Web 2.0 phenomenon. Podcasts are audio and video files that can be downloaded to a desktop computer, iPod, or other portable media player for playback later (Harris & Park, 2008; Potter, 2006). The term "podcast" was first used in 2004 and originates from the combination of the brand name of the then most popular player ("iPod") with "broadcast" (Campbell, 2005). In March 2018 Apple Podcasts, one of the most popular podcast search engines, passed 50 billion all-time episode downloads and streams (Locker, 2018). According to Apple, it is home to over 525,000 active shows, with more than 18.5 million episodes available, including content in over 100 languages (Locker, 2018; Lopez, 2018) and as of last year 50% of all US homes are podcast fans (Newman, 2017). While podcasts were initially created primarily by amateur producers (Affleck, 2005), mainstream media now provides a variety of podcast content. Commercial radio stations turned to podcasting as a way to increase listener involvement (Menduni, 2007) and even non-media organizations like NASA and the U.S. Pentagon are producing podcasts (Demopoulos, 2006). The evolving demographic of content creators in podcasting has meant that amateur science communicators now compete with large well-funded corporations like the British Broadcasting Corporation and the American National Public Radio.

Despite the large number of content consumers who listen to podcasts, reaching an audience is not guaranteed. Many of the factors that describe how a YouTube channel becomes popular can be used to describe how a podcast becomes popular, the difference being that instead of views, podcasts have plays or listens, but they fundamentally describe the same process. Reaching an audience and achieving success is a function of how popular a channel and its videos become, as measured by the number of subscribers and views received (Burgess and Green, 2009). The popularity of any given video is a function of the video's content factors, content-agnostic factors and YouTube's video recommendation system (Borghol et al., 2012; Figueiredo et al., 2014). YouTube's recommendation system both identifies what is popular and

creates what is popular in a rich-get-richer popularity scenario (Szabo and Huberman, 2010; Zhou et al., 2010; Figueiredo et al., 2011). That is, the recommendation system recommends popular videos to viewers, which in turn increases the popularity of those videos (Zhou et al., 2010). The same is true for podcasts but instead of YouTube's recommendation system any number of recommendation systems exist in the form of wherever the audience sources their podcasts, some of the most popular websites and apps being iTunes, Spotify, Google Podcasts, and Player FM. Content factors are the stylistic and informational characteristics of a video (e.g. topic, duration or delivery style), whereas content-agnostic factors relate to characteristics external to the video (e.g. the creator's social network or video upload date and time). Although a growing body of literature has independently addressed content and content-agnostic factors of podcasts broadly, few studies have examined science communication podcasts specifically.

To fill this knowledge gap, the content factors of science communication podcasts were examined for their influence on subscriber popularity. The differences in professionally and user-generated channels was first evaluated, specifically, the number of subscribers, age of the channel and number of podcast episodes created. Then, within the context of PGC and UGC, the impact of podcast length and pace and how the podcast was delivered was considered – delivery being a function of the gender, style, and the continuity of the delivery person(s) between podcast episodes. This was achieved by manually coding content factors of a sample of podcasts and analysing the relationships against a website's popularity metrics. Although manually coding limits the quantity of podcasts that can be sampled, it was necessary to obtain much of the data required. Understanding which podcast content factors contribute to popularity of a podcast and the impact of PGC on UGC, if there is any, will assist content creators to create more engaging and popular science communication content. In the 'Literature review' section, current research on understanding popularity is reviewed, followed by the 'Method' section that will detail the sampling protocols and podcast coding procedures. The 'Results' section follows, divided into channel specific and podcast-specific sections, and finally, the results are discussed and the article concludes by highlighting future research.

Literature review

Podcasts are audio and video files that can be downloaded to a desktop computer, iPod, or other portable media player for playback later (Harris & Park, 2008; Potter, 2006). The process of downloading podcast files as well as the development of podcast programs is known as podcasting (Berry, 2006). Podcasting involves three essential components: capturing, publishing, and distributing electronic media/digital content (Johnson & Grayson, 2006). After recording, content is then published via Really Simple Syndication (RSS), so that it can be picked up by an aggregator for distribution and ultimately distributed to the end user. RSS, also referred to as a web feed, offers the convenience of a subscription service, much like a newspaper or a magazine, and automatically delivers content directly to the subscriber (Johnson & Grayson, 2006; Burns, 2007). To distribute a podcast, a web page with space to store the podcast is needed. A website acts as a contact point for users on the Internet (Johnson & Grayson, 2006). Having published the contents, users need then to subscribe to the podcast to be able to collect the information. Once a user has subscribed to a podcast, the computer takes over the complex task of downloading information and synchronizing a portable media player (Johnson & Grayson, 2006).

Podcasts enable users to time-shift and place-shift content, meaning a user can tune in to the

podcasts they choose, whenever, and wherever she wants. In addition, users can access podcasts by subscribing to feeds which automatically download new podcasts as they become available (Potter, 2006; Burns, 2007; Evans, 2008). Podcasts have an additional advantage, which is largely unexplored in this study, of being a “push” or “subscription” rather than a “pull” technology. That is, the material is delivered directly from the source Internet location to the device, rather than requiring the audience to seek it out and download it (Campbell, 2005).

As there are a limited number of studies that have examined science communication in podcasts, the selection of content factors in this study may seem arbitrary, but this is not the case. Content factors, as opposed to content-agnostic factors, are the focus as they are valuable to understanding drivers of popularity broadly and allow recommendations to be made in the creation of science communication content. As with Wellbourne (2015) upon accepting content factors, the first evaluation is a fundamental separation of professionally generated and user-generated channels and their videos. Expected differences in channel resources between user-generated and professionally generated channels led to an examining of content factors related to the delivery of content. For instance, a channel with large resources may be capable of employing professional creators, which undoubtedly have different skill sets and, therefore, ideas about how a podcast should be presented Wellbourne (2015). Ultimately, the content factors selected provide a baseline for future research to build upon.

Before reviewing content factors, the primary content-agnostic factor that appears to drive video and channel popularity, a channel’s social network, must be addressed. In podcasting, as in YouTube a channel’s social network is the primary content-agnostic factor that influences, and also confounds, video and channel popularity (Burgess and Green, 2009; Juhasz, 2009; Yoganarasimhan, 2012). Podcast listeners are much more active on every social media platform, with 94% being active on at least one, compared with 81% for the entire population (Podcast Stats & Facts, 2018). Therefore the growth of video views is linked to the rich-get-richer effect of the recommendation system (Borghol et al., 2012) and the channel’s social network (Yoganarasimhan, 2012).

As Wellbourne (2015) discusses, despite these findings, social network analysis on YouTube is problematic for two reasons. First, a complete social network of podcasts cannot be attained because not all channels make lists of ‘friends’ or ‘featured channels’ available, and, second, it is not feasible to determine the social network of a channel beyond a podcast subscription service due to difficulties in connecting social networks across platforms (Yoganarasimhan, 2012). Although an analysis of the social network of science communication channels on podcasts is beyond the scope of this article, it is clearly an important consideration in understanding channel popularity generally.

Although the popularity of a podcast is a function of content and content-agnostic factors, content factors appear to be the most informative for understanding broad popularity of a podcast. Broad popularity is meant here as popular among a wide spectrum of viewers, whereas narrow or niche popularity is only popular within a limited audience (Wellbourne, 2015). Figueiredo et al (2014) found that user preferences meant that in many evaluations users could not come to a consensus on which video had the best content, but, in those evaluations where users did come to a consensus, the video identified as having the preferred content was frequently more popular on YouTube. Taken to apply to podcasts, for a podcast to be popular among a broad audience, the content must be broadly appealing. Therefore, understanding the

content factors is vital to understanding what drives popularity broadly.

The type of channel is of particular interest in understanding podcast popularity. Professionally generated channels (i.e. channels that exist to extend commercial branding) often have superior financial resources compared with user-generated channels. Podcast production is relatively inexpensive, which obviates any need for costs to be shifted to the listener (Burns, 2007). However financial resources can allow professionally generated channels to increase the appeal of the channel and/or of specific videos through the creation of regular or large volumes of content and content of high production value (Wellbourne, 2015). Capturing audio content, which is less time-consuming to edit than a video podcast (Armstrong, Tucker, & Massad, 2009), requires some level of expertise (a producer), special software, and the knowledge and skill to record and save the content as an MP3 (Johnson & Grayson, 2006). Although superior resources might allow channels to employ professional video producers and presenters, it has been argued that “in order to operate effectively as a participant in the internet community, it is not possible simply to import learned conventions ... from elsewhere (e.g. from professional television production)” (Burgess and Green, 2009: 69). It is clear then that while regular content assists in engaging an audience, a channel must still host content that the community finds engaging.

In an information-rich world, the limiting factor in consuming content is the consumers' attention (Davenport and Beck, 2001). Podcasting is unique in that listeners often use them to “fill in the air and space ‘while doing tasks such as commuting’” (Heshmat, Yang, & Neustaedter, 2018). In fact when US podcast listeners were polled 22% of podcast listeners consume podcasts while driving, with only 49% of podcast listening being done at home (Podcast Stats & Facts, 2018). Podcast listeners were also found to listen to all or most of each episode, with only 20% saying they don't. A similar percentage of listeners increase the speed of the podcasts they're listening to (19%) (Podcast Stats & Facts, 2018). As discussed previously the audience is quite large and 17% (48 million) of those polled listen to podcasts weekly, up from 15% in 2017, and on average podcast listeners consume 7 different shows per week, up from 5 in 2017 (Podcast Stats & Facts, 2018). Therefore, it logically follows that the majority of popular podcasts being produced currently are listened to in their entirety and do not exceed the consumers' attention.

Due to the limited studies in the field, this study makes a significant contribution to the science communication literature by examining science communication in podcasts more thoroughly.

Method

Podcast procurement

To achieve the aims of this article, it was calculated that a minimum sample of 250 podcast episodes was required. In August 2018, the “most popular” podcast channels were selected from <http://toppodcast.com> using the category of “Science and Medicine” (Top Podcast Charts, 2018). Podcasts in English, with at least 5 episodes and could be defined as science communication (in the context of this study, see definition below) were retained until 50 channels were identified.

Science communication

Science communication in practice is considerably broad, often attracting equally broad definitions in the academic literature (Bryant, 2003; Gilbert and Stockmayer, 2013). As was done in the study by Wellbourne (2015), 'science' was taken as any topic that would be categorised into one of the Scopus science subject areas of physical, life, health or social sciences, excluding the topic of 'Arts and Humanities' (Elsevier, 2017). The tone of communication of these topics can also be quite broad. Hence, 'science communication' in this study was taken to be any video that might be seen as a form of science journalism that is not overtly didactic or instructional, while also not being principally focused on entertainment (Wellbourne, 2015). Although this is somewhat subjective, consistency was maintained as a single author (B.J.K.) reviewed all material for inclusion.

Data coding

The collection of channel data and content factors of the identified podcast channels began in August 2018. Following a method similar to the one used by Wellbourne (2015) data was obtained on episodes and channels using manual coding procedures. In some cases information was also obtained from <http://podbay.fm> (podbay.fm, 2018). The following data was coded for each channel:

- (a) Channel age, as measured from the first upload event;
- (b) Number of episodes at the time of data procurement;
- (c) Channel type, coded as PGC for channels named after corporate entities or as UGC for channels that are amateur derived.
- (d) Colours used in logo, the three main colours used in the logo, using <http://mkweb.bcgsc.ca/color-summarizer/> (Image Color Summarizer, 2018)

The following popularity metrics were extracted for all videos simultaneously:

- (a) "Visit" count;
- (b) Average rating.

Each video was reviewed manually and the following content factors coded.

- (a) Video length (seconds) taken as the complete video duration.
- (b) Pace of content delivery (words per minute) calculated from the audio by taking the average of three 1minute samplings.
- (c) Communicator continuity (binary) identified whether a channel had a continuous science communicator or communicators who delivered content.
- (d) The guest to host ratio, for example where a podcast with no guests is 1:0, and a podcast with one host and two guests is 1:2.
- (d) Gender (male, female, both or no-gender) of the person or persons delivering the science

content.

(e) Podcast style was coded as one of four styles identified while reviewing the dataset –

Monolog: an iconic style where the presenter delivers content as a monolog without anyone else present in the podcast;

Hosted: stylistically similar to monolog where the communicator presents the information; however, other people such as members of the public or interviewees are also part of the podcast content;

Interview: podcasts where the person delivering content is being interviewed by a person who is often the host;

Talkshow: the presenter discusses information with co-hosts or guests, and where everyone speaks roughly equally.

Statistical analysis

All statistical analysis was carried out in the R statistical package version 3.0.2 (Cran Team, 2014). Provided assumptions held and data transformations were suitable, parametric tests were used, otherwise non-parametric tests. Welch's t-test was used in place of Student's t-test where unequal variance was identified using Levene's test for homogeneity of variance. An alpha of .05 was used for significance in all tests. Effect sizes and correlations were described according to Cohen (1988) and Evans (1996).