Science and Islam Videos: Creating a Methodology to Find "All" Unique Internet Videos

Vika Gardner and Salman Hameed

Hampshire College

Abstract:

This methodology article explores the process through which we sought to catalogue videos addressing natural science and Islam on the Internet comprehensively. This data was then used to select videos for inclusion in the Center for the Study of Science in Muslim Societies' Science and Islam Video Portal (www.scienceandislamvideos. com), which evaluates the videos based on their representations of science, Islam and history. As a growing body of research is demonstrating, Internet videos provide a window into the lives of both celebrity and ordinary Muslims and their critics worldwide. The article describes the methodological decisions of what to include and exclude from the study, framing them in a discussion of some of the key terms. We then step through the process of searching for videos and cleaning the data, providing flow charts with details. In the last section of the article, we discuss the results and their ramifications on our continuing research.

Introduction

This methodology article explores the process through which we sought to catalogue videos addressing natural science and Islam on the Internet comprehensively. The project, funded by a grant from the John Templeton Foundation, produced a web site with evaluations of a selection of the videos identified in the cataloguing process (www.scienceandislamvideos. com). The project was to compile the widest possible array of video materials, from educational presentations to spoofs. To keep cataloguing manageable in an English-speaking context, videos were limited to those in English or with English subtitles, although of course one may find videos

in most of the major languages of the world as well. This article presents some of the methodological choices and challenges to lay the groundwork for future studies and to facilitate growth in the field.

As a growing body of research is demonstrating (e.g. Petersen 2016), Internet videos provide a window into the lives of both celebrity and ordinary Muslims and their critics worldwide. The number of people who are able to engage with and through these videos has expanded with the advent of cell phones capable of presenting videos. In addition to entering into wider discourses through creating videos, users may re-upload others' videos to different platforms (web sites) or channels, amplifying the message, join conversations by embedding videos in chat discussions on forums like Reddit, or add comments to existing videos. To begin these discussions, however, we must have some conception of how many different videos, presented by which speakers on what topics, are available on the Internet, if only at one particular moment in time.

We begin our discussion by presenting prior work surveying Internet videos addressing Islam, demonstrating the context for our different approach. We then present the narratives of Islam and science within which the video search was conducted, fitting our work into the disciplines of media studies and applied research on Internet search engines. The methodological decisions of what to include and exclude from the study are framed in a discussion of some of the key terms that we found it necessary to define. We then step through the process of searching and cleaning the data, providing flow charts with details. In the last section of the article, we discuss the results and their ramifications for continuing research.

Prior work on Islamic Internet videos

Given the relative newness of collections of videos on the Internet - YouTube was created in 2005 and currently ranks as the second most popular website in both the US and the world (www.alexa.com/topsites, accessed 8 May 2017) - it is unsurprising that nearly all of the literature on Internet videos and Islam is from the past seven years. While YouTube has expand-

ed its market greatly, it is has also been a target of censorship in countries like Turkey and Pakistan (Vara 2014; Nabi 2013). For this reason the study of online videos needs to include other platforms as well. *Platforms* here may include a wide variety of Internet sites. In addition to YouTube, owned by Google, there are other video hosting sites, such as DailyMotion (popular in Europe), Tune.pk (popular in Pakistan and South Asia), and Vimeo (popular for some for not having the sometimes-onerous community standards issues that YouTube does). These sites make no effort to curate the content, unlike older web sites that sometimes included pages of videos in a wide variety of formats focused on particular topics. By *curated* we mean that a user (or group of users) have shaped the presentation of videos in much the same manner a museum or gallery might (Crick 2016:26). We have found that with the rise in popularity of YouTube and similar platforms, these older curated sites are no longer being supported, and in some cases the content has shifted to individual *channels* (or user-defined spaces) that may shape the content through playlists or *tags* (user-supplied keywords) appended to the video.

Table 1 organizes the research we have identified on videos and Islam. It must be noted that every study is based on material from YouTube alone. The group of articles done by van Zoonen, Vis and Mihelj provides some sophisticated analyses of the videos, including analysis of the user networks that the material they obtained from YouTube provided (van Zoonen, et al. 2010; van Zoonen, et al. 2011; Vis, et al. 2011; Hirzalla, et al. 2013). These studies focused on YouTube video responses to the anti-Islamic film, *Fitna* (2008), a sixteen-minute short film by Geert Wilders (Larsson 2013). Their methodology uses software they developed with Michael Thelwall and examines both the videos and the social networks through which users engaged with the videos (Vis, et al. 2010). Their data includes up to 1,413 videos ("unique uploads") and their attached comments, which they segment in various ways depending on the study. Their data was all uploaded in a four-month window in 2008, material which they downloaded and stored. They have attempted to contact individual users, and obtained 40 responses, but analyses drawn from that additional data have not yet been published as far as we know (Vis, et al. 2010:6).

Author(s)	Date	Platform	Sample Size	Analysis of	Uploaders contacted?
Samuel & Rozario	2010	YouTube	3 speakers/ series	Discussion of major speakers/ creators	no
van Zoonen, Vis & Mihelj	2010	YouTube	776 videos uploaded Feb-May 2008	videos	no
Vis, Thelwall, van Zoonen, Mihelj	2010	YouTube	1,413 unique uploads	video	yes, 40 people
van Zoonen, Vis & Mihelj	2011	YouTube	776 videos uploaded Feb-May 2008	networks among users	no
Vis, van Zoonen & Mihelj	2011	YouTube	63 videos uploaded by 46 women	videos	no
Hirschkind	2012	YouTube	9 videos	Exploration of videos and comments	no
Mosenghvlishvili & Jansz	2013	YouTube	150 videos	valence framing in videos	yes, 15 people
Svensson	2013	YouTube	374 videos, with their comments	videos and comments	no
Al-Rawi	2014	YouTube	3 user channels	comments only	no
Wheeler	2014	YouTube	2 vloggers, 6 videos	Discourse of Muslim women in video	no
Al-Rawi	2015a	YouTube	261 videos, 4,153 comments	"frames"	no

Author(s)	Date	Platform	Sample Size	Analysis of	Uploaders contacted?
Al-Rawi	2015b	YouTube	videos and comments from 7 user channels (3,165 videos, 17,708 comments)	comments, based on word/phrase frequency	no
Pedersen	2016	YouTube	8 videos	Discussion of one vlogger	no

Table 1. Summary of prior literature.

Mosemghvlishvili and Jansz use a similar approach and they also interview some of the users who created the videos (Mosemghvdlishvili and Jansz 2013). Their study focuses on video blogs, examining the framing and motivations of the users who created the videos. They select videos by drawing from two YouTube lists: (1) relevance based on their keyword, from which they draw the top fifty, and (2) a random selection of one hundred videos from those uploaded in one particular month. From this data they approached users and interviewed fifteen of them.

Svensson's analysis of reactions to celebrations of the *mawlid* holiday, which celebrates the birthday of the Prophet Muhammad, is based on searches done within YouTube and sampled in two different ways: (1) from the two sets of one thousand returns each, he selects 324 videos randomly, and (2) from a second sampling of 25 videos that included comments from each of the two return sets (Svensson 2013). The bulk of his discussion comes from his analysis of those comments. These sets of one thousand returns are limits imposed by YouTube.

Ahmed K. Al-Rawi's work also focuses on the YouTube platform and uses software created by Michael Thelwall (Al-Rawi 2015a; al-Rawi 2014;

Al-Rawi 2015b; Thelwall 2009). As noted in Table 1, he limits material by focusing on *channels* (user pages) that are of particular interest based on the topic of the article, such as channels created to cover protests in Bahrain in 2011 or the Danish Muhammad cartoons. He analyzes comments using QDA Miner Wordstat software, which generally points out the "most recurrent words and phrases" (Al-Rawi 2015b:31).

Other researchers, rather than starting with YouTube searches for videos, follow ethnographic discussions with respondents toward a limited number of videos that are represented as exemplars. For example, Hirschkind's research on recorded Friday sermons in mosques (*khutbas*) come from his pursuit of the topic based on information from his respondents while in Egypt (Hirschkind 2012). He discusses videos as a part of the cultural spheres in which the respondents move. In this way his research is not an effort to survey material across the Internet comprehensively. Similarly, Samuel and Rozario's study in Bangladesh starts from the material their human research subjects mention and moves towards the discussion of a few of the video creators, like Zakir Naik and Harun Yahya, who are reported to have had a significant impact (Samuel and Rozario 2010). Pedersen (2016) examines one particular video blogger from the United Kingdom in order to examine how she constructs fashion as both a lifestyle and business; her work discusses the same vloggers presented by Wheeler (2014). Wheeler's study uses the material of only two Western-located female video bloggers, examining 30 videos from each, and then selecting three from each for intensive analysis. This highly focused discussion enables her to dig into how the videos engage the discourses available to them.

One may see that the previous research is in some ways broad and in other ways constrained. Our research is unique for two reasons: (1) we have sought to compile a majority of the videos on a diversely discussed topic, addressing both Islam and science, and (2) we are evaluating the content of the videos and providing those evaluations to the public. Thus one of the fundamental goals of the project rests on our ability to find videos relevant to our topic in a broad spectrum of locations efficiently rather than just YouTube, and compile data to enable us to prioritize the more time-intensive evaluations.

In the next section we will discuss in more detail the discourse of science and Islam that are found in the videos we sought.

Discursive orientation of the project

This project's focus on science and Islam in Internet videos may be seen as a part of a wider discourse on the topic. Science narratives are used in conjunction with religion for a variety of purposes (for a broad overview, see Lindberg and Numbers 1986; Brooke 1991; Dixon, et al. 2010). The reception and reaction to biological evolution is, for example, one of the most common topics discussed in relation to science and religion (for example, Miller 2002; Numbers 2006; Ruse 2006). Science narratives are linked with Islam in narratives that run the spectrum from seeking to prove Islam is true scientifically, through seeking to lampoon individual Muslims' ignorance, to attempting to prove Islam is false using science. For example, for at least fifty years there have been active speakers around the world who seek to read science into and out of the Qur'ān, suggesting that "unknowable" information in the text "proves" that it is divine (Bigliardi 2014b; Guessoum 2011), building upon constructions of the Qur'ān that have been in use for more than one hundred years (Jansen 1974; Yazioglu 2013; Elshakry 2013). These narratives frequently involve Western scholars as scientific authorities, who are presented as impartially discussing the scientific facts and typically converting to Islam (Bigliardi 2012). In addition, starting in the 1980s there was also a push for the Islamization of science or knowledge (Stenberg 1996), which suggested that by contextualizing textbooks and information in an Islamic cultural setting, information and science could advance hand-in-hand with Islam.

The critical study of these narratives has been advancing slowly. In part, this is because thesocial, cultural and historical development of Muslims has an effect on the way debates arise around these issues. Surveys of contemporary landscapes provide introductions to some of the major questions (Guessoum 2011; Edis 2007). The various discourses in the multiple languages and locales of Muslims are influenced by cultural elements like politics, the media, and other local concerns (Clark 2014; Edis 2007;

ElShakry 2013). There are works that look at some of the prominent names in the development of the Islamization of knowledge, as well as interviews of some of the prominent Muslim scientists who discuss these topics (Bigliardi 2011; Bigliardi 2012; Bigliardi 2014a; Stenberg 1996). There have been fewer attempts to study Muslim views of publicly contested ideas like evolution in different geographical locations (Hameed 2008; Everhart and Hameed 2013). Because video engages these discourses both locally and transnationally, they provide insight into questions of the consumption and production of these narrative streams. The intersection of science and Islam is a highly focused lens through which to view this medium and it has enabled us to engage it as has rarely been done before.

The project was initiated in June 2014 and the Science and Islam Video Portal (www.scienceandislamvideos.com) was officially launched in October 2015. The final catalogue contains 1,006 unique videos and to date we have provided evaluations on the Portal for over two hundred videos. The cataloguing project included over a dozen undergraduates from Hampshire College, Mount Holyoke College, and the University of Massachusetts, Amherst, working through the 2014-2015 school year.

Because we were uncertain about what we would find, our questions at the outset were broad: How many videos about Islam and science are there on the Internet? What platforms are they presented on? What roles are curated sites playing in the dissemination of videos? Who makes the videos, and what topics do they address? Which speakers and topics are most popular? To what extent are the videos used to attack Islam or Muslims? To what extent are they used to support particular understandings of Islam that may not be widely held among Muslims as a whole? These questions drove some of our basic methodological choices.

Problems with search strategies in video

Text-based search methodologies on the Internet are fairly well studied (Lewandowski 2015; Thelwall 2015b; Thelwall 2015a). However, video is not text, and comprehensive results are not what most search engines are

designed for. A few works on Internet searching specifically address You-Tube or video generally (Bazzell 2014), but as these are general works, there is not a clear discussion of the pitfalls and issues surrounding video searches, which are more dependent on user tagging than text is. Since search algorithms are proprietary, what additional information, if any, they use to supplement user tagging is not publicly available.

Most people are aware of video search by the Google Search Engine, as well as the Advanced Video Search (www.google.com/advanced_video_search), which provides additional search parameters in an easy-to-use format. During July 2015, we ran several practice searches using a variety of search engines and indices to develop best practices for finding videos. The Application Programming Interface (API) for YouTube limited results by using only YouTube data, which we found undesirable. By limiting its search sphere to only videos within YouTube, the API overlooked videos on curated web sites and other hosting sites. We therefore elected to use the Advanced Video Search instead, with results that will be documented in the final section of this article. Other search engines were also tested, but the output format from Google, which was numbered and vertically linear, made it easy to work with as textual output. Additionally, research has shown that Google outperforms other search engines (Lewandowski 2015:1772). After evaluating the test data, we set initial search parameters and created our online video cataloguing tool (using Google Forms).

The test searches also provided us with data through which we confronted assumptions we had been naively making. We realized that we would need to create definitions for the boundaries of search parameters that reached beyond the technical definition of the MP4 file structure that search engines consider a video file. In order to enable others to consider these assumptions for themselves, we include here not only our definitions, but also research considerations that suggest additional possibilities.

The most basic concept that we defined for the project was what we considered a video. Test searches quickly uncovered the wide variety of materials

that fall into this category. Even if we disregarded elements like video quality and display size, as well as issues of length, language and platform, not everything in an MP4 format, currently the most widely used online video format, was what we classified as a video. The MP4 file format is an envelope containing streams of audio and visual material; the two are not necessarily linked in any way. Since our focus was on how people engage with the topic of science and Islam, our definition of a video stated that the file needed to include both moving images and an audio track that corresponds with the moving images, as in a recording of a lecture, or a documentary with voice-over narration. This excludes both podcasts, which have audio (and may have originally been an audio-only format) but unconnected visuals, and moving images with unrelated audio (in our case, most often Qur'ānic recitations), where the primary focus is apparently to maintain the visual attention of the viewer. Our definition also excludes files with moving images and no audio component. This definition excluded a substantial subset of our search returns on Islam and science, as they were text cards, similar to a PowerPoint display, with little or no audio. (See Figure 1 for an example, Rodzi 2011:min1:40.) These text-card quasi-videos were sometimes up to 20 percent of search returns. We did not use them, but they may be a fruitful source of material for future research.

Figure 1. Example of a video without moving images. (Rodzi 2011:min1:40)

We also defined our parameters for *Islam* and *science* after examining the test data. We included any videos presenting someone who calls what they practice "Islam". This included groups like the Ahmadiyya and the Nation of Islam, although the returns for these minorities were small. A Muslim, for the purposes of cataloguing, is anyone who either presents themself as a Muslim (of whatever subgroup), or might be interpreted as a Muslim by others, whether true or not. Video settings, in religious or secular contexts, was irrelevant. Reference to the Qur'ān was also considered "Islam". Materials might include positive self identifications as well as pejorative constructions. These references might be implicit or explicit, verbal or visual.

Science, as it was the core of our research project, focused on natural science, such as biology, physics, astronomy, chemistry, and medicine. We included mathematics and geometry, as well as the history of science in our searches. We did not specifically include occult sciences, which historically have been understood by Muslims as being a part of natural science, such as letterism, numeracy, and astrology, fields which have recently been receiving more attention (Bigliardi 2015; Binbaş 2016; Melvin-Koushki 2012; Melvin-Koushki 2017; Rapoport and Savage-Smith 2014; Ryan 2011), but we did not specifically exclude them either. Some topics were searched specifically because we knew them to be sites of contestation or authority in the discourse; not all scientific disciplines are productive in these discourses, so some, such as chemistry, did not figure prominently in our search returns. We included videos presenting scientists (either explicitly or implicitly) even when they were not discussing science, although these are much more difficult to capture through ordinary searches.

Our project was to find all, or as close to all as we could come, of the *unique* videos on Islam and science. With video, Islam and science defined, we found we needed to address what *unique* meant as well. It is rare to find a discussion in new media literature of what constitutes a duplicate video (that is, one that is not unique), but it has significant repercussions when examining videos as a form of discourse, and in particular when discussing how popular they are. Multiple copies of the "same" video are not identical: they link different social networks, they have different comments,

and they may appear in different contexts. We defined copies - that is, the not unique - in three different ways: duplicates, multiples, and segments. A *duplicate* is an exact copy of whatever we had as an original. *Original* for this stage of the research was merely whatever was found first in the search returns, and may or may not have been posted by the person(s) who created it. By definition, duplicates are of exactly the same length, although they may be of different resolutions (the number of pixels composing the image). We did not attempt to determine if these were the same precise digital file. Notably, Google's search results do not always show the same still image (thumbnail) for duplicates. Duplicates were excluded through each of the data cleaning stages discussed below. *Multiples* we defined as near-duplicates, but with some minor differences. These differences may arise from the way a video was clipped from its longer source, or because of the addition or a removal of front or back *bumpers* (identifying sequences of a logo or branding animation). Multiples were not removed in the earliest stages; both duplicates and multiples have informed additional research we have done on videos (discussed below).

Segments are pieces of longer videos. These differed from clips in that they were not typically divided in order to capture particular content from the longer original, but rather in order to upload a longer original to YouTube or another platform when these platforms maintained restrictions on the length of videos. Users maneuvered around limitations that YouTube originally placed on video length (limitations which no longer exist) by cutting videos longer than ten minutes into lengths the platform considered acceptable. In some instances these segments were the only versions that could be found on the Internet, so the segments were catalogued. After cataloguing was completed the segments were gathered and replaced by complete versions where possible. Gathering segments after cataloguing was often difficult, as those who uploaded segments might give them a variety of titles, or only upload particular segments. Therefore pieces needed to be found, compared and identified as belonging together. It is also true that Google's search does not necessarily return all the segments of a video. When we needed to use the segments rather than a single video, a subsequent, more targeted search was done to find all the available segments.

One genre of videos met all the above criteria, but was excluded. This genre was conversion narratives. Text-based conversion narratives have a long history in the Islamic world (Bulliet 1979; Davis and Rambo 2009; Knight 2013; Kondo 2015). One may find them in a variety of contexts. For example, sometimes at the frontiers of Islam one finds narratives of a ruler or a leader converting and bringing with him his entire community (DeWeese 1994). We uncovered a fairly lively genre of conversion videos, where an individual is presented as converting to Islam *because* of science, or the focus is on a scientist who becomes a Muslim. There are also videos of individuals leaving Islam because of science as well. Since the primary truth claim being made - the fact of the individual's conversion to Islam (or in a few instances, out of it) - would require research into those individuals' personal relationship with the faith, we decided this was beyond the scope of the project. Given the prevalence with which these claims are made - discussions of the 'conversion' of Maurice Bucaille (d. 1998), author of *The Bible, The Quran and Science* (1976) and a prominent speaker on Islam and science, or Neil Armstrong, the American astronaut, are just two of the many examples in this genre - this may be another fruitful research venture for those studying the sociology of religion.

We used Google's Advanced Video Search, which generally returns videos in which a key term are found in the title, the video's description, or the comments. It is often the case, particularly as one advances more deeply into the search returns, that it was at times unclear why a video was included in the results at all. Few of the videos we examined had searchable transcripts, a process that Google has recently automated, so typically our search returns are not based on what is said within the videos. It should also be noted that the now-automated transcripts produce poor results for those speaking with unusual accents in English or using unusual vocabulary, such as specialized terms about Islam or Muslims from Arabic. In addition, because we sought to find material on natural science and Islam, the materials needed to mention Islam *or* Muslims *and* science. However, this mention might include a visual reference to Muslims, that is, showing people in stereotypically (Seiter 2017) Islamic clothing, referring to a verse of the Qur'ān, or other implicit references.

Vika Gardner and Salman Hameed

These were the most difficult videos to find, and it is likely that search failures, material that was never identified, had only such implicit references to science or Islam.

Many discussions of video content about Islam seek to classify them by genre (Al-Rawi 2015b; Welbourne and Grant 2015:6; Mittell 2017). Test coding, however, demonstrated this to be difficult given the blurry boundaries between the genres. Even something as seemingly straight-forward as a *vlog* (video blog) is difficult to identify when it crosses out of the classic form (talking into a computer's video camera) into someone speaking outdoors, or when a professional lecturer mimics the intimacy of an amateur vlogger but with high production values. *Lectures*, a classic academic genre, were difficult to differentiate when recorded from a single, fixed camera. This eliminates visual or auditory cues about audiences; the size and composition of the audience are important elements for a viewer's interpretation of the authority of the speaker. Because the videos we identified were not only reacting to material, but creating original presentations of it, the spectrum of what we encountered seems to be broader than that of other studies. Therefore, we collected data on the number of speakers, audience, visual setting, and so forth. This allows us to examine these constructs independently, without forcing the videos into genre classifications that might not fit their content or our research well.

The cataloguing process, done online by trained undergraduates through a Google Form, collected over forty data elements. Some were basic facts about the video, such as title, length and upload date. We collected data on the site of the presentation, the speaker(s), languages and accents, how the presentation discussed science and Islam, and a long descriptive field in which the cataloguer included what happened on the screen and what was said. Few videos included every data element, as some locations were impossible to determine, for instance, or a clip from a longer video would not identify the speaker. Ideally we would have spent the time to ensure inter- and intra-rater consistency for the data we collected. Because of the project's time constraints, we did not test for coding consistency.

It should be noted that information not collected in this project includes anything related to audiences for the videos. Generally audience data is private, accessible only by a video's uploader or curator. Although viewership networks might be implied by the names of the commenters, there are various methodological issues involved here as well. In short, using publicly viewable videos as a source for discussing viewers is not a sound methodology at this time.

Given these inclusions and exclusions, we now move on to a more detailed discussion of the search process itself.

Searching for Videos

We conducted test searches to confirm appropriate keywords in our videos of interest. Generally, Islam was represented with key terms *Islam, Muslim, Moslem, Quran, Koran*. Science key terms included *science, astronomy, biology, medicine, mathematics, geometry* as well as more specific terms such as *black hole, Big Bang, expanding universe, evolution, evolutionist, Darwin, Darwinism, aliens, embryology, miracles, mountains, iron*. Very specific search terms, like *embryology*, were used because we know of long-standing debates about these topics in science and Islam discourses. Some terms, like *evolutionist*, came up in early searches and were used to ensure that we compiled exhaustive search returns. Searches were done combining key terms for Islam with one or two key science terms for accuracy.

Searching for materials on natural science was relatively easy, but because of the large number of search terms, generally we searched for materials using the broadest terms (*science*) first, and ran subsequent searches with increasingly specific terms (*evolution, Darwinism, embryology*). The number of keywords we used appears to be substantially more than most of the prior research, which typically used three words or less. In order to ensure new videos were not missed, we re-ran searches monthly from October 2014 to April 2015. We did not download the videos we identified, which has meant occasionally searching again for videos we discover have been removed from the Internet.

Google's Advanced Video Search works best when it is pushed to be as focused as possible. Google has pre-defined video length categories: short (0-3:59 mins), medium (4-19:59 mins) and long (20+ mins). We adopted these length groupings as well. We typically searched for only one video length at a time, and limited the number of web sites by searching separately for videos found on YouTube and outside YouTube. To make sure everything possible was found, we included a final search that excluded all the major hosting sites (YouTube, Vimeo, DailyMotion, Tune.pk). Thus each key term needed to be searched nine times: once for each video length and platform combination. The final search rarely produced videos that had not been collected in the earlier searches, but was especially useful to ensure we did not miss materials on web sites, including blogs and commercial sites, such as those focused on *1001 Inventions* (al-Hassani 2012). Although we specified English as the target language in the search, in some cases the videos returned were not in English and did not have subtitles and were individually excluded. (See discussion on data cleaning below.)

In the end we know that despite all these efforts, we may have missed additional videos for various reasons. One example of a video we might have missed is worth mentioning. Talk Islam, an Australian group, has a video "The Meaning of Life", which includes both images of videos by Muslim creationists, as well as other discussions of science (Talk Islam 2013; Gardner and Hameed, forthcoming; n.d.). The original version under that title was not found in our searches. One user, who re-uploaded the video onto their own YouTube page renamed it "Islam and science" (Ghori 2014). That version, which had only 21 views as of August 2015, was catalogued. Later research uncovered the original Talk Islam copy, which had nearly two million views, making it one of the more popular videos employing science and Islam narratives. This was a small victory: we were fortunate that someone renamed it. Subsequent searches for "the meaning of life" and Islam have turned up many more similar videos by other creators, which discuss science fairly regularly. These are examples of videos that discuss science and Islam but are not tagged as such or being commented on as such, and are therefore unlikely to be found in direct searches.

A simple Google search on *Islam and science videos* today returns close to 6.25 million hits (8 May 2017 without quotes, or 7,350 with quotes). In our search process, we rarely needed to go beyond the first five or six hundred returns before the returns became irrelevant to the search terms. This is consistent with the way search engines work (Lewandowski 2015:1764), as it is likely that algorithms move further and further from the core of the most relevant returns into additional material from related terms on the assumption that users continuing to page through returns have not yet found what they want. This is why tightly focused searches are important, despite the possibility of missing material.
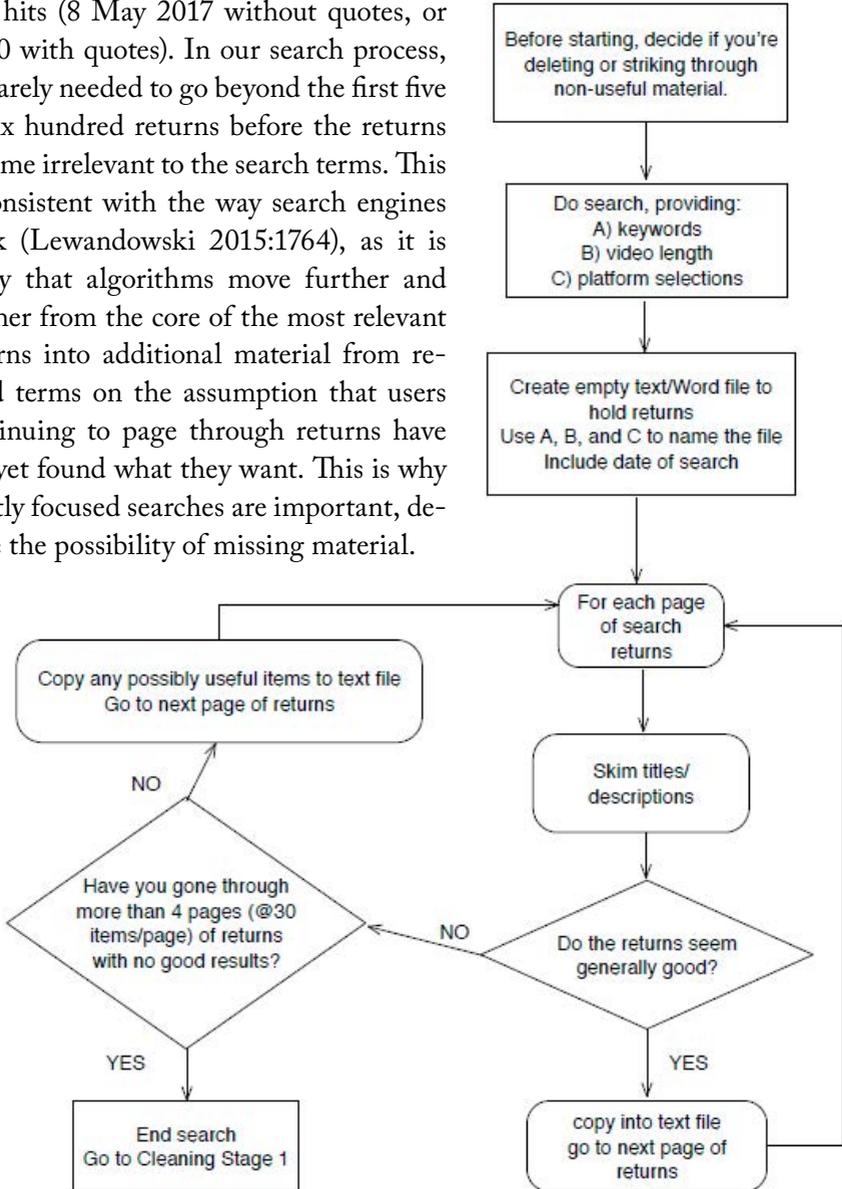

Figure 2. Video Search Flowchart

After acquiring the search returns and placing them in text files, the initial search results required several levels of cleaning. A typical search return goes through the following steps (see Figures 2 and 3): The searcher first checks that the video is active, that the video is in English, and that it is about science and Islam. This quick view through the video also removes conversion narratives and text-card videos. These steps usually take less than 2-3 minutes per video. Despite this, after cataloguing we found that some videos had no science or no Islam/Muslims. This is partly because Google's algorithm includes *religion* when a user searches for *Islam*, resulting in returns that address other faith systems' engagement with science. Ensuring that a video includes (or not) even slight or implicit mentions of Islam in its discussion of religion and science requires watching the entire video. Because of this the cataloguers were instructed to catalogue every video they were given, whether it appeared to be about Islam or not, and the elimination of non-Islam videos was done at a later stage.
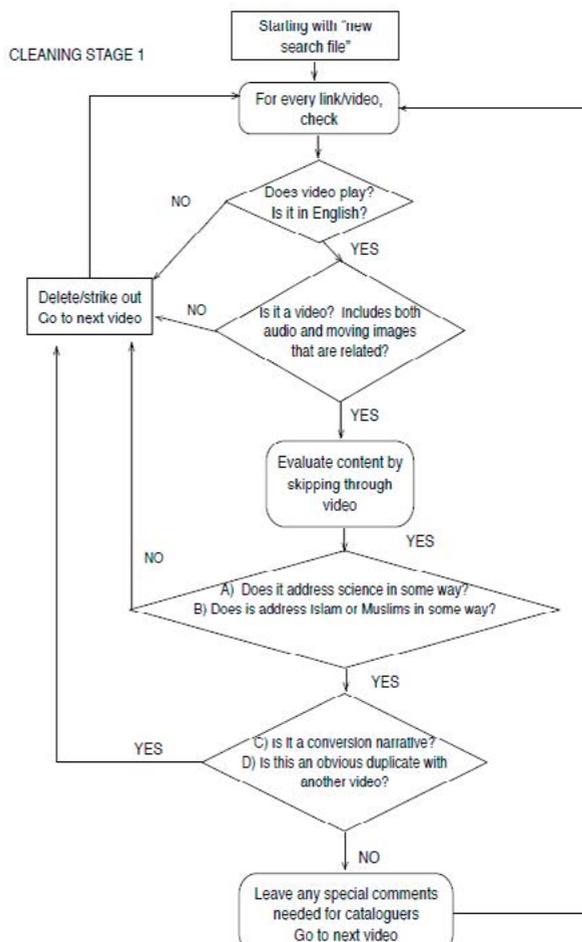

Figure 3. Cleaning Stage 1 Flowchart

A second cleaning pass was done to compare each new search return against videos that had already been identified (See Figure 3). For example, a video found with the key word *science* might also be found with the key word *biology*. We anticipated clips of videos, and believed that clips of the same material but with different lengths ("multiples", see definitions above) would help us to learn what materials individuals consider important from longer videos. As the searching progressed, we also noticed that various groups or individuals add or remove bumpers, making videos longer or shorter. We elected to catalogue these in order to help identify who engages in the process and how, although the material has not yet been analyzed.

Once all the videos were catalogued, a final level of cleaning of the data was done. (See Figure 4.) This eliminated all the videos coded as being without Islam or without science, after checking through the video descriptions and verifying the coding. Duplicates, unnecessary multiples, and unneeded segments were removed. Careful sorting of the individual videos was done in order to ensure that only one copy (without any multiples or duplicates) of each video was included in the final catalogue.
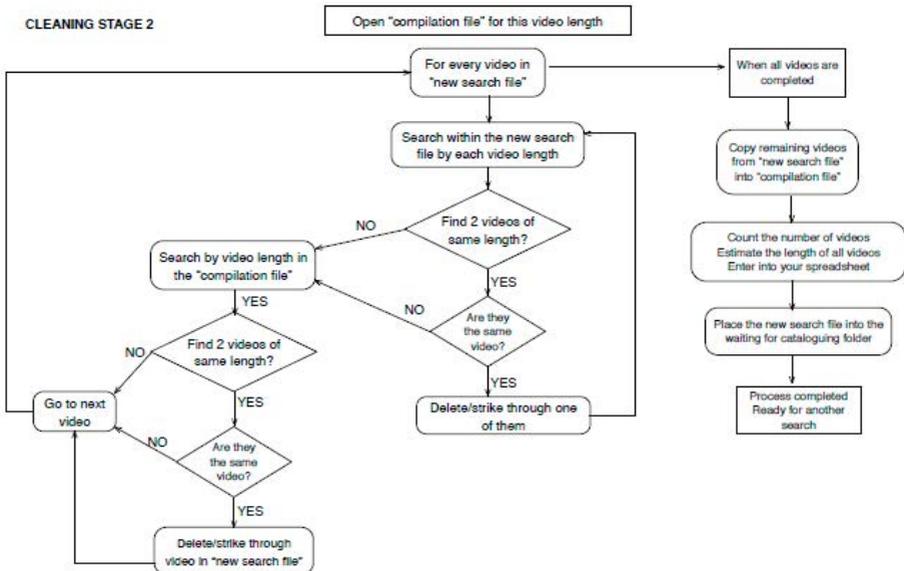

Figure 4. Cleaning Stage 2 Flowchart

The final results people often find surprising: we catalogued 2,282 videos (an estimated 790 hours in length), which resulted in just over a thousand (1,006) videos in the final, cleaned database from which we draw videos for evaluation. These were culled from tens of thousands of Google search results. Although these numbers are not unmanageable, a more narrowly drawn topic would ensure more time to find materials that are not well tagged. It is impossible to know how many of the inadequately tagged or titled videos were not included in the returns.

In order to give a better sense of the range of what we did find, the preliminary results are presented here to provide an overview of this media. More analysis is needed in order to discuss the findings in detail. The number of videos is quite manageable for data analysis. For instance, the primary scientific field (each video may have up to two defined scientific fields in the catalogue) show that the topics of Evolution/Creationism, History of Science, and Medicine combined account for just over half (548/1006) of the videos (Figure 5). Evolution alone has 226 videos, the most numerous scientific field, demonstrating its importance in the discourse by and about Muslims. The variety of scientific fields demonstrates that there is a wealth of material to be analyzed for how it enters discourses about Islam and Muslims.
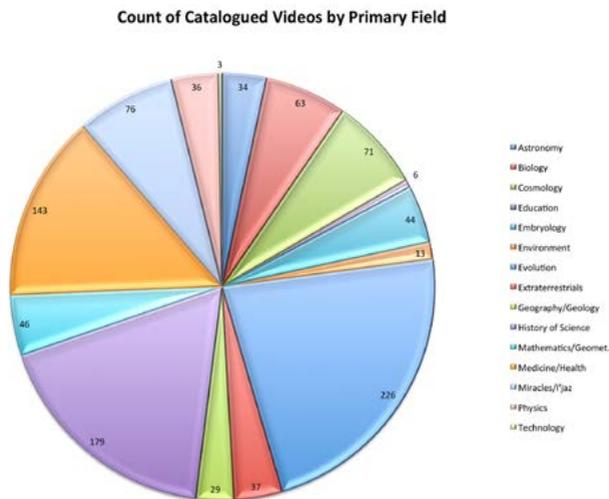

Count of Catalogued Videos by Primary Field

Legend: Astronomy, Biology, Cosmology, Education, Embryology, Environment, Evolution, Extraterrestrials, Geography/Geology, History of Science, Mathematics/Geomet., Medicine/Health, Miracles/I'jaz, Physics, Technology

Figure 5. Count of Catalogued Videos by Primary Field

The number of views a video gathers is often used as a simplistic mea-
sure of popularity (Clarke 2017), a construct we have questioned with later
research (Gardner and Hameed, forthcoming; n.d.). Figure 6 shows the
number of videos with particular ranges of views organized by the various
scientific fields. Evolution/Creationism shows broad bands of videos with
many views, although the History of Science perhaps has the highest per-
centage of videos with more than five hundred thousand views, which were
a rarity in our research. Especially notable are videos on extraterrestrial
life, which has a high proportion of substantial view counts, especially giv-
en the relative scarcity of the topic. As a caution about this data, however,
it should be noted that we did not search each individual title to ensure we
included in our database only the most-viewed copy. Although the assump-
tion is often made that Google Search will return the most-viewed copy,
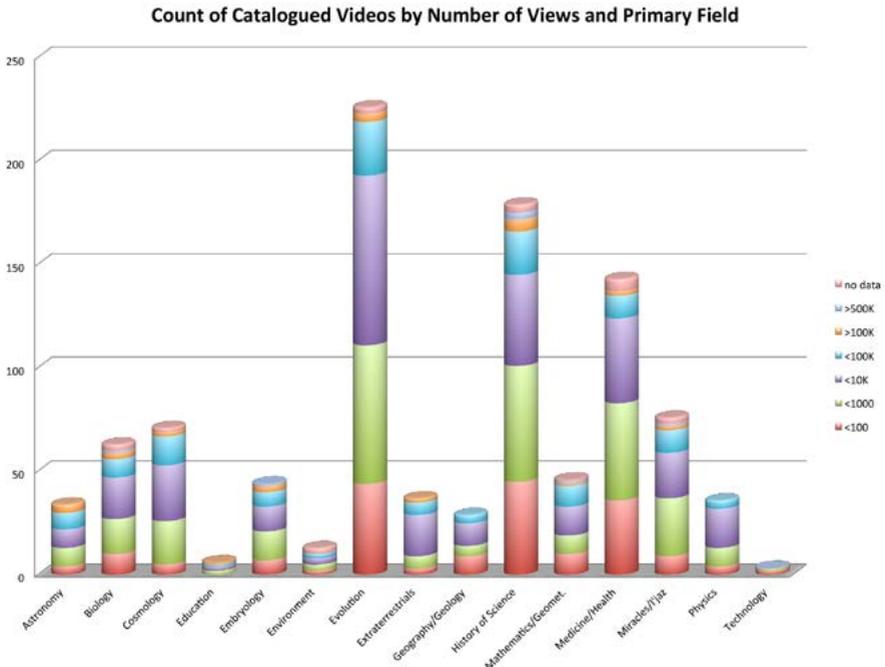this is not accurate.

**Figure 6. Count of Catalogued Videos by Number of Views and Primary Field
(1-99; 100-999; 1000-9,999; 10,000-99,999; 100,000-499,999; more than 500,000,
and those with no data)**

Figure 7 shows the number of videos with particular settings, which in other research might have been used to describe video genre. The substantial number of videos with seated speakers on a stage or soundstage, an atypical format for lectures and vlogs, is likely the result of the frequent use of this format on PeaceTV and shows such as *The Deen Show* and *Islam and Science* (with Zaghloul El Naggar). The category of computer-generated visuals has shown exponential growth over the past few years, and is increasingly being used in spoof or critique videos, which may either spoof ideas that are suggested to be held by Muslims or ideas that are suggested to be atheist.
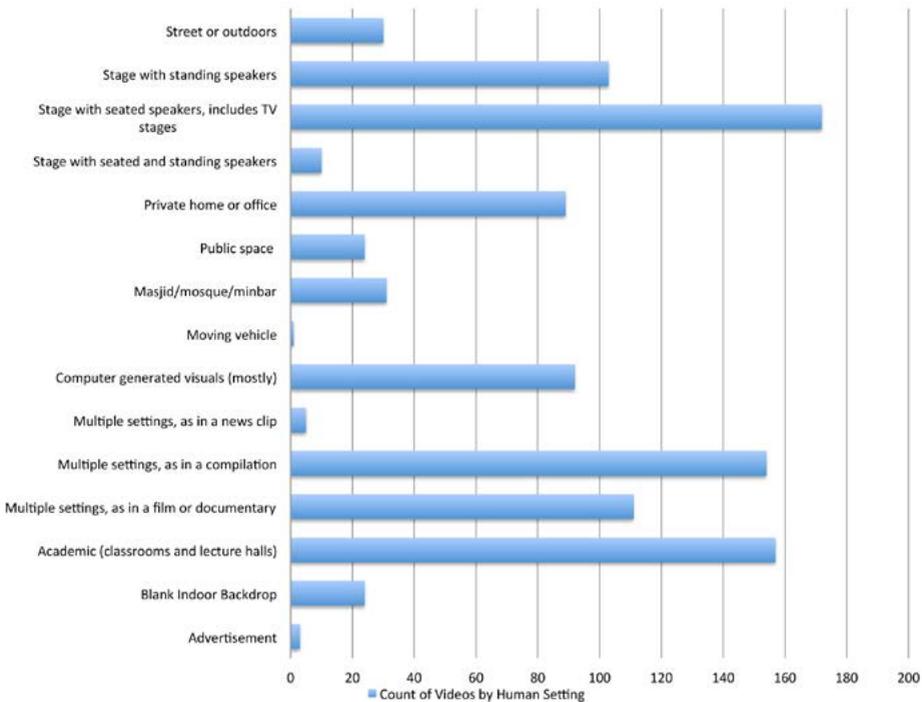

Figure 7. Count of Catalogued Videos by Setting of Video

Table 2 provides a brief summary of the number videos by particular speakers, a category that is complicated to code since not all the video clips identify the speakers, requiring the visual recognition of nearly five hundred individuals who appear in the videos. Seventy-three percent of all videos present at least one named speaker, and of those, 25 percent (184) have additional speakers. The 27 percent without named speakers is high, and may change as additional work is done to identify speakers. Forty-seven percent-nearly half of all the videos with named speakers-are videos from those who have five or more videos in the catalogue. The fifteen speakers with the most unique videos account for nearly a third of all the videos with named speakers, suggesting that these fifteen, which includes several academics, are successfully utilizing videos as a means to distribute their ideas. Foremost among these is Zakir Naik (b. 1965), a Muslim preacher from Mumbai, India, who accounts for 62 videos, far more than any other individual. Naik and his presentations have been a continuing focus for our further research (Gardner, et al. in press; n.d.). Adnan Oktar (b. 1956), a Turkish preacher who is widely known for his creationist stance (Bigliar-di 2014a; Samuel and Rozario 2010; Solberg 2013; Hameed 2008; Hameed 2015; Riexinger 2002), also known as Harun Yahya, appears in 12 videos. This did not include videos that claim to be based on his ideas, since he does not personally appear in them, although these are also numerous.

Speaker Summaries	Total count	Percentage of videos with speakers	Percentage of all videos
Videos with speakers who have >4 videos	348	47.3%	34.5%
Videos of top 15 speakers	236	32.1%	23.4%
Videos with at least one named speaker	736		72.9%
Videos with >1 named speaker	184	25.0%	18.2%
Total number of named speakers	481		
Number of speakers with >4 videos	34		

Table 2. Summary of Counts of Videos by Speaker.

Conclusions

We will conclude this article by discussing some of the findings this methodology suggests about continuing research using Internet videos. Our research demonstrates that collecting and studying videos through ordinary search returns is possible and fruitful. The numbers of returns listed by Google's search engine are wildly unrealistic and should be ignored. Human-powered data collection, which included far more data than could have been scraped from YouTube alone, took only one academic year, and has provided a wealth of data for future research and analysis.
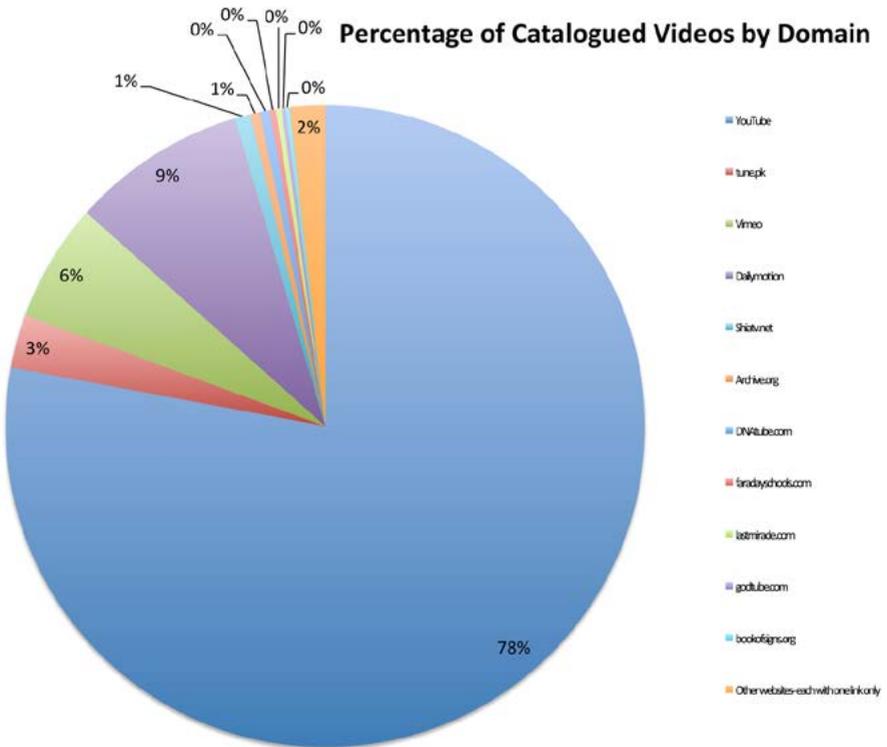

Figure 8. Percentage of Catalogued Videos by Domain

We also note the material outside YouTube, which suggests all the prior research focusing on YouTube alone may be missing important data. The majority of the work on Internet videos is being done using YouTube data. This provides abundant results and ready access to some types of data. It also must be said that if one is interested broadly in the discourse on a topic through video, YouTube's automated removal of both videos and users is pushing some users to other platforms, like Vimeo, which does not have the same sorts of problems, as we have seen when searching for additional copies of videos that have been removed. This automated removal, based on user-reported violations of "community standards", has been documented by our own work, as we revisit our data; YouTube posts a specific message when videos have been removed for this violation. This has been a particular problem for users who create critique or spoof videos, as their work more often receives flags for inappropriate content or use of copyrighted material. Some material on science and Islam is rarely found because its creators rigorously enforce their copyright, such as the *1001 Inventions* organization. We also have been able to find videos from other platforms, like archive. com, which have helped us to go back further in time than YouTube, especially to find complete videos and lectures. Although we ran all the YouTube searches before the non-YouTube searches, nonetheless only 78 percent of our final videos (786/1006) were from YouTube (Figure 8). Research done on YouTube alone is missing that remaining 22 percent, material that does not duplicate the material on YouTube. Particularly when we discuss videos about Islam, the several countries that regularly censor YouTube, perhaps especially Pakistan, with its large English-speaking population, need to be included.

It also seems clear to us from a methodological standpoint that we need to consider the variations of videos -multiples, duplicates and segments -as a part of our data rather than something to screen out. In many instances variations were created by users who altered the original video in some way, such as by adding their own branding in a front or back bumper. This branding information, or the lack of it, is a more important distinction than has been considered in the prior research, which often notes that duplicates were eliminated without detailing how duplicates were defined. When user

networks are studied, these duplicates become still more important, as they extend the reach of the networks through which a particular video communicates, as we will discuss in upcoming articles on Zakir Naik's presentation of evolution. In addition, if we are to speak of the popularity of a video, we must engage with the contact information (views, likes, etc) of *every* copy, since even just the number of copies (and languages into which it is translated) is an indicator of how desirable the video was. As we have demonstrated in our examination of "The Meaning of Life" (Talk Islam 2013; Gardner and Hameed, forthcoming; n.d.), even though a video may start out in English, it does not necessarily stay in that linguistic sphere.

The outcome of this research, in addition to the creation of the Science and Islam Video Portal, has been an increased awareness of the richness of this discourse beyond the celebrity figures, and how readily the discussions move beyond national and linguistic boundaries. For researchers wanting to study the contemporary voices of both ordinary and celebrity Muslims and those debating about Muslims, Internet videos provide a large and relatively untapped source.

Acknowledgements

References

Bazzell, Michael 2014 Open Source Intelligence Techniques: Resources for Searching and Analyzing Online Information. Lexington, KY: CCI Publishing.

Bigliardi, Stefano 2011 Snakes from Staves? Science, Scriptures, and the Supernatural in Maurice Bucaille. Zygon: Journal of Religion and Science 46(4):793-805.

Bigliardi, Stefano 2012 The Strange Case of Dr. Bucaille: Notes for a Re-examination. The Muslim World 102:248-263.

Bigliardi, Stefano 2014a Islam and the Quest for Modern Science: Conversations with Adnan Oktar, Mehdi Golshani, Mohammed Basil Altaie, Zaghloul El-Naggar, Bruno Guiderdoni and Nidhal Guessoum. Istanbul: Swedish Research Institute in Istanbul.

Bigliardi, Stefano 2014b What We Talk About When We Talk About *I'jāz*. Social Epistemology Review and Reply Collective 4(1):38-45.

Bigliardi, Stefano 2015 I'jāz, Conspiracy Theories, and Ufology - Some Suggestions with a Touch of Latour. Social Epistemology Review and Reply Collective 4(11):1-7.

Binbaş, İlker Evrim 2016 Intellectual Networks in Timurid Iran: Sharaf al-Dīn 'Alī Yazdī and the Islamicate Republic of Letters. Cambridge: Cambridge University Press.

Brooke, John Hedley 1991 Science and Religion: Some Historical Perspectives. Cambridge / New York: Cambridge University Press.

Bulliet, Richard W. 1979 Conversion to Islam and the Emergence of a Muslim Society in Iran. *In* Conversion to Islam. N. Levtzion, ed. Pp. 30-51. New York: Holmes & Meier.

Clarke, John 2017 Popular. *In* Keywords for Media Studies. L. Ouellette and J. Gray, eds. Pp. 143-144. New York: New York University Press.

Clark, Kelly James 2014 Religion and the Sciences of Origins: Historical and Contemporary Discussions. New York / Hampshire, UK: Palgrave Macmillan.

Crick, Matthew 2016 Power, Surveillance, and Culture in YouTube's Digital Sphere. Hershey, PA: Information Science Reference / IGI Global.

Davis, Patricia M., and Lewis R. Rambo 2009 Conversion Dreams in Christianity and Islam. *In* Dreaming in Christianity and Islam: Culture, Conflict, and Creativity. K. Bulkeley, K. Adams, and P.M. Davis, eds. Pp. 175-187. New Brunswick, NJ / London: Rutgers University Press.

DeWeese, Devin 1994 Islamization and Native Religion in the Golden Horde: Baba Tükles and Conversion to Islam in Historical and Epic Tradition. University Park, PA: The Penn-

sylvania State University Press.

Dixon, Thomas, Geoffrey Cantor, and Stephen Pumfrey, eds. 2010 Science and Religion: New Historical Perspectives. Cambridge / New York: Cambridge University Press.

Edis, Taner 2007 An Illusion of Harmony: Science and Religion in Islam. Amherst, NY: Prometheus Books.

ElShakry, Marwa 2013 Reading Darwin in Arabic, 1860-1950. Chicago / London: University of Chicago Press.

Everhart, Donald, and Salman Hameed 2013 Muslims and evolution: A study of Pakistani physicians in the United States. Evolution: Education and Outreach 6(2):1-8, https://link.springer.com/article/10.1186/1936-6434-6-2 (accessed 9 May 2017).

Gardner, Vika, and Salman Hameed Forthcoming; N.d. Creating Meaning through Science: Talk Islam's "The Meaning of Life" Video and its Transnational Appeal.

Gardner, Vika, E. Carolina Mayes, and Salman Hameed In press; N.d. Preaching Science and Islam: Dr. Zakir Naik and the Discourses of Science and Islam in Internet Videos. Die Welt des Islams.

Ghori, Ammar 2014 Science and Islam. YouTube, MP4 video file, https://youtu.be/yt-xhNCPiDYU (accessed 8 May 2017).

Guessoum, Nidhal 2011 Islam's Quantum Question: I.B. Tauris.

Hameed, Salman 2008 Bracing for Islamic Creationism. Science 322:1637-1638.

Hameed, Salman 2015 Making sense of Islamic creationism in Europe. Public Understanding of Science 24(4):388-399.

al-Hassani, Salim T. S. 2012 1001 Inventions: The Enduring Legacy of Muslim Civilization Washington, DC: National Geographic / Foundation for Science, Technology and Civilization, Inc.

Hills, Matt 2017 Audience. *In* Keywords for Media Studies. L. Ouellette and J. Gray, eds. Pp. 17-21. New York: New York University Press.

Hirschkind, Charles 2012 Experiments in Devotion Online: The YouTube Khu ba. International Journal of Middle East Studies 44(1):5-21.

Hirzalla, Fadi, Liesbet van Zoonen, and Floris Müller 2013 How Funny Can Islam Controversies Be? Comedians Defending Their Faiths on YouTube. Television & New Media 14(1):46-61.

Jansen, J.J.G. 1974 The Interpretation of the Koran in Modern Egypt. Leiden: E.J. Brill.

Knight, Michael Muhammad 2013 Converts and Conversion. *In* The Cambridge Companion to American Islam. J. Hammer and O. Safi, eds. Pp. 83-97. Cambridge Companions to Religion. New York: Cambridge University Press.

Kondo, Yohei 2015 Ibādī Discussions on Conversion and Commitment. Muslim World 105(2):224-235.

Larsson, Göran 2013 Geert Wilders and the Anti-Muslim Movie Fitna. *In* Muslims and the New Information and Communication Technologies: Notes from an Emerging and Infinite Field. T. Hoffman and G. Larsson, eds. Pp. 147-162. Muslims in Global Societies Series. Dordrecht / Heidelberg / New York / London: Springer.

Lewandowski, Dirk 2015 Evaluating the Retrieval Effectiveness of Web Search Engines Using a Representative Query Sample. Journal of the Association for Information Science and Technology 66(9):1763-1775.

Lindberg, David C., and Ronald L. Numbers, eds. 1986 God and Nature: Historical Essays on the Encounters between Christianity and Science. Berkeley / Los Angeles: University of California Press.

Melvin-Koushki, Matthew 2012 The Quest for a Universal Science: The Occult Philosophy of Sāʾin al-Dīn Turka Isfahānī and Intellectual Millenarianism in Early Timurid Iran, Yale University.

Melvin-Koushki, Matthew 2017 Powers of One: The Mathematicalization of the Occult Sciences in the High Persianate Tradition. Intellectual History of the Islamicate World 5(1-2):127-199.

Miller, Kenneth R. 2002 Finding Darwin's God: A Scientist's Search for Common Ground between God and Evolution. New York: HarperCollins / Perennial Publishers.

Mittell, Jason 2017 Genre. *In* Keywords for Media Studies. L. Ouellette and J. Gray, eds. Pp. 81-83. New York: New York University Press.

Mosemghvdlishvili, Lela, and Jeroen Jansz 2013 Framing and praising Allah on YouTube: Exploring user-created videos about Islam and the motivations for producing them. New Media & Society 15(4):482-500.

Nabi, Zubair 2013 The Anatomy of Web Censorship in Pakistan. 3rd USENIX Workshop on Free and Open Communications on the Internet (FOCI '13), Washington, DC, 2013, pp. 1-7.

Numbers, Ronald L. 2006 The Creationists: From Scientific Creationism to Intelligent Design. Cambridge, MA: Harvard University Press.

Pedersen, Kristin 2016 Performing Piety and Perfection: The Affective Labor of Hijabi Fashion Videos. CyberOrient 10(1), http://www.cyberorient.net/article.do?articleId=9759 (accessed 8 May 2017).

Rapoport, Yossef, and Emilie Savage-Smith, eds. 2014 An Eleventh-Century Egyptian Guide to the Universe: The Book of Curiosities. Leiden: Brill.

Al-Rawi, Ahmed 2015a Online Reactions to the Muhammad Cartoons: YouTube and the Virtual Ummah. Journal for the Scientific Study of Religion 54(2):261-276.

Al-Rawi, Ahmed 2014 Cyber warriors in the Middle East: The case of the Syrian Electronic Army. Public Relations Review 40(3):420-428.

Al-Rawi, Ahmed 2015b Sectarianism and the Arab Spring: Framing the popular protests in

Bahrain. Global Media and Communication 11(1):25-42.

Riexinger, Martin 2002 The Islamic creationism of Harun Yahya. Newsletter of the International Institute for the Study of Islam in the Modern World 11:5.

Rodzi, Mohd Izzat Mohammad 2011 Fasting Islam and Science. YouTube, MP4 video file, https://youtu.be/u09I5Hh0Izg (accessed 8 May 2017).

Ruse, Michael 2006 Darwinism and Its Discontents. Cambridge / New York: Cambridge University Press.

Ryan, Michael A. 2011 A Kingdom of Stargazers: Astrology and Authority in the Late Medieval Crown of Aragon. Ithaca, NY: Cornell University Press.

Samuel, Geoffrey, and Santi Rozario 2010 Contesting science for Islam: The media as a source of revisionist knowledge in the lives of young Bangladeshis. Contemporary South Asia 18(4):427-441.

Seiter, Ellen 2017 Stereotype. *In* Keywords for Media Studies. L. Ouellette and J. Gray, eds. Pp. 184-185. New York: New York University Press.

Solberg, Anne Ross 2013 The Mahdi Wears Armani: An Analysis of the Harun Yahya Enterprise. Stockholm: Södertörns högskola.

Stenberg, Leif 1996 The Islamization of Science: Four Muslim Positions Developing an Islamic Modernity. Lund: Religionshistoriska avdelningen, Lunds universitet.

Svensson, Jonas 2013 ITZ BIDAH BRO!!!!! GT ME?? -- YouTube Mawlid and Voices of Praise and Blame. *In* Muslims and the New Information and Communication Technologies: Notes from an Emerging and Infinite Field. T. Hoffman and G. Larsson, eds. Pp. 89-111. Muslims in Global Societies Series. Dordrecht / Heidelberg / New York / London: Springer.

Talk Islam [Kamal Saleh] 2013 The Meaning of Life | Muslim Spoken Word | HD. YouTube, MP4 video file, https://youtu.be/7d16CpWp-ok (accessed 8 May 2017).

Thelwall, Michael 2009 Introduction to Webometrics: Quantitative Web Research for the Social Sciences: Morgan & Claypool.

Thelwall, Michael 2015a Search Engine Queries for Webometrics, Vol. 2015.

Thelwall, Michael 2015b Webometrics: Search Engine, Web Crawler and Web Link Analysis Information, Vol. 2015.

van Zoonen, Liesbet, Farida Vis, and Sabina Mihelj 2010 Performing citizenship on YouTube: activism, satire and online debate around the anti-Islam video *Fitna*. Critical Discourse Studies 7(4):249-262.

van Zoonen, Liesbet, Farida Vis, and Sabina Mihelj 2011 YouTube interactions between agonism, antagonism and dialogue: Video responses to the anti-Islam film *Fitna*. New Media & Society 13(8):1283-1300.

Vara, Vauhini 2014 The World Cracks Down on the Internet. *In* New Yorker. New York: New Yorker.

Vis, Farida, Michael Thelwall, Liesbet van Zoonen, and Sabina Mihelj 2010 Challenging dominant representations of Islam online: Innovative methods for studying YouTube. *In* Innovative Methods in the Study of Religion Conference. London.

Vis, Farida, Liesbet van Zoonen, and Sabina Mihelj 2011 Women responding to the anti-Islam film Fitna: voices and acts of citizenship on YouTube. Feminist Review 97:110-129.

Welbourne, Dustin J., and Will J. Grant 2015 Science communication on YouTube: Factors that affect channel and video popularity. Public Understanding of Science:1-14.

Wheeler, Kayla Renee 2014 Remixing images of Islam: The creation of new Muslim women subjectivities on Youtube. Heidelberg Journal of Religions on the Internet 6:144-163.

Yazicioglu, Isra 2013 Understanding the Qur'anic Miracle Stories in the Modern Age. University Park, PA: Pennsylvania State University Press.

Vika Gardner and Salman Hameed

Tohami, Ahmed. 2009. Youth Policy in Egypt. EuroMed Youth 111 Programme. http://www.salto-youth.net/downloads/4-17-1864/02-EuroMedJeunesse-Etude_EGYPT-090325.pdf.

United Nations Development Programme. 2010. The Egypt Human Development Report. http://www.undp.org.eg/Portals/0/NHDR percent202010 percent20english.pdf.

Notes

[1] They considered that Egypt, Tunisia, and prospective Palestine comprise the "bully" states of the MENA, while Algeria, Iraq, Libya, Sudan, Syria, and Yemen are the bunker states.

CyberOrient

Online Journal of the Virtual Middle East