

Evaluating YouTube Videos for Young Children

Michelle M. Neumann¹ and Christothea Herodotou²

¹School of Education and Professional Studies, Griffith University, QLD, Australia

²Institute of Educational Technology, Open University, UK

Author Notes

Please direct correspondence to Michelle M. Neumann: School of Education and Professional Studies, Griffith University, Gold Coast QLD 4222, Australia. Telephone: +61(0)7 55529785

E-mail: m.neumann@griffith.edu.au

EVALUATING YOUTUBE VIDEOS

Abstract

YouTube has become a popular digital media platform used by young children. However, concerns have been raised around inappropriate video content and limited quality. A lack of research and theoretical discussion exists on how best to evaluate the quality of YouTube videos made for children. In this study, we reviewed research and developed a set of design principles that informed the production of a YouTube video rubric used to evaluate the quality of YouTube videos targeted at young children 0 to 8 years old. **From this, four key criteria were used to evaluate each video: Age appropriateness, Content quality, Design features and Learning objectives. These criteria demonstrated substantial inter-rater reliability between scorers.** This evaluation tool has the potential to be used by educators to assess the quality of videos for early learning and guide YouTube creators in their production of educational videos for young children.

Keywords: YouTube, young children, video content, **evaluation rubric**, learning

Introduction

The YouTube video platform has become one of the most popular online applications in the world that young children engage with (Marsh et al., 2019; Rideout, 2017), yet compared to other forms of screen media (e.g., television [TV]) this remains rather underexplored (Khan, 2017). There is a paucity of research around video screen media such as YouTube and possible effects on young children's learning and development (Izci, Jones, Ozdemir, Alktebi & Bakır, 2019). However, it is clear that YouTube is a form of screen time and through this media, children watch the behaviour of others. This behaviour has the potential to positively or negatively impact on learning, depending on the quality of videos and images presented on the screen.

When viewed through the lens of Bandura's social cognitive theory, children can learn through "observational learning" – where they imitate certain behaviours through adult and peer modelling (Bandura, Ross, & Ross, 1961). In Bandura's well-known Bobo Doll experiment, children who observed aggressive behaviour (e.g., an adult uses a hammer to hit a Bobo Doll) initiated more aggressive behaviours when they were asked to play with toys than a group of children who observed non-aggressive behaviour (Bandura et al., 1961). Follow-up studies examining the impact of media on children confirmed Bandura's findings that repeated exposure to violent and aggressive video content had a greater influence on children's behaviour, cognition, and emotions than infrequent exposure (Bushman & Huesmann, 2006). Studies also showed that these effects can be moderated by a number of factors: for example, children are more likely to imitate a behaviour when the character is (a) realistic; (b) similar to the child (e.g. age or gender); (3) receives positive reinforcement; and (d) the act can be imitated by the child (Bushman & Huesmann, 2006; Thomas, 2005). For example, action heroes are more likely to have an impact on children's behaviour than villains. Overall, screen media can provide models from which children can learn certain

EVALUATING YOUTUBE VIDEOS

values and ways society works (Bandura, 2001). Children can learn what is perceived as normative behaviour about themselves and about others such as racial groups, and influence ways of behaving, acting or treating others in society.

Another relevant perspective that explains how media such as YouTube and its inbuilt messages may influence children's behaviour and learning is the Cultivation theory (Gerbner, Gross, Morgan, Signorielli, & Shanahan, 2002). This theory proposes that the social perceptions of viewers are shifting overtime towards the screen media version of reality, even if this is not accurate or precise. The more time spent watching YouTube or other screen media such as TV can potentially result in two types of learning: (a) First order learning outcomes relate to the concrete characteristics of the content (e.g. fewer Asian children in sitcoms suggests that there may be fewer Asian children in the real world), (b) Second order learning outcomes derive from implied content. They are messages the viewer can infer from the media for example, the lack of Asian men in sitcoms may imply their lack of humour. The cultivation theory suggests that the way people and social relationships are portrayed in the media may affect children's perceptions about the world creating misleading or inaccurate perceptions of reality and biases. These theoretical perspectives underpin the importance of providing quality video content and media experiences for children that support learning of non-aggressive behaviours and promote positive skills and knowledge about the real world they live in.

To distinguish amongst YouTube videos, general age-related ratings and classifications have been used to categorise video content as suitable or unsuitable for children (Papadamou et al., 2019). For example, these include: *Suitable for children*: General audience classification is where appropriate content for children aged 1 to 5 years, is relevant to young children's interests e.g., child friendly cartoons, songs, children playing, educational videos (alphabet, colours). *Disturbing*: Mature audience classification is where content is

EVALUATING YOUTUBE VIDEOS

inappropriate for children under 15 years of age and contains sexual hints, abusive and inappropriate language, nudity, child abuse, screams, horror, scary scenes or characters (injections, insects). *Restricted*: Restricted audience classification is where it is inappropriate for children under 17 years, containing sexual language, nudity, pornography, violence, inappropriate language, swearing, gambling, drug use, alcohol, distressing situations, and activities. *Irrelevant videos* contain appropriate content for children but would not interest them (e.g., some gaming videos) (Papadamou et al., 2019).

Yet, such classifications are not clearly driven or informed by research that examines how children learn best from interactions with screen media. Also, little work has been done to evaluate YouTube content for young children (Khan, 2017). It is only recently that YouTube introduced restrictions to content targeting children by requiring creators to label whether their videos are made or not for children. Videos made for children will have limited functionality including no personalised advertisements, comments, stories, live chat and notification bell. Yet, the guidelines provided by YouTube as to how to determine whether a video is made for kids are generic and could be interpreted in varied ways: the content should be educational, the video should include child actors or models or characters and toys that appeal to children, the language should be appropriate for children, and the video presents activities for children such as play-acting, songs, poems (see https://support.google.com/youtube/answer/9527654?ref_topic=9257782). If a video meets all the above criteria, it may still be inappropriate for children, for example if negative behaviour such as bullying, or promotion of stereotypes are portrayed and encouraged. In addition, YouTube has recently funded eight YouTube channels to produce high quality educational material recognising the use of the platform as an educational portal for several users (see <https://bit.ly/2MSkMRX>). This study aims to shed further light on this topic by:

EVALUATING YOUTUBE VIDEOS

- 1) Conducting a narrative review of studies on children's interaction with screens and digital media.
- 2) Developing a rubric that can be used to judge the quality of YouTube videos for children aged 0 to 8 years.
- 3) Piloting the proposed evaluation rubric with five popular YouTube videos that target children.

In the following sections, we review relevant studies and use these to inform the development of the evaluation rubric that can be used to guide educators and YouTube creators in selecting and designing good quality educational videos for young children.

Video design features and early learning

As the quality of YouTube videos made for children varies significantly, studies suggest that specific design principles used in the creation of video content for children can influence learning and educational outcomes (Izci et al., 2019; Veblen, Kruse, & Messenger, and Letain, 2018). Video screen media such as TV shows with interactive elements and design features (e.g., characters posing questions to the viewer, searching for clues or solving problems; Crawley, Anderson, Wilder, Williams, Santomero, 1999) have been reported to result in greater learning gains for young children. For example, watching educational TV shows such as Sesame street, Gullah Island, and Blue's Clues actively engaged children's interest (Anderson, Lorch, Field, Sanders, 1981). A series of relevant studies also showed that children pay less attention when they do not understand the language or when the storyline is mixed up which may impact on comprehension.

Certain video design features were used to promote children's interaction with educational content in shows such as Sesame Street and support early literacy skills. For example, in a 'Learning about letters' episode (a) learning goals are orally and explicitly stated at the beginning of the episode, (b) the phoneme-grapheme correspondence is

EVALUATING YOUTUBE VIDEOS

promoted by accompanying the sounds of letters (or words) with the relevant grapheme (written representation), (c) words in a storyline or a song are accompanied by images (representing words) as well as their written representation. In another Sesame Street episode about the moon, the above techniques are blended with (a) songs related to the content presented, (b) animations of facts about the moon (e.g., when one can see the moon in the sky where day is presented alongside the sun), and (c) counting backwards as part of a rocket going to the moon (Wright, St Peters, & Huston, 1990).

Blue's Clues videos provided learning opportunities through a main character on the screen who encourages children to help solve problems during the programme's activities (Crawley et al., 1999). The main character poses questions around solving problems during a treasure hunt, encouraging children to look for clues on the screen or answer math-based questions such as "if Blue and I had to pay one dollar each, how many dollars should we pay to enter the museum?" Each episode is developed around the same storyline or narrative that invites children to take part in a treasure hunt. However, in contrast to Sesame Street, Blue's Clues contains less written text (e.g., words and letters) on the screen which may reduce literacy learning opportunities.

Screen media can encourage children to understand and use new words, yet it is not clear whether they can help them learn more complex language skills such as grammar (Naigles & Mayeux, 2001). Linebarger and Walker (2005) collected data about children's TV viewing every three months, starting at six months old and finishing at 30 months old, and identified that programmes like Dora the explorer, Blue's Cues, or Dragon Tales resulted in greater vocabulary and expressive language, than programmes like Teletubbies, the latter shown to relate to fewer vocabularies and language expression. Sesame Street was only related to smaller expressive language scores and Barney to fewer vocabulary words and more expressive language. These findings can be explained by the fact that shows like Dora

EVALUATING YOUTUBE VIDEOS

the Explorer and Blue Clues encourage children to talk to the screen while Sesame Street does not. Barney and friends aim to engage children with imaginative play and the language is simple, potentially explaining differences in vocabulary word gain. These findings held true even after controlling for parent education, home environment and child's cognitive performance. Watching programmes that stimulate language development through, for example, clear labelling and encouraging repetitions and interactions with on screen characters can foster positive learning outcomes (Moussiades, Kazanidis, & Iliopoulou, 2019).

Another line of research reports significant learning gains in relation to increased imaginative play and creative thinking from shows such as Mister Rogers' Neighbourhood and Barney and Friends; the first one transfers viewers via a trolley to the land of make-believe and the latter is led by a purple dinosaur. This approach of exploring and discovering new ways of doing or making things stimulates curiosity and early learning. These findings align with the "stimulation hypothesis" (Valkenburg, 2001) whereby media stimulates children's imagination and enriches children's ideas when they are engaged in imaginative play or creative tasks. Furthermore, the quality of video screen media is influenced by key content and design features such as questioning, opportunities for child play, problem solving, language and literacy learning through text presented on the screen, and stimulation of curiosity (e.g., Anderson et al., 1981; Crawley, et al., 1999; Moussides et al., 2019). These features should be considered in the design of YouTube video content to support early learning and development.

Other video design features influencing learning

In this section we discuss additional video design features including repetition, fast pace, child age and mediation that have been shown to relate to learning from video screen media. For example, repeated exposure to the same educational content was shown to result

EVALUATING YOUTUBE VIDEOS

in greater learning gains in early numeracy, literacy and vocabulary for four to seven-year-old children compared to children watching an entertainment programme (Baydar, Kağıtçıbaş, Küntay, & Gökşen, 2008). In a similar study, children who watched the same episode of Blue Clues for five times outperformed their peers who watched another episode only once demonstrating greater comprehension and recall of the material in that episode (Crawly et al., 1999). Fast editing in combination with unrealistic content are features of screen media shown to deplete executive functioning in four-year-old children even if the content has been labelled as “educational”. Yet, in a follow-up study in which the pace and content were varied across experimental videos, significant effects of content on executive functioning were found yet not of pacing, suggesting that executive functioning is more likely affected by content than pacing (Lillard, Li, & Boguszewski, 2015). Other studies suggest that overstimulation evoked by fast pace of events (compared to real life) as well as the number of auditory and visual stimuli may be beyond the cognitive capacity of children and may not support learning (Zimmerman & Christakis, 2007).

Another factor shown to influence cognitive gains is the age of children. It has been reported that children younger than three years old learn less from TV than from real-life demonstrations (Barr, 2010). “Video-deficit” has been used to describe the difficulty of under-threes to apply knowledge acquired from “symbolic sources” such as picture books, videos, or mobile applications to real-world situations (Herodotou, 2017a; 2017b). A condition that can minimise the deficit effect is mediation or scaffolding by adults. There are three types of mediation by adults: (a) active mediation that involves conversations about screen content such as questions about the characters’ motivations or understanding of media content that assists children to critically analyse what they watch, (b) restrictive mediation whereby supervision by parents with explicit rules about what they can watch and play with (e.g. online filtering software), (c) co-viewing (active or passive): active co-viewing involves

EVALUATING YOUTUBE VIDEOS

general talk about content (Strasburger, Wilson, & Jordan, 2009). Active mediation or co-viewing can result in enhanced learning outcomes due to parents' engagement with content by pointing and talking about what children are watching (Huston, Wright, Rice, Kerkman & St Peters, 1990).

A significant line of work relevant to this discussion on the importance of video design features for supporting learning is Mayer's (2008) principles of multimedia design. These are summarised as follows: *Coherence*: extraneous material or material not related to the learning goal should be reduced; *Signalling*: essential material or significant information should be highlighted; *Redundancy*: narrated animations should be used without any accompanying on-screen text; *Spatial contiguity*: printed words should be placed next to relevant graphics; *Temporal contiguity*: narration and animation should be presented concurrently; *Segmenting*: animation should be presented in learner-paced segments; *Pretraining*: provide pretraining about key components; *Modality*: present words as spoken text rather than printed text; *Multimedia*: present words and pictures rather than only words; and *Personalisation*: avoid formal wording and use a conversational style e.g. use 'your' than 'the'. These ten principles provide useful directions about how videos should be designed and presented to support instruction in the classroom.

It is also essential that teachers have the tools to evaluate the strengths and limitations of digital media so they can provide children with videos that contain high quality features such as content (developmentally appropriate) and accessibility (used by children with a range of abilities) to maximise educational benefits (Hillman & Marshall, 2009; More & Travers, 2013). Previous studies (e.g., Miller & Kocurek, 2017; Moussiades et al., 2019; Neumann, Wang, Xi, & Neumann, 2019; Papadamou et al. 2019; Veblen, Kruse, Messenger, & Letain, 2018) provide promising guidelines to adopt when evaluating digital media. This is particularly important because YouTube is the most popular social media platform children

EVALUATING YOUTUBE VIDEOS

currently engage with (Marsh et al., 2019). Selecting quality digital media content for young children's learning is critical to ensuring positive learning experiences for children at home and school (Hillman & Marshall, 2009).

Evaluation tools are needed to verify the credibility, accuracy and reasonableness of YouTube resources for young children (Jones & Cuthrell, 2011).

For example, Neumann et al. (2019) developed a scored checklist to test the quality of digital language learning resources for young children. The criteria (e.g., age appropriateness, problem solving, cultural awareness, language and literacy content) were developed to assist teachers to select quality apps for language learning. The criteria and checklist were shown to be a clear and comprehensive tool to evaluate the suitability of these digital resources for pre-school children and are important to consider when assessing the quality of digital content such as YouTube videos.

Based on the studies reviewed, a summary of principles for designing and selecting YouTube content for young children is listed in Table 1. These principles can be utilised to inform video selection or design and production and have been mapped onto five overarching principles: 1) Age appropriateness, 2) Content quality, 3) Design features, and 4) Learning objectives.

1) *Age appropriateness* refers to a child's age and respective developmental maturity and capabilities that relate to understanding or processing content from screen media experiences.

2) *Content quality* refers to meaning making from media or messages and actions screen media may promote such as child friendly and non-violent images.

3) *Design features* refer to a video's structural characteristics and technical features, which represents how the content is presented to children (e.g., animations and text align with a narration or song).

EVALUATING YOUTUBE VIDEOS

4) *Learning objectives* refers to what children may learn or acquire from interacting with the video in relation to cognitive, physical, social-emotional skills and capabilities.

Table 1. *Evidence-based principles for selecting or designing screen media for young children*

Principle	Criterion	Relevant literature
Design screen content that children have the capacity to imitate, is realistic, and shares similarities with the child (age, gender, interests). Positive screen behaviour is modelled and receives positive reinforcement in the video (e.g., feedback, encouragement). Whereas negative screen behaviour is modelled and receives negative reinforcement (e.g. punishment, telling off).	Age appropriateness	Bandura (2001), Thomas (2005)
Ensure that screen media that are not mediated by adults are not used with children under 3 years due to the video deficit effect.	Age appropriateness	Anderson & Pempek, (2005), Barr (2010)
Portray social relationships in the screen content as accurate as possible considering implied messages the child may derive from content.	Content quality	Gerbner et al. (2002)
Ask children to perform or engage in creative tasks, providing alternative or new ideas.	Content quality	Valkenburg (2001)
Ensure the content is clearly and explicitly appropriate for children 3 years old and over.	Content quality	Anderson & Pempek, (2005); Barr (2010)
Present dialogue or narrative that is understandable within a single scene to improve comprehension. Avoid presentation of multiple scenes on the screen.	Content quality	Anderson et al. (1981)
Provide content repetition (e.g. repeat significant concepts or narrative multiple times throughout the video); this can help overcome memory limitations due to the video deficit effect.	Design features	Crawly et al. (1999)
Keep a low pace with infrequent scene and character changes.	Design features	Zimmerman & Christakis (2007)
Insert interactive elements in the screen content encouraging children to interact with content, repeat certain things or solve problems.	Design features	Crawley et al. (1999); Linebarger & Walker, (2005); Kostyrka-Allchorne, Cooper, & Simpson (2017)
Present narration and animation together (e.g., a visual representation should be accompanied by a song or a narrative).	Design features	Naigles & Mayeux (2001); Linebarger & Walker (2005); Mayer (2008)
Avoid formal wording and use a conversational style.	Design features	Mayer (2008)
Present words and relevant pictures (e.g., graphics) together.	Design features	Mayer (2008)
Highlight the learning-related elements of the video	Design features	Mayer (2008)
Ensure that screen media contains content that supports the development of knowledge, skills, and understanding within developmental domains (cognitive, physical, and social-emotional).	Learning objectives	Baydar et al., (2008), Crawley et al. (1999), Linebarger & Walker (2005), Moussiades et al. (2019), Wright et al. (1990).

EVALUATING YOUTUBE VIDEOS

Methodology

The principles presented in Table 1 were used to design a YouTube evaluation rubric (see Table 2) that can assess the educational quality of a video in terms of Age appropriateness, Content quality, Design features, and Learning objectives. These main criteria were divided into 17 sub criteria (questions). A 3-point scoring system was used to assess each question and produce a total numeric assessment per video: No = no evidence (Score = 0 point); Partial evidence (score = 1 point); or Yes = ample evidence (score = 2 points). Total scores were calculated by summing the sub-criteria scores (max score = 17 x 2 = 34). The overall quality rating classification of a given video is: 0 < 17 Not Recommended and 17 to 34 Recommended for viewing by children aged 0-8 years old.

Table 2. *YouTube video evaluation rubric*

Criteria	Sub-criteria (Evaluation Questions)	0 No	1 Partial evidence	2 Yes
-----------------	--	-----------------	-----------------------------------	------------------

EVALUATING YOUTUBE VIDEOS

Age appropriate-ness	1. Can the child imitate the content presented (e.g. repeat a song, make body movements or gestures)?	
	2. Does the content share similarities with the child (e.g., age, gender, interests)?	
	3. Is the behaviour on-screen positive (e.g., ethical, fair, caring, moral, non-violent, non-scary, healthy)?	
	4. Does the on-screen behaviour receive appropriate reinforcement (e.g., positive behaviour is praised or encouraged, and negative behaviour is discouraged)?	
Content quality	5. Are social relationships accurately represented (e.g., gender and cultural stereotypes, power relationships)?	
	6. Does the video encourage children to perform creative tasks, solve problems or provide alternative ideas or ways of doing things?	
	7. Does the video encourage children to repeat content?	
	8. Are the images, audio, sounds and language used appropriately for children (i.e. the child can understand the content)?	
	9. Is each scene clear, logical, and easy to follow?	
Design features	10. Is some content repeated during the video (e.g., to reinforce learning in positive ways)?	
	11. Is there low and gradual pace with infrequent scene and character changes?	
	12. Are pictures/graphics/animations presented alongside words/narration?	
	13. Is conversational style used in wording (oral and written)?	
	14. Are learning elements highlighted in the video?	
Learning objectives	15. Does the video support cognitive development (e.g., language, literacy, math, science knowledge)?	
	16. Does the video support physical development (e.g., gross and fine motor skills)?	
	17. Does the video support socio-emotional development (e.g., fosters positive relationships, communication skills, moral attitudes , resilience, self-regulation, self-confidence)?	
Total Score		
Quality Rating	0 to < 17: <i>Not Recommended</i>; 17 to 34: <i>Recommended</i>	

Note: “Encourage” (e.g., to directly ask a child to repeat or do something).

The proposed evaluation rubric was piloted with five children’s YouTube videos (Ryan’s Slime, Baby Shark, Playdough Sparkle, Ryan’s Giant Egg Surprise, Hamburger Cooking Video; see Table 3). These videos were selected due to their level of popularity which was based on the number of YouTube channel subscribers and most popular categories (genres). In terms of the latter, the most popular categories include songs and nursery rhymes;

EVALUATING YOUTUBE VIDEOS

unboxing; fun toys and play; create and explore (see

<https://www.statista.com/statistics/785626/most-popular-youtube-children-channels-ranked-by-subscribers/> April 2019). Details about each video can be found in Table 3 (i.e., number of views, category, length of video).

Table 3. *YouTube videos used for piloting the proposed evaluation rubric (as at October 2019)*

Video Title and Link	YouTube Channel	No. of Subscribers	No. of views	Genre	Length of video (mins)
Hamburger cooking and baby doll kitchen play house (https://www.youtube.com/watch?v=cGxgkMISNmM)	ToyPuddingTV	25.6M	58,088,402 Jan 22, 2019	Fun toys and play	10:03
Baby Shark (https://www.youtube.com/watch?v=XqZsoesa55w)	Pinkfong! Kids' Songs & Stories	24.1M	3,600,538,732 Jun 17, 2016	Songs and Nursery Rhymes	2:16
GIANT Lightning McQueen Egg Surprise with 100+ Disney Cars Toys (https://www.youtube.com/watch?v=Tldlt2RhrDw)	Ryan Toys Review	21.8M	992,426,874 Jul 1, 2015	Unboxing	7:25
How to Make OOBLEK! DIY Slime at Home with Ryan!!! (https://www.youtube.com/watch?v=kK_i9E-xjeM&t=26s)	Ryan Toys Review	21.8M	1,897,563 Oct 5, 2019	Create and Explore	7:55
Play Doh Sparkle Princess Ariel Elsa Anna Disney Frozen MagiClip Glitter Glider Dolls (https://www.youtube.com/watch?v=I8vzbluvhoo)	FunToys Collector Disney Toys Review	11.7M	594,098,076 Jul 11, 2014	Create and Explore	9:11

Inter-rater reliability

In the first instance, the first and second author co-scored a test video in order to refine and finalise the rubric and scoring criteria. Then each of the five videos in Table 3 was independently scored by author one and two. Of the total 85 items scored across the 5 videos (5 videos x 17 items) there were 70 agreements and 15 disagreements identified. Cohen's Kappa was used to calculate the level of interrater reliability. This was 0.69 which indicates substantial agreement between the scorers. The disagreements were then compared and

EVALUATING YOUTUBE VIDEOS

discussed. The two authors reached agreement on the 15 disagreements. Of the 15 disagreements author one agreed with 6 of author two's scores; and author two agreed with 9 of author one's scores. Therefore, the final agreed scores were used (see Table 4).

Results

Table 4 shows the sub-total and total scores for each video. Ryan's Slime science-based experiment video was the highest scoring video and was recommended for children. The Baby Shark Video focussed on song, dance and movement and also scored highly. The scores of the three remaining videos (Playdough Sparkle; Ryan's Giant Egg Surprise; and the Hamburger Cooking video) were similar and were ranked lower than the top two videos. Although the Playdough Sparkle and Ryan's Giant Surprise Egg videos were recommended (as they scored close to the cut off for being recommended), it is important to select those videos with caution and critically review their potential limitations before use. The Hamburger Cooking video scored the lowest and was not recommended.

Table 4. *Summary of criteria scores of YouTube videos from the evaluation rubric*

YouTube Video	Criteria and *Sub-Criteria	**Total Score
--------------------------	-----------------------------------	--------------------------

EVALUATING YOUTUBE VIDEOS

	<i>Age Appropriateness</i>				<i>Content Quality</i>					<i>Design Features</i>					<i>Learning Objectives</i>			Out of 34
Sub-criteria:	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	
Ryan's Slime	2	2	2	2	2	2	1	2	2	2	2	2	2	1	1	2	2	31
Baby Shark	2	2	2	0	1	0	2	2	2	2	2	2	2	2	2	2	2	29
Playdough Sparkle	1	1	2	0	1	2	0	2	2	2	2	1	0	0	1	2	0	19
Ryan's Giant Egg Surprise	2	2	2	1	1	0	0	1	2	2	2	0	1	0	0	2	0	18
Hamburger Cooking Video	1	2	1	0	1	1	0	1	1	2	2	0	0	0	0	2	2	16

*(See Table 2 for the 17 sub-criteria questions. Score 0 = No; Score 1 = Partial Evidence; Score 2 = Yes)

** (0 to <17 = Not recommended; 17 to 34 = Recommended).

Description and evaluation of the five YouTube videos

How to Make OOBLEK! DIY Slime at Home with Ryan! This video shows a father and son making slime in their kitchen at home. This is framed as a science experiment that is fun and enjoyable. The father models the procedure and process to his son in a fun and engaging way. This video models positive parent-child interactions, can give ideas of how parents can engage with their children, and how they can follow instructions and learn from their parents and each other in positive ways. Laughing and enjoyment were evident in the video along with some audio effects that were used to further enhance the viewer's experience.

The video provided several opportunities for listening and communication as seen through positive language interactions. The father embraces play, fun, and mess and says, "It's ok to be messy!" and encourages his son to participate (e.g., the father says "Ryan touch the slime; and you help me, and I'll help you"). Although the language used was conversational in nature and easy to understand, there was also some math and science

EVALUATING YOUTUBE VIDEOS

vocabulary and words used during the slime making task. For example, some words used were: “pouring, open it up, more, measuring, tell me when to stop, three drops of food colouring, and pressure”. The mother (who is filming) encourages critical thinking by asking questions such as “Is this a solid or liquid or is it hard or soft?” She also provides some health and safety advice to her son such as “Remember never to eat anything when you are doing an experiment”.

The video content follows a logical procedure and provides some engaging digital animation overlays on the screen such as a “volcano cartoon picture” that pops up as Ryan mixes the slime ingredients in the mixing bowl and describes it “like a volcano”. Some science-related words (e.g., “non-Newtonian liquid”) are also added as text on the screen. The activity is extended in a creative way by adding glitter to the slime. Further explanation of content knowledge (e.g. liquids, gases, solids) and provision of additional visual or narrative representation of science concepts would further enhance the video. Based on a score of 31/34 this video is recommended for children.

Baby Shark: This is a video presenting a song with physical actions and gestures. Its theme is set “under the sea” and young children are being chased by a family of animated sharks and they eventually get home safe at last. The video has a simple story line and clear and engaging sound and audio effects along with a fun, happy and joyful musical beat rhythm provided for the dance moves. Words and sentences are presented on the bottom of the screen and align with the singing. Some words and verses are repeated (e.g., “do do do do...”). However, the text is presented rather quickly on screen which may be difficult for young children to follow the words. Yet, this frequency may work well for older children who know how to read from the screen. Clear physical actions are provided by the children dancing in the video that would be easy for young children to imitate (e.g., moving arms up and down like the mouth of a shark and curling fingers in to show grandma shark). There are some

EVALUATING YOUTUBE VIDEOS

gender stereotyped messages communicated in the video such as the colour of costumes (e.g., boys wearing blue and girls wearing pink) and the singing voices of mum and dad (the mum has a gentle voice while the dad has a deeper authoritative voice). Based on a score of 29/34 this video is recommended for children.

Play Doh Sparkle Princess Ariel Elsa Anna Disney Frozen MagiClip Glitter

Glider Dolls: This video is a tutorial on how to make six sparkling dresses out of playdough for six Disney princess dolls. The video begins with showing two hands that pull out the playdough, and a female voice saying each of the dress colours: “We are going to make green, pink, purple, blue, hot pink, yellow play dough dresses”. Two moulding tools (in the shape of hearts and butterflies) are shown on the screen but they are not used in the making of the playdough dresses. The two hands silently model how to flatten the playdough with thumb and fingers and press it neatly around the princess doll’s body.

Limited narration is provided during this process and only some repetitive background music is heard. For example, the narrator says: “Now we are going to use purple”. The two hands go through the process of making each playdough dress and this is repeated six times. There are only a few narrative elaborations heard (e.g. “Now Ariel, and her dress is going to be yellow, I think Ariel looks beautiful in yellow”) throughout the whole video. There was no story line or character building presented and the narration and explanations given were limited. No written text was provided on the screen to explain or describe the task. Some gender bias was evident (towards girls) as clothes were made for Disney princess dolls only (and not for princes). Based on a score of 19/34 this video is recommended for children with careful consideration of its limitations.

Giant Lightning McQueen Egg Surprise with 100+ Disney Cars Toys: This video shows a pre-schooler called Ryan playing with his toys in his bedroom. The first scene begins with Ryan’s mum waking him up from bed in the morning. His mother asks Ryan if he wants

EVALUATING YOUTUBE VIDEOS

to open a giant hollow rubber egg (about one meter in height) filled with toys that is next to his bed. The mother is not visible in the video. Ryan is shown playing with his toy Disney cars as he pulls them out of the giant egg. To open the giant egg Ryan begins by taking a cylinder-shaped inflatable toy and starts to hit the egg with it. After opening the egg, Ryan says, "Let's see what's inside the egg" and begins to pull out large toys from it. The mother says, "Wow what is that?" Ryan says: "It's a racing car track". Ryan plays with the cars and track and he pretends to fill the lightning McQueen car up with gas. He then continues to pull out more cars and trucks from the egg. Ryan says words such as "big", "little", "truck", and "crash" and says: "I want to race all of them". He also asks a few questions during his play for example he asks himself: "What do we have here?".

This video would be easy for children to imitate and contains the element of surprise which appeals to children. It is about a topic that children like such as opening up presents and playing with toys. However, the selected toys (cars and tracks) demonstrate some gender bias as they are mainly targeting boys' interests. A video of boy playing with a mixture of different toys including for example, dolls may overcome this limitation. In addition, there is no story line, narration, or clear character building. There are also no words or text presented on the screen and the language used is not elaborated upon. Based on a score of 18/34 this video is borderline, yet based on our criteria, recommended. However, careful review of its limitations should be considered prior to use.

Hamburger Cooking and Baby Doll Kitchen Play House: This video uses toys to show different ways to prepare food to feed a baby doll. Children are familiar with this topic about making and eating food, for example in the video, hamburger patties are made from brown playdough and they are put into a toy oven using plastic spoons. Human hands are seen on screen pouring drinks and feeding a baby doll sitting in a trolley or highchair. Soft and repetitive background music is heard during the video, but no words are seen or heard,

EVALUATING YOUTUBE VIDEOS

and the baby doll squeals on the rare occasion. The choice of food is not always healthy as the baby doll is fed chips, milk shakes, hot dogs, and sausages which shows unhealthy eating habits and promotion of eating junk food.

There are no words or comments provided about the quality of foods prepared. A female doll looks after a baby doll there are no male figures, and the colours are mainly targeting girls as they are not gender neutral (e.g., lots of pink toys). There is no story line or character development and the video shows continuous and repeated preparation of food. This video lacks encouragement of critical thinking. This may be viewed as an “easy” video for very young children to watch but there is little evidence of how it supports cognitive skills. Based on evaluation using the rubrics criteria (see Table 4) it gained a score of 16/34 and is not recommended.

Discussion

In this paper, we reviewed studies describing how particular design features of screen media may support children’s learning. A set of evaluation criteria was developed that can be used to guide the selection of YouTube videos for children. These evaluation criteria were underpinned by key theories such as sociocultural theory (Vygotsky, 1978) where videos are created and shared globally online, within a social, cultural and historical context; Social-cognitive theory (Bandura, 1961), where children can observe and learn from imitating the behaviours of others through the video content; and Cultivation theory where social perceptions of viewers overtime can shift towards video content versions of reality (Gerbner et al., 2002). The present study is timely given the rising popularity of YouTube videos targeting children and the increasing time children are spending on YouTube (Rideout, 2017; Rideout & Robb, 2019). YouTube compared to other screen media such as TV is rather underexplored (Khan, 2017) except for a few studies examining the design of educational videos and its relationship to behaviour and online engagement (Ten Hove & van der Meij,

EVALUATING YOUTUBE VIDEOS

2015; Shoufan, 2019). Informed by insights from a similar medium (TV), the present study provides a comprehensive set of evaluation criteria that teachers or parents and carers can use to assess the learning quality of YouTube videos, and make informed decisions about which videos their children should watch in order to maximise learning opportunities.

The quality of five popular YouTube videos for children was evaluated using the following criteria: *Age Appropriateness, Content Quality, Design Features, and Learning Objective*. The Create and Explore video (Ryan's slime tutorial) and Songs and Nursery Rhyme (Baby shark) video gained the highest scores indicating an acceptable level of quality with the potential to foster positive behaviour and learning. However, further research is needed to determine the actual learning benefits of watching such videos. In contrast, the videos Create and Explore (Playdough Sparkle Princess), Unboxing (Ryan's Giant Egg Surprise), and Fun Toys and Play (Hamburger Cooking) gained the lowest scores indicating limited quality and the need for careful consideration when these videos are used for educational purposes.

More specifically, two criteria were fully met in all five videos (See Table 4), namely content repetition and low and gradual pace with infrequent scene and character changes. The following criteria were fully satisfied in four of the videos. These were: the video was sharing similarities with the child, had positive on-screen behaviour, was easy to follow scenes, and had socio-emotional objectives. YouTube videos that were limited in quality (See Table 4) lacked the use of reinforcement, encouragement to repeat on-screen tasks, and highlighting of learning elements. Furthermore, four of the videos showed only partial evidence of accurately presented gender and cultural stereotypes suggesting that this issue should be further considered in the design of quality content. A similar concern was identified in a study by Veblen et al. (2018) who found that 81% of the 103 YouTube videos they reviewed on

EVALUATING YOUTUBE VIDEOS

clapping games featured only girls, potentially reinforcing stereotypical perceptions of gender.

Overall, this present analysis suggests that given the popularity of the selected videos, children are more likely to be exposed to both high- and low-quality YouTube video content. Although a greater number of videos need to be evaluated and the rubrics validated before drawing conclusions, the present findings show that design features such as repetition of content, highlighting of learning elements and support of social-emotional development should be considered in the planning and design of video content for young children. A conscious and careful selection of videos should be made by adults or more knowledgeable others to ensure that the time children are spending on YouTube is as beneficial as possible. The evaluation criteria we propose in Table 2 may assist teachers in the process of selection of videos in the classroom and maximise the learning potential of young children's YouTube experiences. In addition, the criteria have the potential to be utilised by people who create and produce YouTube videos for children and help maximise the quality of their videos for early learning, thus addressing the current need for high-quality educational video content in YouTube.

Although speculative, using quality YouTube content has the potential to provide children with valuable opportunities to learn digital literacy and media skills and become producers of media in the future, for example sharing their own video creations on YouTube and broadcasting quality content online (Marsh, 2016). Indeed, Chau (2010) argued that one of the benefits of YouTube as a participatory cultural tool for young people is to foster creativity, self-expression, innovative practices, and collaboration when content is viewed and shared online. To understand the learning impact of YouTube on child development, longitudinal ethnographic studies that follow children through adolescence and adulthood, or

EVALUATING YOUTUBE VIDEOS

studies that examine different aged cohorts of children are needed to determine the long-term effects of YouTube use on learning and development.

In addition, Kostyrka-Allchorne, Cooper, and Simpson (2017) have recently reviewed studies examining the impact of TV on children and concluded that research has failed to consider factors such as individual ability and family context to explain the relationship between time spent on digital media and measured learning outcomes. The importance of video content has been overlooked; for instance, TV content directed to five-year olds differs from content for 10-year olds and when this variation is not considered may result in misleading learning effects and outcomes. These factors should be considered in future research that examines the influence of YouTube videos on early learning and long-term development. Several reports and studies present recommendations for parents and families on how to regulate “screen-time”, that is children’s interactions with screens, such as having non-screen time periods, creating rules, monitoring what children watch online, turning off search functions, and providing opportunities for a wide range of digital and non-digital play activities (Nikken & Janz, 2006; Neumann, 2015).

Elaborating on this line of work, the present pilot study provides an evaluation rubric that has the potential to help families check the quality of YouTube videos and ensure that children are experiencing benefits from use. Adults should also co-view, discuss and help children reflect on the content of YouTube videos to ensure that negative messages (e.g., cyberbullying, violence, and stereotypes) are identified and discouraged, and positive behaviours such as caring and sharing are encouraged and reinforced. Such interactions may contribute to equipping children with the knowledge and skills to be ethical and critical participants in their online community.

Using a conversational and guided interaction approach (Plowman, McPake, Stephen, 2010) when using YouTube videos at home or in the classroom is also important to support

EVALUATING YOUTUBE VIDEOS

early learning (Davidson, Given, Danby, & Thorpe, 2014). For example, using question-answer strategies and embodied actions and gestures (e.g., nodding during co-viewing) can support children's meaning making. Pempek, Demers, Hanson, Kirkorian, and Anderson (2011) highlight the importance of scaffolding as described by Vygotsky's (1978) Zone of Proximal Development; for example, parents who provide positive interactions such as physical and verbal involvement and a positive tone, predicts positive outcomes for the child (Alessandri, 1992). Therefore, the role of the scaffolder or gatekeeper (Izci et al., 2019; Neumann, 2017) and quality of digital resources (e.g., Neumann et al., 2019) is important for young children's early learning with screens and digital media.

Limitations

The findings of this pilot study should not be generalised, or clear recommendations made due to the small sample size of videos assessed. Importantly, an inter-rater agreement check between the two authors showed that the criteria items showed substantial agreement between coders. Yet, there is a need for additional independent raters to check its validity. In particular, the rubric needs to be validated with a group of experts in the field such as classroom teachers. This will provide advice on the content and confirm that the rubric is measuring what it is intended measure. Therefore, the findings of this study should be treated as exploratory and not be generalised due to the small sample of videos assessed and the need for further validity and reliability checks.

Conclusion

Underpinned by theory and design principles, this pilot study has developed an evaluation rubric for assessing the educational quality of YouTube videos created for young children. This criteria can be used to select YouTube videos of good quality and evaluate its usefulness for young children's learning. Future research is needed to examine a wider and diverse range of video content and genres made for kids in order to further refine the

EVALUATING YOUTUBE VIDEOS

proposed rubric. Examining the direct effects of YouTube viewing on learning outcomes would also provide useful insights about the learning benefits and disadvantages of YouTube. The ecological validity of the rubric also needs to be tested with parents and carers at home, teachers in the classroom, and content creators in order to fully assess its effectiveness. This approach will lead to better early learning experiences for young children at school and home and provide opportunities for children to experience good quality YouTube videos.

References

- Alessandri, S. M. (1992). Effects of maternal work status in single-parent families on children's perception of self and family and school achievement. *Journal of Experimental Child Psychology, 54*(3), 417–433. [https://doi.org/10.1016/0022-0965\(92\)90028-5](https://doi.org/10.1016/0022-0965(92)90028-5)
- Anderson, D. R., & Pempek, T. A. (2005). Television and very young children. *American Behavioral Scientist, 48*(5), 505–522. <https://doi.org/10.1177/0002764204271506>
- Anderson, D. R., Lorch, E. P., Field, D. E., & Sanders, J. (1981). The effects of TV program comprehensibility on preschool children's visual attention to television. *Child Development, 52*(1), 151-157.
- Bandura, A., Ross, D., & Ross, S. A. (1961). Transmission of aggression through imitation of aggressive models. *The Journal of Abnormal and Social Psychology, 63*(3), 575–582. <https://doi.org/10.1037/h0045925>
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*, 1-26.
- Baydar, N., Kağıtçıbaşı, Ç., Küntay, A. C., & Gökşen, F. (2008). Effects of an educational television program on preschoolers: Variability in benefits. *Journal of Applied Developmental Psychology, 29*(5), 349-360.
- Barr, R. (2010). Transfer of learning between 2D and 3D sources during infancy: Informing theory and practice. *Developmental Review, 30*: 128–154.
- Bushman, B. J., & Huesmann, L. R. (2006). Short-term and long-term effects of violent media on aggression in children and adults. *Archives Pediatrics and Adolescent Medicine, 160* (4). 348–352. doi:10.1001/archpedi.160.4.348
- Chau, C. (2010). YouTube as a participatory culture. *New directions for youth development, 128*, 65-74. doi: 10.1002/yd.376.

EVALUATING YOUTUBE VIDEOS

- Crawley, A. M., Anderson, D. R., Wilder, A., Williams, M., & Santomero, A. (1999). Effects of repeated exposures to a single episode of the television program Blue's Clues on the viewing behaviors and comprehension of preschool children. *Journal of Educational Psychology, 91*(4), 630-637.
- Davidson, C., Given, L. M., Danby, S., & Thorpe, K. (2014). Talk about a YouTube Video in Preschool: The Mutual Production of Shared Understanding for Learning with Digital Technology. *Australasian Journal of Early Childhood, 39*(3), 76–83.
<https://doi.org/10.1177/183693911403900310>
- Gerbner, G., Gross, L., Morgan, M., Signorielli, N., & Shanahan, J. (2002). *Growing up with television: Cultivation processes*. In J. Bryant, & D. Zillman (Eds.), *Media effects: Advances in theory and research* (2nd edition) (pp. 43-67). Hillsdale, NJ: Lawrence Erlbaum.
- Herodotou, C. (2017a). Young children and tablets: A systematic review of effects on learning and development. *Journal of Computer Assisted Learning, 34*(1) pp. 1–9.
- Herodotou, C. (2017b). Mobile Games and Science Learning: A comparative study of 4 and 5 years old playing the game angry birds. *British Journal of Educational Technology, 49* (1), 6-16. DOI: 10.1111/bjet.12546
- Hillman, M., & Marshall, J. (2009). Evaluation of digital media for emergent literacy. *Computers in the Schools, 26*, 256–270. doi:10.1080/07380560903360186
- Huston, A. C., Wright, J. C., Rice, M. L., Kerkman, D., & St Peters, M. (1990). Development of television viewing patterns in early childhood: A longitudinal investigation. *Developmental psychology, 26*(3), 409.
- Izci, B; Jones, I; Ozdemir, T; Alktebi, L & Bakır, E. (2019). *Youtube and young children: Research, concerns, and new directions*. In edited book: *Children, families and*

EVALUATING YOUTUBE VIDEOS

technology in today's society: What challenges? Which paths? Chapter 7 pp.81-92.

Publisher: Lisbon School of Education.

Jones, T., & Cuthrell, K. (2011). YouTube: Educational Potentials and Pitfalls, *Computers in the Schools*, 28, 75-85, doi: 10.1080/07380569.2011.553149

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

Kostyrka-Allchorne, K., Cooper, N. R., & Simpson, A. (2017). The relationship between television exposure and children's cognition and behaviour: A systematic review. *Developmental Review*, 44, 19-58.

Lillard, A. S., Li, H., & Boguszewski, K. (2015). Television and children's executive function. *Advances in Child Development and Behaviour*, 48, 219-248.

Linebarger, D. L., & Walker, D. (2005). Infants' and toddlers' television viewing and language outcomes. *American Behavioral Scientist*, 48(5), 624–645.

Marsh, J. (2016) 'Unboxing' videos: co-construction of the child as cyberflâneur, *Discourse: Studies in the Cultural Politics of Education*, 37, 369-380, doi: 10.1080/01596306.2015.1041457

Marsh, J., Law, L., Lahmar, J., Yamada-Rice, D, Parry, B., Scott, F., Robinson, P., Nutbrown, B., Scholey, E., Baldi, P., McKeown, K., Swanson, A., & Bardill, R. (2019). *Social Media, Television and Children*. Sheffield: University of Sheffield.

Mayer, R. E. (2008). Applying the science of learning: evidence-based principles for the design of multimedia instruction. *The American Psychologist*, 63(8), 760–769. doi: 10.1037/0003-066X.63.8.760.

Miller, J. L., & Kocurek, C. A. (2017). Principles for educational game development for young children. *Journal of Children and Media*, 11(3), 314-329. <http://dx.doi.org/10.1080/17482798.2017.1308398>

EVALUATING YOUTUBE VIDEOS

- More, C. M., & Travers, J. C. (2013). What's app with that? Selecting educational apps for young children with disabilities. *Young Exceptional Children, 16*, 15–32.
doi:10.1177/1096250612464763
- Moussiades, L., Kazanidis, I., & Iliopoulou, A. (2019). A framework for the development of educational video: An empirical approach. *Innovations in Education and Teaching International, 56*, 217-228, DOI: 10.1080/14703297.2017.1399809
- Naigles, L., & Mayeux, L. (2001). *Television as incidental language teacher*. in Singer, D and Singer J, Handbook of children and the media, London: Sage Publications, pp. 135-152.
- Neumann, M. (2015). Young children and screen time: Creating a mindful approach to digital technology. *Australian Educational Computing, 30*(2). Retrieved from <http://journal.acce.edu.au/index.php/AEC/article/view/67>
- Neumann, M. M. (2017). Parent scaffolding of young children's use of touch screen tablets. *Early Child Development and Care, 188*(12), 1654–1664doi: 10.1080/03004430.2016.1278215.
- Neumann, M. M., Wang, Y., Xi, G., & Neumann, D. L. (2019). An evaluation of mandarin learning apps designed for English speaking pre-schoolers. *Journal of Interactive Learning Research, 30* (2), 167-193.
- Nikken, P., & Jansz, J. (2006). Parental mediation of children's videogame playing: a comparison of the reports by parents and children. *Learning, Media and Technology, 31*, 181–202.
- Papadamou, K., Papasavva, A., Zannettou, S., Blackburn, J., Kourtellis, N., Leontiadis, I., Stringhini, G., & Sirivianos, M. (2019). Disturbed YouTube for kids: Characterizing and detecting inappropriate videos targeting young children. *Computer Science:*

EVALUATING YOUTUBE VIDEOS

Social and Information Networks. Retrieved online from:

<https://arxiv.org/pdf/1901.07046.pdf>

Pempek, T. A., Demers, L. B., Hanson, K. G., Kirkorian, H. L., & Anderson, D. R. (2011).

The impact of infant-directed videos on parent–child interaction. *Journal of Applied Developmental Psychology*, 32(1), 10-19.

<http://dx.doi.org/10.1016/j.appdev.2010.10.001>

Plowman, L., McPake, J., & Stephen, C. (2010). The technologisation of childhood? Young children and technology in the home. *Children and Society*, 24(1), 63–74.

Rideout, V. (2017). *The Common Sense census: Media use by kids age zero to eight*.

San Francisco, CA: Common Sense Media.

<https://www.commonsensemedia.org/research>.

Rideout, V., and Robb, M. B. (2019). *The Common Sense census: Media use by tweens and teens, 2019*. San Francisco, CA: Common Sense Media.

<https://www.commonsensemedia.org/sites/default/files/uploads/research/2019-census-8-to-18-key-findings-updated.pdf>

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

Strasburger, V. C., Wilson, B. J., & Jordan, A. B. (2009). *Children, adolescents, and the media*. 2nd edition. Thousand Oaks, CA: Sage.

Ten Hove, P., & van der Meij, H. (2015). Like it or not. what characterizes YouTube's more popular instructional videos? *Technical Communication*, 62(1), 48-62.

Thomas, R.M. (2005). *Comparing theories of child development* (6th ed.). Belmont, CA: Wadsworth Thomson.

EVALUATING YOUTUBE VIDEOS

- Valkenburg, P. M. (2001). Television and the child's developing imagination. In D. G. Singer & J. L. Singer (Eds.), *Handbook of children and the media* (pp. 121-134). Thousand Oaks, CA: Sage.
- Veblen, K. K., Kruse, N. B., Messenger, S. J., & Letain, M. (2018). Children's clapping games on the virtual playground. *International Journal of Music Education*, 36(4), 547–559. <https://doi.org/10.1177/0255761418772865>
- Vygotsky, L. S. (1978). *Mind and society: The development of higher mental processes*. Cambridge: Harvard University Press.
- Wright, J.C., St Peters, M., & Huston, A. C. (1990). *Family television use and its relation to children's cognitive skills and social behaviour*. In J. Bryant (Ed.), *Television and the American family* (pp. 227-251). Hillsdale, NJ: Lawrence Erlbaum.
- Zimmerman, F. J., & Christakis, D. A. (2007). Associations between content types of early media exposure and subsequent attentional problems. *Pediatrics*, 120(5), 986-992.