YouTube Children's Videos: Development of a Genre under Algorithm

John C. Paolillo
Indiana University
paolillo@indiana.edu

Brian P. Harper
Indiana University
bpharper@indiana.edu

Clara Boothby
Indiana University
crboothb@indiana.edu

David Axelrod
Indiana University
daaxelro@iu.edu

Abstract

YouTube children's video has been claimed to have a preponderance of violent, disturbing or otherwise inappropriate content. To assess this claim, we conduct a content analysis of a sample of children's videos published between January 2016 and December 2018. Our analysis reveals an evolving ecosystem involving a variety of production modes and messages which nonetheless bears the heavy imprint of the algorithm-centered commercial incentives of marketing to children and attracting YouTube advertising. Hence, while content formerly causing public concern appears to be effectively policed at this juncture, algorithmic incentives do appear to distort children's content in potentially unhealthy ways.

1. Introduction

Videos for children old enough to be entertained by moving images on a screen comprise a substantial component of YouTube's catalog of videos. In terms of numbers of videos, time spent viewing them and their economic significance to the platform, this category of video is important enough that YouTube has created a curated version of its service "YouTube Kids" [1]. As with other forms of video – such as state propaganda, terrorism recruitment, reality performance of illegal acts, political protest, etc. – children's videos on YouTube have sometimes been the subject of public outcry. A recurring concern is the potential inappropriateness of YouTube content for children: videos are claimed to encourage immoral, dangerous or bad behavior, to be misleading, to wrongly employ intellectual property, or to otherwise be strange in ways that may lead to developmental harm, especially for younger children.

Concern with children's videos on YouTube heightened considerably in 2017, as news sources [2, 3, 4] first began reporting on videos intended for consumption by children that nevertheless contained disturbingly violent or sexual content. In some cases, the presence of these attributes may have been deliberate rather than an oversight. This content included both live-action and animated videos, often including characters from popular culture. The frequent appearance of Elsa from the Disney film *Frozen* led to some to refer to the 2017 media attention as "Elsagate" [2]. In 2019, Fisher [5] reported that YouTube's algorithms, which are intended to reward engaging content with advertising revenue, effectively fed videos of children to pedophiles who leave sexualizing comments; no effective controls exist to stop such unwanted behavior. YouTube has responded to these concerns as it usually does by adjusting its algorithms for recommending and monetizing and by censure of channels producing offending content. Consequently, YouTube's curation of children's content should also be understood within the context of its manner of managing other problematic content such as white-supremacist videos [6].

Another type of children's video that has drawn attention is one that appears to be algorithmically generated. Such videos characteristically employ a restricted type of script (such as "learn colors", "wrong heads", etc.) in which a set of repeated highly predictable actions occurs with regular precision (e.g., marbles or characters rolling down a track) in a predictable order (red, followed by green, followed by blue; hitting a series of non-matching objects before landing on a matching object; etc.). The narrative culminates in a final, predictable climactic event (e.g. characters engaging in a cheer at the end of the course), and often features characters representing intellectual property for which permission may not have been granted: Spiderman, *Frozen*'s Anna and Elsa, the Hulk, Mickey Mouse, Batman and Superman, and others commonly make appearances. Among these videos, minimally different videos can be found (e.g. learn colors with Spiderman, the Hulk, or cockroach characters), suggesting that some form of A-B testing may be being performed using YouTube as its platform and the public as research subjects. These observations raise a broad spectrum of legal and ethical issues.

These two widely discussed examples of problematic video types represent several common anxieties about children's content on YouTube, but there is less concrete understanding of the precise scale and nature of the issue. As such, there is risk that quieter but more pervasive forms of problematic content are hidden in the wake of more sensational or egregious but otherwise uncommon content. While this study takes an investigation of the prevalence of problematic content as its springboard, we also categorize and investigate the characteristics of content that has been amplified the by algorithms of YouTube to have become broadly available to children. To some extent, the initial research question

addressing whether harmful content is prominent enough on YouTube to merit the media uproar of 2017 developed into a broader exploration of the entanglement of commercial and marketing interests with the video styles most likely to be promoted by YouTube's algorithms.

In spite of the public awareness of the issues around YouTube children's videos, to date there has been no systematic study of this genre, examining the reasons for its particular characteristics, or assessing the degree of the ethical concerns they raise. In part, this is because YouTube is large and any study of its contents is therefore challenging. One needs a suitable frame for sampling and tracking large numbers of videos, as well as an analysis scheme that encompasses the relevant concerns. Moreover, YouTube is international in scope, and one has to consider questions of appropriateness for children within a context of potentially relevant cultural frames. Beyond this, one must recognize that YouTube is a highly dynamic environment, with platform-imposed changes prompting user adaptations over relatively short spans of time. Finally, the subject matter itself is varied and evolving: children's videos involve animation, live actors, game engine simulations other modes of video production; they employ widely-known and lesser-known intellectual properties; they initiate certain fads and respond to others. Consequently, investigating children's videos on YouTube is challenging.

In this paper, we seek to address these challenges by characterizing a substantial body of children's videos published to YouTube between January 2016 and December 2018. These videos were coded for a variety of production and content features, and subsequently submitted to cluster analysis. The clusters were then explored in terms of relevant exemplars, to develop interpretations regarding pressures on the children's video genre over the observed time period.

2. Background

YouTube has proved to be a difficult site to effectively characterize in the literature. It is often described as a place of participatory culture, but discussions of professional and user-generated content [7] are complicated by the fluidity enabled by an ever-changing algorithm [8]. Most scholarship referencing YouTube is oriented towards specific genres like terrorism recruitment videos [9], platform features like multichannel networks [10], or social network effects [11]. Broader studies of YouTube as a whole exist [12,13], but lack the fine detail required to investigate specific incidents.

Studies of children's content on YouTube often focus on YouTube potential to aid or hinder educational

development [14]. A 2011 study [15] directly addressed the question of what children are watching on YouTube, but did so as a survey of children's habits rather than examining the content of the videos directly. Moreover, YouTube has changed so much since 2011 that it is necessary to update such work. Craig and Cunningham [16] produced a detailed discussion of YouTube toy unboxing videos from a media regulation and economics perspective. Nicole and Nansen [17] produced a content analysis of the genre features of toy unboxing, focusing on the conflicting values of professionalization and authenticity.

In examining genres on YouTube, this paper uses the conceptualization of genre by Miller [18] and applied by other scholars [19, 20] to new media platforms, such as blogs and YouTube. Genres "typified rhetorical action based in recurrent situations" [18], which may be analyzed by investigating their formal features and the social function that they fulfill. Previous studies of YouTube genres have tried to link YouTube with TV [21], an approach which works well in some cases, but fails to account for the development of new genres made possible by the affordances of the platform and the algorithmic systems that shape genre development on the platform. Still, work on educational television [22, 23] provides neccessary context to YouTube videos intentionally seek to emulate educational television, and others at least attempt to employ educational framing for their videos.

In addition, we draw on scholarship by Levinovitz [24] for his theoretical work on toy play. Levinovitz highlights some of the common assumptions on play between these disciplines while applying speech act theory to propose his own definition of toy play. Of 1080 total videos examined in this study, 450 feature children or adults unboxing and/or playing with toys. Such videos devoted to different aspects of toy play thus require contextualization and exploration into how videos of toy play affect the dynamics of contemporary toy play.

3. Observation Method

From a large-scale scrape of YouTube videos conducted up to 2019, we used a series of carefully selected search terms to identify a pool of children's videos appearing from 2016 to 2018. We developed coding tags through iteratively assigning free-form tags at the video level. These tags were then used during the formal coding of a randomly selected sample of 1800 videos, distributed evenly across time from January 1st, 2016 to December 31st, 2018.

3.1. Sampling

YouTube has a large number of videos that could be categorized as children's videos; those that YouTube curates and identifies as such lack any publicly-available formal characterization. Instead, loose associations employed by YouTube's proprietary recommendation algorithm appear to drive their categorization. This makes acquiring an appropriate sample of children's videos a difficult and critical first step for conducting the systematic study of children's videos on YouTube. One must anticipate a large number of potentially relevant creators and videos, and these may not have a clear intention for their content or may not have large numbers of followers or viewers. We aimed for our sample to reflect both the scope of children's videos, while representing more popular videos that would have a more appreciable effect on the culture of YouTube.

With these considerations in mind, we employed a previously-captured, broad sample of YouTube scraped from July 2015 to March 2017 [13], which was updated in entirety in March 2018 and again specifically for the purpose of this study in January 2019. For the second update, we identified channels in the database that had at least one video with over 1 million views at some point since January 2016 and whose title or description matched a specific set of search terms (Table 1). As this is an unusually large number of views, channels with one or more such videos can be considered to enjoy some success, which in turn generally implies deliberate action on the creator's part. This requirement also avoids collecting data from less popular creators whose content is less engaging for (child) YouTube viewers. The channels identified were queried using the YouTube Public Data API for their complete catalogue of videos, up to January 1, 2019. These were later manually culled to remove channels that had no child-related content.

Our search terms identified through a combination of browsing children's videos, including ones similar to those identified as suspect in Bindle [2], and our own observations of the genre. Children's nursery rhymes, Minecraft "Let's Play" videos toy and unboxing videos were salient kinds of children's content we noted. Vlad and Nikita, a highly popular live-action show featuring Russian actors and having a fraught YouTube history, frequently appeared to have been pirated, appearing with titles labeled *Johny Gombal* (apparently after the song "Johnny, Johnny, Yes Papa") or the Lithuanian *vaikai vaikams* ("children for children"). Learn colors was a similarly salient theme, and many suspect animated videos feature Marvel and Disney characters, saliently listing their names in the titles. Finally, a set of

the suspect videos concerns appropriations of Disney's Mickey Mouse characters. The search strategy was intended to allow us to find at least some of the reportedly problematic content, while giving us some sense of how it was situated within YouTube channels; as will be shown below, the resulting sample was very different from what was expected when constructing the query. Despite this focus on terms linked to potentially problematic content, the collection of all the videos from identified channels provided a far broader range of children's content than the limited search terms might suggest.

**Table 1.** Search terms for collecting children's videos

Terms	Explanation
children & nursery & rhyme, minecraft & kids, unboxing & toys	Common themes
vaikai & vaikams, johnny & gombal, masha & bear	Title words (Lithuanian and English) from Russian-origin videos
learn & colors, kinder & farben, aprender & colores, uczyć & kolory	Common title words for suspect videos: English, German, Spanish and Polish
spiderman & elsa, marvel & mcqueen, peppa & pig, thomas & tank & engine	Intellectual property terms found in some suspect videos
micky, Микки, Маус, ميكي & ماوس, मिकी & माउस	Variants of Disney's "Mickey Mouse" found in suspect videos

Our query gave us a pool of 75,788 children's videos published between Jan. 1, 2016 and Dec. 31, 2018 with a mean view count greater than ten thousand. As this sample is large for manual coding, and since the breadth of the videos strains categorization, we opted for a randomly-selected subsample of videos that would allow coders time to give appropriate scrutiny to each video. This resulted in a coding sample of 1,800 videos or 50 videos per month, covering a sample period of 36 months.

3.2. Coding

Formal coding was conducted on the remaining 30 videos per month (1080 total) in the full sample. These were divided up among observers so that all videos would be coded by at least two observers, any two observers from the four would be paired with each other for some set of videos, and a substantial subset would coded by all four in common. This assignment was arranged using a month-wise coin-flipping process, resulting in the two months January and July being coded by all four observers. All four observers held weekly meetings to clarify definitions of tags, anomalous videos, and other questions arising during coding. To develop the coding categories, we used a training set of 20 videos per month that were randomly selected from our coding sample. In this exploratory period, observers iteratively assigned free-form tags describing the content of each video in the same set and discussed in conference observations on each video until a consensus was reached about the major characteristics to be noted. While this method entailed a growing number of codes from iteration to iteration requiring disambiguation and in some cases culling, it also allowed us to adapt our coding system to fit the varied and complex range of children's videos we found.

Our ultimate set of codes was divided into five main categories: activity type, production method, language, actors, and subject matter. These categories for the coding were separated from each other using color in the coding interface to help ensure that each category of code was applied when appropriate (see Figure 1). Multiple codes from each category were permitted if a video had any elements appropriate to them. This allowed for hybrid categories of videos to be coded (e.g. those assembled from both animated and live-action elements). When necessary, no codes from a given category might also be assigned, e.g., as sometimes happened with the actor and subject categories if there were no particular actors or subjects present. The interface also allowed for the entry of new, free-form codes to describe previously unencountered elements just in case they might need to be elevated to a regular code, or for later retrieval of specific videos. Observers otherwise made every effort to assign only pre-established codes during the formal coding stage. Some content still defied consistent coding due to the sheer variety present in our sample and the different ways such content could be interpreted by an individual viewer.
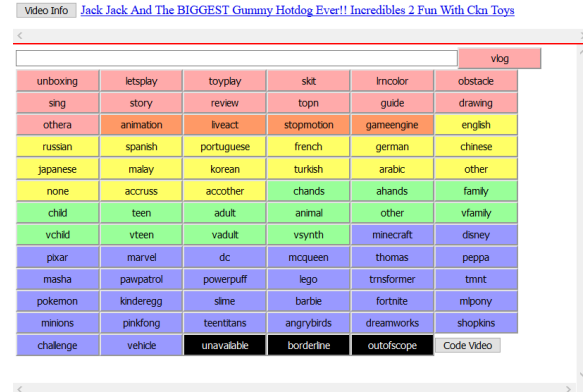
Figure 1. Video coding interface used for this research project, showing the codes used and their categories: activity type (pink), production method (orange), language (yellow), actors (green) and subjects (blue).

4. Analysis Method

Analysis of the coding data must accomplish two things: it must correct for discrepancies across observers, and it must express the relationships found among the codes in a way that is interpretable in terms of genre variation. To accomplish this, we employed Principal Components Analysis of an incidence matrix of observers, codes and videos, followed by two separate cluster analyses on the principal components for codes and videos respectively, using Gaussian Mixture Models. This allows us to address both analytical needs in a principled and unified way.

Our coding resulted in 175 distinct codes applied to one video or another, and hence a total of 700 different potential observer-code combinations to consider. Because the size of the code set and the number of observers, traditional inter-coder reliability tools such as Krippendorff's alpha [25, 26] are unusable. We opt instead for a two-stage approach of culling and cluster analysis based on the correlations of code use across observers. Hence, our first task was to cull codes from the complete set that were unlikely to behave consistently across observers, such as any code used only once, or used by a single observer.

For this, we constructed the 175 code × 4 observer matrix and subjected it to PCA[1], yielding a one-component solution accounting for 92.2% of the variance, meaning the four observers shared some overall level of use of some set of codes, which turned out to be those with a PC 1 score less than zero. The remaining codes were replaced with a single code 'xotherx', so they could still be included in the analysis in case they provided any useful information, i.e. if two or more observers agreed that some additional code was needed for some set of videos but did not happen to agree on the label. Only 277 code instances out of the total 10,303 observations were replaced, leaving a set of 66 codes used by at least two observers each.

To address inter-observer coding variation and identify code co-occurrence in the videos, we constructed the full 1080 video × 259 observer-code incidence matrix. Of these, 180 represented videos coded in common by all four observers; this sub-matrix was submitted to (q-mode) PCA, to provide the basis space for distribution of videos and codes. A four-component solution accounting for 19.0% of the variation in the sub-matrix was chosen; the remaining PCs from 5 onward showed spherical structure and were not interpretively useful [27]. We then computed factor scores for the full set of 1080 videos, treating the 900 videos coded by only two observers as having missing data from the other two coders. The rotation matrix of the PCA and the scores matrices were used for cluster analysis of the observer-code combinations and videos, respectively.

We performed a cluster analysis for gaussian mixture models using the R [28] package mclust [29]. This approach was deemed better than more familiar methods [30] because it permits a variety of model specifications useful for multivariate normal mixtures and offers a principled way to assess the fit and complexity of alternative solutions within a Maximum Likelihood framework [31, 32, 33]. For the observer-code cluster analysis, we permitted solutions with anywhere from 1 to 20 components, with the full suite of cluster shape-volume-orientation (complexity) options available in mclust [29]. The optimal cluster analysis for observer-code combinations had six components with equal volume and variable shape and orientation. Observer-code clusters were interpreted by inspecting their memberships; when all four observers' use of a given code was assigned to the same cluster, we considered it to reliably characterize the cluster for interpretation. Since clusters are multivariate gaussians, some consideration must be given to the variance of the cluster in terms of how it should be interpreted.

For videos, we first identified an optimal clustering based on the PC scores of the in-common set of videos. This resulted in a five-cluster equal-volume, equal-shape, variable orientation solution. We then sought a cluster solution for the whole set of 1080 videos that matched these characteristics. This was done so as to avoid over-fitting the video cluster model to groups of videos coded by only two observers, and expecting that the centers, shape and orientation of the new clusters would match those of the in-common subset. We then listed each of the videos belonging to a cluster in order of its distance from the cluster center as active links in a web page, permitting their re-inspection for interpretation, treating those closest to the centers as "archetypes." Comparison across the two cluster analyses permitted matching codes with corresponding videos. Finally, variation across the different video clusters and their relative overlap were interpreted in terms of potential genre hybridization and evolution.

5. Cluster Analyses

The observer-code clusters resulted in the code clusters represented in Table 2; Figure 2 presents PC plots on PCs 1 & 3 (left panels) and 2 & 4 (right panels) of the cluster classifications for both codes (top panels) and videos (bottom panels). This arrangement allows the different clusters to appear more separated and readable (95% confidence interval ellipses enclose the centers of each cluster in the plots). In addition, where the observer-code and video clusters correspond closely, they are drawn in the same colors, and distinct colors are used where they do not. In this way, by scanning between top and bottom panels, one can match code clusters that contribute to the clusters of videos.

6. Interpretation

Interpreting the results requires understanding both the video and coder sets of clusters. These map roughly to each other, but show a few important differences. A characterization of each follows, using V to refer to video clusters and C to refer to coder clusters.

---

[1] Using a ln(x+1) transform, to scale large values and address zero values, and with column-wise centering and scaling.
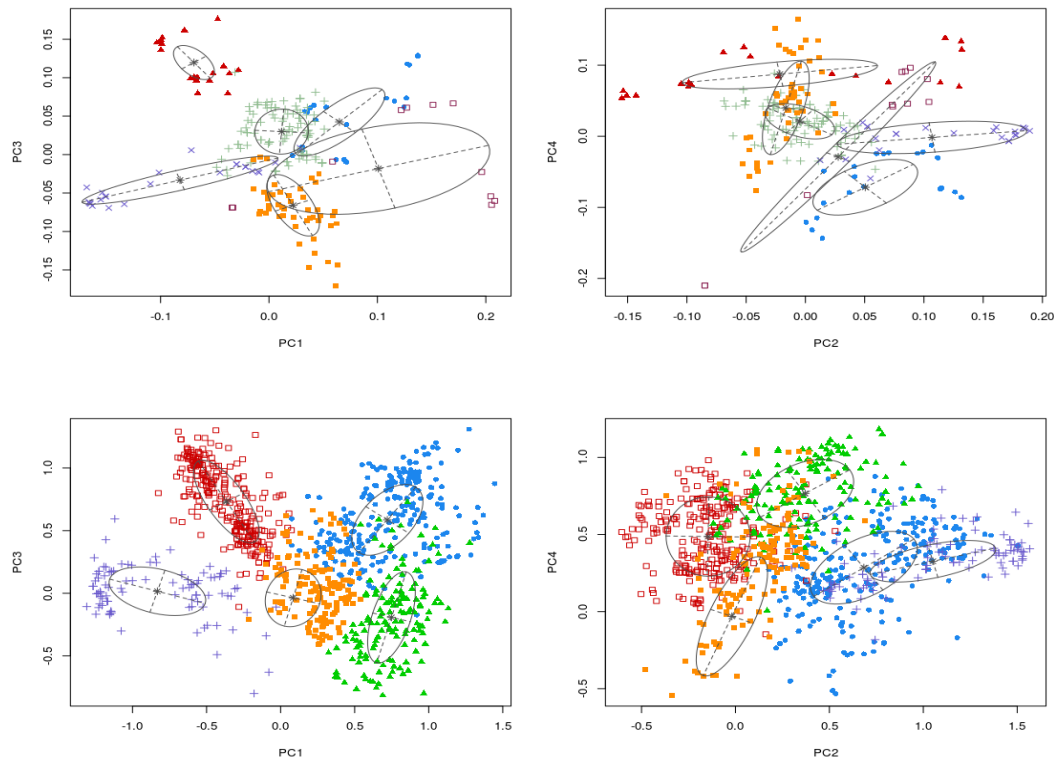
Figure 2. Cluster graphs. Video clusters on bottom. Coder clusters on top. PCs 1 & 3 left. PCs 2 & 4 right.

6.1 Video Clusters

The videos form five highly distinct clusters; these are plotted on PCs 1 & 3 (left panels of Figure 2) and PCs 2 & 4 (right panels), as these present the clusters with the maximum clarity, though all combinations are broadly interpretable.

V1 (blue) is comprised of live-action videos focusing specifically on the toys as the primary object. Shots are predominantly close-ups of toys, where human hands are usually visible, but no other features of the content creator are seen. Multiple styles are present within this cluster, but hands playing with toys are ubiquitous and definitional. Adult hands are more frequent in this cluster than that of children's hands. This cluster includes adults or children playing with toys with or without a connecting narrative, unboxing toys from their packaging, opening Kinderegg-like objects containing toys, and drawing videos focusing on content like fire trucks or cartoon characters. As the toys are the primary focus of these videos, branded products are regularly seen. Aside from generic toy vehicles, the toys in these videos are typically easily identifiable by brand, and in

the case of the unboxing videos, the packaging makes brand identification highly visible.

Language-elements are optional in this cluster, as the emphasis on the toys themselves allows for a purely visual style of video, though it might contain various noises and grunts from the creator. The title and video description will contain one or more languages, but lacking a definite language might broaden the appeal of the video to larger markets. The absence of language fits well into the toy play, toy unboxing, and drawing videos in this cluster, as the close focus shifts emphasis away from the content creator. Content creators in this cluster have a reduced need to project a clear personality in their videos, and are either performing model work by displaying toys or are acting as the toys themselves rather than present a consistent personality to the channel.

V2 (red) consists of animated video content, often of shorts and songs, or alternatively a compilation of many different shorts and songs, on occasions exceeding an hour in length. Alongside shorts and songs, hybridity is even more common, as characters often shift from spoken narrative to songs and vice versa. The content of these videos covers a spectrum between education and entertainment, with comedic entertainment being far

more common, though videos attempting to teach numbers, objects, and social norms also occur.

Both 2D and 3D animation are quite common in **V2**, along with some stop-motion videos, which are otherwise more similar to the videos in **V1** than the rest of the cluster, and display a similar prevalence of branding. The 2D and 3D video shorts and songs in **V2** rarely display obvious brands unless those brands are directly affiliated with the channel (e.g., the official Barbie channel).

V2 also contains formulaic animated features like the "learn color" and "missing head" pieces as mentioned in the introduction section. Such videos often have brands attached to them like Marvel and Disney characters, though this is more common of 2D videos. 3D videos usually use unbranded and likely free animation software assets. Despite these types of videos partially motivating this study and being subjectively prominent in the coding process, they did not appear to profoundly influence the clustering, and are in the periphery of **V2**.

V3 ("green") is very similar to **V1**, speaking to the overlap between these in the video cluster plots of Figure 2. However, instead of disembodied hands holding the central object of toys, children and adults are the central focus, albeit still playing with toys. This is the defining difference between **V1** and **V3**. Toy play and unboxing are still common features in this cluster, with branded toys as regularly mentioned objects. Not only are children more common as actors in these videos, the personalities of the adults and children playing with and opening these toys are far more central to their appeal, and the content creators in the videos speak directly to the audience more frequently than in **V1**.

Another difference from **V1** is the presence of videos focused directly on actors performing a skit rather than strict focus on toys. Branded toys and other objects are still common within these skits, but the narrative between the actors is given greater priority. Minor animation flourishes and editing sound effects are common in these skits, and some of them are quite elaborate and heavily scripted, and are rarely something that children are likely to have much involvement in planning beyond their acting roles. In addition to these more formal skits, more improvisational content also exists in this cluster, though this blurs into **V5**. This content is otherwise similar to the live-action skits, but is filmed in a style closer to a vlog with a phone camera and purports to show more authentic and unscripted content such as family trips, celebrations, and shopping, though some activities are more complex, such as watching some children delight in their home being filled with sponge cubes and jumping around in them.

V4 (lilac) is quite separate from the rest of the clusters and is composed of Let's Play videos (videos showing someone playing through a video game with audio commentary and optional video feed of their face). Though these videos are intended for younger audiences than the average Let's Play on YouTube, aside from avoiding mature language and focusing on more child-friendly games, they are otherwise very similar to conventional Let's Plays. Both the content creator's personality and the game in question have considerable impact on the style and content of the video. The most common video game shown in this cluster is Minecraft, with great variety in the type of videos. Some are of more conventional playthroughs, while others introduce their own narrative elements to the Let's Play. The ability for the content creator to express their own style of video seems important to the choices of games that ended up in this group, and so creative and building oriented games like Minecraft and Roblox are common.

In addition to more freeform games, there were also a fair number of mobile game playthroughs in this cluster. These games are ones explicitly intended for children, and approach the styles of animation seen in **V2**. They lack the content creator personality elements of the Minecraft Let's Plays, often playing through a game in a fashion that simply replicates an animation, albeit an animation with mobile UI elements. **V5** (orange) is the category of last resort in this analysis. The unavailable videos and those that were deemed not applicable to our analysis ended up here. Additionally, videos that did not easily fit into the other categories ended up here, so some videos that blur animation and live-action are in **V5**. Additionally, **V5** blurs with **V3** where the topic of vlogs is concerned. Videos in **V3** might have been filmed like a vlog, but a formal, sitting vlog where a content creator simply talks to the camera and there are no other skit-like elements appear in **V5**.

6.2 Code Clusters

The code clusters generally replicate the results of the video clusters and will not be discussed in great detail except where they differ. **V1**≈**C1**, **V2**≈**C3**, **V4**≈**C6**, and **V5**~**C5**. The colors between the coder and video clusters have been kept consistent for clarity where possible. **V3** ("green") does not have a clear analogue, so the code cluster uses different colors: violet (**C2**) and light green (**C4**) to represent these noncorresponding clusters.

Table 2. Code cluster code lists

C1 (blue)	C4 (light green)	C5 (orange)
unboxing	learn colors	vlog
drawing	stop motion	other activity
adult hands	French	adult actor
Peppa Pig	German	teen actor
Shopkins	Spanish	animal
	Turkish	child actor
C2 (violet)	child voice	family actors
	Barbie	Teen Titans
toy play	Disney	borderline
live action	Lego	out-of-scope
unavailable	Marvel	
	Masha	
C3 (red)	McQueen	**C6 (lilac)**
	Minions	
sing	My Little Pony	game engine
story	Thomas	let's play
animation	vehicle	Minecraft
English	xotherx	Pokemon
adult voice		

Fortunately, the logic of these cluster correspondences remains comprehensible, especially cross-referencing Table 2 with Figure 2. **C1** (blue) shows the same focus on hands drawing and unboxing toys, with emphasis on certain brands that frequently were unboxed. **C2** (violet) contains the toy play and live-action terms which in the video clusters were split between **V1** and **V3**, but here are broken out into their own cluster. **C3** (red) contains the animated songs and stories comparable to **V2**.

C4 (light green) contains many brand codes and has no corresponding video cluster. As brands were present in both **V1** and **V3**, there is logic to the code cluster breaking them into a distinct cluster which overlaps with the others. **C5** (orange) is a similar category of last resort to **V5**, though it could also be described as a blend with **V3**, as it contains the terms for video with fully visible children, adults, and families, like in the live-action skits of **V3**. **C6** (lilac) is the Let's Play cluster, and like in the video cluster, it remains separate from the rest of the clusters.

7. Discussion

The cluster analysis above shows the children's videos to be grouped in several broad genres, and characterizes some of the major genre categories. However, there is also hybridization between genres, and the grouping of certain video types within clusters helps to indicate areas of overlap. Toy play and unboxing seem to have a similar orientation: vicarious playing with toys connected with unboxing new toys.

The entertainment central to both types of toy oriented videos is premised on vicarious experience of escapist fantasies. The key difference found in our clusters is of two different fantasies: the fantasy of possessing a toy and the fantasy of living within an idealized family. Both of these fantasies have antecedents in children's television content, but are amplified on YouTube. The videos focusing centrally on toys are not dissimilar to toy commercials, although of far greater duration and lacking obvious commercial artifice. Similarly, the fantasy family videos are somewhat similar to live-action television shows of children, but with far greater apparent authenticity. These fantasy families also convey a fantasy of wealth, privilege, and parental love from the abundant toys and attention provided to the children in these videos, despite the regular labor that must go into their production.

While previous work on children's unboxing videos has argued that "children's mimetic and imitative play intensifies when brought into contact with virtual spaces" [17], this is largely only possible for toy fantasies and not for parental fantasies. While most toys featured in unboxing videos can be purchased with relative ease, changing the dynamics of one's home and family to align with a fantasized version viewed on YouTube is much more difficult or impossible. Additionally, while it has been suggested that children's videos feature loving families in order to alleviate concerns of child exploitation [17], the image of the loving family is also for the benefit of the child-viewer. This is especially true if parents play a more limited role in curating their child's time on YouTube.

Since adults often act as gatekeepers for children's access to the devices required to view these videos [15], it should be noted that several video features may serve as markers for adults to indicate content appropriate for their children. In addition to authenticity [17], one such feature is the abundance of characters from shows and movies familiar to adults. Another feature, perhaps even more important, is pseudo-educational components to these videos. Such content is mostly learn color videos (or other subjects such as numbers or animals) and sing-along videos centered on moral or safety lessons. In both cases, the educational aspects are more likely to grab the attention of a child's guardian rather than the intended viewer and serve to legitimate the video in the eyes of the guardian.

Unboxing and toy play videos, like other genres, follow a standardized script. Both genres are defined by the presence of toy objects, and as mentioned above, the unboxing genre also anticipates later toy play. Given this online-offline relationship, it is reasonable to inquire as to how the two genres can affect later play offline. As previous scholarship has noted, toy play is fundamentally different from game-play, and the main differentiation is the greater agency afforded in establishing the rules of play [24]. Utilizing this interpretive framework, toy play content on YouTube often amounts to scripts that convey "correct" ways to engage with a given toy. As a result, the ability for a child to gain the benefits of play is severely infringed upon by these videos' own interpretation of the toy. This reduced agency means that the mode of play is shifted increasingly towards game-play or imitative play [24].

Excessively scripted toy play also sets goals within these videos but that are not inherent to these items. By providing specific goals for play, beyond anything an advertisement could convey, toy play videos function increasingly like Let's Play videos featuring video game play with its in-game goals. As with Let's Play videos, by submitting oneself to another person's script and goals, the child is on the one hand able to participate in a communal act of play, and at the same time forfeiting the opportunity to act as a demiurge in relation to the play-object. Furthermore, if conventional gender roles are exhibited in tightly scripted videos, these forms of play may likewise constrains the possibilities of play for children to adhere to gender constrained and heteronormative roles. However, these questions require additional investigation.

By shifting the source of variability and the nature of the affected toy play increasingly towards game-play or imitative play, the room for children's imagination is decreased in one of the premier arenas for developmental experimentation and creative thinking. The developmental implications of this trend need to be examined further, but the toy play genre may well be an instance where YouTube's profit motive has produced negative externalities that have escaped public scrutiny to date.

While these videos follow established scripts within the genre, there are also certain forms of variety and blurring with other genres. While some literature on the dynamics of mimesis among children's YouTube videos has argued that competition for viewership requires differentiation [17], the innovation we have observed is largely limited to different ways of merging existing scripts. The live-action family outing videos blur with the styles of adult vlogs (which were treated as out-of-scope for this study), a trait most obvious in the code clusters, but one still present in the video clusters. This hybridity serves multiple possible functions: providing variety to both content creator and the audience, incorporating other genre traits that might be recognized and appreciated as audience's tastes change, and chasing the desires of YouTube's algorithms. In particular, the potentially transformative role of recommendation algorithms calls into question the notion that content creators must strive to stridently differentiate themselves as in other markets. While differentiation is indeed important in other markets, YouTube's recommendation algorithm, by promoting content similar to previously viewed content, likely changes the dynamics of attracting viewership on their platform. While the viewer may ultimately still gravitate towards novel content, YouTube's recommendation algorithm serves as a gatekeeper before parental or viewer preferences can kick in.

8. Conclusions

Although children's content faces many of the same concerns regarding the misuse of copyrighted properties as the rest of YouTube, our cluster analysis shows that the use of branded materials appears most frequently in contexts (toy play and unboxing clusters) where showing branded toys is not generally regarded as a copyright violation that would elicit concern. As for other types of content that concern the public, we did not find very many truly upsetting videos in our sample, in spite of attempting to direct sampling toward Elsagate-styled content.

Qualitatively, we have noticed patterns where video creators posting potentially questionable materials (e.g., Vlad and Nikita live-action skits under copyright dispute) are subsequently removed by YouTube, sometimes in a matter of weeks. Our sample contained 161 links to similar videos that had been removed from YouTube since their publication less than three years ago, which throws into relief the ways that we are attempting to study a medium that erases its own past. It may be that any problematic video will have a shorter life online, which makes it difficult for researchers to capture and track with stable links, with remaining videos more likely to be formulaic or adhering to patterns that are "safe" from YouTube's algorithm-directed efforts to purge problematic content.

This study raises several questions about the effect of YouTube genres on child development. The nature of child's play may be impacted by the rigidity of the toy play genre's scripting of play. At the same time, children are exposed to a barrage of seemingly innocuous advertising normally restricted in other spheres. Other questions remain yet to be explored, such as the evolution of

these video genres over time, and the extent to which they might influence child development in other ways, such as with respect to gender.

9. References

[1] B. Burroughs. YouTube kids: The app economy and mobile parenting. *Social Media+ Society*, *3*(2), 2056305117707189. 2017.

[2] Bridle, J. Something is wrong on the internet. *Medium*. 2017. https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2

[3] Subedar, A., and Yates, W. The disturbing YouTube videos that are tricking children. *BBC Trending*. 2017. https://www.bbc.com/news/blogs-trending-39381889

[4] Masheshwari, S. On YouTube Kids, startling videos slip past filters. *The New York Times*. 2017. https://www.nytimes.com/2017/11/04/business/media/youtube-kids-paw-patrol.html

[5] Fisher, M., and Taub, A. On YouTube's Digital Playground, an Open Gate for Pedophiles. *The New York Times*. 2019. https://www.nytimes.com/2019/06/03/world/americas/youtube-pedophiles.html

[6] Roose, K. The Making of a YouTube Radical. *The New York Times*. 2019. https://www.nytimes.com/interactive/2019/06/08/technology/youtube-radical.html

[7] J. Kim. The institutionalization of YouTube: From user-generated content to professionally generated content. *Media, Culture & Society*, *34*(1), 2012. 53-67.

[8] Burgess, J., and Green, J. *YouTube: Online video and participatory culture*. John Wiley & Sons. 2018.

[9] J. Klausen, E. Barbieri, A. Reichlin-Melnick and A. Zelin, "The YouTube Jihadists: A social network analysis of Al-Muhajiroun's propaganda campaign", Perspectives on Terrorism, 2012, 6(1).

[10] S. Cunningham, D. Craig, and J. Silver. YouTube, multichannel networks and the accelerated evolution of the new screen ecology. *Convergence*, *22*(4), 2016. pp. 376-391.

[11] X. Cheng, C. Dale and J. Liu, "Statistics and social network of youtube videos." In: IWQoS 2008, 16th International Workshop on Quality of Service. IEEE Computer Society, Los Alamitos, CA, 2008, pp. 229-238.

[12] J Paolillo, "Structure and network in the YouTube core", In: Proceedings of the 41st Annual Hawaii International Conference on System Sciences, Los Alamitos, CA, 2008.

[13] J. Paolillo, S. Ghule, and B.P. Harper. A Network View of Social Media Platform History: Social Structure, Dynamics and Content on YouTube. In Proceedings of the 52nd Hawaii International Conference on System Sciences. 2019.

[14] T. Jones, and K. Cuthrell. YouTube: Educational potentials and pitfalls. *Computers in the Schools*, *28*(1), 2011. pp. 75-85.

[15] M. Buzzi. What are your children watching on youtube?. In *International Conference on Advances in New Technologies, Interactive Interfaces, and Communicability* Springer, Berlin, Heidelberg, 2011. pp. 243-252.

[16] D. Craig, and S. Cunningham. Toy unboxing: living in a (n unregulated) material world. *Media International Australia*, *163*(1), 2017. pp. 77-86.

[17] B. Nicoll, and B. Nansen. Mimetic Production in YouTube Toy Unboxing Videos. *Social Media + Society*. 2018. https://doi.org/10.1177/2056305118790761

[18] C.R. Miller. Genre as social action. *Quarterly journal of speech*, *70*(2), 1984. pp. 151-167.

[19] S.C. Herring, L.A. Scheidt, S. Bonus, and E. Wright, E. Bridging the gap: A genre analysis of weblogs. In *37th Annual Hawaii International Conference on System Sciences, 2004. Proceedings of the*. IEEE. 2004. pp. 1-11.

[20] T.M. Simonsen. Categorising YouTube. *MedieKultur: Journal of media and communication research*, *27*(51), 2011. pp. 72-93.

[21] H.K. Ekenel, and T. Semela. Multimodal genre classification of TV programs and YouTube videos. *Multimedia tools and applications*, *63*(2), 2013. pp. 547-567.

[22] Fisch, S. M. *Children's learning from educational television: Sesame Street and beyond*. Routledge. 2014.

[23] Lesser, G. S. *Children and television: Lessons from "Sesame Street."* New York: Random House. 1974.

[24] A. Levinovitz. Towards a theory of Toy-Play. *Human Studies*, 40, 2017. pp. 267-284

[25] Krippendorff, K. *Content Analysis, an Introduction to Its Methodology, 2nd Edition*. Thousand Oaks, CA: Sage Publications. 2004.

[26] K. Krippendorff. Computing Krippendorff 's Alpha-Reliability. 2011. http://repository.upenn.edu/ asc_papers/43

[27] Basilevsky, A. T. *Statistical factor analysis and related methods: theory and applications* (Vol. 418). John Wiley & Sons. 2009.

[28] R Core Team, "R: A language and environment for statistical computing", R Foundation for Statistical Computing, Vienna, Austria.

[29] L. Scrucca, M. Fop, T.B. Murphy and A.E. Raftery. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal 8/1*, 2016. pp. 205-233

[30] Romesburg, C. *Cluster analysis for researchers*. Lulu.com. 2004.

[31] C. Fraley, and A. E. Raftery. MCLUST: Software for model-based cluster analysis. *Journal of classification*, *16*(2), 1999. pp. 297-306.

[32] C. Fraley and A.E. Raftery. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association*, *97*(458), 2002. pp. 611-631.

[33] C. Fraley, and A.E. Raftery. Model-based methods of classification: using the mclust software in chemometrics. *Journal of Statistical Software*, *18*(6), 2007. pp. 1-13.