Received: 25 February 2022 | Accepted: 16 September 2022

DOI: 10.1002/pan3.10416

RESEARCH ARTICLE

PEOPLE AND NATURE
BRITISH ECOLOGICAL SOCIETY

The viewer doesn't always seem to care—response to fake animal rescues on YouTube and implications for social media self-policing policies

Lauren A. Harrington[1] | Angie Elwin[2] | Suzi Paterson[2] | Neil D'Cruze[1,2]

[1]Wildlife Conservation Research Unit (WildCRU), Department of Biology, University of Oxford, The Recanati-Kaplan Centre, Abingdon, Oxfordshire, UK

[2]World Animal Protection, London, UK

Correspondence
Lauren A. Harrington
Email: lauren.harrington@biology.ox.ac.uk

Funding information
World Animal Protection

Handling Editor: Russell Hitchings

Abstract

1. Animal-related content on social media is hugely popular but is not always appropriate in terms of how animals are portrayed or how they are treated. This has potential implications beyond the individual animals involved, for viewers, for wild animal populations, and for societies and their interactions with animals.
2. Whilst social media platforms usually publish guidelines for permitted content, enforcement relies at least in part on viewers reporting inappropriate posts. Currently, there is no external regulation of social media platforms.
3. Based on a set of 241 'fake animal rescue' videos that exhibited clear signs of animal cruelty and strong evidence of being deliberately staged (i.e. fake), we found little evidence that viewers disliked the videos and an overall mixed response in terms of awareness of the fake nature of the videos, and their attitudes towards the welfare of the animals involved.
4. Our findings suggest, firstly, that, despite the narrowly defined nature of the videos used in this case study, exposure rates can be extremely high (one of the videos had been viewed over 100 million times), and, secondly, that many YouTube viewers cannot identify (or are not concerned by) animal welfare or conservation issues within a social media context.
5. In terms of the current policy approach of social media platforms, our findings raise questions regarding the value of their current reliance on consumers as watch dogs.

KEYWORDS
animal exploitation, animal welfare, corporate social responsibility, human-animal interactions, sentiment analysis, social media platforms

1 | INTRODUCTION

As of 2022, there were 4.6 billion active social media users, amounting to 58.4% of the current global population (Statista, 2022a). YouTube, currently the biggest online video platform worldwide (featuring a wide variety of corporate and user-generated content ranging from music and gaming videos to DIY and educational clips), had almost 2.6 billion users in January 2022 (Statista, 2022b) that streamed 694,000 h of video content every internet minute (Statista, 2022c). The global reach, popularity and influence of social media means that it has considerable potential to educate and inform people about the natural world and the wild animals within it (e.g. Pavelle & Wilkinson, 2020). Nevertheless, across platforms (and topics, not limited to the natural world), misinformation, fake and inappropriate content are commonplace (e.g. Menczer & Hills, 2020; White et al., 2018).

Animal-related content on social media (involving wild as well as domestic animals) is hugely popular (Hallinan et al., 2021) but sometimes portrays situations that may (intentionally or unintentionally) cause stress, injury, distress or death to the animals involved (including behind the scenes, e.g. poor captive conditions off-camera), or involve individuals that have been removed from threatened wild populations (and that may or may not be illegal under national legislation). None of these potential issues are easily discernible to non-expert viewers. Viewers of videos of slow lorises *Nycticebus* spp., for example, were more likely to 'like' videos where a slow loris displayed signs of stress or was kept in the light (slow lorises are nocturnal primates and so the presence of light is opposite to their behavioural needs and likely to negatively impact their health, Nekaris et al., 2015). The consequences of this type of misunderstanding can extend beyond the individual animal in the video because such imagery, and the resultant public response to it (especially when it is shared widely), may influence public perceptions, societal behaviour and social norms regarding appropriate treatment of wild animals (e.g. Riddle & MacKay, 2020; Thomas-Walters et al., 2020). Ultimately (whilst there is currently little research on the causal nature of these relationships), how wild animals are portrayed online may, for example, perpetuate misperceptions that certain wild animals are suitable pets (e.g. Freund et al., 2021; Leighty et al., 2015; Ross et al., 2011; Schroepfer et al., 2011) or that it is acceptable, humane or safe (for the animal or the human) to touch and hold wild animals (for example, in Wildlife Tourist Attractions, e.g. Carder et al., 2018; Osterberg & Nekaris, 2015; Van Hamme et al., 2021). There are dangers for people associated with imitation of video presenters (or vloggers; e.g. approaching a dangerous animal in a National Park too closely, or attempting to pick up a venomous species, e.g. Neme, 2010) and, for species consistently portrayed in the media as 'savage' or 'evil', there may be negative impacts on public support for necessary conservation, or human-wildlife coexistence, efforts (Cermak, 2021; Neme, 2010).

Recently, a particular type of online wildlife video, portraying an animal (either wild or domesticated) being rescued by a human from a predator attack, has attracted attention in the mainstream news media (Doward, 2020; Ewe, 2021; Knowles, 2021; Maron, 2021). Among a variety of videos available online showing people rescuing animals in a number of different scenarios, these particular videos appear to be filmed in Southeast Asia and most, but not all, are framed in the context of a reportedly 'primitive' life, entitled, for example, 'Primitive boy saves family chickens from python attack'. Various aspects of these short videos, including, among others, atypical predator behaviour, appearance of the same identifiable individual animals on multiple occasions, and varied camera angles, suggest that they are deliberately staged (Maron, 2021), and that they subject both prey and predator to considerable stress. Awareness of the prevalence of these 'fake animal rescue videos' (e.g. Social Media Animal Cruelty Coalition [SMACC], 2021) has raised questions regarding YouTube's management of this and other types of animal cruelty content.

Predator–prey interactions are commonplace in nature and an essential component of normal ecosystem function but are rarely observed directly by people. In attempting to capture such moments on film, wildlife filmmakers have been accused of audience deception (using captive animals and portraying unnatural behaviours; Boboltz, 2015), and commonly used practices, such as placing predator and prey animals together in an enclosure, are considered cruel and unacceptable in the wildlife film making industry (Neme, 2010; Palmer, 2011). Unlike wildlife documentaries (e.g. Pollo et al., 2009; Somerville et al., 2021), YouTube videos currently have no particular obligation to be informative or realistic (depending on the source and stated intent of the video). Even YouTube channels run by respected zoos tend to post entertaining rather than educational material (Llewellyn & Rose, 2021). YouTube does, however, have an obligation towards social responsibility and to protect against animal cruelty on its own platform. YouTube's Community Guidelines include policies against the depiction of violent or graphic content, and, for animals specifically, prohibit '*Content where animals are encouraged or coerced to fight by humans*' and '*Content featuring animal rescue that has been staged and places the animal in harmful scenarios*' (Google, 2021; updated in June 2021 to refer explicitly to fake animal rescue videos). Whilst YouTube does use people and machine learning to monitor content uploaded to its platform, the scale of the process means that detection of policy violations is also dependent on viewers reporting or flagging content that '*they find inappropriate*' (Google, 2021). There is currently no external regulation pertaining to the use of wildlife on YouTube or any other social media platform (Esmail et al., 2020).

Determining whether this type of self-regulated viewer-reliant system is likely to be effective in detecting and removing potentially harmful animal-related content online depends on an understanding of viewer perceptions of the videos, and their attitudes towards the treatment of the animals involved. To gain relevant insight, we investigate viewer response to a set of fake animal rescue videos posted on YouTube between 2018 and 2021. We focused specifically on those involving rescue from a predator attack portrayed by content providers in the context of 'primitive people', as a case study. To assess the potential influence of these videos, we used publicly available video metrics to quantify their reach and popularity, and the extent to which they were engaged with. We also used content and sentiment analysis to describe and quantify the perceptions and attitudes of viewers as reflected in their comments on the videos. Sentiment analysis (sometimes referred to as 'opinion mining') analyses people's opinions, attitudes and emotions from a text (Liu, 2015) and is able to quantify the relative polarity of a text (e.g. Fang & Zhan, 2015); it is used increasingly in various fields using social media data (e.g. Chauhan et al., 2021; Moloney et al., 2021; Nemes & Kiss, 2021; Piedrahita-Valdé et al., 2021; Poecze et al., 2019). Content analysis, in contrast (in its simplest terms and as used here), is used to determine the presence of particular words, themes or concepts in the text (see e.g. Neuman, 1997; Neuendorf, 2017; although the term itself has a broader meaning, cf. Macnamara, 2005) and thus (in the context of video comments) provides insight on which aspect of the videos viewers are commenting on, and, potentially, why (or what) viewers like or do not like (about) the video. We were interested in whether viewers responded negatively or positively to fake animal rescue videos; in particular, we were interested in whether viewer response to the videos, as illustrated by viewer comments, suggested that viewers were aware of the likely fake nature of the videos, or of the cruelty to the animals involved.

2 | METHODS

2.1 | Video selection

Fake animal rescue videos were initially collated using the search function in YouTube (www.youtube.com) and the search terms 'primitive man saves' and 'primitive boy saves' (in English only), in May, June and July 2021. An additional set of similar videos was obtained from a database held by Animals for Asia (AfA, www.asiaforanimals.com), from which we included all YouTube videos entered under the theme 'fake rescue'. The AfA database (which contains data on several online cruelty content themes) was compiled between July 2020 and August 2021 by members of SMACC and a team of volunteer researchers. All videos, from both sources, were manually screened for relevance; we included all those that involved direct interaction between two species (where one species attacks the other, hereafter the 'predator' and the 'prey'), and intervention by a human (we made no attempt to assess whether the rescue shown was genuine). We excluded videos that involved humans rescuing an animal in other types of scenarios, that did not include human intervention, or that involved only humans. Whilst we cannot be certain that all videos included were artificially set-up rather than genuine rescue attempts, all exhibited the same cinematic characteristics—multiple camera angles, a dramatic soundtrack, formulaic series of events, and an expression of surprise or shock by the human rescuer upon discovering the animal under attack. We included reposted videos on different YouTube channels because we were interested in the number of videos (or 'posts') available to potential viewers rather than the number of unique videos per se. All videos that conformed to our inclusion criteria were saved in an unlisted playlist on YouTube.

2.2 | Data extraction and analysis

Data were extracted from all videos in the playlist between 24 June 2021 and 02 August 2021 using the TUBER package (Sood, 2020) in R (version 4.1.0, R Core Team, 2021). For each video, we extracted the date the video was posted, the number of views, 'likes', 'dislikes' and comments. We also viewed each video and recorded the 'predator' and 'prey' species featured in the video (the animal attacking, and the animal being attacked, respectively), and the YouTube channel that the video was posted on. The 'Home' and 'About' pages of each YouTube channel that posted a fake animal rescue video were then manually checked and the number of subscribers, total number of views and channel location recorded. Predator and prey were identified to species level from video screenshots where possible, and corresponding data on threat status were collated from the IUCN Red List of Threatened Species (hereafter IUCN Red List; IUCN, 2021).

To characterise each video, we defined three parameters: 'exposure' (views), 'popularity' (likes) and 'engagement' (comments), quantified as the total number of views, the ratio of likes:views as well as dislikes as a percentage of total likes and dislikes, and the ratio of comments:views, respectively. Likes and comments were expressed as ratios to account for the effect of differences in exposure (number of views). To provide context (as in Harrington et al., 2019), we compared video metrics with the following published benchmark figures for YouTubers (referred to as marketing 'metrics of success'; Robertson, 2014): a like:view ratio of 0.04 (or 4 likes per 100 views), a comment:view ratio of 0.005 (5 comments per 1000 views) and a percentage of dislikes no greater than 40%. Correlations among channel and video metrics were assessed using Pearson's product–moment correlation coefficient. Throughout we refer to the posting of a video on a particular channel as a 'video' (regardless of whether the same video had been posted elsewhere) and, because viewers (and thus viewer response) will differ among channels, we treated these as independent units although there was some element of overlap due to repeat posting. A preliminary analysis of the influence of video content (taxa of the 'predator' and 'prey' featured in the video) and video channel (specifically the number of subscribers to the YouTube channel that the video was posted on), on the exposure and popularity of, and engagement with, resultant video posts is provided in Appendix S1.

2.3 | Content and sentiment analysis

To assess the attitudes and perceptions of viewers that commented on videos, we analysed the content and sentiment of video comments. We restricted our analysis to comments of those videos that had at least 1000 comments to ensure meaningful sample size and avoid bias due to over-representation of response to a small number of videos. Because there were only five videos that met this criterion, we analysed comments to each video separately. For each video, we identified the most frequently occurring words in the comments as an indicator of comment content (seeking specifically to identify content indicative of viewer perception, or awareness of fake video content and/or cruelty to the animals involved), and, using a lexicon-based approach, calculated sentiment scores based on the sentiment of those words (taking account of their frequency of occurrence). We counted strings of emojis in the text as 'words' because they are known to enhance, and modify the meaning of, the text (Novak et al., 2015) and their inclusion improves the accuracy of sentiment scoring compared to using the linguistic text alone (Gupta et al., 2020; Tian et al., 2017). This type of approach (based on words rather than sentences) is somewhat reductive, and is less sensitive than a sentence-level approach, but was considered most suitable in this case due primarily to the multiple languages within the text.

For each of the five videos with >1000 comments, we first extracted the full text of all comments (also using the TUBER package) and exported the text for each video to a text file prior to further analysis. A corpus was then created in R for manipulation and cleaning of the text using the text mining package TM (Feinerer et al., 2008; Feinerer & Hornik, 2018). Common symbols (/ @ ~ < > # & =) were removed, the text was transformed to lower case, and numbers and common English, French, Spanish and Portuguese stopwords were removed. We then carried out an iterative cleaning process to remove frequently occurring words that were in the title of the video (or that named the animals in the video), as well as profanities, words related to YouTube ('com', 'href', 'http/s', 'www', 'youtube', 'video', 'watch'), greetings ('hello', 'hi', 'bro') and words that had little meaning out of context or that added little to the understanding of the text (e.g. 'can', 'just', 'know', 'made'; a full list of words removed is in Appendix S2). We quantified the frequency of occurrence of words in the text and repeatedly inspected the 50 most frequently appearing words; the cleaning process was terminated when we reached a set of approximately 50 non-removable words that had been cited a minimum of three times in the text (we included more than 50 words where multiple words ranked 50th by frequency of occurrence). As foreign language words were highlighted during the cleaning process, they were translated using Google Translate (https://translate.google.co.uk/) to check their meaning and were removed or retained in the text (in their original language) in accordance with the same rules. Strings of emojis were also retained in their original form and were treated as if they were words. Within the final sets of 'cleaned' words, we identified the 10 most frequently cited for comparison among the five videos; the full set of 50 most frequently cited words for one of the videos, as an example, were visualised using a wordcloud, constructed using the WORDCLOUD package (Fellows, 2018). Before applying sentiment scores, foreign language words were translated (using Google Translate), and emojis were replaced with their word equivalents using the 'replace_emoji' function in the TEXTCLEAN package in R (Rinker, 2018, Appendix S2; single word equivalents were used regardless of the number of emojis in the string to avoid over-inflating sentiment scores). Sentiment scores were then applied to individual 'words' within each set of (*c*. 50) cleaned and translated words, using the SYUZHET package in R (Jockers, 2015) and the 'afinn' lexicon (Nielsen, 2011), and mean and total sentiment scores calculated for each video by summing the scores for each word, weighted by their frequency of occurrence. The afinn lexicon was deemed most suitable for this dataset because it was developed based on social media language (i.e. it provides a sentiment score for commonly used online slang terms, such as 'lol'). The system provides a score for each word between −5 (negative sentiment) and + 5 (positive sentiment), according to words and scores already compiled in the afinn lexicon; words that could not be scored received a neutral (0) score. To avoid possible misinterpretation due to discarding potential valence shifters or modifiers (e.g. the words 'do not', 'really') during the cleaning process, and reliance on single words, we checked significant word associations post hoc (using the TM package) for all frequently occurring words that could be misinterpreted (e.g. 'good' might have occurred in the full comment text as 'good' or 'not good'). For detection of word associations, we used a minimum correlation limit of 0.3 (where 1 means the words always occur together, and 0 means they never occur together) on the basis that 0.3 is generally cited as the minimum value for a correlation coefficient representing a 'moderate' linear relationship (Ratner, 2009).

We used a one-sample t-test to test whether mean sentiment scores differed significantly from zero, and an exact binomial test to test for significant orientation among the polarised words in the comment text (ie. departure from a 1:1 ratio in the relative proportion of negative and positive words, on the assumption that a polarised but balanced text would contain an equal number of positive and negative words). To test whether the overall popularity of these five videos differed from other fake animal rescue videos we used a two-sample t-test to compare mean % dislikes, we included only those videos with >100 likes and dislikes combined to reduce the possibility of bias due to low sample size, and used % dislikes rather than like:view ratios to reduce the influence of view count.

Finally, because this analysis only incorporated emojis that occurred in commonly used string lengths (in this case, strings of one to three, see Results) and did not fully represent overall emoji use in the text, we quantified emoji use (as individual symbols), separately, to provide as complete a picture as possible of the sentiment contained within the comments. Whilst it is preferable to combine text and emojis for sentiment analysis (e.g. Gupta et al., 2020), the frequency of occurrence of individual emojis (which can be very high, i.e. strings of >100 emojis) is not directly comparable to that of individual words. Conversely, some comments contain only a long string of emojis, so excluding them risks losing considerable information. For each of the five sets of video comments, we created a subset of comments that included only those that contained emojis, and then extracted the emojis using the EMOJI package in R (Hvitfeldt, 2022) to create a new 'text' file containing only emojis. A corpus was created in R, and the frequency of occurrence of each emoji was quantified using the TM package (as above). We retained and ranked all individual emojis that had appeared in the text at least 10 times (excluding emojis of the animals in the videos, as for the word analysis). Sentiment scores for each emoji were obtained from Novak et al. (2015), and an exact binomial test was used to test for departure from a 1:1 ratio in the relative proportion of negative and positive emojis taking account of their frequency of occurrence in the comment text (sentiment scores were not available for some of the newer emojis, in which case we assumed their polarity based on similarities with other emojis). For context, we quantified the number of comments that included emojis, and summarised the number of emojis per comment, for each of the five videos (using the EMOJI package).

2.4 | Ethical considerations

Our study involved covert observation (Thompson et al., 2021); however, we used only data (videos, video metrics, and comments) that were publicly available on the YouTube platform and posted on public accounts. Given that comments were posted as public feedback to the video creators/posters, we assume that commenters intended to contribute to the commentary associated with the video, and thus expected and wanted their comments to be read by others. Accordingly, and as in Townsend and Wallace (2016), use of these data did not require informed consent. In addition, commenters were not identified individually, and data deriving directly from comments are presented here only in aggregate form (single words and sentiment scores). We did not engage in deceptive practice and did not engage directly or otherwise with YouTube users (either those posting videos or commenting on videos). To protect the identity of the individuals posting the videos and those commenting on videos, web addresses and channel names are not reported here (see, e.g. Zook et al., 2017).

3 | RESULTS

3.1 | Number of videos

Using the specified search terms, and relevant links in the AfA database, we identified a total of 241 video posts portraying apparently fake animal rescues, between September 2018 and July 2021 (Appendix S3). Videos were posted on 160 different YouTube channels (three of which were the same channels renamed), located in all global regions, in at least 26 countries (predominantly the USA, where 23 channels were reportedly located; location was not available for 74 channels, Appendix S3). We did not attempt to quantify the number of unique videos, but note, based on the number of unique 'predator–prey' pairings (Appendix S3) and other differences in titles (e.g. man vs. boy; not reported for privacy reasons), that there were a minimum of 70 different videos posted, and that some videos were reposted 20–30 times.

3.2 | Video content

Videos depicted various domestic (e.g. chickens *Gallus gallus domesticus*, rabbits *Oryctolagus cuniculus*, cats *Felis catus*, and dogs *Canis familiaris*) and wild species in apparently natural habitats, where they were approached and attacked by a wild predator species before being rescued by a human who appeared to encounter the attack by chance, or by a person searching for, for example, their dog (Figure 1). There was no, or limited (often muffled), dialogue from the people in the video, and no voice-over narration. Wild species under attack ranged from a mouse (Muridae) to adult crocodiles (*Crocodylus* spp.) and monitor lizards (*Varanus* spp.), the latter often referred to in the video title as 'the family [animal]'...(e.g. 'the family crocodile'). In total, we identified at least 11 wild predator species and 16 wild prey species, in addition to a pack of domestic dogs as predators, and various domestic species under attack. Snakes (predominantly two python species: Burmese python *Python molurus bivittatus* and reticulated python *Malayopython reticulatus*) were the most frequently portrayed wild predator (Table 1). Other wild predators included several raptor species (see Table 1), the Bengal monitor lizard *Varanus bengalensis*, Siamese crocodile *cf. Crocodylus siamensis*, and (in one instance) a box turtle *cf. Cuora amboinensis*. With the exception of the Burmese python and changeable hawk eagle *cf. Nisaetus cirrhatus*, all 'predator' species were also depicted as the animal being attacked in different videos (albeit snakes relatively rarely, Table 1). Snakes attacking various fowl species (28.2% [n = 68] videos), followed by snakes attacking cats, dogs or rabbits (16.2% [n = 39] videos), or monitor lizards (14.9% [n = 36] videos) were the most commonly portrayed scenario (Figure 1). Six of the wild species identified in the videos are categorised as threatened or Near Threatened on the IUCN Red List: the elongated tortoise *Indotestudo elongata*, Siamese crocodile *Crocodylus siamensis*, amboina box turtle, Burmese python, Mekong snail-eating turtle *Malayemys subtrijuga* and Bengal monitor lizard (Table 1).

3.3 | Channels posting fake animal rescue videos

On average, the channels posting fake animal rescue videos appeared to be relatively small-scale (median number of subscribers = 23.5, median number of channel views = 8572, Appendix S1), belonging to individuals or small local groups. Thirty-five percent of channels (n = 57 of 160) had fewer than five subscribers and a few hundred channel views; 21.8% (n = 35) of channels had no subscribers. However, videos were also posted on a few large popular accounts with more than a million subscribers (n = 3; 6 channels had >200,000 subscribers), some of which (n = 5) received over a hundred million channel views (correlation between channel subscribers and channel views, Pearson's product–moment correlation, r = 0.97, p < 0.001).

Individual channels posted between 1 and 13 fake animal rescue videos (most posted 1–4 videos, n = 132 [83.5%] posted one video; four channels posted >4 videos, n = 9, 10, 13 and 13, respectively).

3.4 | Video metrics

Video metrics (indices of exposure, popularity and engagement) were all strongly left-skewed but covered a broad range such that the vast majority of videos scored relatively low on all three measures whilst a small number of videos scored particularly high (summary data

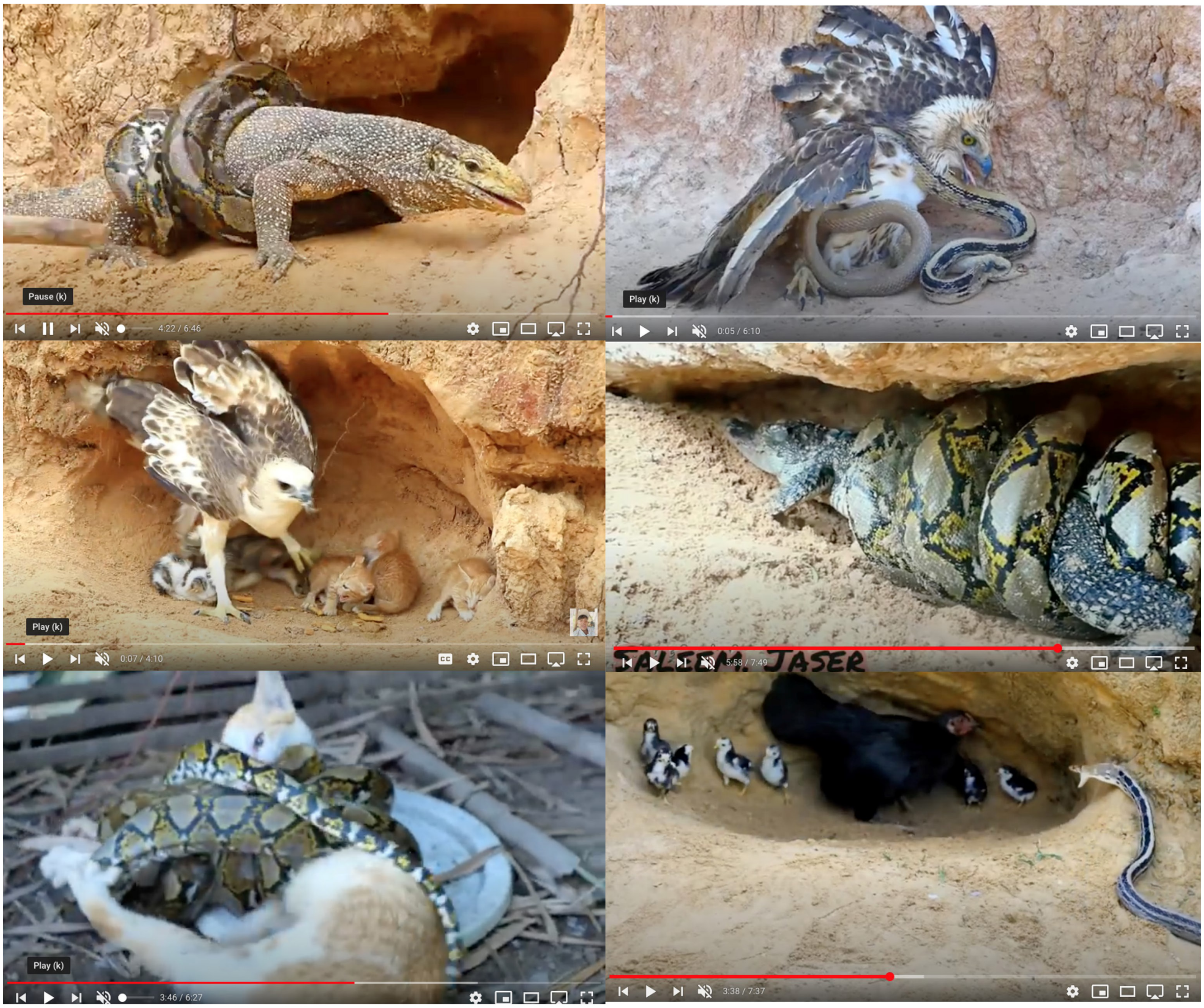

FIGURE 1 Screenshots of 'fake animal rescue' videos downloaded from YouTube. Videos were posted between September 2018 and July 2021, and screenshots taken between June and August 2021. All but two of the videos included in the study had been removed by YouTube at the time of writing (see Discussion).

in Appendix S1). Over half (61.8%, $n = 149$) of the videos received fewer than 100 views (minimum = 7, median = 58, $n = 241$), but five (2.1%; one of which was posted only two months prior to the study) received more than one million views and one (posted in December 2018) received over 100 million views (these are the same five videos that contained sufficient comments for content analysis, below). Similarly, 78.4% ($n = 189$) of videos failed to meet the benchmark like:view (popularity) value of 0.04 or more (median = 0.008, or eight 'likes' per 1000 views), and 84.2% ($n = 203$) videos failed to meet the benchmark comment:view (engagement) value of 0.005 or more (median = 0; 66.8%, $n = 161$, videos received no comments). High values for popularity and engagement were, in most case, based on low numbers of views: only three of the 52 videos with a like:view ratio of >0.04, and only one of the 38 videos with a comment:view ratio >0.005, received >1000 views. Similarly, high percentage (>40%) dislikes were, in most cases, based on low response rates: only two of the 32 videos that received more than 40% dislikes, received >100 likes and dislikes combined (and even for these videos, maximum percentage dislike was only 46%). More than a quarter of all videos (27.8%, $n = 67$, all of which had <1000 views) received no likes or dislikes.

Notably, whilst the five videos with high exposure rates received the highest absolute numbers of likes, all scored low on *relative* popularity (like:view ratios 0.001–0.004, i.e. <median and benchmark scores). Conversely (and in common with the majority of videos in the study), there was no evidence that these five videos were *actively* disliked (none received 40% or more dislikes, as a percentage of total likes and dislikes). Across all videos, the number of video likes was strongly positively correlated with the number of views (Pearson's correlation coefficient $_{[\text{likes vs. views}]}$: $r = 0.959$, $p < 0.001$), but this was also true for the number of dislikes (Pearson's correlation coefficient $_{[\text{dislikes vs. views}]}$: $r = 0.939$, $p < 0.001$).

TABLE 1 Species identified in videos. Conservation status is based on the IUCN Red List of Threatened Species categories (version 2021-2; IUCN, 2021): CR = critically endangered, EN = endangered, VU = vulnerable, NT = near threatened, LC = least concern. Native refers to SE Asia (Y = yes, N = no).

Class	Family	Species (common)	Species (latin)	Conservation status	Native	Number of videos species appears in	
						as predator	as prey
Reptilia	Pythonidae	Burmese python	*Python molurus bivittatus*	VU	Y	69	–
		Reticulated python	*Malayopython reticulatus*	LC	Y	73	1[a]
		Unidentified	–	–	–	31	–
	Colubridae	Radiated rat snake	*cf. Coelognathus radiatus*	LC	Y	26	3
	[*Serpentes* spp.]	–	–	–	–	2	1
	Varanidae	Bengal monitor lizard	*Varanus bengalensis*	NT	Y	11	33
		Asian water monitor	*Varanus salvator*	LC	Y	–	3
		Unidentified	*Varanus* spp.	–	–	–	4
	Crocodylidae	Siamese crocodile	*cf. Crocodylus siamensis*[b]	CR	Y	3	8
		Saltwater crocodile	*Crocodylus porosus*	LC	Y	–	1
		Unidentified	*Crocodylus* spp.	–	–	2	10
	Testudinidae	Elongated tortoise	*Indotestudo elongata*	CR	Y	–	13
	Geoemydidae	Mekong snail-eating turtle	*cf. Malayemys subtrijuga*	NT	Y	–	1
		Amboina box turtle	*Cuora amboinensis*	EN	Y	1	1
	[*Testudines* spp.]	–	–	–	–	–	7
Amphibian		Asian Common Toad	*cf. Duttaphrynus melanostictus*	LC	Y	–	1
Aves	Accipitridae	Short-toed Snake eagle	*cf. Circaetus gallicus*	LC	Y[c]	7	2
		Changeable hawk eagle	*cf. Nisaetus cirrhatus*	LC	Y	2	–
		Crested serpent eagle	*cf. Spilornis cheela*	LC	Y	1	1
		White-bellied sea eagle	*cf. Haliaeetus leucogaster*	LC	Y	5	–
		Unidentified	–	–	–	3	1
	Strigidae	Buffy fish owl	*Ketupa ketupu*	LC	Y	–	4
		Collared scops owl	*Otus cf. lettia*	LC	Y	–	2
		Unidentified owl	*Tyto* spp.	–	–	4	2
	Phasianidae	Domestic chicken	*Gallus gallus domesticus*[d]	(LC)	(Y)	–	57
	Anatidae	Domestic duck/goose	Various	–	–	–	23
	Columbidae	Pigeon/dove	*cf. Columba livia domestica*	–	–	–	4
Mammalia	Canidae	Domestic dog/s	*Canis familiaris*	–	–	1	20
	Felidae	Leopard cat	*Prionailurus bengalensis*	LC	Y	–	1
		Domestic cat	*Felis catus*	–	–	–	20
	Leporidae	Domestic rabbit	Provence unknown	–	–	–	9
	Muridae	Mouse	Unknown	–	–	–	1
	Bovidae	Domestic goat	*Capra hircus*	–	–	–	1
	Suidae	Pig	*Sus scrofa (domesticus)*	(LC)	(Y)	–	6

[a]Questionable as to which species was predator and which was prey in this video.

[b]Video screenshots did not always provide clear images of the nape region to allow for observation of the paired post-occipital scales that are a distinguishing feature of this species, in most images the animal was considered most likely to be *C. siamensis*.

[c]Resident in India and Myanmar, vagrant in Cambodia; CITES Appendix II under listing for all falconiformes.

[d]Some images were tentatively identified as the wild red junglefowl *Gallus gallus* but it was not possible to confirm.

3.5 | Comment content and sentiment

The 10 most frequently occurring words (hereafter 'top ten') in the comments to all five videos included a number of potentially conflicting, primarily English-language, words (Table 2), suggesting widely differing attitudes and perceptions among commenters. Words such as 'nice', 'love', 'like', 'good', 'super' and 'ótimo' ('excellent' in Portuguese), for example, suggest enjoyment of the video,

TABLE 2 Most frequently cited words in the five most commented-on 'fake animal rescue' videos; showing the 10 most frequently occurring words ('top ten') for all words that occurred at least 10 times (more than 10 words are shown where there were multiple words that occurred the same number of times in tenth place). *n* is the number of comment lines analysed; Freq is the number of times the word occurred in the entire document and was not necessarily the same as the number of comments that the word occurred in. Only videos with at least 1000 comments were included. Percentage 'likes' and 'dislikes' provides further context on the views of all viewers since not all viewers will have commented on the video.

Video 8[a] (*n* = 1161)		Video 14[b] (*n* = 1877)		Video 143[c] (*n* = 4810)		Video 144[d] (*n* = 2903)		Video 145[e] (*n* = 1332)	
Word	Freq	Word	Freq	Word	Freq	Word	Freq	Word	Freq
Nice	31	Fake	65	God	245	Ears	367	Fake	152
Wow	17	Good	47	Life	242	Fake	349	Nice	68
Love	14	Like	35	Spirit	113	Like	238	Good	51
Good	12	Nice	34	Motherless	112	Rescue	236	Like	21
Fake	10	Wings	30	Nice	92	Stop	168	Scripted	21
Like	10	Mother	24	Help	73	Hold	150	Save	20
—		Fly	21	Like	71	Right	147	Acting	19
—		Scripted	20	Good	70	Poor	139	Wild	17
—		Staged	19	Fake	68	Cruelty	135	Super	17
—		Attack Real Love	18	Family	64	Eat	129	الله ótimo	16
Likes	Dislikes	Likes	Dislikes	Likes	Dislikes	Likes	Dislikes	Likes	Dislikes
76.2%	23.8%	82.4%	17.6%	74.6%	25.4%	66.7%	33.3%	79.6%	20.4%

Note: Translations: الله = Allah (Arabic), ótimo = excellent (Portuguese).

[a]Video 8: Video showing a boy rescuing a Bengal monitor lizard (referred to in the video title as a Komodo dragon) from a reticulated python attack, posted 29/04/2021 with 26 million views as of June 2021.

[b]Video 14: Video showing a boy rescuing domestic chickens from a changeable hawk eagle attack, posted 01/012/2019 with 28 million views as of June 2021.

[c]Video 143: A previous posting of video 8 on a different YouTube channel, posted 26/12/2018 with 100 million views as of July 2021.

[d]Video 144: Video showing two boys rescuing a domestic rabbit from a Burmese python attack (referred to in the video title as an anaconda attack), posted 15/09/2018 with 4 million views as of July 2021.

[e]Video 145: Video showing a man saving a white pigeon (Columbidae) from a radiated rat snake attack, posted 09/02/2019 with 17 million views as of July 2021.

whilst other viewers clearly identified the video as a fake. 'Fake' appeared in the top 10 words in all five videos, and in video 145 was the most frequently occurring word (cited 152 times in 1332 comments). The word 'scripted' also appeared in the top 10 words in video 14 and 145, together with 'staged' in video 14 and 'acting' in video 145. Some viewers recognised the cruelty inflicted on the animals involved: the word 'cruelty' appeared in the top 10 words in video 144 and in the full set of all most frequently occurring words in the comments to video 14, as well as acknowledgement that the bird's wings were clipped ('wings', 'fly', 'clipped'), and that the bird might have been 'injured', in video 14 (Table 2, Appendix S4). The words 'ears' and 'hold' in the top 10 words (cited 367 and 150 times, respectively, in 2903 comments) in video 144 (Figure 2; Table 2), similarly, related to holding the rescued rabbit by the ears. The words 'brave' and 'save' (Figure 2) suggest that a proportion of viewers that recognise the potential for animal suffering in the scenarios shown, were unaware that the situation shown was staged. The word 'report', in the context of commenters saying that they had reported the video for animal cruelty or that they were encouraging others to do so, appeared in the most frequently used words in the comments to video 143, 144 and 145 (Figure 2; Appendix S4).

FIGURE 2 Wordcloud illustrating the most frequently occurring words in the comments to one of the five most commented-on 'fake animal rescue' videos (Video 144). Font size is relative to the frequency at which each of the words appears in the comments; stopwords and profanities have been removed (see Methods).

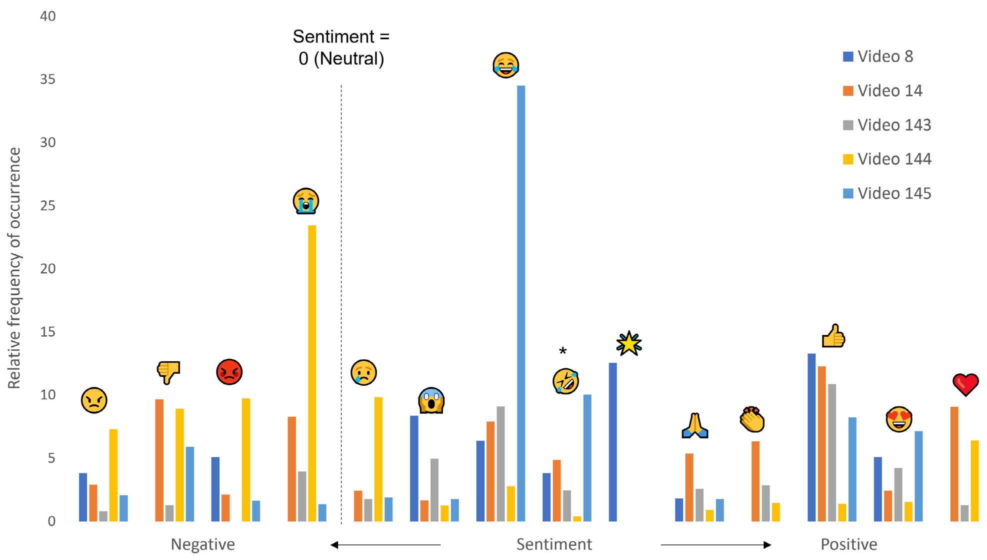

FIGURE 3 Relative frequency of occurrence of emojis (depicted at the top of the bars) as used in the comments to each of the five most commented-on 'fake animal rescue' videos (see key). Emojis ranked from the most negative (left) to the most positive (right), based on Novak et al. (2015) emoji sentiment scores; exact sentiment scores not shown. Emojis marked with an asterix are new and are not included in Novak et al. (2015) sentiment scores, their relative position on the negative–positive scale is estimated based on apparent similarities with other emojis. Only those emojis with a relative frequency of occurrence of at least 5% in at least one video are shown (others are included in Appendix S4).

Languages detected within the most frequently occurring words included Arabic, Estonian, Hindi, Indonesian, Portuguese, Spanish, Swedish and Vietnamese (Appendix S4).

Per video, between 8.5 and 18% of comments contained emojis, with a maximum of 53 to 264 emojis per comment (median number of emojis per comment for all five videos = 3.0). Five different emojis ('grinning face', 'thumbs up', 'face with tears of joy', 'rolling on floor laughing' and 'face screaming in fear'; emoji word equivalents from the Full Emoji List, v14.0, Unicode CLDR Project; https://unicode.org/emoji/charts/full-emoji-list.html) appeared among the most frequently occurring 'words' in the comments to video 8, 14 and 145 in strings of one to three (Appendix S4). An additional 82 emojis were identified as occurring in the comment text a minimum of 10 times when counted as individual symbols. As for the most frequently occurring words, emojis suggested both positive and negative emotions (Figure 3).

The word 'like' was associated with 'look/s/ed' in videos 143, 8 and 145 (with correlations 0.49, 0.5, and 0.38, respectively), indicating that in at least some cases, in these videos, 'like' was used in the context of 'looks like' and not in the positive use of the word. Similarly, in video 8, 'love' was associated with 'deserve' (correlation 0.68) suggesting that viewers were referring to the animals 'deserving love' rather than 'loving' the video. However, word associations also confirmed the positive intent of other frequently occurring words: for example, in video 8, 'nice' and 'good' were associated with 'very' (correlations 0.5, and 0.45, respectively), and, in video 14, 'love' was associated with 'it' (as in 'love it', correlation 0.36).

Overall, there was considerable heterogeneity in the sentiment of the most frequently occurring words and emojis in the comment text within and among videos: individual word scores for each video ranged between −3 or − 4 and + 3 or + 4 (i.e. comments for all five videos contained both negative and positive words), and mean sentiment scores for the five videos ranged between −0.28 and + 1.53 (Figure 4). Four of the videos had comments that were, on average, positive (video 8, 14, 143 and 145; mean sentiment score > 0, t-test, $p < 0.001$ in all cases, one-tailed test, Figure 4) and comprised (among the most frequently occurring words and emojis) predominantly positive words and symbols (64.1–85.0% of polar words [34.0–61.4% of all words] and 74.1–89.0% of emojis were positive; binomial test, $p < 0.001$ in all cases, one-tailed test, Figure 3). The reverse was true for one of the videos, albeit less strongly: video 144 had a mean sentiment score statistically significantly lower than zero (t-test, $p < 0.001$, one-tailed test, Figure 4) and comprised (among the most frequently occurring words and emojis) predominantly (57.3% of polar words [30.7% of all words], 66.4% of emojis) negative words and symbols (binomial test, $p < 0.001$ in all cases, Figures 2 and 3).

Mean percent dislike for these five videos did not differ significantly from that of all videos (mean for the five videos analysed = 24.1, SD = 6.0, mean for all videos with >100 likes and dislikes combined = 24.6, SD = 12.5, n = 20; two sample t-test: t = −0.13, df = 13.2, $p = 0.899$).

4 | DISCUSSION

Our study provides unique insights into viewer attitudes when watching fake animal rescue videos which may help inform an appropriate policy response to online animal cruelty and define potential associated human behaviour change initiatives. We examined viewer response to 241 videos that exhibited clear signs of animal cruelty

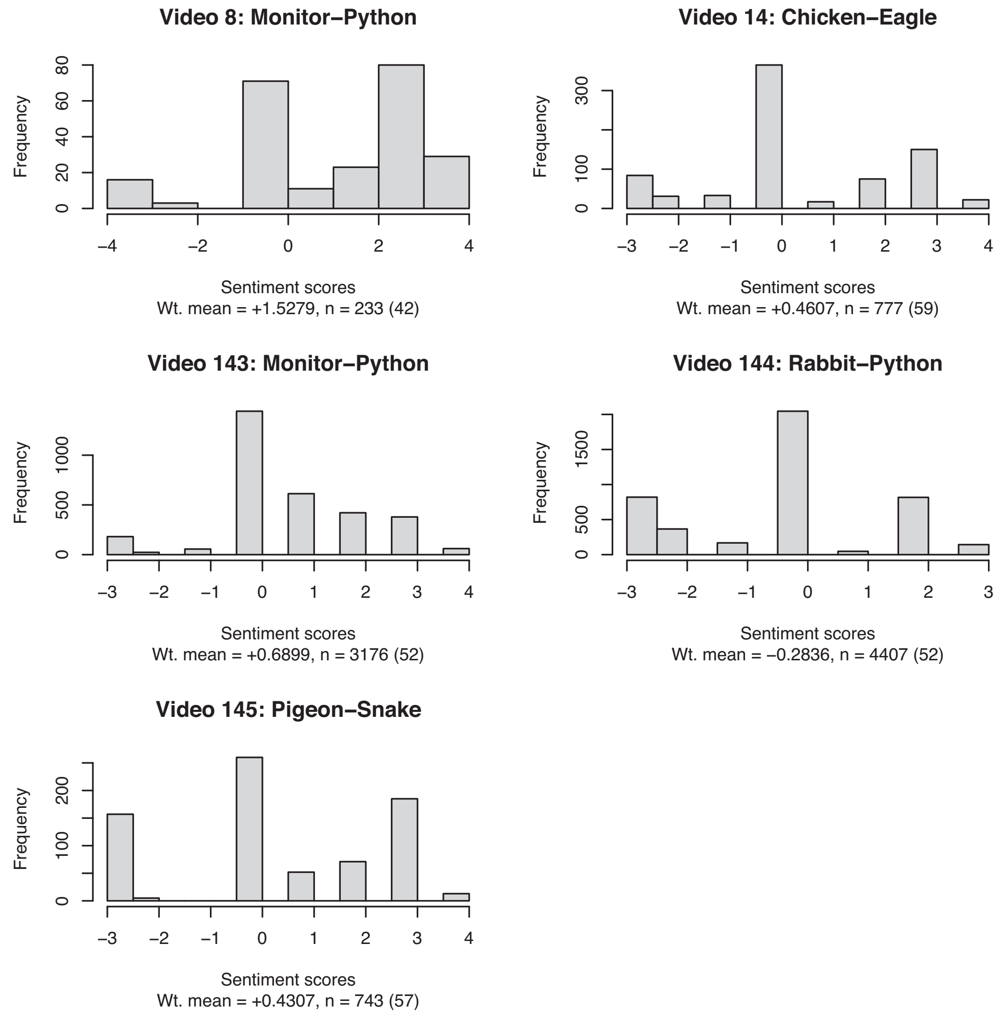

FIGURE 4 Distribution of sentiment scores for the 'top 50' most frequently cited words in the five most commented-on 'fake animal rescue' videos, summarised by video. Sentiment scores were assigned using the SYUZHET package in R, and the afinn lexicon, with a maximum possible range of −5 to +5 (see Jockers, 2015). Wt. mean = mean sentiment score, weighted by the frequency of occurrence, n = total number of word occurrences (number of unique words used in brackets). Monitor = Bengal monitor lizard, python = reticulated python in videos 8 and 143, and Burmese python in video 144, eagle = changeable hawk eagle, snake = radiated rat snake, chicken, rabbit, and pigeon were domestic or feral species (see footnotes to Table 2).

(although notably only two videos were preceded by a graphic content warning, Appendix S3), and strong evidence of being deliberately staged, but found little evidence that viewers disliked the videos. Overall, viewer response appeared to be mixed in terms of awareness of the fake nature of the videos, and attitudes towards the welfare of the animals involved. Stated country locations of the channels that posted the videos, and the variety of languages detected in viewer comments, suggest that interest in these types of videos is global.

We did not attempt to formally assess the welfare state of the animals in the videos to avoid potentially inaccurate and/or inconsistent assessments that might arise as a result of variable camera angles and incomplete coverage (i.e. not all parts of the animal were always shown). However, we note that many videos depicted the prey animals in the grip of the predator before being 'rescued' (most notably in attacks by pythons, Figure 1), a situation that is undoubtedly extremely stressful for the animal being attacked, and one that may have long-term negative physiological and behavioural impacts (Zanette & Clinchy, 2020). The attacking animals were also often shown being bitten, pecked, and scratched by the 'prey', and prodded with sticks and otherwise roughly handled by the human during the rescue. The snakes depicted in videos often had bleeding snouts and damaged rostral scales, and the raptors had missing feathers and clipped wings. Various atypical behaviours, particularly by the predator in response to the intervening human (i.e. no attempt to escape or attack the human), suggest that the wild predator and prey species (as well as the domestic species under attack) were likely maintained in captivity, but it was not possible to determine whether the wild animals had been captive bred or were originally sourced from the wild.

During the course of the study, all but two of the 241 videos included in the study were removed by YouTube following a request to investigate by National Geographic (Maron, 2021) and SMACC coalition members. Forty-four of the 160 channels that had posted videos were also terminated (38 of them due to '*multiple or severe violations of YouTube's policy on violence*'; Appendix S3).

4.1 | Video exposure, popularity and engagement

Precisely what defines a 'viral' video is complex and ill-defined (France et al., 2016, and references therein); however, on the basis of number of views and length of time that the videos had been available, at least two of the video posts in our study might be considered to have 'gone viral' (a coarse definition of a viral video is one that receives at least five million views per week; Nalts, 2011). These two videos (Video 8 and 143, see Table 2) were in fact the same video (showing a monitor lizard being attacked by a python), posted on different channels in two different years (Appendix S3). What might have driven the apparently viral behaviour in some of these videos but not others is not clear. Our preliminary analysis of the effect of taxa (predator or prey) and the number of channel subscribers on video metrics (details in Appendix S1) failed to detect any discernible effect due to either, and although the number of channel subscribers was correlated with the number of channel views, there was no evidence of a consistent relationship between the number of channel subscribers and the number of views of individual videos posted on the channel (cf. Bakshy et al., 2011). As in other studies of YouTube videos (e.g. Bärtl, 2018; Harrington et al., 2019; Morgan et al., 2014) and social media posts on alternative platforms (e.g. Harrington et al., 2018), most video posts included in our study received little attention in terms of either exposure (views) or popularity (likes). Nevertheless, five videos received over a million views and one was viewed 100 million times, a pattern also in keeping with previous studies whereby the majority of views are restricted to a small minority of videos or posts (e.g. Bärtl, 2018).

Video popularity was difficult to ascertain. Whilst popularity scores appeared to be low for most videos, there was also little evidence of widespread dislike of the videos. Further, although response rates (both likes and dislikes) were positively correlated with video views, even the most viewed videos (those with >1 million views) received responses (both positive and negative) lower than benchmark values. The same was true of engagement: few videos received any comments at all. Whether or not this result indicates genuine ambivalence (i.e. most viewers neither liked nor disliked the video), or simply lack of motivation to respond to the video, is not clear. In this context, for example, it is well known that viewers are more likely to comment on videos or other social media posts that they feel strongly about (e.g. Berger & Milkman, 2012; France et al., 2016).

4.2 | Perceptions and attitudes of viewers

4.2.1 | A contradiction of terms

Based on the most frequently cited words in video comments, it appears that sentiment contained within the comments to four of the five videos we examined was, on average, positive. This is reflected by the prominent appearance of clearly positive words such as 'good', 'nice' and 'love' (Table 2; alebit recognising that not all occurrences of these words were positive). One of the videos received an overall negative score, perhaps because it featured an attack on a domestic rabbit whereas the other four videos featured attacks on a monitor lizard, domestic chickens, and a pigeon (cf. the 'cute response'; Borgi & Cirulli, 2016, and references therein). However, mean absolute sentiment scores, for all videos, were relatively low (< |1| for four of the videos), and the distribution of scores, which ranged between highly positive and highly negative, indicates the diverse nature of individual comments for all five videos (Figure 4). Comments for all videos contained both positive and negative words, and comments for even the most negative video (based on mean sentiment score)—Video 144—included the word 'like' (associated with the word 'look', as in 'looks like', but only in approximately half of occurrences). The contradictory nature of these single terms is particularly clear in the comments to Video 144 (Figure 2), which included the words 'rescue', 'save' and 'kind' but also 'abuse', 'abusing', 'hurt' and 'kill', among the most frequently used words. In short, although mean scores were predominantly positive (albeit low), the full picture of the sentiment contained within the comments is revealed only by examining the distribution of all scores, which, for all videos, was approximately normally distributed and ranged from highly negative to highly positive.

4.2.2 | Fake or funny?

Comment content suggested that at least some viewers of all five videos recognised the fake nature of the material. Beyond the word 'fake' itself (referring to something that is not genuine, or is presented as something that it is not, Oxford Languages, 2022), the words 'staged' (an event or situation that is planned, organised or arranged in advance, and that is intended to seem otherwise, Oxford Languages, 2022), 'acting' (performing a fictional role, Oxford Languages, 2022), and 'scripted' (referring to something that was written or arranged in advance, Cambridge Dictionary, 2022) also suggest some level of perception of 'pretence', and particularly 'pre-planning', by the video-makers. Collectively, these terms (all of which appeared among the top 10 most frequently cited words in one or more of the five most commented on videos, see Table 2) suggest that these videos were perceived by viewers to be showing something other than a natural series of events. Given the almost complete lack of dialogue in the videos or of narration added post-production, the word 'scripted' (which would normally refer to a speech, a discussion, or to words spoken aloud) appears misplaced but is presumably (bearing in mind that we are unable to infer what viewers actually meant by the comments that they made) used here in the context of something that was arranged in advance (as in 'staged'). Notably all three terms ('staged', 'acting', and 'scripted') are pre-production terms that presumably refer to the animals (most likely captive animals) being placed in a particular location, and/or in close proximity to one another, and the person/people 'pretending' to encounter the situation unexpectedly (other similar terms used in the comments included 'set', 'actor/s', and 'script', Appendix S4)

– these differ from post-production terms such as 'edit' or 'edited' (that appeared either rarely, or not at all, in the comments for these five videos, Appendix S4). Whilst post-production terms might also be suggestive of viewer deception (for example, when used to alter the natural sequence of events for various dramatic purposes, cf. Somerville et al., 2021) they involve manipulation of the recorded video rather than of the animals during filming, and thus lack the cruelty element of the latter (i.e. both are 'fake', but the latter is also 'cruel').

Some viewers (albeit for only some of the videos) clearly recognised the animal cruelty involved. Other viewers appeared to enjoy the videos, were impressed (the word 'wow' appeared frequently in three of the videos, and in Arabic as well as in English, Appendix S4), and/or found the content humorous (indicated by 'lol' ['laugh out loud'] in the comments, and the 'rolling on floor laughing' emoji). Emoji use, in particular, reflected an element of anger (seen in the use of the 'angry face' emoji) but also the enjoyment (and apparent humour) expressed by many of the viewers (seen in the use of various heart symbols, 'clapping hands' and 'laughing' emojis, Figure 3). Whilst this is consistent with the contradictory nature of comment sentiment described above, it is also true that neither 'fakeness' nor cruelty necessarily precludes humour or awe. People have long been drawn to watching animal fights ('blood sports' e.g. Davis, 2021), and there is an extensive literature (beyond the scope of this paper) exploring why people enjoy and partake in such pastimes (e.g. Barber, 2022; Iliopoulou & Rosenbaum, 2013; Kalof & Taylor, 2007, and references therein). Viewers may thus be impressed with the actual fight depicted; alternatively, they might be impressed with the apparent heroism shown by the human rescuers (a version of biopower discussed by von Essen et al., 2021), or simply be impressed with the theatrical ability of the video producers in producing such a scene. Equally, the apparent humour expressed by some of the viewers does not necessarily suggest that viewers find the cruelty funny per se, but might be directed at the obviously fake nature of the videos (although even the latter suggests a disregard of the well-being of the animals involved). Without directly questioning viewers, it is not possible to determine the precise meaning of their comments—this is particularly the case for emojis, where, for example, a 'laughing' emoji (or even an 'angry face' emoji) might be meant ironically, or to tone down the seriousness of the disapproval expressed in the text (Tian et al., 2017). A detailed more in-depth study would be required to better understand why viewers enjoy these and/or any other types of videos that portray animal suffering; here, the question posed was simply whether viewers enjoyed the videos, and whether they recognised that they were either fake or cruel. We tentatively conclude that enjoyment (interpreted broadly) was indeed a predominant theme in video comments (based on both word and emoji use) although we cannot be certain whether viewers who expressed enjoyment genuinely found the videos funny (likeable, or loveable) or whether their reaction was intended to be ironic. Importantly, neither case (whether due to lack of awareness or of simply not caring) suggests any particular concern for animal cruelty.

4.2.3 | Limitations

Beyond the complexities outlined above regarding inferring the intended meaning of social media comments, our study has some limitations associated with the nature of the dataset used. First, comments to YouTube videos are not a random sample of viewer response because commenters self-select, and they may not be fully independent because individual commenters may be responsible for multiple comments, and are potentially influenced by comments made by other viewers (through the formation of echo chambers; Cinelli et al., 2021). Consequently, we cannot be sure that the sentiments detected in video comments are representative of the views of all viewers, or of all fake animal rescue videos. That there were more than 1000 comments to the five videos that we were able to analyse in detail and no more than 36 comments for any other videos in our study set suggest, for example, that there may have been something specific about these five videos that prompted viewers to comment, or that these particular sets of viewers held stronger opinions (and thus were more likely to comment) than others. However, with the exception of view and comment counts, there were no apparent differences in video metrics that might point to a difference in strength of opinion between the viewers of these five videos compared with those of other videos in the dataset. Secondly, our sample of videos contains some repetition due to the occurrence of the same videos shown on different channels, including two of the five videos for which we analysed comments (video 8 and video 143). Across all videos, the high level of variability within video metrics suggests that repetition within the sample had minimal (albeit unknown) effect. We did not test statistically for differences between video 8 and video 143 but observations of the most frequently cited words (Table 2) and distribution of sentiment scores (Figure 4) for the two videos do not suggest any greater similarity than among the other videos. Nevertheless, for future studies of this kind, it would be useful to explore the extent to which the same videos shown on different channels (or indeed on different social media platforms) represent independent units for analysis.

Sampling limitations aside, we reiterate that caution is required in drawing conclusions based on precise sentiment scores (due to uncertainties and complexities in interpretation; e.g. Puschmann & Powell, 2018); however, the large sample size that we were able to obtain (a total of 13,066 comments), together with the relative consistency (in terms of diversity and conflicting views) among videos and between words and emojis, lends weight to the validity of the overall patterns revealed.

4.3 | Policy implications

All of the videos included in this study violate YouTube's new (2021) community guidelines. All but two have now been removed. However, in addition to the two remaining, an ad hoc search of YouTube using our original search terms in November 2021 located a further 22 fake animal rescue videos (all of which had been posted prior to the YouTube removals) suggesting that the mechanisms currently used by YouTube to detect and remove these videos are not sufficiently thorough. Further, actions taken by YouTube thus far appear to be driven largely by targeted NGO advocacy efforts (SMACC, 2021), and attention in the public media (Maron, 2021; Talamo, 2021). That at least 13 of the videos had been online for two years or more prior to the campaign, and videos posted recently had received tens of millions of views without being removed, suggest that YouTube's self-regulatory system and their reliance on viewers is generally ineffective in policing against this type of online animal cruelty. As of October 2021, the platform is the subject of a lawsuit filed by Lady Freethinker (LFT) for failing to enforce their own rules against fake animal rescues and other forms of online animal abuse (e.g. Beals, 2021).

Although the videos included in this study had, collectively, been viewed almost 180 million times, the overall lack of viewer response suggests that most viewers are unlikely to report or 'flag' a video. Even if we assume that a viewer that strongly dislikes the video will report it, our results demonstrate that 'dislike' was not a universal response. Perceptions and attitudes towards fake animal rescue videos appeared to be highly variable and many viewers (perhaps unaware of the welfare costs for the animals in the videos, or that the rescue was fake) liked the videos, were impressed by, and/or enjoyed the content. From a wildlife conservation perspective, none of the viewers of Video 8, 143 or 144 recognised (presumably due to lack of expert knowledge), or thought worthy of mentioning, the threatened status of the Bengal monitor lizard or the Burmese python shown in the video, or (where the videos were recognised as fake) questioned the legality of keeping either of these species as pets, or using them for entertainment. These are complex issues that are affected by taxonomic details, provenance of the individual, geographic location, and relevant local, national and international legislation, suggesting the need for expert assessment.

YouTube is not the only social media platform to rely, at least in part, on viewers to police content (Esmail et al., 2020). But our findings suggest that alternative pro-active approaches may need to be adopted by social media platforms if they are to adequately protect wild animals in line with their own community guidelines and any relevant national legislation. With respect to the latter, the UK's draft Online Safety Bill (currently passing through parliament; Bill CP 405, HMSO, 2021) and the EU's proposed Digital Services Act (EC, 2020; EPRS, 2020) are relevant, and both may provide future opportunities for amendments incorporating animal welfare violations within the scope of intended offences, alongside the current focus on content that is illegal and/or harmful to people. The UK's Online Safety Bill, for example, if adopted, would impose a duty of care, in relation to illegal and/or harmful content, on providers of internet 'user-to-user services' and 'search services', and would confer power of oversight and enforcement on the Office of Communications (Anon., 2021). Existing methods used to identify online wildlife crime and animal-related content that violates platform rules include key word and imagery detection, triangulation of information across user posts, automated detection systems, and manual investigation informed by expert knowledge (e.g. Davies et al., 2021; Di Minin et al., 2018; Harrington et al., 2019), and it has been suggested that a combination of these approaches should become integral to the moderating practices of social media companies to prevent harmful content (Davies et al., 2021). Given that users may be already active on, and have the ability to readily shift to other social media platforms, it has also been suggested that social media platforms should work collectively to eliminate harmful content. For example, by adopting standardised definitions of what constitute cruel animal content and parallel policies to prohibit them (SMACC, 2021).

Although our research suggests that user policing alone is unlikely to be an effective approach, efforts to raise public awareness and human behaviour change could still help to reduce the amount of cruel content posted on social media platforms. For example, whilst recognising that expertise may often be required to determine key information on whether a post violates platform rules, some advocacy efforts have drawn attention to key aspects to inform their judgement, including, in the case of 'fake animal rescues', checking for signs of physical injury to wild animals before the 'rescue', unnatural settings (i.e. microhabitats that the species shown do not typically use in the wild), and the occurrence of multiple similar posts by the same user (World Animal Protection, 2021). Similarly, human behaviour change initiatives targeted at the makers, or the channel owner posting them, in particular raising awareness of the welfare and potential conservation or legal issues involved, might also be another option (cf. Wallen & Daut, 2018). However, for these initiatives to be fully effective, information regarding the motivations of the producers and distributors (e.g. Rare and The Behavioural Insights Team, 2019), and messaging that will deter them, as well as the viewers (consumers), from viewing, sharing or posting harmful content in future is required (cf. TRAFFIC and the Behavioural Insights Team, 2019).

An estimated 700,000 h of new content is uploaded to YouTube every day, watched by global users that increase in number by approximately 5% per year (Statista, 2022d). Inevitably, online videos and other social media posts will include a proportion of inappropriate and unacceptable material: a recent study, for example, estimated that 23% of children in a region of England had been exposed to online animal abuse (RSPCA, 2018). In short, in the absence of effective mechanisms to detect and prevent the spread of content that involves cruelty to animals, this issue will likely only grow. Whatever the approach taken, the precise details of emerging legislation, and the role played by other actors at different points along the chain (from producers to distributors to viewers), there is an urgent need for the major social media companies to take responsibility for the content hosted on their platforms (see also Davies et al., 2021) with respect to the appropriate treatment of animals, promotion of an appropriate relationship between wild animals and people, and a long-term view towards the well-being of both animals and people.

AUTHOR CONTRIBUTIONS

Neil D'Cruze conceived the study; Angie Elwin, Suzi Paterson and Lauren A. Harrington contributed to data collation; Lauren A.

Harrington designed the analysis and wrote the first draft of the paper; and all authors contributed to data interpretation and final editing of the manuscript.

ACKNOWLEDGEMENTS

This work is part of a larger research project on animal cruelty on social media, funded by World Animal Protection. We thank SMACC volunteers for initially collating fake animal rescue videos, Mark Auliya and Alix Harrington for assistance with species identification, and Leila Harrington for collating YouTube channel data.

CONFLICT OF INTEREST

This study was funded by an animal welfare organisation; three of the authors are employed by the same organisation (NDC holds the position of Head of Research). Our intent in carrying out this study was to provide insight that would help inform recommendations and advice for YouTube (and other social media platforms) with respect to an appropriate policy response to online animal cruelty. Our results pertaining to viewer response were in no way influenced by either the funding source, or our own personal views on animal welfare.

DATA AVAILABILITY STATEMENT

To ensure full protection of users' privacy and compliance with General Data Protection Regulation (GDPR 2016/679), raw data used in the study cannot be made publicly available. Data summaries are available in a Dryad online repository linked to this manuscript (https://doi.org/10.5061/dryad.q573n5tn6).

ORCID

Lauren A. Harrington https://orcid.org/0000-0002-7212-2336
Angie Elwin https://orcid.org/0000-0001-8583-3295

REFERENCES

Anon. (2021). *Online safety bill explanatory notes*. Department for Digital, Culture, Media & Sport. https://www.gov.uk/government/publications/draft-online-safety-bill

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on Twitter. In *WSDM'11: Proceedings of the fourth ACM international conference on Web search and data mining* (pp. 65–74). Association for Computing Machinery. https://doi.org/10.1145/1935826.1935845

Barber, N. (2022). The strange allure of blood sports. *Psychology Today*, January 20 2022. https://www.psychologytoday.com/gb/blog/the-human-beast/202201/the-strange-allure-blood-sports

Bärtl, M. (2018). YouTube channels, uploads and views: A statistical analysis of the past 10 years. *Convergence*, *24*(1), 16–32. https://doi.org/10.1177/1354856517736979

Beals, M. (2021). Animal rights group sues YouTube over abuse videos, says bans not enforced. *The Hill, 19 October 2021*. https://thehill.com/policy/technology/577393-animal-rights-group-sues-youtube-over-abuse-videos-says-bans-not-enforced (accessed 14 February 2022).

Berger, J., & Milkman, K. L. (2012). What makes online content viral? *Journal of Marketing Research*, *49*(2), 192–205. http://www.jstor.org/stable/23142844

Boboltz, S. (2015). These are some of the sketchy ways nature documentaries are actually filmed. *Huffington Post*, 13 May 2015. https://www.huffingtonpost.co.uk/entry/nature-documentary-tricks_n_7249528

Borgi, M., & Cirulli, F. (2016). Pet face: Mechanisms underlying human-animal relationships. *Frontiers in Psychology*, *7*, 298. https://doi.org/10.3389/fpsyg.2016.00298

Cambridge Dictionary. (2022). Cambridge University Press. https://dictionary.cambridge.org/dictionary/english/scripted

Carder, G., Plese, T., Machado, F. C., Paterson, S., Matthews, N., McAnea, L., & D'Cruze, N. (2018). The impact of 'selfie' tourism on the behaviour and welfare of brown-throated three-toed sloths. *Animals*, *8*, 216. https://doi.org/10.3390/ani8110216

Cermak, I. (2021). Jumping the shark: White shark representations in great White serial killer lives—The fear and the (pseudo-)science. *Journalism and Media*, *2*, 584–604. https://doi.org/10.3390/journalmedia2040035

Chauhan, P., Sharma, N., & Sikka, G. (2021). The emergence of social media data and sentiment analysis in election prediction. *Journal of Ambient Intelligence and Humanized Computing*, *12*, 2601–2627. https://doi.org/10.1007/s12652-020-02423-y

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The echo chamber effect on social media. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(9), e2023301118. https://doi.org/10.1073/pnas.2023301118

Davies, A., Hinsley, A., Nuno, A., & Martin, R. O. (2021). Identifying opportunities for expert-mediated triangulation in monitoring wildlife trade on social media. *Conservation Biology*, *36*, e13858. https://doi.org/10.1111/cobi.13858

Davis, J. (2021). History of animal fights and blood sports. In M. Roscher, A. Krebber, & B. Mizelle (Eds.), *Handbook of historical animal studies* (pp. 587–602). De Gruyter Oldenbourg. https://doi.org/10.1515/9783110536553-043

Di Minin, E., Fink, C., Hiippala, T., & Tenkanen, H. (2018). A framework for investigating illegal wildlife trade on social media with machine learning. *Conservation Biology*, *33*(1), 210–213. https://doi.org/10.1111/cobi.13104

Doward, J. (2020). YouTube must remove videos of animal cruelty, says charity. *The Guardian*, 19 December 2020. https://www.theguardian.com/world/2020/dec/19/youtube-must-remove-videos-of-animal-cruelty-says-charity

EC. (2020). Proposal for a regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC. Brussels, 15 December 2020. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0825&from=en

EPRS. (2020). *Reform of the EU liability regime for online intermediaries*. European Parliament. https://doi.org/10.2861/08522

Esmail, N., Wintle, B., 't Sas-Rolfes, M., Athanas, A., Beale, C., Bending, Z., Dai, R., Fabinyi, M., Gluszek, S., Haenlein, C., Harrington, L. A., Hinsley, A., Kariuki, K., Lam, J., Markus, M., Paudel, K., Shukhova, S., Sutherland, W., Verissimo, D., … Milner-Gulland, E. J. (2020). Emerging illegal wildlife trade issues: A horizon scan. *Conservation Letters*, *13*, e12715. https://doi.org/10.1111/conl.12715

Ewe, K. (2021). Fake rescues and animal torture are disturbingly common on social media. *Vice*, 1 September 2021. https://www.vice.com/en/article/akg9pz/fake-rescues-and-animal-torture-are-disturbingly-common-on-social-media

Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*, 5. https://doi.org/10.1186/s40537-015-0015-2

Feinerer, I., & Hornik, K. (2018). *tm: Text mining package*. R package version 0.7-6. https://CRAN.R-project.org/package=tm

Feinerer, I., Hornik, K., & Meyer, D. (2008). Text mining infrastructure in R. *Journal of Statistical Software*, *25*(5), 1–54. https://doi.org/10.18637/jss.v025.i05

Fellows, I. (2018). *wordcloud: Word clouds*. R package version 2.6. https://CRAN.R-project.org/package=wordcloud

France, S. L., Vaghefi, M. S., & Zhao, H. (2016). Characterizing viral videos: Methodology and applications. *Electronic Commerce Research and Applications*, *19*, 19–32. https://doi.org/10.1016/j.elerap.2016.07.002

Freund, C. A., Heaning, E. G., Mulrain, I. R., McCann, J. B., & DiGiorgio, A. L. (2021). Building better conservation media for primates and people: A case study of orangutan rescue and rehabilitation YouTube videos. *People and Nature*, *3*, 1257– 1271. https://doi.org/10.1002/pan3.10268

Google. (2021). *YouTube terms of service*. Community Guidelines. https://www.youtube.com/howyoutubeworks/policies/community-guidelines/. [Update announcement available here: https://support.google.com/youtube/thread/115480979/youtube-policy-update-animal-abuse-violent-or-graphic-content?hl=en]

Gupta, S., Singh, A., & Ranjan, J. (2020). Sentiment analysis: Usage of text and emoji for expressing sentiments. In M. Kolhe, S. Tiwari, M. Trivedi, & K. Mishra (Eds.), *Advances in data and information sciences. Lecture notes in networks and systems* (Vol. *94*). Springer. https://doi.org/10.1007/978-981-15-0694-9_45

Hallinan, B., Kim, B., Scharlach, R., Trillò, T., Mizoroki, S., & Shifman, L. (2021). Mapping the transnational imaginary of social media genres. *New Media & Society*. https://doi.org/10.1177/14614448211012372

Harrington, L. A., D'Cruze, N., & Macdonald, D. W. (2018). Rise to fame: Events, media activity, and public interest, in pangolins and pangolin trade, 2005-2016. *Nature Conservation*, *30*, 107–133. https://doi.org/10.3897/natureconservation.30.28651

Harrington, L. A., Macdonald, D. W., & D'Cruze, N. (2019). Popularity of pet otters on YouTube: Evidence of an emerging trade threat. *Nature Conservation*, *36*, 17–45. https://doi.org/10.3897/natureconservation.36.33842

HMSO. (2021). *Draft online safety bill. CP 405*. Department for Digital, Culture, Media & Sport. https://www.gov.uk/government/publications/draft-online-safety-bill

Hvitfeldt, E. (2022). *Emoji: Data and function to work with Emojis*. R package version 0.1.0. https://CRAN.R-project.org/package=emoji

Iliopoulou, M., & Rosenbaum, R. (2013). Understanding bloodsports. *Journal of Animal & Natural Resource Law*, *IX*, 125–140.

IUCN. (2021). *The IUCN red list of threatened species*. Version 2021-2. https://www.iucnredlist.org

Jockers, M. L. (2015). *Syuzhet: Extract sentiment and plot arcs from text*. https://github.com/mjockers/syuzhet

Kalof, L., & Taylor, C. (2007). The discourse of dog fighting. *Humanity and Society*, *31*, 319–333. https://doi.org/10.1177/016059760703100403

Knowles, T. (2021). YouTube bans fake animal rescue videos. *The Times*, 25 March 2021. https://www.thetimes.co.uk/article/youtube-bans-fake-animal-rescue-videos-l5vwtnxls

Leighty, K. A., Valuska, A. J., Grand, A. P., Bettinger, T. L., Mellen, J. D., Ross, S. R., Boyle, P., & Ogden, J. J. (2015). Impact of visual context on public perceptions of non-human primate performers. *PLoS One*, *10*, 1–6. https://doi.org/10.1371/journal.pone.0118487

Liu, B. (2015). *Sentiment analysis: Mining opinions, sentiments, and emotions*. Cambridge University Press. https://doi.org/10.1017/CBO9781139084789

Llewellyn, T., & Rose, P. (2021). Education is entertainment? Zoo science communication on YouTube. *Journal of Zoological and Botanical Gardens*, *2*(2), 250–264. https://doi.org/10.3390/jzbg2020017

Macnamara, J. (2005). Media content analysis: Its uses, benefits and best practice methodology. *Asia Pacific Public Relations Journal*, *6*(1), 1–34.

Maron, D.F. (2021). *How fake animal rescue videos have become a new frontier for animal abuse*. National Geographic, June 30, 2021. https://www.nationalgeographic.com/animals/article/how-fake-animal-rescue-videos-have-become-a-new-frontier-for-animal-abuse

Menczer, F., & Hills, T. (2020). The attention economy. *Scientific American*, *323*(6), 54–61. https://doi.org/10.1038/scientificamerican1220-54

Moloney, G. K., Tuke, J., Dal Grande, E., Nielsen, T., & Chaber, A. L. (2021). Is YouTube promoting the exotic pet trade? Analysis of the global public perception of popular YouTube videos featuring threatened exotic animals. *PLoS One*, *16*(4), e0235451. https://doi.org/10.1371/journal.pone.0235451

Morgan, J. S., Barjasteh, I., Lampe, C., & Radha, H. (2014). The entropy of attention and popularity in YouTube videos. *ArXiv*, *abs/1412.1185*. https://arxiv.org/abs/1412.1185

Nalts (pseudonym). (2011). *How many views do you need to be viral?* Will Video for Food. Online-Video marketing. http://willvideoforfood.com/2011/05/06/how-many-views-do-you-need-to-be-viral/

Nekaris, K. A. I., Musing, L., Vazquez, A. G., & Donati, G. (2015). Is tickling torture? Assessing welfare towards slow lorises (*Nycticebus* spp.) within web 2.0 videos. *Folia Primatologica*, *86*, 534–551. https://doi.org/10.1159/000444231

Neme, L. (2010). The secrets of animal TV: Many nature shows rely on unethical tactics. *Mongabay*, 14 December 2010. https://news.mongabay.com/2010/12/the-secrets-of-animal-tv-many-nature-shows-rely-on-unethical-tactics/

Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. *Journal of Information and Telecommunication*, *5*(1), 1–15. https://doi.org/10.1080/24751839.2020.1790793

Neuendorf, K. (2017). Defining content analysis. In *The content analysis guidebook* (pp. 1–35). SAGE Publications, Inc. https://doi.org/10.4135/9781071802878.n1

Neuman, W. L. (1997). *Social research methods: Qualitative and quantitative approaches* (3rd ed.). Allyn and Bacon.

Nielsen, F. A. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. *Proceedings of the ESWC2011 Workshop on Making Sense of Microposts: Big things come in small packages*, 93–98. ArXiv, abs/1103.2903. https://arxiv.org/abs/1103.2903

Novak, P. K., Smailović, J., Sluban, B., & Mozetič, I. (2015). Sentiment of Emojis. *PLoS ONE*, *10*(12), e0144296. https://doi.org/10.1371/journal.pone.0144296

Osterberg, P., & Nekaris, K. A. I. (2015). The use of animals as photo props to attract tourists in Thailand: A case study of the slow Loris *Nycticebus* spp. *Traffic Bulletin*, *27*(1), 13–18. https://www.traffic.org/site/assets/files/3008/traffic_pub_bulletin_27_1.pdf#page=17

Oxford Languages. (2022). Oxford University Press. https://languages.oup.com/google-dictionary-en/

Palmer, C. (2011). Into the wild, ethically: Nature filmmakers need a code of conduct. *Documentary Magazine*, *17 May 2011*. https://www.documentary.org/feature/wild-ethically-nature-filmmakers-need-code-conduct

Pavelle, S., & Wilkinson, C. (2020). Into the digital wild: Utilizing twitter, Instagram, YouTube, and Facebook for effective science and environmental communication. *Frontiers in Communication*, *5*, 82. https://doi.org/10.3389/fcomm.2020.575122

Poecze, F., Ebster, C., & Strauss, C. (2019). Let's play on Facebook: Using sentiment analysis and social media metrics to measure the success of YouTube gamers' post types. *Personal and Ubiquitous Computing*, *26*, 901–910. https://doi.org/10.1007/s00779-019-01361-7

Pollo, S., Graziano, M., & Giacoma, C. (2009). The ethics of natural history documentaries. *Animal Behaviour*, *77*(5), 1357–1360. https://doi.org/10.1016/j.anbehav.2009.01.022

Puschmann, C., & Powell, A. (2018). Turning words into consumer preferences: How sentiment analysis is framed in research and the news media. *Social Media & Society*, *4*(3). https://doi.org/10.1177/2056305118797724

R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rare and The Behavioural Insights Team. (2019). *Behavior change for nature: A behavioral science toolkit for practitioners*. Rare. www.rare.org

Ratner, B. (2009). The correlation coefficient: Its values range between +1/−1, or do they? *Journal of Targeting, Measurement and Analysis for Marketing*, *17*, 139–142. https://doi.org/10.1057/jt.2009.5

Riddle, E., & MacKay, J. R. D. (2020). Social media contexts moderate perceptions of animals. *Animals*, *10*(5), 845. https://doi.org/10.3390/ani10050845

Rinker, T. W. (2018). *textclean: Text cleaning tools version 0.9.3*. https://github.com/trinker/textclean

Robertson, M.R. (2014). *3 metric ratios to measure channel success*. tubularinsights.com, September 18 2014. https://tubularlabs.com/blog/3-metrics-youtube-success/

Ross, S. R., Vreeman, V. M., & Lonsdorf, E. V. (2011). Specific image characteristics influence attitudes about chimpanzee conservation and use as pets. *PLoS ONE*, *6*, 1–5. https://doi.org/10.1371/journal.pone.0022050

RSPCA. (2018). *23 percent of kids have seen animal abuse online*. https://www.rspca.org.uk/-/16_10_18_genkind

Schroepfer, K. K., Rosati, A. G., Chartrand, T., & Hare, B. (2011). Use of 'entertainment' chimpanzees in commercials distorts public perception regarding their conservation status. *PLoS ONE*, *6*, 1–8. https://doi.org/10.1371/journal.pone.0026048

Social Media Animal Cruelty Coalition (SMACC) (2021). Making money from misery. In *How social media giants profit from animal abuse*. Asia for Animals Coalition. https://www.asiaforanimals.com/, https://www.asiaforanimals.com/smacc-report

Somerville, K., Dickman, A., Johnson, P. J., & Hart, A. G. (2021). Soap operas will not wash for wildlife. *People and Nature*, *3*, 1160–1165. https://doi.org/10.1002/pan3.10202

Sood, G. (2020). *tuber: Access YouTube from R*. R package version 0.9.9. https://cran.r-project.org/web/packages/tuber/tuber.pdf

Statista. (2022a). *Social media – Statistics and Facts*. https://www.statista.com/topics/1164/social-networks/

Statista. (2022b). *Most popular social networks worldwide as of January 2022, ranked by number of active users (in millions)*. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/

Statista. (2022c). *Media usage in an internet minute as of August 2021*. https://www.statista.com/statistics/195140/new-user-generated-content-uploaded-by-users-per-minute/

Statista. (2022d). *Hours of video uploaded to YouTube every minute as of February 2020*. https://www.statista.com/statistics/259477/hours-of-video-uploaded-to-youtube-every-minute/

Talamo, L. (2021). UPDATE: YouTube announces ban on fake 'rescue' videos following LFT campaign. *Lady Freethinker*, March 29, 2021. https://ladyfreethinker.org/victory-youtube-to-ban-fake-rescue-videos-following-lft-campaign/

Thomas-Walters, L., McNulty, C., & Veríssimo, D. A. (2020). A scoping review into the impact of animal imagery on pro-environmental outcomes. *Ambio*, *49*, 1135–1145. https://doi.org/10.1007/s13280-019-01271-1

Thompson, R. M., Hall, J., Morrison, C., Palmer, N. R., & Roberts, D. L. (2021). Ethics and governance for internet-based conservation science research. *Conservation Biology*, *35*, 1747–1754. https://doi.org/10.1111/cobi.13778

Tian, Y., Galery, T., Dulcinati, G., Molimpakis, E., & Sun, C. (2017). Facebook sentiment: Reactions and Emojis. In *Proceedings of the fifth international workshop on natural language processing for social media, W17-1102*. Association for Computational Linguistics. https://aclanthology.org/W17-1102

Townsend, L., & Wallace, C. (2016). Social media research: a guide to ethics. Workshop proceedings, 'Social Media, Privacy and Risk: Towards More Ethical Research Methodologies', February 2016, University of Aberdeen. https://ahrecs.com/resources/

TRAFFIC and the Behavioural Insights Team. (2019). *Reducing demand for illegal wildlife: Designing effective messages*. TRAFFIC International. www.traffic.org

Van Hamme, G., Svensson, M. S., Morcatty, T. Q., Nekaris, K. A. I., & Nijman, V. (2021). Keep your distance: Using Instagram posts to evaluate the risk of anthroponotic disease transmission in gorilla ecotourism. *People and Nature*, *3*, 325–334. https://doi.org/10.1002/pan3.10187

von Essen, E., Turnbull, J., Searle, A., Jørgensen, F. A., Hofmeester, T. R., & van der Wal, R. (2021). Wildlife in the digital Anthropocene: Examining human–animal relations through surveillance technologies. *Environment and Planning E: Nature and Space*. https://doi.org/10.1177/25148486211061704

Wallen, K. E., & Daut, E. F. (2018). The challenge and opportunity of behavior change methods and frameworks to reduce demand for illegal wildlife. *Nature Conservation*, *26*, 55–75. https://doi.org/10.3897/natureconservation.26.22725

White, C. M., Cutello, C. A., Gummerum, M., & Hanoch, Y. (2018). A cross-cultural study of risky online self-presentation. *Cyberpsychology, Behavior and Social Networking*, *21*(1), 25–31. https://doi.org/10.1089/cyber.2016.0660

World Animal Protection. (2021). Views that abuse: The rise of fake 'animal rescue' videos on YouTube. https://www.worldanimalprotection.us/blogs/views-abuse-new-report-highlights-cruel-fake-animal-rescues-youtube

Zanette, L. Y., & Clinchy, M. (2020). Ecology and neurobiology of fear in free-living wildlife. *Annual Review of Ecology, Evolution, and Systematics*, *51*(1), 297–318. https://doi.org/10.1146/annurev-ecolsys-011720-124613

Zook, M., Barocas, S., Boyd, D., Crawford, K., Keller, E., Gangadharan, S. P., Goodman, A., Hollander, R., Koenig, B. A., Metcalf, J., Narayanan, A., Nelson, A., & Pasquale, F. (2017). Ten simple rules for responsible big data research. *PLoS Computational Biology*, *13*(3), e1005399. https://doi.org/10.1371/journal.pcbi.1005399

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Harrington, L. A., Elwin, A., Paterson, S., & D'Cruze, N. (2023). The viewer doesn't always seem to care—response to fake animal rescues on YouTube and implications for social media self-policing policies. *People and Nature*, *5*, 103–118. https://doi.org/10.1002/pan3.10416