Enhancing Audiences' Engagement with Health Education Videos Visual Call or Textual Call?

Tong, Jingjing; Xu, Jingjun (David)

Published in:
Proceedings of the 57th Hawaii International Conference on System Sciences

Published: 01/01/2024

Document Version:
Final Published version, also known as Publisher's PDF, Publisher's Final version or Version of Record

Publication details:
Tong, J., & Xu, J. (2024). Enhancing Audiences' Engagement with Health Education Videos: Visual Call or Textual Call? In T. X. Bui (Ed.), *Proceedings of the 57th Hawaii International Conference on System Sciences* (pp. 3800-3809) https://scholarspace.manoa.hawaii.edu/items/493eefd3-bc0b-400e-b8d8-4e30bed66e47

Enhancing Audiences' Engagement with Health Education Videos: Visual Call or Textual Call?

Jingjing Tong
City University of Hong Kong
Jingjtong4-c@my.cityu.edu.hk

Jingjun (David) Xu
City University of Hong Kong
davidxu@cityu.edu.hk

Abstract

With the rising popularity of video-based social media, more and more doctors are utilizing them to provide health education videos and employing the call-to-action strategy to enhance video engagement. Although the effectiveness of the call-to-action has been established in customer-brand communication literature, it remains unknown how it affects engagement with health education videos. Drawing on multiple literature streams, we propose that visual call positively affects video engagement, while textual call has a negative effect. Additionally, the medical term use weakens the positive impact of visual call and strengthens the negative impact of textual call. Our analysis results support the positive impact of visual call and the negative effect of textual call on audiences' liking and collecting behavior, which are contingent upon medical term use. Our study contributes to research on technology-mediated patient education and the literature on call-to-action and provides implications for doctors and platform managers.

Keywords: Health education, video-based social media, medical term use, call to action

1. Introduction

Today's society is witnessing a process of change in which social media platforms shift from text-centric to visual-oriented experiences (Li & Xie, 2020). Recently, a multitude of video-based social media platforms, such as Tik Tok, have experienced rapid growth (Xie et al., 2019). These platforms enable users to consume, participate, and create videos that can range in duration from a few seconds to several minutes (Omar et al., 2021). Due to the increasing popularity of video-based social media, many doctors are utilizing these platforms to offer health education, thereby enhancing public health literacy and boosting their offline appointments (Zhu et al., 2020). According to a report from Byte Dance (2023), more than 35,000 doctors on Tik Tok create approximately 21,000 health education videos every day, which are viewed by 200 million users on a daily basis.

However, active engagement with videos, manifested by social media behavior, such as leaving likes, is required to fully realize the benefits of health education videos for the general public because engagement is a proxy of audiences' accepting and utilizing medical information (Liu et al., 2020). On the other hand, audiences' engagement is associated with doctors' platform performance because videos that receive numerous likes or comments are more likely to be recommended by platforms, attracting more audiences, and enhancing doctors' online reputation. This online exposure can have positive spillover effects, resulting in an increase in both doctors' offline appointments and revenues (Lu & Wu, 2019). Given the importance of video engagement, an enduring crucial query for platform managers and doctors revolves around how to drive engagement with health education videos. The "call-to-action" strategy, which involves providing incentives to motivate users to perform certain behaviors (Oltra et al., 2022), has been utilized by doctors from diverse fields (Ju et al., 2023).

In this study, we conceptualize the call-to-action as explicit suggestions from doctors that ask audiences to interact with their health education videos by either offering likes, leaving comments, or adding videos to collections. This strategy is described as the most popular and actionable practice in social media adopted by practitioners (Ju et al., 2023). Recent studies have demonstrated its effectiveness in driving customer-brand engagement in social media (Moran et al., 2020; Oltra et al., 2022). However, it is still unknown whether this strategy is beneficial in enhancing engagement with health education videos. This strategy may also enhance video engagement as audiences are reminded, increasing the likelihood of their engagement with the video as a way of showing appreciation for doctors' unpaid service. However, unlike in customer-brand communication contexts, where audiences are aware that brands are profit-

URI: https://hdl.handle.net/10125/106842
978-0-9981331-7-1

oriented, and their engagement benefits the focal organization, audiences in this context tend to believe that delivering health education via social media is a non-profit behavior (Zhu et al., 2020). A call-to-action from doctors may raise doubts among audiences regarding the doctors' intention of providing health education, which is typically considered as welfare-driven behavior. These doubts may lead audiences to believe that doctors are trying to benefit themselves, which contradicts the stereotype of doctors (Howe et al., 2019), thereby reducing the motivation to engage with the video. The unique characteristics of this context make it challenging to generalize prior research findings to this context.

In addition, prior research primarily focused on the textual call-to-action (abbreviated as textual call) that is inserted into the message included in post texts (Jung et al., 2020; Moran et al., 2020). In this context, doctors are shown in the video delivering medical information, and therefore they can convey visual call-to-action (abbreviated as visual call[1]) by stating, for instance, "Please give me a like." We distinguish between the textual call and visual call based on the communication modality used to deliver the call. The textual call is delivered through written messages displayed in the video without verbal cues or doctors' attendance, whereas the visual call is verbally expressed by the doctor with dynamic facial expressions visible. However, it is unknown how textual and visual calls may differently affect audiences' engagement with health education videos. Therefore, our first research question is: *how do visual and textual calls affect audiences' engagement with health education videos on video-based social media platforms?*

Furthermore, previous research on the impact of call-to-action on engagement behavior has mainly focused on customer-brand communication, where the communication content is easily understood by the audience (Moran et al., 2020). Thus, limited research attention has been given to understanding how the call-to-action strategy affects engagement when there are communication barriers between the two parties. The healthcare field is characterized by high professional barriers and is rich in terminologies, which can hinder efficient doctor-patient communication (Xu, 2021). The use of medical terms can negatively affect patients' evaluations of medical information, understanding of their health conditions, treatment adherence, and perceived communication quality (Dahm, 2012; Wernick et al., 2016). Thus, the medical term use may weaken positive associations between the call-to-action and video engagement. However, the use of terminology is also considered as a signal of competence, increasing the credibility of information senders (Moldovan, 2022; Xu, 2021). Hence, it is also reasonable to expect that medical term use could strengthen the positive relationship. This leads to our second research question: *how does medical term use moderate the relationship between call-to-action and audiences' engagement with health education videos?*

2. Theoretical foundation and hypothesis development

2.1. Related literature

We conducted a literature review on the impact of the call-to-action on social media engagement behavior and achieved two primary observations.

First, previous research on the call-to-action mainly focused on contexts where organizations utilize social media to drive customers' engagement, with limited attention given to understanding how this strategy works in expert-layperson communications (Moran et al., 2020; Oltra et al., 2022). The findings from research on customer-brand communication cannot be generalized to the context of health education because organizations are widely accepted as profit-oriented, while healthcare professionals are seen as welfare-oriented. The difference in perceptions may lead to different reactions to the call-to-action. For example, when audiences receive a call-to-action, they may suddenly realize that doctors' health education is not solely for welfare purposes but also a tactic to gain popularity on the platform, which may induce negative attitudes towards the doctor and their videos. Additionally, the complex nature of medical information may make it difficult for audiences to understand the communication content (Wernick et al., 2016), which is not observed in customer-brand communication. The difficulty in understanding may affect the impact of the call-to-action on engagement. Thus, it is imperative to explore how the call-to-action works in this context.

Second, although the call-to-action strategy has been extensively studied, there has been no research comparing the differences between visual and textual calls in driving social media engagement behavior. As mentioned earlier, existing studies mainly focused on textual call and examined how the framing of the call influences its effects (Jung et al., 2020; Oltra et al.,

[1] In this context, "verbal call-to-action" is not considered because this new form of call-to-action encompasses both verbal cues and visual cues, such as doctors' facial expression and gestures. Therefore, we refer to it as "visual call-to-action."

2022). For example, Jung et al. (2020) found that when firms frame call-to-action for referral programs, the framing emphasizing the benefits of senders' friends outperforms framings emphasizing senders' benefits or emphasizing both the sender and his/her friends' benefits. To the best of our knowledge, no study has explored how visual call affects social media engagement behavior. One possible reason for this gap in the literature is that the call-to-action research is dominated by customer-brand communication on social media, where organizations have limited opportunities to use visual call due to the way of information presentation. In contrast, health education videos are typically created and posted by doctors who appear in the video to discuss medical information, facilitating the utilization of visual call. In other words, this unique context allows us to investigate and compare textual and visual calls. To address these research gaps, we build a research model, illustrated in **Figure 1**.

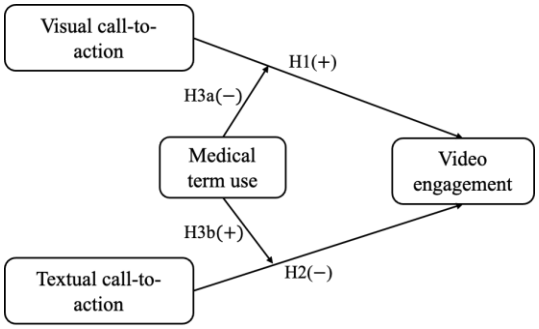

Figure 1. Research model.

2.2. Hypothesis development

We employed the media richness theory and the persuasion knowledge model to theorize the relationship between the call-to-action and video engagement. The media richness theory was initially developed to understand the selection of communication media in organizational settings to deal with task uncertainty and equivocality (Daft & Lengel, 1986; Kahai & Cooper, 2003). According to this theory, media vary in their abilities to emulate real-life communications, which is determined by four dimensions: multiple cues, feedback immediacy, personalization, and language variety (Daft & Lengel, 1986; Kahai & Cooper, 2003). Therefore, in the context of video-based health education, visual media is considered richer than textual media as it allows for the inclusion of multiple cues, such as facial expressions, tone of voice, and body gestures (Kahai & Cooper, 2003). McGrath and Hollingshead (1993) extended the media richness theory by proposing the task-media fit hypothesis, which states that the fit between task requirements in information processing and a medium's ability to convey information can enhance task performance (Wang & Doong, 2014).

The persuasion knowledge model, originating from the advertising literature, posits that individuals possess knowledge structures about persuasion that can be used to recognize and respond to persuasion attempts (Friestad & Wright, 1994). The persuasion knowledge encompasses individuals' beliefs about marketers' motivations, strategies, and skills to handle marketers' persuasion (Friestad & Wright, 1994). Marketers' actions can activate individuals' persuasion knowledge, leading to resistance against persuasion. However, the activation of persuasion knowledge requires cognitive resources, and thus individuals' cognitive resources determine the extent to which persuasion knowledge is activated when exposed to persuasion attempts (Jing Wen et al., 2020). In this context, the call-to-action strategy may have a negative impact as it has the potential to activate persuasion knowledge by inducing doubts among audiences regarding the intention of doctors delivering health education. However, the medium used to deliver the call can influence the activation of persuasion knowledge. We argue that visual media is better suited for the engagement request task, while the use of textual media may lead to task failures.

On one hand, visual media is well-suited for requesting engagement behavior for two reasons. Firstly, when making requests, visual media allows doctors to utilize multiple cues, which helps make the audience feel the actual presence of the doctor (Chidambaram & Jones, 1993) and fosters a socio-emotional environment (Kahai & Cooper, 2003). When individuals perceive a higher social presence from others, they are more likely to engage in prosocial behavior because social presence is positively associated with psychological closeness and perceived warmth (Kahai & Cooper, 2003). Secondly, sending engagement requests through visual media is less likely to activate audiences' persuasion knowledge that can have a negative impact on subsequent social media engagement behavior (Chen et al., 2020). According to the persuasion knowledge model, the activation of persuasion knowledge is influenced by individuals' cognitive resources (Jing Wen et al., 2020). Specifically, when individuals have limited cognitive resources to deal with persuasion attempts, their persuasion knowledge is less likely to be activated. Visual media, with its multiple cues, competes for audiences' cognitive resources, leaving limited cognitive resources available for the activation of persuasion knowledge. Previous research has

demonstrated that online influencers who utilize richer media to communicate with their audiences are perceived as more authentic and genuine, leading to an increased intention among audiences to follow their advice (Chidiac & Bowden, 2022). Combining the two underlying mechanisms, we propose that:

H1: The use of visual call-to-action has a positive effect on video engagement.

On the other hand, textual media is not suitable for the task of requesting audience engagement for two reasons. First, textual media lacks the ability to deliver multiple cues, resulting in a low social presence (Chidambaram & Jones, 1993). Previous research has shown that individuals are less likely to engage in prosocial behavior when they perceive lower social presence (Kahai & Cooper, 2003). Additionally, the call-to-action is typically displayed at the end of the video. Displaying textual call using terms, such as "giving likes and comments", may give the impression that the doctor is psychologically distant and has put less effort into requesting help from the audience. Second, using textual media to deliver the engagement request is more likely to activate audiences' persuasion knowledge. The limited contextual cues presented in textual call requires fewer cognitive resources. Consequently, the audience will have more cognitive resources available to process the engagement request, increasing the likelihood of activating persuasion knowledge (Jing Wen et al., 2020). Combining the two mechanisms, we propose the following hypothesis:

H2: The use of textual call-to-action has a negative effect on video engagement.

Although the use of medical terms is very likely to be considered a signal of competence (Moldovan, 2022; Xu, 2021), in the context of health education, doctors' competence can be signaled through multiple cues, diminishing the impact of terminology use on perceptions of competence. For instance, doctors' identities (e.g., being a chief doctor in a top-tier hospital) are usually displayed when they appear in the video. As a result, the relationship between medial term use and perceptions of doctors' competence may be unobservable in this context.

In contrast, medical term use can diminish the positive effect of visual call and amplify the negative effect of textual call through two mechanisms. First, a higher level of medical term use can make the content more difficult to understand, which may evoke negative emotions towards the health education video (Dang, 2020). Previous research found that individuals may suffer from information confusion when they find the information hard to comprehend (Walsh et al., 2007). When faced with information confusion, audiences may be hesitant to perform the engagement behavior that helps the popularity of the video, even if they are motivated by visual calls. Likewise, information confusion can reinforce negative attitudes towards the video content triggered by textual calls, resulting in reduced video engagement.

Second, a higher level of medical term use may give the impression that the doctor has put less cognitive effort into making the video comprehensible for the layperson. This perception could induce negative attitudes towards the doctor in health education videos. Research on patient-doctor communication has demonstrated that patient-friendly language is crucial for effective communication and building a positive patient-doctor relationship (Howe et al., 2019). Therefore, we propose the following hypotheses:

H3a: Medical term use weakens the positive relationship between visual call-to-action and video engagement.

H3b: Medical term use strengthens the negative relationship between textual call-to-action and video engagement.

3. Research methods

We chose Douyin (Tik Tok) as our research platform because it is the largest and most popular video-based social media platform in China (Yang et al., 2022). We collected health education videos in January 2022 by using the keyword searching function of the platform. About 200 keywords related to diabetes were used in the retrieval (see Liu et al., 2020). We selected diabetes as our research case due to its large patient population and extensive influence on society. After filtering out videos that were not created by doctors and did not contain doctors or educational content, we were left with 6,193 videos.

Our independent variables, visual call and textual call, were identified through manual coding. Specifically, we considered a video to have visual call if the doctor in the health education video called on the audience to engage with the video by stating phrases like "Please give me a like." The presence of textual call was determined by examining whether the video displayed texts that requested the audience to perform engagement behaviors. In this study, we did not differentiate between specific engagement behaviors (i.e., comment, like, and collection) because prior research has shown that requesting any type of engagement behavior has a near universal impact on all engagement behaviors (Moran et al., 2020). Research assistants conducted the content analysis, and one author randomly selected 120 videos to assess

the consistency. The calculated Cohen's kappa values for visual and textual calls were 0.80 and 0.84, respectively, indicating a substantial agreement in the coded results (Viera & Garrett, 2005).

The measurement of medical term use was based on the ratio of the number of unique medical terms to the total number of terms in a video. To perform the annotation task, we first converted the video content into text using speech-to-text software. Then, we adopted the text analysis software Linguistic Inquiry and Word Count (LIWC) to extract all the terms from these videos. By processing the transformed text files with LIWC, we obtained a list of terms that appeared in all the health education videos. Subsequently, one author manually coded each term as either medical or non-medical by referring to a Chinese Medical Subject Headings (MeSH) translation. MeSH is the most significant medical vocabulary thesaurus published by the National Library of Medicine, and it has been translated into many languages (Lu et al., 2008). Finally, we calculated the number of unique medical terms and the corresponding level of medical term use for each video based on the coded results. The dependent variable, video engagement, was measured using the number of times a video is liked, commented on, or collected by platform users (Shen & Pritchard, 2022). Collection is a typical social media behavior that allows users to add videos to their personalized video list for later access.

We controlled for video-level and doctor-level characteristics that may affect video engagement. We measured doctors' platform features following prior studies (Han et al., 2020; Li et al., 2021)-namely, the number of followers a doctor has as a measure of a doctor's popularity, the number of users the doctor is following as a measure of doctors' gregariousness. Besides, the percentage of male doctors in a hospital is typically more than that of female doctors, particularly for senior positions (Mast & Kadji, 2018). In addition, prior research showed individuals with a younger age are more competent in utilizing new technologies (Tams, 2022). Thus, we considered doctors' gender, age, and job title as potential factors influencing video engagement. Furthermore, the topic of health education videos may also affect video engagement. Therefore, we controlled for the effect of doctor fields, which largely determine video topics. Doctors' gender, job title, and field were identified through doctors' profiles on Tik Tok. As some doctors do not display their age on their profiles, we treated doctors' age as a category variable. Research assistants manually coded age as 0 if they perceived the doctor in videos to be less than 35 years old, and 1 otherwise. One author randomly selected 120 videos to assess the consistency, and the Cohen's kappa value was 0.88, indicating a high agreement (Viera & Garrett, 2005).

To account for the impact of video-level features, we considered videos' collection time, length, and emotions. Since our dataset is cross-sectional, videos posted earlier on the platform may have had more time to accumulate likes, comments, and collections. Therefore, we controlled for the number of days that each video was displayed on the platform to mitigate any potential bias. Video length was also taken into account as it can influence video engagement. Shorter videos may provide limited medical information, potentially reducing engagement. Emotions also play a crucial role in patient-doctor communication. Positive emotions can encourage and energize the audience towards disease management, while negative emotions make the audience feel alert, leading to increased attention to health education content (Li et al., 2021). Thus, we controlled for emotions in the video as well. The emotions present in each video were calculated using LWIC by inputting the text derived from the video content. The LWIC can help us assess the emotional valence of the healthcare information by computing the ratio of positive and negative emotion-related word counts, respectively. To remove hospital-level effects, we specifically selected videos posted by doctors from grade-A tertiary hospitals, the highest-level hospitals in China. On Tik Tok, doctors' profiles display information about the hospitals they work at, which is used for Google searches to determine the level of the hospital. This step filters out 75 (1.2%) videos created by doctors from non-grade A hospitals, reducing our sample size from 6,193 to 6,118.

4. Results

Out of the total 6,118 videos, 677 (11.1%) videos only utilize visual call, 898 (14.7%) videos only employ textual call, 353 (5.8%) videos incorporate both visual and textual call, and 4,190 (68.5%) videos includes neither visual nor textual calls. Among the 6,118 videos, 2,180 videos were posted by female doctors, and 3,938 video were posted by male doctors. 448 videos were created by doctors aged below 35 years old, and 5,670 videos were posted by doctors aged above 35 years old. 3,611 videos were posted by chief doctors, and 2,507 were posted by non-chief doctors. **Table 1** presents a summary of the average, maximum, minimum, and standard deviations of other variables in this study.

Table1. Summary statistics of variables.

Variables	Mean (S.D.)	Max	Min
MTU	.06 (.04)	1.65E6	0.00
Likes	6,676 (61,317)	88,000	0.00
Collections	683 (21,949)	2.15E7	0.00
Comments	271 (2,245)	9,436	0.00
Follower	5.51E5 (1.52E6)	900	10.00
Following	426 (902)	320	18.00
VL (seconds)	55 (45)	.30	.00
DP (days)	168 (66)	2.65E6	2.00
Positive emotions	3.14 (2.13)	1.65E6	0.00
Negative emotions	3.54 (2.47)	88,000	0.00

***Note**: MTU, medical term use. VL, video length. DP, days in platform. S.D. standard deviations.*

4.1. Impacts of call-to-action

We employed a negative binomial model to test our hypotheses for two reasons. First, the dependent variable, video engagement, was measured as count data. Thus, Poisson regression and negative binomial regression are appropriate methods (Khurana et al., 2019). Traditional OLS regression may yield biased results when the dependent variables are count data. Second, Poisson regression requires the mean-variance equality assumption, which is not satisfied by our count data. Prior to conducting the negative binomial regression analysis, we performed a VIF test to examine whether there was a multi-collinearity issue. The results revealed that the VIF values associated with each variable were lower than the conventionally recommended threshold (VIF<10), with the highest VIF value at 1.14. Thus, there is no multi-collinearity problem.

Table 2 shows our regression analysis results. We found that visual call had a significant and positive effect on liking (β=.315, p<.001) and collecting (β=.372, p<.001) behavior, but its impact on commenting behavior was not significant (β=-.021, p=.717). Interpreting the regression coefficient in terms of the incidence rate ratio, we found that, on average, by employing visual call, doctors could achieve 37% (=e^0.315-1) higher likes than not using visual call. Similarly, by adopting visual call, doctors could, on average, attain 45% (=e^0.372-1) higher collections than not using this strategy. Therefore, H1 is partially supported. In contrast, textual call significantly decreased the number of likes (β=-0.658, p<.001), collections (β=-0.817, p<.001), and comments (β=-0.423, p<.001) that videos received. Hence, H2 is supported.

Table2. Negative binomial regression results.

Variable	Model 1 Like	Model 2 Collect	Model 3 Comment
Visual call	**.315***** (.055)	**.372***** (.063)	-.021 (.059)
Textual call	**-.658***** (.050)	**-.817***** (.060)	**-.423***** (.052)
MTU	-4.699*** (.522)	-2.127*** (.626)	-7.812*** (.530)
Follower	8.4e-07*** (3.3e-08)	9.7e-07*** (3.9e-08)	6.2e-07*** (3.2e-08)
Following	-.000 (.000)	-.000 (.000)	-.000 (.000)
Gender	.450*** (.045)	.515*** (.054)	.506*** (.047)
Age	-.276*** (.079)	-.422*** (.097)	-.199* (.079)
Job title	-.048 (.044)	-.101 (.052)	-.059 (.045)
VL (seconds)	.002*** (.001)	.005*** (.001)	.001 (.001)
DP (days)	.000 (.000)	-.001*** (.000)	.000 (.000)
Positive emotions	-.004 (.009)	-.041*** (.011)	-.007 (.009)
Negative emotions	.050*** (.008)	.014 (.010)	.036*** (.008)
Field	Included	Included	Included
Number	6,118	6,118	6,118

***Note**: MTU, medical term use. VL, video length, DP, days in platform. Standard errors are in parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.*

We conducted additional analyses to examine the robustness of the positive effect of visual call and the negative effect of textual call. First, we estimated the Poisson model as a robustness check. The results showed that visual call had significant positive effects on the number of likes (β=.551, p<.001) and collections (β=2.087, p<.001), but the impact on the number of comments was insignificant (β=.076, p>.05). Additionally, textual call had significant negative effects on liking (β=-.836, p<.001), collecting (β=-2.130, p<.001), and commenting (β=-.545, p<.05) behavior. The results are consistent with our primary analysis.

Second, some videos employ both visual and textual calls, which may have an interaction effect on video engagement. For instance, the presence of textual call may reduce the positive impact of visual call on video engagement. To address this concern, we removed videos with both visual and textual calls from the dataset and then performed negative binomial regression and propensity score matching (PSM) as robustness checks (N=5,765). The negative binomial regression results showed that visual call significantly

increased the number of likes (β=.372, p<.001) and collections (β=.422, p<.001), but it had no effect on commenting behavior (β=-.032, p=.636). Moreover, textual call significantly decreased the number of likes (β=-.577, p<.001), collections (β=-.737, p<.001), and comments (β=-.363, p<.001). These findings align with our main analysis. In the PSM, we considered videos employing visual/textual call as the treatment group and those without visual/textual call as the control group. To account for the substantial variances in the number of likes, collections, and comments, a log transformation with the offset value of one was applied to the dependent variables (Fan et al., 2022). Doctor-level and video-level control variables were used for matching. The results revealed that visual call had a significant positive effect on liking (ATT=.105, t=2.00) and collecting (ATT=.055, t=2.06) behavior, while it had no impact on commenting behavior (ATT=-.043, t=.90). In contrast, textual call had negative effects on liking (ATT=-.166, t=-4.15), collecting (ATT=-.192, t=-4.65), and commenting (ATT=-.111, t=-2.85) behavior. The results of the PSM analysis are also in line with our main analysis.

4.2. Moderating effect of medical term use

In this section, we tested the moderating effect of medical term use on the relationship between visual/textual call and video engagement. The results are shown in **Table 3**. We found that medical term use significantly decreased the positive relationship between visual call and the number of likes (β=-4.269, p<.001) and collections (β=-8.365, p<.001), but medical term use had no significant moderating effect on the relationship between visual call and the number of comments (β=2.265, p>.05). The estimated coefficients suggests that if the value of medial term use rises by 0.01, the visual call strategy provides an additional 4.4% (=e^0.04269-1) decrease in the number of likes and an additional 8.7% (=e^0.04269-1) decrease in the number of collections. The value of medical term use in our dataset ranges from 0 to 0.30. The overall impact of visual call on the number of likes and collections can be negative when the medical term use is greater than 0.15 (0.616-4.269*0.15<0; 0.928-8.365*0.15<0). Thus, H3a is partially supported.

Likewise, the medical term use significantly increased the negative relationship between textual call and the number of likes (β=-3.548, p<.01) and collections (β=-6.473, p<.001), but medical term use had no significant moderating effect on the relationship between textual call and the number of comments (β=.914, p>.05). The coefficient estimated for the interaction terms showed that if the value of medical term use rises by 0.01, the textual call strategy provides an additional 3.6% (=e^0.03548-1) decrease in the number of likes and an additional 6.7% (=e^0.06473-1) decrease in the number of collections. Hence, our H3b is partially supported.

Table3. Moderating effects of medical term use.

Variable	Model 1 Like	Model 2 Collect	Model 3 Comment
Visual call	.616***	.928***	-.185
	(.105)	(.119)	(.115)
Textual call	-.441***	-.421***	-.479***
	(.090)	(.105)	(.092)
MTU	-3.253***	.897	-8.292***
	(.620)	(.770)	(.592)
Visual call*MTU	**-4.269*** **	**-8.365*** **	2.265
	(1.316)	(1.531)	(1.433)
Textual call*MTU	**-3.548** **	**-6.473*** **	.914
	(1.221)	(1.428)	(1.284)
Follower	8.4e-07***	9.9e-07***	6.2e-07***
	(3.3e-08)	(3.9e-08)	(3.2e-08)
Following	2.78e-06	-1.37e-05	-3.88e-05
	(2.03e-05)	(2.32e-5)	(2.11e-05)
Gender	.460***	.535***	.504***
	(.045)	(.054)	(.047)
Age	-.252***	-.359***	-.206**
	(.079)	(.096)	(.079)
Job title	-.038	-.090	-.065
	(.044)	(.052)	(.045)
VL (seconds)	.002***	.005***	.001
	(.001)	(.001)	(.001)
DP (days)	-.000	-.002***	.000
	(.000)	(.000)	(.000)
Positive emotions	-.002	-.034**	-.008
	(.009)	(.011)	(.009)
Negative emotions	.050***	.015	.035***
	(.008)	(.010)	(.008)
Field	Included	Included	Included
Number	6,118	6,118	6,118

***Note**: MTU, medical term use. VL, video length. DP, days in platform. Standard errors are in parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.*

Finally, to examine the robustness of the interaction effect of medical term use and visual/textual call on video engagement, we conducted two additional analyses. First, in our main analysis, medical term use was measured by dividing the number of unique medical terms by the total number of terms in a video. Although prior research used unique medical terms to measure medical information (Liu et al., 2020), repeated incomprehensible terms may also create difficulties in content understanding. To alleviate this concern, we conducted a robustness check in which we measured medical term use by dividing the number of all medical terms by the total number of terms. The results of negative binomial

regression revealed that medical term use significantly attenuated the positive relationship between visual call and the number of likes (β=-3.275, p=.002) and collections (β=-4.674, p<.001). However, the interaction term of medical term use and visual call had no significant effect on the number of comments (β=.855, p=.445). Similarly, medical term use significantly strengthened the negative relationship between textual call and the number of likes (β=-1.717, p=.039) and collections (β=-3.372, p<.001). Still, it had no effect on the relationship between textual call and the number of comments (β=.210, p=.806). These results align with our main analysis.

Second, we divided videos into two groups based on the level of medical term use. Videos with a medical term use higher/lower than the median were assigned to the high/low group. Analysis of variance was used to examine the interaction effect. The results showed that the interaction effect of medical term use and visual call on the number of likes (F=9.754, p<.01) and collections (F=18.325, p<.001) was significant, while the interaction effect on commenting behavior (F=1.651, p>.05) was insignificant. Likewise, medical term use and textual call had a significant interaction effect on liking (F=5.001, p<.05) and collecting (F=5.426, p<.05) behavior, but the interaction effect on commenting behavior was insignificant (F=3.510, p>.05). The results are also consistent with our main analysis.

5. Discussion

5.1. Discussion of results

In this study, we investigated how visual call and textual call affect video engagement and how medical term use moderates the relationship between visual/textual call and video engagement. We conducted an analysis of archival data and examined three engagement metrics, the number of likes, collections, and comments. Our results showed that visual call significantly increases the number of likes and collections, while textual call significantly decreases the number of likes, collections, and comments. However, we found that visual call has no effect on commenting behavior, which is inconsistent with our hypotheses. One possible explanation is that performing the commenting behavior requires audiences to type a response and thus demands more time and effort (Moran et al., 2020). In contrast, to perform the liking and collecting behavior only requires one simple click (Moran et al., 2020). Therefore, commenting behavior may be more difficult to activate than liking and collecting, leading to the insignificant relationship between visual call and the number of comments.

Furthermore, our study demonstrated that medical term use decreased the positive relationship between visual call and the number of likes and collections and increased the negative relationship between textual call and the number of likes and collections. However, medical term use has no significant moderating effect on the relationship between visual/textual call and the number of comments. This could be because commenting requires more time and effort, making it less susceptible to contextual factors than liking and collecting behavior (Moran et al., 2020). For example, when an audience is prepared to write a comment for a health education video, the medical term use could not make the audience abandon the commenting behavior since time and effort are already spent on framing sentences.

5.2. Theoretical contributions

This study has two primary contributions. First, we contribute to the field of technology-mediated patient education by demonstrating the positive impact of visual call and the negative impact of textual call on engagement with health education videos, deepening our understanding of the interaction process between doctors and video audiences. Prior research primarily focused on patient-doctor interactions within online-health communities (Fan et al., 2022), neglecting the emerging form of interaction where doctors deliver health education to the general public via video-based social media with an intention to increase online popularity and offline appointment numbers. Our study offers a foundation to motivate future research into this new mode of interaction. In addition, while the call-to-action strategy is widely recognized as an effective practice and has already been employed by doctors (Ju et al., 2023), limited research attention has been given to understanding how this strategy works in doctor-layperson communications. We found that visual call and textual call play opposite roles in driving health education video engagement, and their effects are contingent on the medical term use of doctors. This study paves the way for future research to investigate other contextual factors that may affect the impact of the call-to-action strategy.

Second, we contribute to the literature on call-to-action by not only testing its effects in a communication context between experts and non-experts, but also by demonstrating the importance of the communication modality used to deliver the call. Prior research on the call-to-action strategy primarily focused on textual call-to-action strategy, and no study has explored visual call (Moran et al., 2020). The

context of health education via video-based social media allows us to examine the effect of visual call and compare the differences between the two types of call-to-action. This study found that visual call increases video engagement, while textual call decreases video engagement. This finding can motivate future research to investigate the potential distinct effects of the call-to-action strategy caused by different communication modalities.

5.3. Practical implications

This study has practical implications for doctors and video-based social media platform managers. First, our results suggest that using visual call significantly increases engagement with health education videos, while textual call has a negative effect on video engagement. Therefore, doctors who want to obtain more video engagement should use visual call and avoid textual call. The platform can use our findings to develop video-making guidelines for doctors. For example, the platform can recommend that popular health education videos should include visual call when doctors are creating them.

Second, the use of medical terms can reduce the positive effect of visual call. Therefore, doctors should use layperson-friendly language to deliver health education content. Platform managers can also use our findings to assist in providing health education videos. For example, the platform can employ strategies to help doctors assess whether the health education videos doctors want to post are too complicated to understand for the layperson and provide feedback on the content modification to reduce the number of terminologies. In addition, if the platform detects that health education videos are using too many medical terms, it can adjust its recommendation strategies.

5.4. Limitation and future research

This study has several unaddressed limitations. First, the use of archival analysis does not allow us to confirm the causal relationship between call-to-action and video engagement. However, we controlled for variables at the video, doctor, and hospital levels that could affect video engagement and conducted additional analyses to ensure the robustness of our findings. Future studies can conduct experiments to isolate the causal impact of the call-to-action on video engagement by directly manipulating the presence of call-to-action.

Second, this study did not consider the effect of doctors' appearances on video engagement, although we controlled the impacts of doctors' gender and age. It is reasonable to expect that audiences are more likely to engage with videos in which doctors are perceived as beautiful or handsome. Future studies could examine the impact of doctors' appearances.

Third, our findings showed that visual call do not affect commenting behavior, and the use of medical terms does not moderate the relationship between the call-to-action and commenting behavior. One possible reason for this is that commenting behavior requires more time and effort compared to liking and collecting behavior. However, this study did not formally test this underlying mechanism. Thus, a future direction is to examine the impact of the cost of engagement behavior on the relationship between call-to-action and engagement behavior.

6. Acknowledgements

The work described in the paper was partially supported by grants from the Strategic Research Grants of the City University of Hong Kong (Grant Nos. 7005595 and 9680052) and the National Natural Science Foundation of China (Project No. 72271210).

7. References

Chen, T. Y., Yeh, T. L., & Chang, C. I. (2020). How different advertising formats and calls to action on videos affect advertising recognition and consequent behaviors. Service Industries Journal, 40(5), 358–379.

Chidambaram, L., & Jones, B. (1993). Impact of communication medium and computer support on group perceptions and performance: A comparison of face-to-face and dispersed meetings. MIS Quarterly, 17(4), 465–491.

Chidiac, D., & Bowden, J. (2022). When media matters: the role of media richness and naturalness on purchase intentions within influencer marketing. Journal of Strategic Marketing, 1-21.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. Management Science, 32(5), 554–571.

Dahm, M. R. (2012). Tales of time, terms, and patient information-seeking behavior-An exploratory qualitative study. Health Communication, 27(7), 682–689.

Dang, V. T. (2020). Information confusion and intention to stop using social networking site: a moderated mediation study of psychological distress and perceived novelty. Information Technology and People, 33(5), 1427–1452.

Fan, W., Zhou, Q., Qiu, L., & Kumar, S. (2022). Should doctors open online consultation services? An empirical investigation of their impact on offline appointments. Information Systems Research. Article in advance.

Friestad, M., & Wright, P. (1994). The persuasion knowledge model: how people cope with persuasion attempts. Journal of Consumer Research, 21(1), 1–31.

Han, Y., Lappas, T., & Sabnis, G. (2020). The importance of interactions between content characteristics and creator characteristics for studying virality in social media. Information Systems Research, 31(2), 576–588.

Howe, L. C., Leibowitz, K. A., & Crum, A. J. (2019). When your doctor "gets it" and "gets you": the critical role of competence and warmth in the patient–provider interaction. Frontiers in Psychiatry, 10, 475.

McGrath, J.E., & Hollingshead, A.B, (1993). Putting the" group" back in group support systems: Some theoretical issues about dynamic processes in groups with technological enhancements. Macmillan Publishing Company.

Jing Wen, T., Kim, E., Wu, L., & Dodoo, N. A. (2020). Activating persuasion knowledge in native advertising: the influence of cognitive load and disclosure language. International Journal of Advertising, 39(1), 74–93.

Ju, R., Jia, M., & Cheng, J. (2023). Promoting mental health on social media: A content analysis of organizational Tweets. Health Communication, 38(8), 1540–1549.

Jung, J., Bapna, R., Golden, J. M., & Sun, T. (2020). Words matter! Toward a prosocial call-to-action for online referral: Evidence from two field experiments. Information Systems Research, 31(1), 16–36.

Kahai, S. S., & Cooper, R. B. (2003). Exploring the core concepts of media richness theory: The impact of cue multiplicity and feedback immediacy on decision quality. Journal of Management Information Systems, 20(1), 263–299.

Khurana, S., Qiu, L., & Kumar, S. (2019). When a doctor knows, it shows: An empirical analysis of doctors' responses in a Q&A forum of an online healthcare portal. Information Systems Research, 30(3), 872–891.

Li, L., Tian, J., Zhang, Q. (2021). Influence of content and creator characteristics on sharing disaster-related information on social media. Information & Management, 58(5), 103489.

Li, Y., & Xie, Y. (2020). Is a picture worth a thousand words? An empirical study of image content and social media engagement. Journal of Marketing Research, 57(1), 1–19.

Liu, X., Zhang, B., Susarla, A., & Padman, R. (2020). Go to YouTube and call me in the morning: Use of social media for chronic conditions. MIS Quarterly, 44(1), 257–283.

Lu, W., & Wu, H. (2019). How online reviews and services affect physician outpatient visits: content analysis of evidence from two online health care communities. JMIR Medical Informatics, 7(4), e16185.

Lu, W.-H., Lin, R. S., Chan, Y.-C., & Chen, K.-H. (2008). Using Web resources to construct multilingual medical thesaurus for cross-language medical information retrieval. Decision Support Systems, 45(3), 585–595.

Mast, M. S., & Kadji, K. K. (2018). How female and male physicians' communication is perceived differently. Patient education and counseling, 101(9), 1697-1701.

Moldovan, A. (2022). Technical Language as Evidence of Expertise. Languages, 7(1), 41.

Moran, G., Muzellec, L., & Johnson, D. (2020). Message content features and social media engagement: evidence from the media industry. Journal of Product and Brand Management, 29(5), 533–545.

Oltra, I., Camarero, C., & San José Cabezudo, R. (2022). Inspire me, please! The effect of calls to action and visual executions on customer inspiration in Instagram communications. International Journal of Advertising, 41(7), 1209–1234.

Omar, B., Kim, D., Montag, C., Yang, H., & Elhai, J. D. (2021). On the psychology of TikTok use: A first glimpse from empirical findings. Frontiers in Public Health, 9, 641673.

Shen, Z., & Pritchard, M. J. (2022). Cognitive engagement on social media: A study of the effects of visual cueing in educational videos. Journal of the Association for Information Science and Technology, 73(9), 1253-1267.

Tams, S. (2022). Helping older workers realize their full organizational potential: A moderated mediation model of age and IT-enabled task performance. MIS Quarterly, 46(1), 1–34.

Viera, A. J., & Garrett, J. M. (2005). Understanding interobserver agreement: the kappa statistic. Family Medicine, 37(5), 360–363.

Walsh, G., Hennig-Thurau, T., & Mitchell, V.-W. (2007). Consumer confusion proneness: scale development, validation, and application. Journal of Marketing Management, 23(7–8), 697–721.

Wang, H.-C., & Doong, H.-S. (2014). Revisiting the task-media fit circumflex: A further examination of negotiation tasks. Information & Management, 51(6), 738–746.

Wernick, M., Hale, P., Anticich, N., Busch, S., Merriman, L., King, B., & Pegg, T. (2016). A randomized crossover trial of minimizing medical terminology in secondary care correspondence in patients with chronic health conditions: impact on understanding and patient reported outcomes. Internal Medicine Journal, 46(5), 596–601.

Xie, X. Z., Tsai, N. C., Xu, S. Q., & Zhang, B. Y. (2019). Does customer co-creation value lead to electronic word-of-mouth? An empirical study on the short-video platform industry. The Social Science Journal, 56(3), 401–416.

Xu, Z. (2021). I don't understand you, but I trust you: using computer-aided text analysis to examine medical terminology use and engagement of vaccine online articles. Journal of Communication in Healthcare, 14(1), 61–67.

Yang, Z., Luo, X., Jia, H., Xie, Y., & Zhang, R. (2022). Personal narrative under nationalism: Chinese COVID-19 vaccination expressions on douyin. International Journal of Environmental Research and Public Health, 19(19), 12553.

Zhu, C., Xu, X., Zhang, W., Chen, J., & Evans, R. (2020). How health communication via Tik Tok makes a difference: A content analysis of Tik Tok accounts run by Chinese provincial health committees. International Journal of Environmental Research and Public Health, 17(1), 192–205.