Health Talk: Understanding Practices of Popular Professional YouTubers

Thanh-Trung Phan*
Fulbright University Vietnam
Vietnam
trung.phan@fulbright.edu.vn

Chloe Michoud
Institute of Psychology, University of Lausanne
Switzerland
Chloe.Michoud@unil.ch

Lucia Volpato
Institute of Psychology, University of Lausanne
Switzerland

Maria del Rio Carral
Institute of Psychology, University of Lausanne
Switzerland
maria.delriocarral@unil.ch

Daniel Gatica-Perez
Idiap Research Institute and EPFL
Switzerland
gatica@idiap.ch

Abstract

Practices related to health are circulated widely on YouTube. With a health psychology perspective, we present a study to understand health and wellbeing-related practices of a group of popular, professional YouTubers from the audio-visual content they produce. We first identify, via polytextual thematic analysis, six thematic health-related categories, and use them to label a set of 2500 YouTube videos. Agreement among three independent annotators was acceptable for these health-related categories. We then present an analysis of speech transcriptions and visual content, demonstrating that distinctive patterns exist for these health-related categories. These include linguistic markers and specific scene types and objects. Finally, with an interpretability focus, we study the feasibility of classifying health-related video categories in a binary setting, and compare performance across features, finding best accuracy for linguistic features (74-87%), and various patterns of linguistic and visual relevance used for the classification of health categories. The results shows promise to support mixed-methods research in health psychology, combining manual analysis and data-driven methods. More generally, our work contributes to the understanding of current health practices shared and promoted on social video.

CCS Concepts

• **Human-centered computing** → **Ubiquitous and mobile computing**; *Ubiquitous and mobile computing design and evaluation methods.*

Keywords

YouTube, social media, health psychology, LIWC, deep learning, audio-visual analysis

ACM Reference Format:
Thanh-Trung Phan, Chloe Michoud, Lucia Volpato, Maria del Rio Carral, and Daniel Gatica-Perez. 2022. Health Talk: Understanding Practices of Popular Professional YouTubers. In *21th International Conference on Mobile and Ubiquitous Multimedia (MUM 2022), November 27–30, 2022, Lisbon, Portugal.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3568444.3568453

1 Introduction

In psychology, individual behavior has been widely studied in connection to their role in health and illness [3, 36, 54, 64, 69, 71]. Individual health-related behaviors such as diet, physical activity, smoking habits, and rest, have been conceptualized as social practices, that is, as patterns of activity that are socially shared within a given group or community at a given historical period, and thus always contextualized [10, 25]. Combining computational methods within this perspective can bring valuable insights to understand the communication of health-related behaviors.

Social media gives opportunities for users to create and share content about health. YouTube, as the quintessential ubiquitous and mobile multimedia platform, has modified the ways in which health-related information spreads, by allowing video makers to directly share health experiences with their audiences via spoken and visual content. Qualitative research on YouTube content has studied specific subjects related to personal experiences with illness and chronic conditions, both in health research [19, 86] and human-computer interaction (HCI) research [42] . A particular case of interest is that of professional YouTubers, i.e., users who talk about their experiences and opinions and produce YouTube videos as a job, and who can have a large impact on their (often young) audiences regarding the perception and adoption of health practices for a variety of issues, including eating, exercise, and rest. Research has began to study the practices disclosed and promoted by professional YouTubers based on linguistic analyses, either with health [39, 70, 97] or HCI angles [97], and sometimes in connection to marketing [74].

*This is the corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
MUM 2022, November 27–30, 2022, Lisbon, Portugal
© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-1-4503-9820-6/22/11…$15.00
https://doi.org/10.1145/3568444.3568453

Starting from a health psychology perspective, in this paper we set out to understand health and wellbeing-related practices of professional YouTubers using both a qualitative approach for data labeling and computational analyses of speech transcripts, video titles and tags, and visual content. We address two research questions:

RQ1: Given a sample of health-related videos generated by professional YouTubers, what health categories can be extracted by using polytextual thematic analysis? Given this categorization, what are the observable similarities and differences among health categories in terms of audio-visual content?

RQ2: If such differences exist, how can audio-visual descriptors of health videos be used to classify videos according to health categories for further use in health psychology research?

The contributions of the paper are the following:

1. We identify health-related categories in videos from popular, professional YouTubers. For this, we used a dataset collected in our previous work [26]. This sample contains 2500 videos from 50 English-speaking YouTubers. We use polytextual thematic analysis [35] to discover and manually annotate videos with respect to six health-related categories, namely 'Food & Nutrition', 'Self-Development', 'Bodycare & Beauty', 'Physical Activity', 'Companionship', and 'Tech-Free & Rest'. These categories span a number of themes that have been studied in other social media platforms, and therefore allow to frame our findings in the context of the larger social media literature.

2. We show that linguistic patterns exist within and across health-related categories, by using text analysis on titles, tags, and speech transcripts, using a combination of embedded text vectors and Linguistic Inquiry and Word Count (LIWC). This analysis provides several insights, regarding the differentiated use of certain LIWC categories depending of the health-related theme, e.g., physical activity videos have the highest values of future orientation linguistic markers.

3. We extract visual features by applying deep learning models on videos to characterize them by the depicted scenes and by the main objects that are present. The analysis based on these generic semantic descriptions of videos shows specific visual patterns per health category, and indicates a degree of homogeneity in terms of visual scene composition for a given a health-related theme. This raises various questions about the motivations and intent to produce this specific kind of content.

4. With an emphasis on interpretability, we use a machine learning framework to classify the six health-related video categories in a binary classification setting, and compare the performance of the various types of features. The results shows that using LIWC features produces the highest accuracies (74-87%) for all health-related videos. Meanwhile, the same task using visual features produces lower accuracies compared to LIWC features (7-12%). We show results of separate and joint modalities (speech transcriptions vs. visual scene descriptions). An analysis of feature relevance reveals meaningful patterns for visual and linguistic features used for each health category. These results represent an initial step towards developing a functionality to support further mixed-methods research in health psychology, by proposing health-related category tags that facilitate access to and search on larger data samples.

The paper is organized as follows. Section 2 discusses previous work. Section 3 describes the dataset and the health categorization process. Section 4 presents the audio-visual content analysis of health videos. Section 5 presents the video classification tasks and results. Section 6 discusses the findings and limitations. Section 7 concludes the paper.

2 Related Work

2.1 Health psychology

Health psychology has a long tradition of studying the role that behaviors play in health and illness, through the analysis of practices that concern diet, physical activity, alcohol consumption, and smoking [24]. There is often a gap between the intention of adopting a health-promoting behavior and its concrete implementation [24]. For this reason, health psychology research has been concerned with the understanding of the discrepancies between intention and adoption of health behaviors [81]. Researchers have developed a wide array of models, such as the Health Belief Model [36], the Transtheoretical Model [69], the Theory of Planned Behaviour [3], and the Self-Determination Theory [72]. However, such models have shown limited effectiveness [55, 63]. To understand the discrepancies between intention and implementation of behavior, critical approaches in health psychology examine how people experience health and illness from their own perspective. These approaches are thus interested in capturing meaning that individuals attach to their health behaviors, within their specific social and cultural contexts [51].

2.2 Health-related work in social media

Leading a healthy life through specific practices has become a trend in Western societies [22], and social media provides channels to communicate health issues as well as potential ways to improve health outcomes [57]. Social media allows patients, health professionals, and the public to share their experiences and opinions about health, based on scientific knowledge, but also on individual beliefs [28, 84]. Research in health and social media is too wide to review here in depth, but spans studies on Facebook [2, 7, 29, 33, 38, 73], Twitter [18, 40, 75, 76, 79, 90], Instagram [4, 13, 15, 23, 31, 48, 52, 58, 62, 80], and YouTube [42, 85, 86, 91].

The use of social media for health communication has been investigated from many perspectives, including both benefits and limitations. For example, Facebook has been studied as a health educational medium [2], and also regarding disease diagnosis and management [29, 33, 38]. Literature on Twitter has studied how the platform is used to communicate health issues [18, 75, 79]; to identify antibiotic misuse or misunderstanding [76]; or to track weight-loss interventions [90]. Instagram has been studied from many angles, including acute illness, alternative medicine, chronic illness, nutrition, exercise, or skincare [4, 13, 20, 23, 48, 58, 67, 68, 80, 96], and also mental health [15, 31, 52]. Social media has also been also analyzed from the perspective of accountability and authenticity [38]. Several of these themes appear in our work.

2.3 YouTube Vlogs and Wellbeing/Health

YouTube is a popular platform to share health and wellness content [85], much of which is consumed on mobile devices – 63% of

YouTube viewing time across the globe, according to [1]. Qualitative research in both health and HCI literature has studied YouTube content related to organ donation [86], personal experiences among cancer survivors and caregivers [19], and discussions about chronic conditions like HIV [42]. Unrelated to health, YouTube conversational videos have been used in research, including analysis of speaking dynamics [5] and perceived personality from conversational vlogs [8], sentiment analysis from movie reviews [92], or analysis of public speakers [82]. More recent work focused on health has studied approaches to identify general knowledge level [49] and misinformation [41] in YouTube medical videos, the latter with a multimodal interaction angle.

A progressive shift from amateur video makers to professional YouTubers (who make YouTube videos as a job) has consolidated on the platform, accompanied by features such as product placement [32, 85]. The presentation of health content by professional YouTubers has been studied from the perspective of its effect on viewers [13, 37, 74, 97]. However, relatively little research has paid attention to understanding how health practices are displayed and promoted by professional YouTubers [39, 70]. Ratwatte et al. studied the health practice of exercising [16], with regard to the 'fitspiration' trend, also popular in other platforms [70]. Harris used a mixed-methods approach (surveys, interviews, focus groups, netnography) to conclude that young people see professional YouTubers as a valid source of health information [39]. This is important as some YouTubers might not have certifications as health professionals. Besides the need to address the problem of health-related misinformation circulating in YouTube [50] while identifying trustworthy social media in health communities [30], further research is needed to provide an in-depth understanding of the health practices promoted by YouTubers. It is thus relevant to develop systematic methodologies for analysis of YouTube health content, tackling both the linguistic and visual dimensions of their videos.

Our work extends the current literature on the understanding of practices by professional YouTubers who produce health-related content, and more generally about the understanding of social video as a platform to discuss health [11, 41, 42, 49]. Table 1 summarizes the most closely related work.

3 Dataset

3.1 Data Collection

The data used in this paper was collected in our previous work [26]. The data was collected in early 2020 following an interactive methodology. One researcher in our team [26] typed keywords on YouTube's search bar, and opened the first relevant video in a new tab, then watched and listened to parts of the video, while searching for health content on the corresponding YouTube channel to decide its inclusion in the database. The following inclusion criteria were defined: videos about health or well-being spoken in English; YouTubers aged over 18 years old; and users having at least 100K subscribers, which set a baseline level of popularity. The initial keyword list was created by the research team and included typically addressed topics around healthy habits and lifestyle (e.g., 'healthy habits', 'healthy lifestyle', 'healthy living'). Tags in video results are potentially new keywords (e.g., anti-anxiety, ketosis, etc.) to explore more videos for each category. However, more specific conditions or disorders were considered out of the scope of the study. The database was expanded by iterating this process until no further relevant results were found. Channels that were less health-related (e.g. focusing on music, fashion, or home decoration) were not included (for more details see [26]). We note that health-related videos on YouTube are not from medical doctors, but from YouTubers talking about well-being themes with their fans or audiences. The study was approved by the ethical commission of the University of Lausanne. The final dataset consists of 2500 videos (50 users, 50 videos per user), including video title, tags, speech transcripts, and video content. Mean (resp. median) video duration is 10.5 min. (resp. 8.5 min.) Mean values for likes and dislikes are 206 and 10, respectively. Manual inspection of the dataset showed that 82% (41/50) of the users self-defined as professionals, including coaches, entrepreneurs, and other roles.

3.2 Health-related Themes: Definition and Annotation

We first performed a polytextual thematic analysis [35] to identify the main recurring visual topics in the videos. Polytextual thematic analysis is a qualitative technique defined by the iterative and systematic manual annotation of visual data [35]. The following procedure was conducted. First, we led a frame-by-frame manual annotation across fifteen videos, based on a diverse range of health behaviors. Following this summarized description of the videos' visual content, preliminary themes were defined [35]. Secondly, an iterative comparison across identified themes was performed. This procedure was repeated until saturation and inter-coder consistency were achieved. This analysis led to the definition of eight preliminary visual themes. After a comparison with themes stemming from computational linguistic analyses (not discussed here for space reasons), two of the preliminary themes were discarded. This procedure resulted in six final themes: Bodycare & Beauty (e.g., skincare), Companionship (e.g., spending time with family), Food & Nutrition (e.g., cooking or nutrition in food), Physical Activity (e.g., workout techniques), Self-Development (e.g., a plan to achieve a particular target), and Tech-Free & Rest (e.g., meditation). Note that the themes are not mutually exclusive, i.e., a given video can discuss more than one theme.

We used these themes to annotate the full dataset. We implemented an annotation task by asking annotators to read the titles and tags of each video, and to assign three possible values based on the presence or absence of each of the visual content themes in each video: (1) if the theme was not present, (2) if not sure, and (3) of the theme was present. For instance, a hypothetical video "How To Reduce Weight in Three Months" with tags "Weight loss, Workout, Gym tips, Nutrition, How often to exercise" would get a label '3' for Food & Nutrition, Physical Activity, and Self-Development, while the rest categories would get a label '1'. Prior to the full annotation task, we conducted a pilot test on a sample of 50 random videos from the dataset. These videos were coded by three annotators from our research team. We then computed Intraclass Correlation Coefficients (ICC(2,k)) as recommended in [45] to assess reliability (agreement among annotators). The ICC results of annotated health-related categories on sampled 50 videos (shown in Table 2) show that ICC(2,k) for four of the health-related themes is very good

Table 1: Comparison between previous work and our work.

Work	Goal	Data	Task	Finding
[26] (2021)	Use critical perspective to analyze health content.	2500 videos from professional YouTubers.	Define theoretical framework and approach to blend narrative and visual qualitative components.	Analysis points towards an individualized pursuit of health.
[70] (2019)	Understand fitspiration videos on YouTube.	Videos by fitness YouTubers.	Thematic analysis of videos.	Several patterns were found; unhealthy fitness practices were promoted by creators and audiences.
[39] (2019)	Understand how professional YouTubers relate to youth health behavior.	931 questionaires; 133 health-related YouTube videos; semi-structured interviews.	Four-stage mixed-methods approach	YouTubers are perceived as legitimate in terms of health knowledge, yet non-expert advice involves risks.
[41] (2019)	Develop automatic assessment of accuracy of health infornation in videos.	250 YouTube prostate cancer videos, annotated for misinformation.	Classify misinformation videos using linguistic, acoustic, and user engagement features	Classification accuracy: 74% (76.5% precision; 73.2% recall)
[49] (2019)	Develop automatic video classifer with respect to level of medical knowledge.	600 YouTube videos labeled with (1) medical terns and (2) high vs. low medical knowledge	Classify high vs. low medical knowledge from text + visual cues	Classification F-measure performance up to 84.0%.
[11] (2019)	Understand practices of parents sharing YouTube videos of children with developmental disabilities.	300 YouTube videos.	Semi-structured interviews; analysis of YouTube content.	Parents share videos for various reasons (authenticity, advocacy, link with others.)
This work	Understand health-related practices of professional YouTubers from audio-visual content.	2500 YouTube videos from [26], enriched with manual health-relatec labels on six categories.	Investigate possible patterns in 6 health categories using video transcriptions and visual content; classify health categories using textual and visual features.	Health-related videos are well characterized by linguistic categories, scene types, and scene objects. Binary classification of health categories can be achieved with accuracy 74-87% by using linguistic features.

(above 0.80), while the remaining two themes had good reliability (0.60-0.80). This result motivated the use of a single annotator for the rest of the videos. The final results are shown in Table 2. We observe that Food & Nutrition, Self-Development, Bodycare & Beauty, and Physical Activity are the most frequent categories, compared to Companionship and Tech-Free & Rest. Table 2 also shows statistics of duration and view count for each of the categories. The video duration of the dataset is in line with trends reported in previous literature [17]. The mean view count of Physical Activity (231K) and Food & Nutrition (187K) indicate that these two topics gather more attention compared to the other health-related categories, yet all categories are popular (with mean view count above 145K).

To visualize the co-occurrence of the six health-related categories, we show a chord diagram in Figure 1. Each category is represented by a colored segment on the outer part of the circular layout. The size of the arc between categories is proportional to the amount of co-occurrence. We observe that all categories include videos that have only one theme (hill-like shapes on the circle), as well as co-occurrence of health categories.

4 Health-related Content Analysis (RQ1)

In this section, we analyze the health-related videos using textual content (titles, tags, and transcripts), and visual content (extracted frames from video). As stated in the Introduction, our goal is to focus on models that can support the work of psychologists from the perspective of interpretation. This motivates our choices as presented below.

Table 2: Statistics of videos per health-related category.

Categories	N	Mean Duration ±sd (mins)	Mean View Count ±sd	ICC
Food & Nutrition	1062	10.6 ±8.8	187K ±31K	0.96
Self-Development	695	10.6 ±8.7	169K ±25K	0.75
Bodycare & Beauty	666	10.5 ±8.6	153K ±19K	0.93
Physical Activity	440	11.4 ±9.7	231K ±31K	0.91
Companionship	186	10.6 ±8.5	145K ±16K	0.89
Tech-Free & Rest	158	10.6 ±8.7	167K ±22K	0.78

4.1 Textual Content

Using the various sources of textual content (video titles, tags, and transcripts), we first characterize the health-related video categories according to their content. Then, we investigate more specific linguistic trends using LIWC.

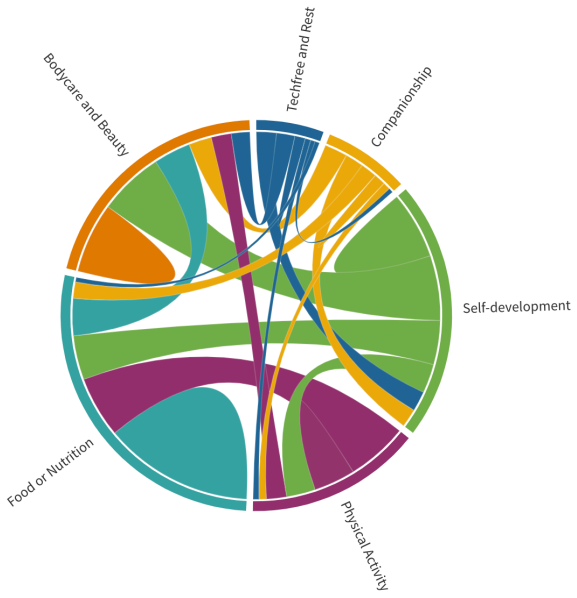

Figure 1: Co-occurrence of the six health-related labels.

Characterizing health-related categories from titles, tags, and transcripts. We first set out to understand the main linguistic trends for each of the six categories of health-related videos. We use the complementary value of video titles and tags (which as sources of curated text are meant to increase discoverability via search) and the full transcripts (which contain the details of the intended message). We applied pre-processing steps on the transcripts and combined text of titles with tags. We convert text into lowercase, remove non-alphabet characters, and remove both short and common words in English. Titles and tags are concatenated and embedded using the fastText library [56]. FastText is a pre-trained model, trained on English text of Common Crawl and Wikipedia using unsupervised methods. We use fastText instead of other standard techniques used in text representation (e.g., TF-IDF or topics). We obtain a representation by using vectors from FastText built on the aforementioned large English dataset (in contrast, representations like TF-IDF would reflect how important a word is in a video using simple statistics computed on the video dataset itself.) This model outputs 300-dimensional vectors. We then compute the cosine distance between the vector of each word in the transcripts and the vector of the corresponding video's titles and tags, to estimate the similarity between individual terms spoken in the videos and their curated metadata descriptions.

Table 3 shows the top words in the transcripts for the six health-related categories, in terms of the frequency with which each embedded word has a cosine distance below 0.5 w.r.t. the corresponding video's title embedding. As the table shows, Food & Nutrition includes words like 'eat', 'food', 'vegan', or 'diet'. Self-Development includes words like 'life', 'self', 'personal', and 'confidence'. Bodycare & Beauty includes words related to physical and mental health. Physical Activity includes words like 'workout', 'exercise', and 'fitness'. Companionship videos have words like 'love', 'family', 'relationship', 'mom', and 'child'. Finally, Tech-Free & Rest videos contain words like 'meditation', 'yoga', or 'relaxation'. The six health-related categories overlap due to co-occurrence of words in the same videos. For instance, the word "mental" exists in multiple categories, i.e., Bodycare & Beauty, Companionship, Self-development, and Tech-free & Rest.

LIWC-based analysis of transcripts. LIWC computes the degree with which various language dimensions are used in a text. LIWC2015 is the most recent version [65]. The outputs of LIWC2015 are distributions of language variables (e.g., Tone, Authentic, etc.), linguistic dimensions (e.g., I, We, You, etc.), psychological processes (e.g., FocusPast, FocusPresent, and FocusFuture, etc.), and others. LIWC categories contain subcategories, with the full dictionary composed of 6,400 words and word stems. For instance, the present time orientation category ('FocusPresent') has 421 words (e.g., today, is, now, etc.), while the past time orientation category ('FocusPast') has 341 words (e.g., ago, did, etc.). The LIWC distribution represents the percentage of matched keywords for each category for the total number of words in a video transcript.

Table 4 shows descriptive statistics of LIWC categories for the dataset, and a comparison with previously reported results for text blogs [77] and natural speech [65]. The Table compares five feature groups including text analytic, pronouns, emotion, time orientation, and language. Text analytic includes 'Tone' (a higher number indicates a more positive emotional tone) [21]; 'Authentic' (a higher number corresponds to a larger use of personal or honest words) [61]; 'Clout' (the higher the number, the larger use of influential language) [44]; and 'Analytic' [66] (a higher number indicates a higher use of analytical thinking words). Compared to previous results using LIWC, blogs and natural speech tend to have the highest and lowest mean values, compared to the ones observed in the video dataset. Physical Activity videos have the highest 'Authentic' mean value, while Companionship videos have the lowest one.

In related work in the HCI literature, Borgo-Rodriguez et al. [11] analyzed how parents share videos about their children with developmental disabilities on YouTube. This work found that parents share this content to create a sense of authenticity, and to connect to others. Our results show that that the Authentic LIWC category has high values across health-related categories, especially for (P)hysical Activity videos. Tech-Free & Rest videos have the highest value for Sixltr (Words > 6 letters), which suggests the use of more complex language [83].

Regarding the use of pronouns, the first person of the singular ('I') has a higher mean value than the second person of the singular ('You'), and also higher than the first person of the plural ('We'). This trend is similar to previously reported results on text blogs and natural speech. Interestingly, Tech-Free & Rest videos have the lowest mean value for 'I' and the highest mean value for 'You', compared to all other video categories. Meanwhile, Companionship videos have the highest mean value for 'We'.

Regarding the use of emotion-related terms, positive emotion mean values for all health-related categories are higher than for negative emotion, in both the video dataset (e.g., 3.9 vs. 0.9 for the whole video sample). This results also matches with previously published results on text blogs (3.66 vs. 2.06), and natural speech (5.31 vs. 1.19). Among the six health-related categories, Companionship

Table 3: Top words per health-related category for the full video transcripts, using embeddings between single words and video titles/tags. The numbers in brackets correspond to the number of videos for which the cosine distance between the embeddings of single words and video titles is below 0.5.

Category	N	Top words (ranked by #videos)
Food & Nutrition	1062	('eat', 206), ('food', 206), ('vegan', 166), ('eating', 160), ('diet', 130), ('healthy', 115), ('meal', 106), ('delicious', 84), ('fat', 66), ('meat', 53), ('day', 53), ('nutrition', 50), ('calorie', 48), ('veggie', 45), ('smoothie', 36), ('gluten', 36), ('vegetable', 34), ('recipe', 34), ('health', 33), ('weight', 32), ('snack', 30), ('fasting', 27), ('ate', 26), ('life', 25), ('nutritious', 25), ('protein', 24), ('carbohydrate', 24), ('carb', 24)
Self-Development	695	('life', 83), ('self', 58), ('personal', 42), ('confidence', 41), ('eating', 30), ('way', 27), ('healthy', 26), ('diet', 24), ('day', 24), ('confident', 24), ('love', 23), ('get', 22), ('food', 20), ('eat', 20), ('feeling', 19), ('want', 19), ('desire', 17), ('work', 17), ('learning', 17), ('good', 14), ('workout', 14), ('weight', 14), ('fat', 14), ('anxiety', 13), ('exercise', 13), ('give', 13), ('help', 13), ('makeup', 13), ('mental', 12), ('know', 12)
Bodycare & Beauty	666	('sex', 50), ('life', 45), ('mental', 45), ('depression', 37), ('eating', 37), ('anxiety', 36), ('health', 35), ('orgasm', 30), ('food', 28), ('makeup', 28), ('woman', 28), ('healthy', 26), ('diet', 23), ('masturbation', 20), ('want', 18), ('feeling', 18), ('eat', 17), ('self', 17), ('illness', 16), ('exercise', 16), ('meditation', 15), ('thing', 15), ('way', 15), ('sexual', 15), ('get', 14), ('love', 14), ('day', 14), ('disorder', 13), ('mind', 13)
Physical Activity	440	('workout', 56), ('exercise', 48), ('fitness', 38), ('body', 26), ('weight', 26), ('gym', 25), ('day', 23), ('eat', 21), ('diet', 21), ('fasting', 20), ('food', 19), ('health', 17), ('muscle', 16), ('healthy', 14), ('eating', 14), ('life', 13), ('fat', 12), ('cardio', 10), ('nutrition', 10), ('sport', 9), ('vlog', 9), ('leg', 8), ('training', 8), ('meal', 7), ('calorie', 7), ('lose', 6), ('lifestyle', 6), ('routine', 6), ('ketosis', 6), ('exercising', 5)
Companionship	186	('life', 21), ('love', 14), ('girl', 12), ('eat', 11), ('day', 11), ('mental', 10), ('eating', 10), ('vegan', 10), ('health', 9), ('food', 9), ('self', 9), ('anxiety', 8), ('healthy', 8), ('family', 8), ('sex', 8), ('thing', 8), ('want', 7), ('relationship', 7), ('get', 7), ('way', 7), ('good', 7), ('time', 6), ('one', 6), ('mom', 6), ('baby', 6), ('woman', 6), ('child', 6), ('help', 6), ('confidence', 6), ('know', 5)
Tech-Free & Rest	158	('meditation', 19), ('life', 15), ('sleep', 11), ('yoga', 11), ('relaxation', 7), ('mind', 7), ('get', 6), ('anxiety', 5), ('eating', 5), ('self', 5), ('way', 5), ('meditate', 5), ('love', 5), ('time', 5), ('sleeping', 5), ('stress', 4), ('depression', 4), ('mindfulness', 4), ('meditating', 4), ('diet', 4), ('feeling', 4), ('relationship', 4), ('mental', 4), ('personal', 4), ('deprivation', 4), ('exercise', 3), ('food', 3), ('healthy', 3), ('peace', 3), ('feel', 3)

and Tech-Free & Rest have the largest mean values for both positive and negative emotion.

Regarding time orientation, Physical Activity videos have the highest mean value for 'FocusFuture'. Manual inspection of some of these videos show that they encourage audiences to keep practicing to achieve fitness results. Finally, regarding informal and health language use based on the dictionary of health-related categories in LIWC, Tech-Free & Rest videos have the lowest mean value for informal language, and all health categories have lower mean values than previously reported results on natural speech. For health language use, all six health categories have (unsurprisingly) higher mean values compared to previously reported results on natural speech and text blogs; Physical Activity videos have the highest mean values. As a summary, Physical Activity videos have the highest percentage of authentic, future orientation, and health words. Meanwhile, Tech-Free & Rest videos have the highest percentage of use of complex words and the 'You' pronoun, and the lowest percentage of use of the 'I' pronoun, the 'FocusPast' time orientation, and of 'Informal' language. Finally, Companionship videos have the highest percentage of clout and of the 'We' pronoun. The other three health video categories have intermediate values for most LIWC categories.

Table 5 shows the Pearson correlation across a subset of LIWC categories for all video transcripts (N=2500). We only show correlations above 0.20 and p-value <0.001. For pronouns, 'We' is uncorrelated to most categories except 'Clout'. Meanwhile, 'I' has a positive correlation with 'Authentic' and a negative correlation with 'We', 'You', 'Clout', 'Analytic', 'Sixltr'. Previous work studied LIWC variables on multiple datasets, including blogs, expressive writing, novels, natural speech, and newspapers [65, 77] . The reported correlation between 'I' and 'Authentic' is similar to the one found on the video data data, i.e., the more 'I', the more 'Authentic' across the previously investigated sources of text data. In addition, 'We' and 'You' have a positive correlation with 'Clout'. Finally, 'Clout' has a positive correlation with 'Analytic' and 'Sixltr', but also a negative correlation with 'Authentic'.

4.2 Visual Content

We examined the visual content of the dataset with the goal of identifying visual elements and possible differences across health-related videos made by professional YouTubers. We are interested in understanding the overall visual environment of videos, including the kind of scenes and objects present in the videos. This is motivated by the qualitative analysis from Section 3, which highlighted that health-related videos are shot in specific settings (e.g. a Tech-Free & Rest video in a relaxing outdoor scene), and with deliberate intent.

Visual feature extraction. We are interested in visual representations that are interpretable for further work in health psychology. We first sampled one frame every four seconds for each video. We then applied two different pretrained deep neural network models for (1) scene type detection (also referred to as places in the rest

Table 4: Mean of LIWC categories for YouTube video transcripts (for all videos, and six health-related video categories), and for previous works. For each LIWC category, the highest value is denoted with (**) and the lowest value is denoted with (*). P: Physical Activity; F: Food & Nutrition; S: Self-Development; T: Tech-Free & Rest; B: Bodycare & Beauty; C: Companionship.

Group	Categories	Our dataset							Blogs[77]	Natural Speech [65]
		All	P	F	S	T	B	C		
Text - Analytic	Word count	1.8K	1.6K	1.6K	1.8K	1.4K	1.7K	1.7K	3.2K(**)	794.1(*)
	Tone	74.3	73.85	75.83	74.69	71.96	70.57	74.20	54.50(*)	79.29(**)
	Authentic	59.2	69.42(**)	55.45	62.96	59.84	59.68	52.84(*)	60.93	61.32
	Clout	62.2	63.59	57.26	64.19	70.70	61.01	73.84(**)	47.87(*)	56.27
	Analytic	30.4	29.74	32.06	28.74	34.78	26.08	25.15	49.89(**)	18.43(*)
	Sixltr	13.0	12.64	12.66	14.13	14.77(**)	13.08	13.10	14.38	10.42(*)
Pronouns	You	3.9	4.54	3.42	4.24	4.76(**)	3.91	3.97	1.32(*)	4.04
	I	5.3	5.57	5.51	5.33	4.24(*)	5.55	4.73	6.26	7.03(**)
	We	1.1	1.20	1.02	1.14	1.30	1.00	1.54(**)	0.91	0.87(*)
Emotion	Posemo	3.9	3.89	3.89	3.99	4.21	3.96	4.39	3.66(*)	5.31(**)
	Negemo	0.9	0.88	0.81(*)	1.03	1.31	1.29	1.32	2.06(**)	1.19
Time Orientation	FocusPast	2.4	2.35	2.43	2.36	1.97(*)	2.39	2.72	4.25(**)	3.78
	FocusPresent	14.2	14.66	13.66	14.57	14.27	14.64	15.15	10.95(*)	15.28(**)
	FocusFuture	2.0	2.25(**)	2.21	1.83	1.61	1.89	1.93	1.06(*)	1.45
Language	Informal	1.5	1.80	1.55	1.14	0.92(*)	1.41	2.19	2.09	7.10(**)
	Health	0.9	1.34(**)	0.90	0.88	0.87	0.90	0.69	0.61	0.38(*)

Table 5: Pearson correlation (p-value <0.001) across LIWC categories for full video transcripts (N=2500 videos). Entries marked with (*) correspond to |correlation| <0.20 or p-value >0.001.

	I	We	You	Clout	Authentic	Analytic	Sixltr
I	1	-0.32	-0.53	-0.84	0.70	-0.54	-0.41
We		1	*	0.46	*	*	*
You			1	0.71	-0.30	*	*
Clout				1	-0.56	0.32	0.30
Authentic					1	-0.33	-0.33
Analytic						1	0.27
Sixltr							1

of the paper); and (2) semantic scene segmentation on each frame (also referred to as scene parts or objects in the rest of the paper). We then aggregated this frame-based representation over all frames (using a mean vector estimate) to obtain a representation for a full video. For semantic scene segmentation, we used a Residual Network model, namely Resnet50 trained on the ADE20K dataset [95]. The output of this model on each frame is a 150-dimensional vector which represents an image by 150 scene parts or objects, each described by their relative size in the frame. Using the mean vector for each video, we estimated the mean distribution for all videos in each of the six health-related categories. For place detection, we use Resnet18, a model pretrained on the Place-365 dataset [94]. This model outputs a 365-dimensional vector of scene class probabilities. This model also yields out indoor and outdoor values for each frame. The implementation uses grid computing, and is based on Python using Tensorflow (Tensorflow 2020) and GluonCV (GluonCV 2020) for place detection and scene segmentation, respectively.

Place-based analysis. To analyze the visual content of the health-related categories, we show the top detected places in Figure 2(a). This shows both similarities and differences among the video categories. A few categories were removed on purpose from the analysis because they were not part of the scope of analysis (the 'person' category), they were common to all videos (e.g. the 'wall' category), or they were clearly incorrect from a semantic perspective (e.g. the 'jail cell' category). These steps are important as significant recent work has demonstrated the limitations of existing pre-trained visual models given the datasets they were trained on [9, 34, 43, 78].

Each health-related category contains characteristic scenes. More specifically, Physical Activity videos contain images of gyms or fitness places (e.g., 'gymnasium/indoor', 'martial arts gym', 'ballroom'); Food & Nutrition videos contain kitchen-related places (e.g., 'ice cream parlor', 'kitchen', 'coffee shop'); Tech-Free & Rest videos feature 'bamboo forest'; and Companionship feature 'library/indoor' and 'bookstore'. Interestingly, 'car interior' is also a common element to all videos, as YouTubers often record while being in a car. Based on the inferred values for 'indoor' and 'outdoor', we generated a box plot for the full dataset, shown in Figure 3. We observe that videos for all categories are likely recorded indoors. However, Tech-Free & Rest has both indoor and outdoor cases, suggesting

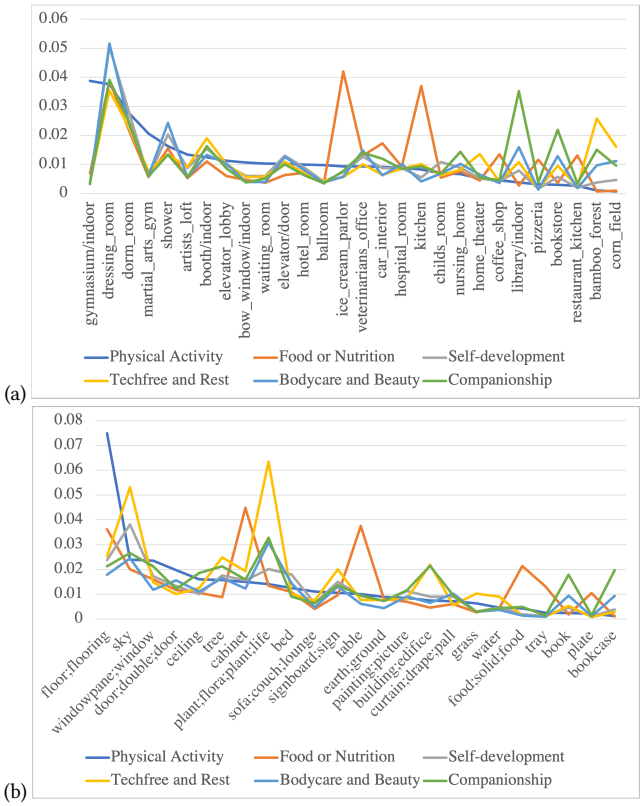

Figure 2: Top 15 detected places and scene parts. The unit of the y-axis is probability.

that this category is distinctive w.r.t. featuring a mixture of both indoor and outdoor, compared to the other health categories.

Scene part-based analysis. Figure 2(b) shows the top scene parts for each of the video categories. The values are based on the percentage of the image covered by the corresponding scene parts. We observe that all health-related videos have characteristic scene parts. For example, 'floor' has the highest percentages in Physical Activity videos; Food & Nutrition features 'cabinet', 'table', or 'food' as top scene parts; and 'sky' and 'plant' are the top scene parts for Tech-Free & Rest videos.

Overall, the results of analysis based on scene types and scene parts complement each other, and suggest that the visual content produced by professional YouTubers follows a rather concrete, and to some extent homogeneous, visual scene content given a health-related theme. One question is whether such regularity would allow for the identification of health-related video categories, which is presented in the next section.

5 Classifying Health-Related Videos (RQ2)

We investigated the feasibility of classifying videos into health-related categories. As videos can belong to one or more categories, an appropriate setting is a binary one, e.g., Food & Nutrition vs. non-Food & Nutrition, and likewise for the other categories. We use the representations described in Section 4 to train a Random Forest (RF), including textual (LIWC from transcripts) and visual (150

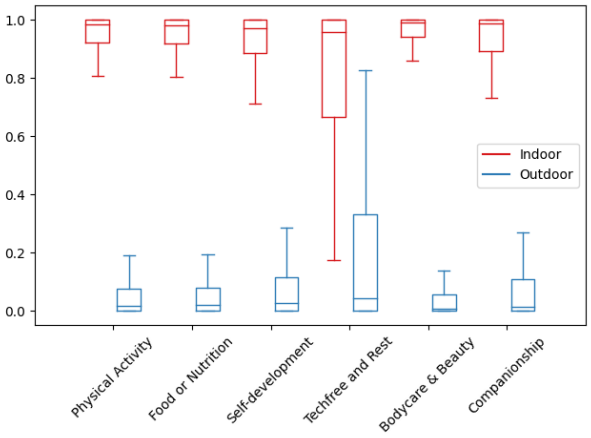

Figure 3: Box plot of indoor/outdoor probabilities. Red: indoor; Blue: outdoor. The unit of the y-axis is probability.

scene parts and 360 places). Importantly, while more sophisticated representations and classifiers could be used for this task, (e.g. word embeddings for text, larger-dimensional layers of the deep networks for images, or fine-tuning of pretrained models), we are interested in results that are fully interpretable given our motivation to potentially build useful tools for health psychology work.

5.1 Experimental Setting

As recommended by [47], for the Random Forest, we set ntree=500 and mtry=3. We used 10-fold cross validation over 5 repeated times for accuracy evaluation. For the classification experiments, we use a different number of data points N_c for each health-related category, which includes 50% of the positive class, and 50% of a random sample from the negative class (coming from the rest of the videos in the dataset). These balanced datasets correspond to a naive random baseline accuracy equal to 50%. The number of data points used is shown in Table 6.

5.2 Classification Results

Table 6 presents the classification results. We observe that the Random Forest tends to performs better for higher N_c values. In addition, all feature groups have performance improvements (15-37%) compared to the random baseline (50%). We observe that visual features (places and scene parts) give comparable accuracies, but lower than those obtained with LIWC features, which provide the best performance (accuracies in the range 0.74-0.87). We also observed that feature fusion (S +L) did not improve performance.

To understand the importance of features in binary classification, Figure 4 shows the top 30 features used by the Random Forest trained with the combination of LIWC features, scene parts, and places for four of the health categories. Note that a higher mean decrease indicates higher variable importance. We observe that for Food & Nutrition, the top features are a combination of LIWC (e.g., ingest, bio), scene parts (e.g., food/solid food, tray, etc.), and places (e.g., ice cream parlor, pizzeria, etc.). For Self-Development, LIWC categories related to cognitive processes and insights are relevant, in addition to other common ones (ingest and bio.) For Bodycare

Table 6: Classification results (accuracy as a 0-1 fraction)

Binary class	N_c	Baseline	(S)cenes 150-dim	(P)laces 360-dim	(L)IWC 87-dim	S + L
Food & Nutrition	2120	0.5	0.80	0.80	0.87	0.84
Self-Development	1374	0.5	0.69	0.69	0.75	0.75
Bodycare & Beauty	1328	0.5	0.71	0.72	0.78	0.76
Physical Activity	878	0.5	0.72	0.76	0.78	0.79
Companionship	370	0.5	0.65	0.69	0.77	0.76
Tech-Free & Rest	316	0.5	0.68	0.66	0.74	0.72

& Beauty, LIWC categories related to ingest, body, and perception, and motion are the most relevant features, as well as visual cues related to food and food places. For Physical Activity, places like gymnasium indoor or basketball court, and LIWC categories related to health, leisure, and motion are the most relevant features. For Techfree & Rest, LIWC categories related to future time orientation and religion are relevant, in addition to features (both linguistic and visual) related to eating. Finally, 13 LIWC features are among the 15 features relevant to classify Companionship. It is interesting to see that LIWC categories are consistently among the top 1-2 features to classify all health-related categories. This highlights the importance of verbal features, which as shown in the results presented in Table 6, and produces the highest classification results.

6 Discussion

Health videos by professional YouTubers are multifaceted. In section 3, we found that that over one quarter of all videos combine multiple categories. While 72.1% of videos are about one single category, 21.3% belong to two categories, and 6.6% to three or more categories. Overall, Food & Nutrition, Self-Development, and Bodycare & Beauty are categories present in over 25% of videos. Physical Activity and Food & Nutrition also have the largest audiences (measured by their mean view count), possibly due to their direct appeal to viewers [53]. In light of this result, it is interesting to point out that the large-scale YouTube-8M video dataset released by YouTube [93], consists of millions of videos encompassing multiple categories, e.g., Art & Entertainment, Game, Sports, Food & Drink, etc. The top five YouTube categories (measured by their video count) include Sport (858K videos) and Food & Drink (683K videos), which are conceptually similar to our health-related categories extracted via polytextual thematic analysis, i.e., Physical Activity and Food & Nutrition.

Semantic patterns emerge in linguistic content of health-related videos. LIWC-based analysis of YouTube video blogs is not new [8], but its application to health videos of professional YouTubers brings a novel angle. In Section 4, we showed that video categories have specific patterns with respect to LIWC categories. For instance, Physical Activity videos use the highest average of words about health, focus on the future, and authenticity, compared to the rest of the videos. The second single pronoun (i.e., 'You') and longer words have the highest average values in Tech-Free & Rest videos. Companionship videos have the highest use of the 'We' pronoun. All health-categories have a larger proportion of positive emotion terms than negative ones, the highest percentage corresponding to the Companionship category. We also found some similarities and differences with respect to other forms of text anaylyzed by LIWC.

Visual content of health-related videos generally matched expectations. The results in Section 4 also showed specific visual patterns related to scene types and scene parts, which are largely consistent with our expectations. For instance, a higher relative proportion of gym scenes appear on Physical Activity videos have; kitchen scenes on Food & Nutrition videos, and library scenes on Companionship videos. Most health-related videos were recorded indoor. In contrast, the environments of Tech-Free & Rest reveals outdoor aspects related to restorative environments [12, 59, 88, 89]. The relative visual consistency of videos by professional YouTubers open some questions related both to performative aspects and marketing aspects, i.e., to the motivations that lead to the creation of videos in the first place. Answering this question would require future qualitative research.

Applications in health psychology. The classification results in Section 5 can be seen as an feasibility study towards a video tagging framework to support health psychology research. Such tool could be used to explore video data for hypothesis formation, but also to validate hypotheses generated with qualitative methods in psychology concerning smaller data samples. The characterization in sections 3-4 could support interpretative analysis, e.g., for visual semantic characterization. We believe that HCI research could play an important role from the perspective of human-centered design and evaluation of such functionalities. More generally, a mixed-methods approach that combines qualitative and quantitative analyses could be useful to provide nuanced analyses regarding the complexity of health-related information in YouTubers' videos. For instance, the finding that influential YouTubers follow a form of 'standardized model' that affects their audiences [60] has implications for health psychology research. YouTubers develop unprecedented forms of health expertise through specific textual and visual means facilitated by YouTube. This new expertise [87] can be scientifically based but not associated to that of health professionals or scientific experts [6]. Future research in health psychology could analyze the semantic patterns specific to health-related videos on YouTube to better understand how experience-based knowledge is promoted by YouTubers to prevent chronic illness, such as cardiovascular disease through regular exercise. Furthermore, future research perspectives may explore YouTube content in non-Western contexts. This is another area where HCI research could be very valuable [46].

Non-expert health advice from YouTubers. Professional YouTubers are often viewed as role models by youth [39, 74]. However, some may give advice on health promotion or illness prevention based on their own experience, beyond scientific expertise [39]. For instance, one of the YouTubers explained in a video that after being diagnosed with a chronic illness, she made 'junk food' unavailable to her to avoid unhealthy food temptations. Another example of this kind of experience-based advice concerned the need to stay

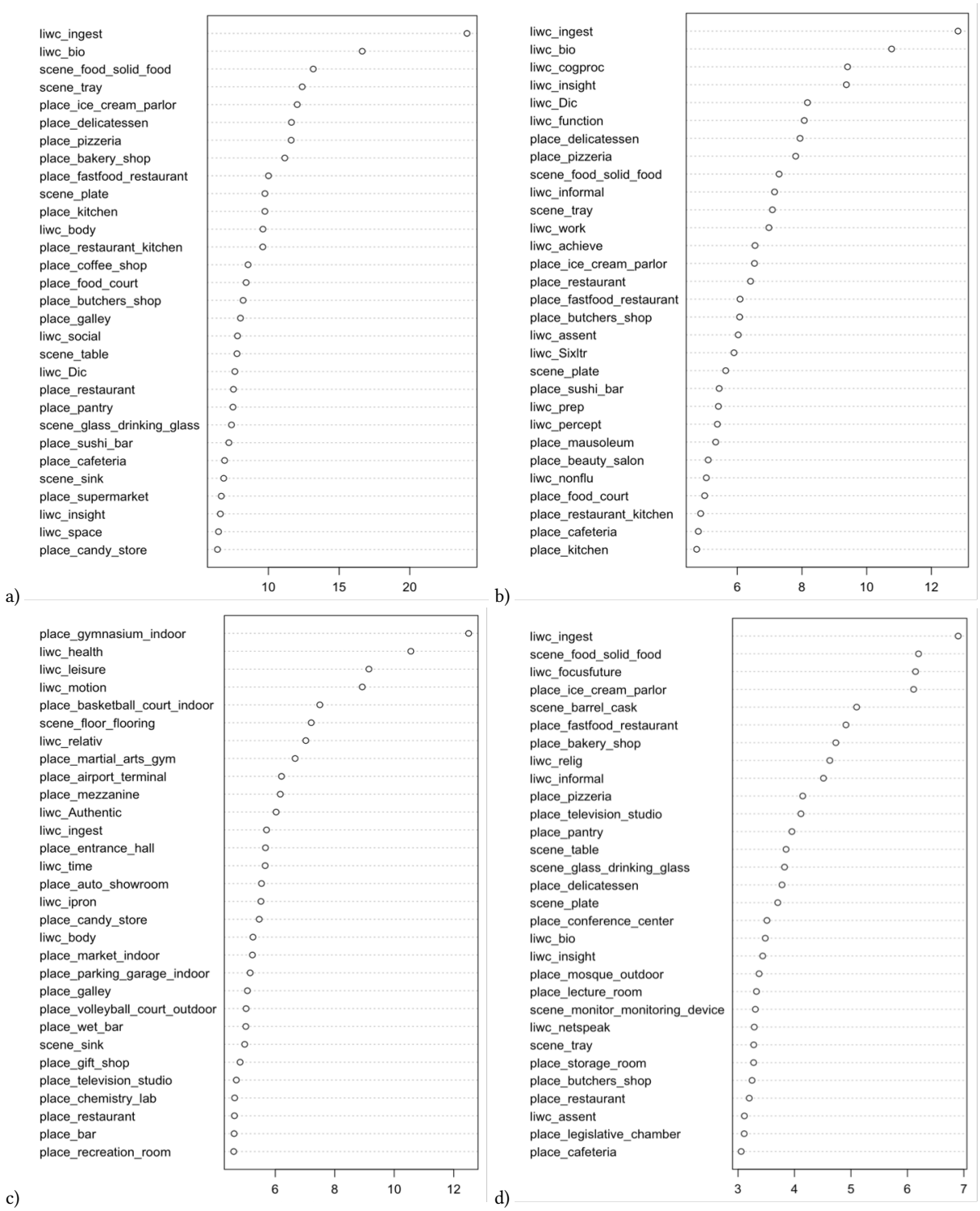

Figure 4: 30 most important features for binary classification: (a) Food & Nutrition, (b) Self-Developemnt, (c) Physical Activity, and (d) Tech-Free & Rest. The x-axis is a Mean Decrease in Gini. A higher value indicates higher variable importance.

hydrated to improve physical and cognitive performances. While this clearly goes beyond the scope of the paper, it is important to further investigate the impact of such advice upon viewers, which may involve the analyses of viewers' interpretation and contexts of use of such videos. Furthermore, professional YouTubers often financially rely on product placement and advertisement of specific health-related goods. The influence of commercial purposes around health videos deserves further attention. Beyond the potential beneficial uses of YouTube health videos among viewers, misinformation is an important issue that needs to be studied in future research [41].

Ethical considerations. There are several issues that follow from our study. First, sensitive topics, e.g., related to mental health are discussed. There are risks associated to linking such topics to individuals, even if they themselves disclose such information on YouTube. By design, all results presented here are in anonymous and aggregated form. Second, certain controversial health practices might be also talked about, specifically related to eating and exercising. Recent research has started to study these issues in the specific context of YouTube [70]. Similar issues have been raised in the context of other social media platforms by social computing and HCI researchers [14]. The fact that non-experts can freely give advice on health to a potential large audience has additional ethical and health implications [27]. Finally, data protection regulations vary across different world regions, and the terms of use of social media platforms also change, sometimes rapidly.

Generality of findings. One key limitation has to do with the generality of the presented results. Our work analyzed content of 50 popular, professional, English-speaking YouTubers, which represents a small sample that was accessible through the (largely unknown) search mechanisms built in YouTube. While a systematic procedure was followed to define the sample (Section 3), clearly there are limitations in terms of representation. We cannot make any assertions on whether the observed trends would apply to other YouTubers who talk about health. Starting from the analysis presented here, future work would be needed to understand whether less popular video makers who talk about health, or video makers from other world regions, follow similar trends as discussed in this paper. Investigating this aspect is critical to account for user diversity, and would be part of future research.

7 Conclusion

In this paper, we presented a comparative study of the characteristics of six health-related categories on videos produced by professional YouTubers, including manual health coding of videos, textual and visual feature extraction, and classification of health-related videos. We conclude by answering the two research questions we posed:

RQ1: What health-related categories can be extracted from polytextual thematic analysis, and what are the observable similarities and differences among such health categories in terms of audio-visual content? After collecting health-related videos, we analyzed the patterns from textual content (titles/tags and transcripts) and visual content (places and semantic object segmentation) on the six health-related categories. We observed linguistic patterns through LIWC categories in mean values and correlation. We showed that linguistic patterns in health-related videos are differently formed w.r.t. pronouns, emotion, or time orientations. Finally, we used pre-trained deep learning models to extract places and scene parts to represent videos. Most of these descriptors relevantly characterized health-related videos.

RQ2: How can audio-visual descriptors of videos be used to classify them according to health categories for further use in psychology research? Regarding classification, textual features from LIWC were used for binary classification with the highest accuracy (and less computational effort) compared to visual features. The total number of videos in our dataset and the relative number of videos per health-related category could be a limitation for model training for some classes. However, our results show that health categories can be classified with accuracy in the [0.74, 0.87] range compared to a random baseline of 0.50. This could be used to generate health-related category tags for health videos, so researchers in health psychology use them for data exploration, problem definition, and hypothesis validation.

Acknowledgments

The authors acknowledge the support of the Swiss National Science Foundation through the Health Vlogging project (Spark call, Grant 190387). D. Gatica-Perez acknowledges the additional support provided through the AI4Media project, funded by the European Commission (Grant 951911) under the H2020 Programme ICT-48-2020.

References

[1] 2022. *Global YouTube viewing time share by device 2021.* Retrieved Aug. 2022 from https://www.statista.com/statistics/1173543/youtube-viewing-time-share-device/
[2] Osman Hassan Ahmed, S John Sullivan, Anthony G Schneiders, and Paul Mccrory. 2010. iSupport: do social networking sites have a role to play in concussion awareness? *Disability and rehabilitation* 32, 22 (2010), 1877–1883.
[3] Icek Ajzen. 1985. From intentions to actions: A theory of planned behavior. In *Action control*. Springer, 11–39.
[4] Einas Al-Eisa, Asma Al-Rushud, Ahmad Alghadir, Shahnawaz Anwer, Bashayer Al-Harbi, Noha Al-Sughaier, Noha Al-Yoseef, Reem Al-Otaibi, and Hanadi Ali Al-Muhaysin. 2016. Effect of motivation by "Instagram" on adherence to physical activity among female college students. *BioMed research international* 2016 (2016).
[5] Oya Aran, Joan-Isaac Biel, and Daniel Gatica-Perez. 2014. Broadcasting Oneself: Visual Discovery of Vlogging Styles. *IEEE Transactions on Multimedia* 16, 1 (2014), 201–215.
[6] Stephanie Alice Baker and Chris Rojek. 2020. The Belle Gibson scandal: The rise of lifestyle gurus as micro-celebrities in low-trust societies. *Journal of Sociology* 56, 3 (2020), 388–404.
[7] Jacqueline L Bender, Maria-Carolina Jimenez-Marroquin, and Alejandro R Jadad. 2011. Seeking support on facebook: a content analysis of breast cancer groups. *Journal of medical Internet research* 13, 1 (2011), e16.
[8] Joan-Isaac Biel, Vagia Tsiminaki, John Dines, and Daniel Gatica-Perez. 2013. Hi YouTube! Personality impressions and verbal content in social video. In *Proceedings of the 15th ACM on International conference on multimodal interaction*. 119–126.
[9] A. Birhane and V. Uday Prabhu. 2021. Large datasets: a pyrrhic win for computer vision?. In *Proc. IEEE/CVF Winter Conference on Applications of Computer Vision*.
[10] Stanley Blue, Elizabeth Shove, Chris Carmona, and Michael P Kelly. 2016. Theories of practice and public health: understanding (un) healthy practices. *Critical Public Health* 26, 1 (2016), 36–50.
[11] Katya Borgos-Rodriguez, Kathryn E Ringland, and Anne Marie Piper. 2019. Myautsomefamilylife: Analyzing parents of children with developmental disabilities on YouTube. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.
[12] Ralf C Buckley, Paula Brough, and Diane Westaway. 2018. Bringing outdoor therapies into mainstream mental health. *Frontiers in public health* 6 (2018), 119.
[13] Elise R Carrotte, Alyce M Vella, and Megan SC Lim. 2015. Predictors of "liking" three types of health and fitness-related content on social media: a cross-sectional study. *Journal of medical Internet research* 17, 8 (2015), e205.

[14] Stevie Chancellor, Michael L. Birnbaum, Eric D. Caine, Vincent M. B. Silenzio, and Munmun De Choudhury. 2019. A Taxonomy of Ethical Tensions in Inferring Mental Health States from Social Media. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 79–88.

[15] Stevie Chancellor, Jessica Annette Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. 2016. # thyghgapp: Instagram content moderation and lexical variation in pro-eating disorder communities. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 1201–1213.

[16] Jordaan Charlotte. 2019. #fitgirl: The influence of vloggers on health behavior To what extent do social media influencers affect the acceptance of health messages? *Bachelor thesis Communication and Information Science Specialization, Tilburg University, Tilburg* (2019).

[17] Xu Cheng, Cameron Dale, and Jiangchuan Liu. 2008. Statistics and social network of youtube videos. In *2008 16th Interntional Workshop on Quality of Service*. IEEE, 229–238.

[18] Cynthia Chew and Gunther Eysenbach. 2010. Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PloS one* 5, 11 (2010), e14118.

[19] Wen-Ying Sylvia Chou, Yvonne Hunt, Anna Folkers, and Erik Augustson. 2011. Cancer survivorship in the age of YouTube and social media: a narrative analysis. *Journal of medical Internet research* 13, 1 (2011), e7.

[20] Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A Munson. 2017. When personal tracking becomes social: Examining the use of Instagram for healthy eating. In *Proceedings of the 2017 CHI Conference on human factors in computing systems*. 1674–1687.

[21] Michael A Cohn, Matthias R Mehl, and James W Pennebaker. 2004. Linguistic markers of psychological change surrounding September 11, 2001. *Psychological science* 15, 10 (2004), 687–693.

[22] Robert Crawford. 1980. Healthism and the medicalization of everyday life. *International journal of health services* 10, 3 (1980), 365–388.

[23] Rachel G Curtis, Jillian C Ryan, Sarah M Edney, and Carol A Maher. 2020. Can Instagram be used to deliver an evidence-based exercise program for young women? A process evaluation. *BMC public health* 20, 1 (2020), 1–9.

[24] Denise de Ridder. 2015. Health Psychology. In *International Encyclopedia of the Social Behavioral Sciences (Second Edition)* (second edition ed.), James D. Wright (Ed.). Elsevier, Oxford, 674–678. https://doi.org/10.1016/B978-0-08-097086-8.14099-1

[25] María del Río Carral and Antonia Lyons. 2020. Embodying health behaviours in everyday life: the social and gendered practices of female senior managers. *Psychology & health* 35, 10 (2020), 1249–1267.

[26] María del Río Carral, Lucia Volpato, Chloé Michoud, Thanh-Trung Phan, and Daniel Gatica-Perez. 2021. Professional YouTubers' health videos as research material: Formulating a multi-method design in health psychology. *Methods in Psychology* 5 (2021), 100051.

[27] Daniel Diethei, Ashley Colley, Julian Wienert, and Johannes Schöning. 2022. Different Length, Different Needs: Qualitative Analysis of Threads in Online Health Communities. *arXiv preprint arXiv:2201.01688* (2022).

[28] Nina Eyrich, Monica L Padman, and Kaye D Sweetser. 2008. PR practitioners' use of social media tools and communication technology. *Public relations review* 34, 4 (2008), 412–414.

[29] Adam D Farmer, CEM Bruckner Holt, MJ Cook, and SD Hearing. 2009. Social networking sites: a novel portal for communication. *Postgraduate medical journal* 85, 1007 (2009), 455–459.

[30] Luis Fernandez-Luque, Randi Karlsen, and Genevieve B Melton. 2012. HealthTrust: a social network approach for retrieving online health videos. *Journal of medical Internet research* 14, 1 (2012), e22.

[31] Jessica L Feuston and Anne Marie Piper. 2018. Beyond the coded gaze: Analyzing expression of mental health and illness on instagram. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–21.

[32] Florencia García-Rapp. 2017. Popularity markers on YouTube's attention economy: the case of Bubzbeauty. *Celebrity Studies* 8, 2 (2017), 228–245.

[33] Maria Teresa Garcia-Romero, Fernanda Prado, Judith Dominguez-Cherit, Maria Teresa Hojyo-Tomomka, and Roberto Arenas. 2011. Teledermatology via a social networking web site: a pilot study between a general hospital and a rural clinic. *Telemedicine and e-Health* 17, 8 (2011), 652–655.

[34] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. 2021. Datasheets for Datasets. *Commun. ACM* 64, 12 (Decembe 2021), 86–92.

[35] Kate Gleeson. 2012. 21 Polytextual Thematic Analysis for Visual Data. *Visual methods in psychology: Using and interpreting images in qualitative research* (2012), 314.

[36] Gaston Godin and Gerjo Kok. 1996. The theory of planned behavior: a review of its applications to health-related behaviors. *American journal of health promotion* 11, 2 (1996), 87–98.

[37] Victoria A Goodyear, Kathleen M Armour, and Hannah Wood. 2018. Young people and their engagement with health-related social media: New perspectives. *Sport, education and society* (2018).

[38] Jeremy A Greene, Niteesh K Choudhry, Elaine Kilabuk, and William H Shrank. 2011. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. *Journal of general internal medicine* 26, 3 (2011), 287–292.

[39] J Harris. 2019. *A mixed methods study examining the role of professional YouTubers in young people's health behaviours in the UK: implications for health interventions*. Ph.D. Dissertation. Liverpool John Moores University.

[40] N Heaivilin, B Gerbert, JE Page, and JL Gibbs. 2011. Public health surveillance of dental pain via Twitter. *Journal of dental research* 90, 9 (2011), 1047–1051.

[41] Rui Hou, Veronica Perez-Rosas, Stacy Loeb, and Rada Mihalcea. 2019. Towards Automatic Detection of Misinformation in Online Medical Videos. In *2019 International Conference on Multimodal Interaction (ICMI '19)*. Association for Computing Machinery, New York, NY, USA, 235–243. https://doi.org/10.1145/3340555.3353763

[42] Jina Huh, Leslie S Liu, Tina Neogi, Kori Inkpen, and Wanda Pratt. 2014. Health vlogs as social support for chronic illness management. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 4 (2014), 1–31.

[43] Ben Hutchinson, Andrew Smart, Alex Hanna, Emily Denton, Christina Greer, Oddur Kjartansson, Parker Barnes, and Margaret Mitchell. 2021. Towards Accountability for Machine Learning Datasets: Practices from Software Engineering and Infrastructure. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. 560–575.

[44] Ewa Kacewicz, James W Pennebaker, Matthew Davis, Moongee Jeon, and Arthur C Graesser. 2014. Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology* 33, 2 (2014), 125–143.

[45] Terry K Koo and Mae Y Li. 2016. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine* 15, 2 (2016), 155–163.

[46] Neha Kumar, Vikram Kamath Cannanure, Dilrukshi Gamage, Annu Sible Prabhakar, Christian Sturm, Cuauhtémoc Rivera Loaiza, Dina Sabie, Md. Moinuddin Bhuiyan, and Mario A. Moreno Rocha. 2020. HCI Across Borders and Sustainable Development Goals. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems (CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–8.

[47] Andy Liaw, Matthew Wiener, et al. 2002. Classification and regression by randomForest. *R news* 2, 3 (2002), 18–22.

[48] Sam Liu, Megan Perdew, Alexander Lithopoulos, and Ryan E Rhodes. 2021. The Feasibility of Using Instagram Data to Predict Exercise Identity and Physical Activity Levels: Cross-sectional Observational Study. *Journal of Medical Internet Research* 23, 4 (2021), e20954.

[49] Xiao Liu, Bin Zhang, Anjana Susarla, and Rema Padman. 2018. YouTube for Patient Education: A Deep Learning Approach for Understanding Medical Knowledge from User-Generated Videos. arXiv:cs.CV/1807.03179

[50] Alto S Lo, Michael J Esser, and Kevin E Gordon. 2010. YouTube: a gauge of public perception and awareness surrounding epilepsy. *Epilepsy & Behavior* 17, 4 (2010), 541–545.

[51] Antonia C Lyons and Kerry Chamberlain. 2017. Critical health psychology. In *The Palgrave handbook of critical social psychology*. Springer, 533–555.

[52] Lydia Manikonda and Munmun De Choudhury. 2017. Modeling and understanding visual attributes of mental health disclosures in social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. 170–181.

[53] Michael Marmot, T Atinmo, T Byers, J Chen, T Hirohata, A Jackson, W James, L Kolonel, S Kumanyika, C Leitzmann, et al. 2007. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. (2007).

[54] Joseph D Matarazzo. 1980. Behavioral health and behavioral medicine: frontiers for a new health psychology. *American psychologist* 35, 9 (1980), 807.

[55] Frances Mielewczyk and Carla Willig. 2007. Old clothes and an older look: the case for a radical makeover in health behaviour research. *Theory & Psychology* 17, 6 (2007), 811–837.

[56] Tomas Mikolov, Edouard Grave, Piotr Bojanowski, Christian Puhrsch, and Armand Joulin. 2018. Advances in Pre-Training Distributed Word Representations. In *Proceedings of the International Conference on Language Resources and Evaluation (LREC 2018)*.

[57] S Anne Moorhead, Diane E Hazlett, Laura Harrison, Jennifer K Carroll, Anthea Irwin, and Ciska Hoving. 2013. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. *Journal of medical Internet research* 15, 4 (2013), e85.

[58] Volha Murashka, Jiaying Liu, and Yilang Peng. 2020. Fitspiration on Instagram: Identifying Topic Clusters in User Comments to Posts with Objectification Features. *Health Communication* (2020), 1–12.

[59] Michael Mutz and Johannes Müller. 2016. Mental health benefits of outdoor adventures: Results from two pilot studies. *Journal of adolescence* 49 (2016), 105–114.

[60] Vaibhavi Nandagiri and Leena Philip. 2018. Impact of influencers from Instagram and YouTube on their followers. *International Journal of Multidisciplinary Research and Modern Education* 4, 1 (2018), 61–65.

[61] Matthew L Newman, James W Pennebaker, Diane S Berry, and Jane M Richards. 2003. Lying words: Predicting deception from linguistic styles. *Personality and*

social psychology bulletin 29, 5 (2003), 665–675.

[62] Ferda Ofli, Yusuf Aytar, Ingmar Weber, Raggi al Hammouri, and Antonio Torralba. 2017. Is Saki delicious? The Food Perception Gap on Instagram and Its Relation to Health. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 509–518. https://doi.org/10.1145/3038912.3052663

[63] Jane Ogden. 2003. Some problems with social cognition models: A pragmatic and conceptual analysis. *Health psychology* 22, 4 (2003), 424.

[64] Jane Ogden. 2012. *Health psychology: A textbook: A textbook*. McGraw-Hill Education (UK).

[65] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015*. Technical Report.

[66] James W Pennebaker, Cindy K Chung, Joey Frazee, Gary M Lavergne, and David I Beaver. 2014. When small words foretell academic success: The case of college admissions essays. *PloS one* 9, 12 (2014), e115844.

[67] Thanh-Trung Phan and Daniel Gatica-Perez. 2017. Healthy# fondue# dinner: analysis and inference of food and drink consumption patterns on instagram. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia*. ACM, 327–338.

[68] Ladislav Pilař, Lucie K Stanislavská, Roman Kvasnička, Richard Hartman, and Ivana Tichá. 2021. Healthy Food on Instagram Social Network: Vegan, Homemade and Clean Eating. *Nutrients* 13, 6 (2021), 1991.

[69] James O Prochaska and Wayne F Velicer. 1997. The transtheoretical model of health behavior change. *American journal of health promotion* 12, 1 (1997), 38–48.

[70] Priyanjali Ratwatte and Emily Mattacola. 2019. An Exploration of 'Fitspiration'content on YouTube and its impacts on consumers. *Journal of health psychology* (2019), 1359105319854168.

[71] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.

[72] Richard M Ryan and Edward L Deci. 2000. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55, 1 (2000), 68.

[73] Kamran P Sajadi and Howard B Goldman. 2011. Social networks lack useful content for incontinence. *Urology* 78, 4 (2011), 764–767.

[74] MD Nazmus Sakib, Mohammadali Zolfagharian, and Atefeh Yazdanparast. 2020. Does parasocial interaction with weight loss vloggers affect compliance? The role of vlogger characteristics, consumer readiness, and health consciousness. *Journal of Retailing and Consumer Services* 52 (2020).

[75] Marcel Salathé and Shashank Khandelwal. 2011. Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS Comput Biol* 7, 10 (2011), e1002199.

[76] Daniel Scanfeld, Vanessa Scanfeld, and Elaine L Larson. 2010. Dissemination of health information through social networks: Twitter and antibiotics. *American journal of infection control* 38, 3 (2010), 182–188.

[77] Jonathan Schler, Moshe Koppel, Shlomo Argamon, and James W Pennebaker. 2006. Effects of age and gender on blogging.. In *AAAI spring symposium: Computational approaches to analyzing weblogs*, Vol. 6. 199–205.

[78] S. Shankar, Y. Halpern, E. Breck, J. Atwood, J. Wilson, and D. Sculley. 2017. No Classification without Representation: Assessing Geodiversity Issues in Open Data Sets for the Developing World. In *Proc. NeurIPS Workshop on Machine Learning for the Developing World*.

[79] Alessio Signorini, Alberto Maria Segre, and Philip M Polgreen. 2011. The use of Twitter to track levels of disease activity and public concern in the US during the influenza A H1N1 pandemic. *PloS one* 6, 5 (2011), e19467.

[80] Amy Slater, Neesha Varsani, and Phillippa C Diedrichs. 2017. # fitspo or# loveyourself? The impact of fitspiration and self-compassion Instagram images on women's body image, self-compassion, and mood. *Body image* 22 (2017), 87–96.

[81] Stephen Sutton. 2015. Health Behavior, Psychosocial Theories of. In *International Encyclopedia of the Social Behavioral Sciences (Second Edition)* (second edition ed.), James D. Wright (Ed.). Elsevier, Oxford, 577–581. https://doi.org/10.1016/B978-0-08-097086-8.14153-4

[82] M. Iftekhar Tanveer, Samiha Samrose, Raiyan Abdul Baten, and M. Ehsan Hoque. 2018. Awe the Audience: How the Narrative Trajectories Affect Audience Perception in Public Speaking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173598

[83] Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: LIWC and computerized text analysis methods. *Journal of language and social psychology* 29, 1 (2010), 24–54.

[84] Rosemary Thackeray, Brad L Neiger, Carl L Hanson, and James F McKenzie. 2008. Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. *Health promotion practice* 9, 4 (2008), 338–343.

[85] Leslie Thomson. 2017. "In my humble opinion…": Serious YouTubers' self-presentations, roles, and authority. *Proceedings of the Association for Information Science and Technology* 54, 1 (2017), 811–813.

[86] Yan Tian. 2010. Organ donation on Web 2.0: content and audience analysis of organ donation videos on YouTube. *Health communication* 25, 3 (2010), 238–246.

[87] Marika Tiggemann, Isabella Anderberg, and Zoe Brown. 2020. # Loveyourbody: The effect of body positive Instagram captions on women's body image. *Body image* 33 (2020), 129–136.

[88] Margarita Triguero-Mas, Payam Dadvand, Marta Cirach, David Martínez, Antonia Medina, Anna Mompart, Xavier Basagaña, Regina Gražulevičienė, and Mark J Nieuwenhuijsen. 2015. Natural outdoor environments and mental and physical health: relationships and mechanisms. *Environment international* 77 (2015), 35–41.

[89] Margarita Triguero-Mas, David Donaire-Gonzalez, Edmund Seto, Antònia Valentín, David Martínez, Graham Smith, Gemma Hurst, Glòria Carrasco-Turigas, Daniel Masterson, Magdalena van den Berg, et al. 2017. Natural outdoor environments and mental health: Stress as a possible mechanism. *Environmental research* 159 (2017), 629–638.

[90] Gabrielle Turner-McGrievy and Deborah Tate. 2011. Tweets, Apps, and Pods: Results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. *Journal of medical Internet research* 13, 4 (2011), e120.

[91] Christopher J Warren, Rohan Sawhney, Tejash Shah, Sara Behbahani, and Hossein Sadeghi-Nejad. 2021. YouTube and Men's Health: A Review of the Current Literature. *Sexual Medicine Reviews* (2021).

[92] Martin Wollmer, Felix Weninger, Tobias Knaup, Bjorn Schuller, Congkai Sun, Kenji Sagae, and Louis-Philippe Morency. 2013. YouTube Movie Reviews: Sentiment Analysis in an Audio-Visual Context. *IEEE Intelligent Systems* 28, 3 (May 2013), 46–53. https://doi.org/10.1109/MIS.2013.34

[93] YouTube. 2017. *YouTube-8M Dataset*. Retrieved Dec 22, 2021 from https://research.google.com/youtube8m/

[94] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2017. Places: A 10 million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).

[95] Bolei Zhou, Hang Zhao, Xavier Puig, Sanja Fidler, Adela Barriuso, and Antonio Torralba. 2017. Scene parsing through ade20k dataset. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 633–641.

[96] Joy Zhou and Lionel Bercovitch. 2018. Instagram and the dermatologist: an ethical analysis. *Journal of the American Academy of Dermatology* 78, 6 (2018), 1226–1228.

[97] Maria Zimmermann and Regina Jucks. 2019. With a view to the side: YouTube's sidebar and YouTuber's linguistic style as hints for trust-related evaluations. *International Journal of Human–Computer Interaction* 35, 14 (2019), 1279–1291.