Video consumption and viewer engagement with the Royal Institution's science videos

A report for the Royal Institution

October 2015

Foreword

On behalf of Pro Bono Economics, I am delighted to introduce this report for the Royal Institution.

For over two hundred years, the Royal Institution has been committed to the sharing of science knowledge and learning with the general public in Britain and beyond. Educational and research institutions in the UK are not exempt from the less than certain funding environment that many charitable organisations now find themselves in. This report is an excellent example of how Pro Bono Economics is enabling organisations to make best use of their scarce resources without compromising on core activities. With changing times, the Ri now focuses a significant proportion of its resources on educational videos, and thus wanted to know what impact these were having on engaging the public with the sciences.

FTI economists analysed data provided by the Ri to shed some light on the demographics of their video viewers and the most engaging content according to the number of comments, likes and shares. Findings show that the Ri's online audience is young (with over two thirds of views coming from those younger than 35) and skewed towards a male audience. Since the collection of data, the Ri have been working to grow their female audience. In terms of format, Ri talks and events receive the highest levels of engagement.

As a patron of PBE and having worked as an economist for many years, I am keen to see the skills of our profession employed for the good of society. The work of Pro Bono Economics does just that. This report will help the Royal Institution pioneer new methods of digital engagement with the sciences in a way that extends the reach and impact of their videos, whilst providing insight into the most engaging content. But as FTI economists have identified, more work needs to be done, particularly in the way data is collected and kept to allow for further research and proper econometric analysis. This would help the Ri better track its impact and make more informed decisions on strategy in this area.

Vicky Pryce

Chief Economic Adviser at CEBR
Patron of Pro Bono Economics

Terms of reference

This report has been prepared by FTI Consulting LLP ("FTI Consulting") for the Royal Institution of Great Britain (the "Ri"), in connection with the Ri's use of online video content in furthering public engagement with science. We were introduced to the Ri by Pro Bono Economics ("PBE"), and have produced this report on a *pro bono* basis.

The Ri has asked for our advice on measures of video consumption and viewer engagement with its online science videos. It will use this information to inform the development of its digital strategy, to track the impact of its digital content and to support fundraising initiatives.

We have worked closely with the Ri to develop relevant metrics, to examine how they vary between different types of video and viewer, and to make recommendations for further analysis. We presented our full findings to the Ri in a separate (private) report. This report presents a shorter summary of our findings.

Authors

Tim Battrick is a Director at FTI Consulting. His work focuses on the valuation of shares, businesses and other assets. Tim has helped advise companies and law firms on economics and finance issues across a range of sectors.

Ravi Kanabar is a Senior Consultant at FTI Consulting. Ravi specialises in applying economic, statistical and financial analyses to advise public sector organisations, companies and law firms in a range of sectors.

Acknowledgements

The authors would like to thank Frances Cheesman for her help in compiling the dataset used in this report, and Dugald Young and Tom Quigley for their contributions to the analysis.

Disclaimer

This report has been prepared solely for the benefit of the Ri for the purposes above. FTI Consulting accepts no liability or duty of care to any person (other than to the Ri and PBE under the relevant terms of our engagement) for the content of this report and disclaims all responsibility for the consequences of any person (other than to the Ri and PBE on the above basis) acting or refraining to act in reliance on the report or for any decisions made or not made which are based upon the report.

This report contains information obtained or derived from a variety of sources, including a detailed dataset provided by the Ri. FTI Consulting has not sought to establish the reliability of those sources or verified the information provided. No representation or warranty of any kind (whether express or implied) is given by FTI Consulting to any person (other than to the Ri and PBE on the above basis) as to the accuracy or completeness of this report.

Contents

1.   Introduction

Background

1.1   The Ri is a registered charity, at the forefront of public engagement with science. It was founded in 1799, with the aim of "*introducing new technologies and teaching science to the general public*".[1] Its charitable purpose is to "*encourage people to think more deeply about the wonders and applications of science*",[2] which it does by funding:

- educational activities, including the 'CHRISTMAS LECTURES', its flagship series of lectures, broadcast on national television every year, and most recently, the digital educational activities that are discussed below; and

- heritage work, by maintaining and exhibiting collections of books, scientific instruments, and furniture relating to the history of science, at the Faraday Museum in London.[3]

1.2   As part of its educational activities, the Ri launched an online video platform in late 2011, called the Ri Channel.[4] The Ri Channel is a website showcasing the "*very best science videos from the Ri and around the web*", including full recordings and highlights of recent Ri talks, re-digitised footage from the Ri archive, and other science videos made by the Ri and other science institutions. Many of the same videos are also published on the Ri's YouTube page.[5] The Ri's videos are released under a 'Creative Commons licence', and so they can be freely copied, edited, and distributed – for example, by being 'embedded' on other websites.[6]

1.3   A key element in the Ri's strategy relates to these videos: it aims to "*increase audience reach and impact with specific focus on…pioneering new methods of digital engagement through the Ri Channel*".[7] It is this aspect of the Ri's work that we are advising on.

---

[1]   http://www.rigb.org/our-history

[2]   http://www.rigb.org/about/mission-and-vision

[3]   http://www.rigb.org/visit-us/faraday-museum

[4]   http://www.richannel.org/

[5]   https://www.youtube.com/user/TheRoyalInstitution/featured

[6]   https://creativecommons.org/licenses/

[7]   2013 summary information return to the Charity Commission, http://apps.charitycommission.gov.uk/SIR/ENDS38/0000227938_SIR_20130930_E.PDF

1.4   We have been asked to advise the Ri on measures of video consumption and viewer engagement with the videos on the Ri Channel and its YouTube page. These measures are to be used to inform the development of the Ri's digital strategy, to track the impact of this digital content, and to support the Ri's fundraising initiatives. The Ri requested a report providing a snapshot of the data at a specific point in time, as well as help and advice about using the report to support research that will analyse reach and educational impact on an ongoing basis. We presented our full findings to the Ri in a more detailed report.

1.5   Greater engagement with the Ri's science videos may lead to greater engagement with science in general – for example, through an increase in visitor numbers to the Ri's Faraday Museum, or wider benefits such as an increase in the number of students studying science at university. It is difficult to distinguish the effect on such wider forms of engagement of (1) the Ri's videos, from (2) all other factors that may also influence it. In this report, we focus on engagement with the videos themselves.

Structure of the rest of this report

1.6   This report is structured as follows:

- in Section 2, we introduce the Ri's online video collection and present measures of video consumption and viewer engagement for the Ri's videos;

- in Section 3, we discuss how video consumption and viewer engagement vary between different types of video and viewer; and

- in Section 4, we recommend some next steps and further analysis for the Ri's ongoing research.

2.   The Ri's online video collection

The Ri's dataset and video catalogue

2.1   The Ri has provided us with some detailed datasets collected between October and December 2014, containing information on 757 videos. Our analysis focusses on 342 (45%) of these videos: those produced by the Ri and active on the Ri Channel website, YouTube, or both – which we refer to as "the Ri's active catalogue".[8]

2.2   The chart below illustrates the Ri's active catalogue.

Figure 2-1: Ri videos by platform and availability
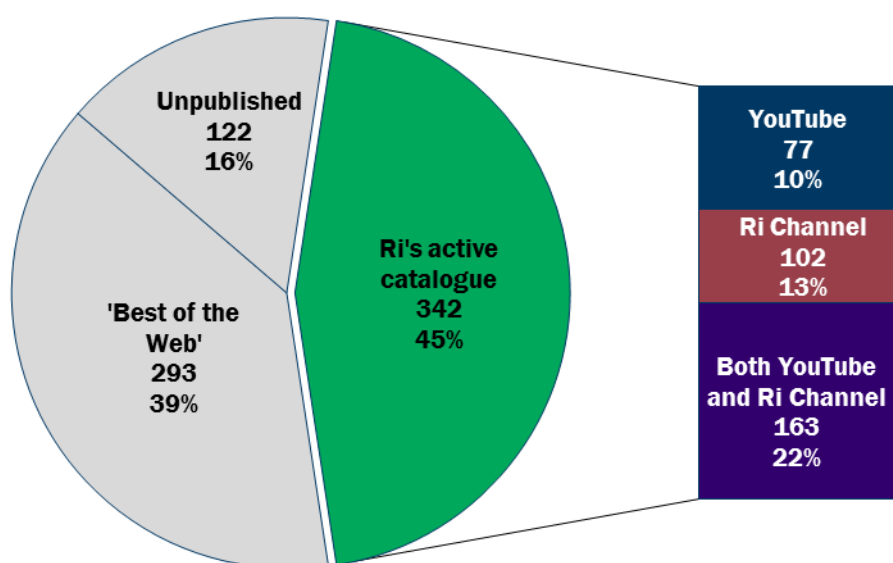
2.3   Videos in the Ri's active catalogue are classified according to their:

- Collection (e.g. CHRISTMAS LECTURES, Ri Shorts, Tales from the Prep Room); and

- Subject (e.g. Biology, Chemistry, Physics);

- Theme (e.g. 'Being human', 'Space & Time', 'Talking Science');

- Format (e.g. Documentary, Demo, Interview).

2.4   These categories are summarised in the table below.

---

[8]   A further 122 (16%) of videos are unpublished, and 293 (39%) are 'Best of the Web' videos posted on the Ri Channel but not produced by the Ri. Our analysis does not consider these videos because they are not produced by the Ri, and because complete data is not available for them.

Table 2-1: Categories of video

Collection	Subject	Theme	Format
Advent 2012	Biology	Being Human	Animation
Advent 2013	Chemistry	Engineering	CHRISTMAS LECTURE
CHRISTMAS LECTURES	Engineering	Environment	Demo
Crystallography	Maths	Materials	Documentary
Event video	Natural World	Maths	Event
ExpeRimental	Physics	Natural World	Interview
Ri Shorts	Space	Space & Time	Talk
Tales from the prep room		Talking Science	
		Technology	

Source: Ri dataset.

Video consumption

2.5 'Video consumption' is the extent to which the Ri's videos are watched. It can be measured using a range of metrics, which differ in their:

- unit of measurement – we use the number of views and the minutes watched as our primary units of measurement;

- level of aggregation – the metrics can be summarised for the entire catalogue of videos, for the active videos only, for YouTube videos only, for certain categories of videos, or for individual videos; and

- time period – metrics can be in cumulative terms (e.g. number of views since publication), or between two specific points in time (e.g. number of views last month), or an average over time (e.g. average number of views received per day, in the month after publication).

2.6 We have focused on the following metrics for individual videos:

- total views, since publication – i.e. the cumulative number of views of a particular video, since publication, up to late 2014; and

- average views per day – i.e. the average number of views per day of a particular video. This reduces bias against new videos, which tend to have fewer views simply because they have not been available for as long as older videos.

We have calculated these statistics for each individual video and also calculated the average of these statistics across specific sets of videos, such as the Ri's entire YouTube catalogue and videos addressing a specific topic.

2.7 We calculate equivalent metrics for minutes watched.

2.8   Videos in the Ri's active catalogue have received more than 7 million views in total, between the dates they were each published, and late 2014. The majority of these views are via YouTube (6.7 million views on YouTube, compared to 0.4 million views on the Ri Channel).[9] These views have resulted in over 50 million minutes watched (approximately 97 years of video).

2.9   We have also examined these metrics on a per video and per day basis. The metrics suggest that the Ri's videos hosted on YouTube appear to receive far better exposure than videos hosted on the Ri Channel. On average, videos on YouTube receive 75 views per day, while videos on the Ri Channel receive just 2. However, the Ri's dataset is likely to understate the true number of Ri Channel views, because it does not reflect views of Ri Channel videos which have also been published on players other than YouTube's, such as Vimeo and the AOL On platform. In the rest of our analysis, we focus on YouTube metrics.

Viewer engagement

2.10   The Ri's videos might encourage public engagement with science in a number of different ways. We focus on metrics than be calculated from the Ri's existing data, concerning whether the videos:[10]

(1)   inspire conversation. That is, do Ri videos receive large numbers of comments and get discussed on forums and in news articles? What proportion of this conversation is 'on topic'?

(2)   result in applause. Viewers are often invited to express their satisfaction (or dissatisfaction) with online content, for example, by clicking a 'like' button. How much applause do Ri videos receive? What proportion of views result in such applause?

(3)   get shared. That is, do individuals broadcast the content to those in their own social networks (who may, or may not already be familiar with the Ri)? What proportion of views result in such sharing?

2.11   Metrics that describe conversation, applause and sharing go beyond the simple 'number of views' and 'minutes watched' measures of video consumption, to describe what viewers do after viewing the content. Before we turn to specific metrics, we note that there is a distinction between:

---

9   For the 163 videos that are available on both YouTube and the Ri Channel, the dataset suggests that 84 have no views on the Ri Channel. On closer inspection, these videos appear to play not in the Ri's own video player, but using some other player. For example, some are embedded YouTube videos. It may be that users are viewing these videos on the Ri Channel website, but that their views are registered as a view on YouTube. This may bias the "YouTube views" upwards, and the "Ri Channel views" downwards.

10   These three categories are similar to metrics that can be used to measure engagement on social media pages, such as Twitter, Facebook of Google+. See www.truesocialmetrics.com.

- the overall level or quantity of engagement (which could be proxied by, for example, the total *number* of comments on a video), and

- 'how engaging' the video is, i.e. its 'engagement quality' (which could be proxied by, for example, the *proportion of views* that result in a comment).

2.12 The overall quantity of engagement of a video is the product of its level of consumption and its engagement quality. Two videos can therefore achieve the same overall level of engagement (e.g. 100 comments in total), but with very different levels of consumption:

- a relatively unengaging/low engagement quality video (e.g. with only 1 in every 10,000 viewers adding a comment), which is nevertheless very popular (e.g. with 1,000,000 views in total), or

- a very engaging/high engagement quality video (e.g. with 1 in every 10 viewers feeling compelled to leave a comment), but with limited consumption (e.g. only 1,000 views).

2.13 The overall quantity of engagement can therefore be improved by increasing the level of consumption and/or by producing more engaging videos. In Table 2-2 we list metrics that can be calculated with the Ri's dataset, for both the quantity of engagement, and its engagement quality.

Table 2-2: Engagement metrics

Type	Metrics for quantity of engagement	Metrics for engagement quality
Conversation	Total comments	Comments per 1000 views
Applause	Total likes Total favourites (Total dislikes)	Likes per 1000 views Favourites per 1000 views (Dislikes per 1000 views) (Likes per dislike)
Sharing	Total shares	Shares per 1000 views
Other	Subscribers added (Annotation data)	Subscribers per 1000 views Percentage of video watched

Note: we discuss these metrics in a more detailed report provided to the Ri. We do not focus on the metrics in parentheses, for reasons discussed in that report.

2.14 The summary statistics for these engagement metrics, across all YouTube videos in the Ri's active catalogue show that:[11]

- the average Ri video on YouTube has received over 70 comments (about 3 in every 1,000 views) and 350 likes (about 14 in each 1,000 views);

- it has been added to a registered user's 'favourites' playlist over 70 times (about 2 in every 1,000 views), and has been shared using YouTube's built in functions about 50 times (about 3 in every 2,000 views);

---

[11] Videos hosted only on the Ri Channel are excluded from this analysis, as we do not have the data required to calculate these metrics.

- about 3 in every 1,000 views result in the viewer subscribing to the Ri's YouTube channel; and

- on average, viewers watch about 56% of a video.

2.15   However, there is significant variation in engagement between videos. We examine this variation further, in the next section.

3. How do video consumption and viewer engagement differ between videos and viewers?

3.1 Video consumption and viewer engagement vary significantly between different types of video and viewer: some video characteristics (subject, format, duration) are associated with better consumption and engagement than others, and different types of viewer (age, gender, geographical location, and viewing devices) have different video consumption and engagement patterns.

3.2 We consider this variation in this section but these findings must be interpreted with care, because they may not represent causal relationships and they may not generalise to other settings. They are a first step in further analysis, which we recommend the Ri considers in Section 4.

Correlation and causation

3.3 We identify below the factors that are correlated with better consumption and engagement. These factors may not cause better consumption and engagement. Before concluding that a particular characteristic of a video (e.g. its subject, format, or duration) is responsible for that video's success (or lack thereof), it is important to recognise the contribution of other factors too. Video consumption and viewer engagement will be influenced by a range of factors, including:

- the characteristics of the videos including any specific topics, formats, durations etc.;

- other factors within the Ri's control including how the videos are marketed, when and where they are published, etc.; and

- factors outside of the Ri's control including whether a related topic has been reported in the news recently, or the efforts of other competing or complementary organisations to stimulate interest in science, or whether the topic of the video is so niche that there are a limited number of potential viewers, or the extent to which the video has been embedded on third party sites, including news sites and blogs.

3.4 Specialist statistical techniques (econometrics) can sometimes be used to isolate the impact of particular characteristics, so as to distinguish the causes of a video's success. We discuss this in our section on further work.

3.5 A related question is whether the associations identified in this section (setting aside whether they are causal relationships or not) can be generalised to other settings – for example, to any new videos that the Ri produces in the future, or to videos produced by other organisations.

3.6 If the Ri chooses to produce more videos with the positive characteristics we identify, it *may* find that these videos receive similarly high levels of consumption and viewer engagement. However:

- even if these videos have had high levels of consumption and engagement in the past, there may be limited benefit to producing more such videos, especially if they are on niche topics, with a limited audience, and little scope for achieving a wider viewership; and

- videos that are relatively unpopular may still be worth producing because they relate to important issues that some viewers are interested in and for which there are limited other sources of information.

Variation by type of video

Subject

3.7 The majority of videos in the Ri's active catalogue relate to Biology, Physics, or Chemistry. However, there is significant variation in consumption and engagement metrics by subject:

- the three traditional science subjects are responsible for the vast majority of YouTube views and minutes watched. This is to be expected, given these subjects' dominance in the Ri's active catalogue;

- on a views per day basis, individual Physics and Chemistry videos are more popular than Biology videos. Physics and Chemistry also enjoy a high quantity of viewer engagement (in terms of conversation, applause and sharing), although this is driven by the high numbers of views rather than the quality of engagement;

- videos on the subjects of Space and Engineering have lower levels of viewer engagement. These results are driven both by the relatively small number of views of these videos, and those viewers that do watch them being less inclined to comment or click like; and

- Maths videos are a relatively new addition to the Ri's catalogue, but they have particularly high consumption metrics (receiving 126 views per day, significantly higher than all other subjects on average) and consistently strong viewer engagement, with high rates of comments, likes, favourites and shares.

Format

3.8 The Ri's videos are mainly recordings of Demos, Events or excerpts of CHRISTMAS LECTURES, with a smaller number of Interviews, Documentaries and Talks. The consumption and engagement metrics vary between these formats:

- Events and Talks are particularly popular in terms of the number of views they receive, and also enjoy high levels of engagement, with Talks being especially well 'shared';

- although there are only five Animations in the catalogue, these videos are relatively popular and generate high levels of engagement, although this appears to take the form of 'lower level' applause, and not discussion; and

- videos that use the Interview or Documentary format do not perform as strongly as the other formats, either in terms of consumption or engagement.

Duration

3.9   The average video in the Ri's active catalogue is 14 minutes long – this is much longer than the typical popular YouTube video, which is likely to reflect the nature of the content of the Ri's videos.[12] However, the majority of Ri videos fall into three 'clusters': shorter than 10 minutes (60% of videos), 35 to 40 minutes (5%), and 55 to 60 minutes (15%).

3.10   There is significant variation in consumption and engagement, by duration:

- videos shorter than 10 minutes, are responsible for the majority of YouTube[13] video consumption and viewer engagement. However, this mainly reflects the sheer number of short videos that exist. On a per video basis, they are not as popular as longer videos. On a per view basis, they are not 'as engaging' as longer videos either;[14]

- medium length videos (between 10 and 40 minutes long) have poor consumption and engagement metrics; and

- longer videos have strong levels of engagement, driven not only by their popularity, but by their viewers' tendency to comment, like, favourite and share the videos. The 22 videos between 50 and 70 minutes in length are alone responsible for recruiting over a third of all subscribers to the Ri's YouTube channel.

Variation by type of viewer

Device

3.11   Figure 3-1 shows that the majority of the Ri's YouTube videos are viewed on a desktop or laptop computer (70%), with mobile phones (16%) or tablets (10%) representing most of the remainder.

---

[12]   YouTube does not publish statistics for the average duration of its videos, although studies based on samples of videos (for example, the 100 most popular videos at a point in time), suggest that the average duration is around 5 minutes.

[13]   These comments consider only those videos available on YouTube.

[14]   There is significant variation within the 0-10 minute segment, in both usage and in engagement. In general, the longer videos within the segment (e.g. 5-10 minutes, as opposed to 0-5 minutes) are far more highly consumed, both on a per video per day basis, and in aggregate (despite there being fewer such videos); and have greater levels of engagement, although this appears to be driven by their high level of usage (and not users finding them particularly 'more engaging').

Figure 3-1: YouTube views, by device

Other, 4%
Tablet, 10%
Mobile, 16%
Computer, 70%
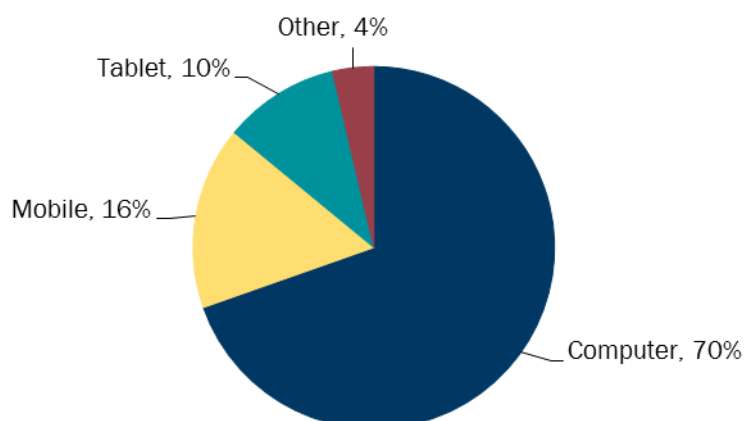
Note: (1) this data was collected by the Ri on 2 December 2014, and reflects views received up until that date. The proportion of views on tablets and mobiles may have increased over time; (2) the 'Other' category includes TVs, games consoles and unknown devices.

3.12   Users tend to watch similar subject videos across devices. However, big screen devices (TVs, desktops, game consoles) are used to watch videos for longer than smaller screen devices (tablets, and to a greater extent, mobile phones).

3.13   For YouTube as a whole, 50% of views are on mobile devices.[15] This is much higher than equivalent figure for the Ri's active catalogue (26% on mobiles or tablets), and is likely to reflect differences in the types of videos produced by the Ri, including that the Ri's videos tend to be longer, and therefore better suited to big screen devices.

Age and gender

3.14   For each video, we have data on:

- age group and gender, for views by registered and logged in users;[16] and

- country, for all views, regardless of whether the user is registered and logged in or not.

3.15   The Ri's audience (as measured by the number of views by these registered users) is young (with over two thirds of views coming from those younger than 35).

3.16   Male and female viewers appear to have similar tastes in video subjects, although males are slightly more interested in Chemistry, Physics and Space videos than females, while females prefer Biology and Natural World videos.

---

[15]   https://www.youtube.com/yt/press/en-GB/statistics.html

[16]   The analysis of age group and gender should be interpreted with care, because we do not know: (a) how representative *registered and logged in* viewers are of *all* viewers, and (b) how truthfully users enter their age group and gender information when registering their account.

3.17   Minor differences also exist by age, with Biology and Chemistry videos being more popular with younger viewers, and Maths, Space, Physics and Engineering videos having a comparatively older audience.

Country

3.18   Each time a video is viewed, the Ri's dataset records which country it was viewed from. This data shows that the Ri's viewership is international. Viewers in the United States account for over a third of the Ri's YouTube views; more than double the number of views from the United Kingdom, in second place with 16.2% of views.

3.19   Table 3-1 lists the top 10 countries, in terms of their number of views. It also shows these countries' contribution to the various indicators of engagement (including comments, likes and number of subscribers to the Ri's YouTube Channel).

Table 3-1: Countries contributing the most views

	Country	Views	Minutes watched	Comments	Likes	Subscribers
1	United States	33.5%	35.3%	35.6%	31.7%	31.2%
2	United Kingdom	16.2%	15.0%	19.0%	12.0%	13.7%
3	Canada	5.3%	5.7%	6.2%	4.5%	5.4%
4	Germany	4.0%	4.6%	3.6%	4.4%	3.7%
5	Australia	3.8%	4.1%	4.8%	3.3%	4.2%
6	Poland	2.4%	2.1%	0.9%	1.6%	1.2%
7	Netherlands	2.2%	2.4%	2.7%	2.0%	1.9%
8	India	1.9%	1.6%	1.6%	2.0%	2.8%
9	Sweden	1.5%	1.8%	1.8%	1.8%	1.3%
10	Brazil	1.3%	1.09%	1.2%	3.3%	2.2%
	Top 10	**72.0%**	**73.5%**	**77.3%**	**66.6%**	**67.7%**

Note: the metrics should be read as, for example, "35.6% of all comments come from viewers in the US".

4.   Recommendations for next steps and further analysis

4.1   In this section, we set out our main recommendations for further analysis that the Ri may consider pursuing.

Value for money analysis

4.2   Measures of 'value for money' ("VFM") can be constructed by combining: (1) video consumption and viewer engagement metrics; and (2) data on the costs incurred by the Ri to produce videos.

4.3   These statistics could be used to quantify the benefits that donations to the Ri have achieved in the past, to inform the Ri's video production strategy (so as to focus on those types of videos that provide the most value for money), and to support the Ri's further fundraising activities.

4.4   To calculate these metrics, it is first necessary to attribute video production costs, to specific videos.[17] Costs should be attributed carefully to identify incremental video production costs. For example, when an existing talk or event is recorded and placed online, the additional costs associated with the video might be minimal compared to the costs associated with the event (which would be incurred even if it were not filmed). In addition, incremental video costs should be distinguished from general corporate overheads that would again be incurred even if no videos were produced. In this way, it would be possible to calculate how much video consumption (in terms of the number of views, or minutes watched) and viewer engagement (in terms of the number of likes, or comments, or shares) results from an investment of £10,000 (for example), and compare these statistics between videos of different subjects, formats or durations.

4.5   A full analysis of VFM would consider consumption and engagement that has *already* been achieved (as described in earlier sections), as well as the consumption and engagement that existing videos will provide in the *future,* at minimal incremental cost. Statistics that give no credit for future consumption, will likely understate the VFM.

4.6   VFM comparisons should also be made on a like-for-like basis. For example, it would not be fair to compare simple VFM metrics between videos produced in different years, because videos that have been available for longer are likely to have generated greater consumption and engagement than those that have been only recently released. Statistics that give no credit for the age of a video, will likely overstate the VFM of the oldest videos, relative to the newer videos.[18] We have set out our recommendations in a separate report provided to the Ri.

---

[17]   The Ri does not currently have such granularity in its cost data, and so we have not calculated these statistics.

[18]   This is because once the initial production costs have been incurred there is minimal cost associated with keeping the video available online. Comparisons should therefore be made on a consistent basis – for example, by considering consumption and engagement in the first 6 months after a video's publication only.

External benchmarking

4.7   Our analysis focuses on the Ri's own videos. It may be possible to benchmark performance against videos produced by other like-minded organisations. We have discussed our recommendations with the Ri.

Econometric analysis

4.8   Econometric analysis can be used to estimate the causal relationship between one factor (e.g. the subject of a video, which the Ri can decide) and an outcome variable (e.g. the number of comments the video receives per day), by controlling for the other factors that may also determine the outcome (e.g. the duration of the video, or the format of the video).

4.9   For example, we have found that Maths videos have particularly high levels of engagement, as measured by the average number of comments per day. We also know that Maths videos are relatively long, and that long videos tend to receive high levels of engagement too. Which factor is causing these videos to have such high levels of engagement? Is it their subject, or their duration? Or a combination of both? Or another factor (for example, they happened to be of a format that is particularly engaging)?

4.10   Econometric analysis seeks to answer questions such as this. However, the usefulness of such approaches depends on the size and nature of the available dataset. Our review suggests that the Ri's dataset, in its current form, is unlikely to be amenable to econometric analysis. We have set out our recommendations for the Ri to develop its dataset in a separate report.

Selecting which type of engagement to target

4.11   In performing our work, we found that videos that are successful in terms of one metric (e.g. conversation), might be relatively unsuccessful at encouraging other forms of engagement (e.g. sharing).

4.12   In order to target future analysis, it may be helpful for the Ri to consider how much weight it would place on the different types of engagement. For example, would the Ri prefer to produce videos that encourage discussion and debate, even if they receive fewer 'likes'? Or would it prefer videos to collect a large number of likes, and get shared on social media, even if they don't accumulate a significant number of comments?

Analysing the content of comments

4.13   The engagement data we have analysed so far provides little information about the *content* of the engagement. For example, we have not received data on the content of the comments beneath YouTube videos, and so have not considered whether the comments demonstrate genuine interest and engagement with the video, or instead contain off topic discussion or spam.

4.14　One simple indicator could be the extent to which comments are 'in reply' to pre-existing comments or receive 'likes'. Replies are less likely to be ad hoc, off topic, or spam, and are more likely to be part of a debate or discussion. The existing dataset does not contain this information, but it is likely to be collected by YouTube, and so may be available on request. Another simple indicator could be the number of characters included in comments.

4.15　We understand that the Ri is also investigating the possibility of text mining the comments to identify key words and phrases that reflect genuine interest and engagement with the video.