Manuscript: egusphere-2022-494
1st Author: Emily E. Zawacki (eezawack@asu.edu)
Title: "Exploring TikTok as a promising platform for geoscience communication"

Dear Dr. Mohadjer,

We would like to thank you for your management of our manuscript, and we thank the two anonymous reviewers for their constructive and helpful comments. Based on the reviewer comments, we have worked to resolve the following key issues:

1. It is unclear what research gap the paper is addressing (including literature review).
2. It is unclear what the plotted relationship are looking at and why they are analyzed.
3. The study lacks statistical analyses and rigor regarding data interpretation.
4. Components of the manuscript are too US-centric.

We have addressed these issues by updating the manuscript in these ways:

1. We have better clarified that this is the very first study to examine geoscience communication on TikTok and that this work provides a foundation for future studies.
2. We have more thoroughly explained all the parameters and relationships that we evaluate, both in the text and in figure captions.
3. We performed correlation analyses to determine relationships and the statistical significance of the data we evaluate.
4. We have included additional context and summary so that the manuscript is better applicable to a broader, worldwide audience.

Line-by-line detailed responses to reviewer comments are included in this document. Line numbers refer to the original submitted version of the manuscript. A complete documentation of all changes made is shown in the file "TikTokManuscript_GC_Revised_TrackChanges.pdf."

The manuscript is now ~8,100 words with 9 figures and 4 tables, in addition to Supplementary Materials ("TikTokSupplementaryMaterials.pdf").

Sincerely,

Emily E. Zawacki

Emily E. Zawacki, corresponding author (eezawack@asu.edu)

Arizona State University, School of Earth and Space Exploration
PO Box 876004, Tempe, AZ 85287-6004

Manuscript: egusphere-2022-494
1st Author: Emily E. Zawacki (eezawack@asu.edu)
Title: Exploring TikTok as a promising platform for geoscience communication

REVIEWER COMMENTS

Anonymous Reviewer #1:

Old Text:	"Exploring TikTok as an effective platform for geoscience communication"
Reviewer Comment:	The title is somewhat misleading. The study does not test the effectiveness of TikTok as a platform for geoscience communication. An important part of effective communication is what happens on the receiving side (e.g., information retention). This study looks solely at reach and basic form of engagement.
Author Reply:	We have adjusted our title to "Exploring TikTok as a promising platform for geoscience communication." This change still largely conveys our intent (we demonstrate the exceptional reach and low-barrier to entry that TikTok provides, which is extremely promising for science communication efforts), and future studies can build on to more fully evaluating qualitative effectiveness on the side of the viewer. [Alternatively, the title could be simplified to "Exploring TikTok as a platform for science communication"]
New Text:	"Exploring TikTok as a promising platform for geoscience communication"
Line:	16
Old Text:	"We found that short videos (< 20 s) had a high viewer retention 20 rate, but…"
Reviewer Comment:	The term "retention" is a tad misleading. Until it was explained, I took it as a measure for how much of the video was remembered.
Author Reply:	'Viewer retention' is a commonly used term for videos on social media, however, we have changed this term to "average video view duration" throughout for clarity.
New Text:	"We found that short videos (< 30 s) had a high average view duration, …"
Line:	20
Reviewer Comment:	What is the "For You" page? If it's a page with personal video suggestions, maybe use a more general description of it, such as "suggested viewing page ("For You" page on TikTok)" for those unfamiliar with TikTok. It's explained in section 2, but I'd use a more explanatory and general term for it in the abstract so it doesn't require knowledge of TikTok.

Author Reply:	The "For You" page is the Proper Noun name for this feed on TikTok. In the abstract, we already define it as "TikTok's algorithmic recommendation feed."
Line:	22
Old Text:	"Our videos that went the most viral featured content that was related to…"
Reviewer Comment:	"Went the most viral" – I'm aware that such terminology has established itself fairly well by now (among specific demographics at least), but I nevertheless recommend rephrasing that for an audience that is unfamiliar with it yet have an interest in exploring such media platforms as a communication tool.
Author Reply:	We have adjusted and simplified this sentence.
New Text:	"Our videos that received the highest number of views featured content that was…"
Line:	53–55
Old Text:	"We analyzed the reach (how many individuals saw the video), the engagement (the number of interactions with the content, e.g., "likes" or "comments"), and the viewer retention rate (how much of the video is watched on average) of each video we posted to determine the qualities of a successful educational geoscience TikTok."
Reviewer Comment:	It's unclear what the purpose and nature of the analysis is. If it is to determine factors of a video that help maximise reach (line 55), why were engagement and view duration analysed at all? Usually, an introduction outlines the broad questions, goals and hypotheses tackles by a study. This helps guide the reader through the manuscript. At this point, I still don't know what these are or what analyses were conducted on the 3 metrics and for what specific purpose. (Unfortunately, I also miss clarification of this down the line in the following sections).
Author Reply:	We have added on an additional statement to help clarify why we analyze and care about these metrics.
New Text:	"We analyzed the reach (how many individuals saw the video), the engagement (the number of interactions with the content, e.g., "likes" or "comments"), and the average view duration of each video we posted to determine the qualities of a successful educational geoscience TikTok video. The longer that someone watches a video and the more engagement the video receives demonstrates interest in the video and

	the impact of the science communication, which will likely be rewarded within the algorithm with expanded video reach.."
Line:	58–77
Reviewer Comment:	I appreciate the introduction to TikTok (section 2) and think it's important. I would add some comments on TikTok's reach. Who are the viewers, how are they distributed globally? This may be very important information for communicators who want to target specific communities.
Author Reply:	We already include TikTok's reach in the Introduction (1 billion monthly active users), and the number of app downloads we list are global numbers. We have added another sentence to this regard. TikTok has a fully global, large audience (and it would not be relevant here to specifically list the number of TikTok users by each country).
Added text:	"During the first quarter of 2022, TikTok was the most downloaded app worldwide (Sensor Tower, 2022)."
Line:	67–68
Existing Text:	"TikTok uses AI algorithmic recommendations to determine what videos are shown on a user's For You page, which are often based on the user's profile settings, their location, and their activity on the app (Smith, 2021)."
Review Comment:	This is a tad vague. What are these recommendation algorithms based on exactly? The authors list "user's profile settings", "location" and "activity". Can users choose topics of interests in the profile settings? Since the study's results are, to a degree, sensitive to this, I think a closer look at it and a more elaborate explanation is warranted. This will also help communicators plan their activities on TikTok better.
Author Reply:	TikTok provides no information regarding their AI algorithm or how it works, hence why studies like this are important. The information we provide in this section is essentially the extent of all information TikTok publicly makes available. Users can indicate they are 'not interested' in a video, but there is no way to choose topics of interest other than watching videos on the app (or skipping them) and liking videos that they enjoy. There is no further information we can provide regarding TikTok's algorithm.
Line:	77
Existing Text:	"Videos additionally can be viewed on a creator's profile page or by searching video hashtags or sounds."

Reviewer Comment:	It's unclear what is meant by searching sounds. It reads as though sounds are used as hashtags are. Is that accurate?
Author Reply:	We recommend downloading TikTok for a brief user experience. Sounds are background audio clips that can be re-used by multiple users and are separate from hashtags. If you click on a 'sound' (audio clip/song), you can see all the videos that were made using that sound.
Line:	85–93
Reviewer Comment:	What is the longevity of such hashtags on TikTok? On twitter, for example, they may rapidly change and quickly lose popularity and reach. Is #LearnOnTikTok ongoing and intended as a "stable" hashtag? Was the associated campaign was a one time event or something ongoing and sustainable? If it's ongoing and the hashtag serves almost like an education "channel" on TikTok now, this should be communicated. As a communicator, this would be very useful for me then. If it's not expected to survive long as a hashtag, on the other hand, it doesn't warrant as much explanation.
Author Reply:	The lifetime of a TikTok hashtag is essentially infinite. You can click on a hashtag and see all videos that have ever been created that use that hashtag. Any person can use any hashtag at any time. As we discuss later in the paper, hashtags on TikTok have a fundamentally different 'function' than they do on platforms like Twitter. People generally do not 'search' for hashtags on TikTok like they do on Twitter. We hypothesize that hashtags are primarily used in TikTok's AI to provide additional information and context for the video when deciding who to show it to with the algorithm.
Line:	104–105
Existing Text:	"Unlike chemistry and physics, the geosciences do not as easily lend themselves to experiment-based content, but do allow for unique, hands-on demonstrations as well as the potential for "field trip" style presentations."
Reviewer Comment:	I disagree with the general statement that geoscience does not lend itself to experiment-based content as much. On the contrary, since geoscience is effectively physics and chemistry applied to the Earth, many classic physics and chemistry experiments can be used to demonstrate how something works on Earth. In atmospheric sciences alone, there are plenty of pressure and moisture experiments that can be conducted (at home) to demonstrate important concepts. On the

	other hand, for Earthquake science, experiments involving concepts of friction and liquefaction come to mind.
Author Reply:	We do not say that it is impossible to have experiment videos related to geosciences, merely that they are generally less common and/or easy to produce compared to videos of chemistry and physics experiments. A survey of educational science content on TikTok demonstrates an abundance of videos of chemistry and physics experiments and near total absence of geology-related experiments (which we find is more so related to the ease with which at home or classroom chemistry/physics experiments can be filmed and shared on the app).
Line:	123–124
Existing Test:	"Our videos featured three rotating front-facing hosts and were typically done in a "lecture-style" format, where the host explains a geoscience topic while using engaging visuals. Two of the hosts are female and presented the majority of the videos, and one is male."
Reviewer Comment:	Was this by design and are there studies of how the metrics used here (reach, engagement and viewing time) vary depending on the gender of the presenter? Given prevailing differences in perception of scientists of different genders, this may be important information for geoscientists who consider TikTok as a communication platform.
Author Reply:	The video presenters are the existing communications staff at our respective organizations. The male presenter was only able to create two videos during this duration, hence we are unable to analyze video metrics based on the gender of presenter, although this is work we would like to do in the future.
New Text:	
Line:	Section 4 (general)
Reviewer Comment:	It's unclear what determined the choices in video characteristics and how these choices were made with a specific experiment/goal in mind. This is essentially the "experiment setup" section, but I cannot see how the choices made here serve the testing of a broader hypothesis or help answer the study's overarching question(s). For example, if one were to determine the optimal video length to maximise the engagement rate, one would vary the factor of interest and reduce the number of other free parameters (such as upload time). Reversely, If I wanted to determine the best time to upload a video, I would vary this parameter and keep others constant. I don't see this type of experiment-based and hypothesis-guided thinking in this section.

Author Reply:	As it is very difficult to vary just a singular variable when creating TikTok videos, we use the GMV (ground motion visualization) videos as a way to most clearly isolate variable trends with posting. All the Terra Explore videos were created during a pilot project, in which the goal was to test and produce different types of geoscience content on TikTok (different topics, durations, hashtags used, time and day posted, etc.). Although there was not a systematic design behind every single video posted, that does not make the analyses of these videos any less meaningful.
Line:	175
Old Text:	"Videos were uploaded at various times throughout the day, and the videos that received the most views were uploaded either in the morning or early afternoon Mountain Standard Time (MST)."
Reviewer Comment:	Please express time using the primary standard UTC (or list in in brackets). As someone who's not from North America, I was completely unaware of MST.
Author Reply:	We have included the corresponding UTC time in all of our discussions of time zone, in addition to MST. (The figures are kept on an MST axis as we want to express morning vs evening upload for our primarily US-based audience).
New Text:	"Videos were uploaded at various times throughout the day, and the videos that received the most views (>90,000) were uploaded either in the morning or early afternoon Mountain Standard Time (MST) (UTC-7)."
Line:	Figure 2
Reviewer Comment:	I think the text-based summary of views is sufficient. The figure does not show any pattern or provides any valuable additional information and thus seems a tad redundant.
Author Reply:	Recommended change made, figure deleted.
Line:	Figure 3
Reviewer Comment:	Unless most viewers are known to be from North America and ~ the MST time zone, I would change this to UTC. In fact, if the viewer location does not coincide with the MST zone, discussing times of the day may be misleading. Furthermore, I question the merit of presenting this data at all, since (a) there is no clear pattern except maybe an increase in views towards the evening, and (b) global viewers will likely be spread across different time zones.

Author Reply:	~85% of viewers of Terra Explore are within the United States, hence the majority of views come from 'local' time zones. All our most highly viewed videos (>90,000 views) were posted between 7 am - 2 pm MST (UTC-7). The theory of posting early in the day local to your audience is that it allows more time for the content to be live and viewed before the number of active users drops off overnight. However, we found no statistically significant relationship between the number of views and the time a video was published.
Line:	Figure 6
Reviewer Comment:	The ratio remains ~constant over time, making the figure rather redundant. It's sufficient to simply state the ratio, as already done in the text.
Author Reply:	Recommended change made, figure deleted.
Line:	215
Old Text:	6.2 Viewer retention and engagement rate
Reviewer Comment:	Please change the term "retention" to something more intuitive like "view duration". You use the latter term already. I recommend simply sticking to that, also to avoid confusion with viewer information retention (memory).
Author Reply:	Recommended change made, referred to it as the 'average video view duration (%)' throughout.
New Text:	6.2 Video view duration and engagement rate
Line:	Figure 7
Reviewer Comment:	It is unclear why video views is plotted against video "view duration". Is a relationship expected? Why? What would be the meaning of it?
Author Reply:	Traditionally, yes, you would expect to see a relationship between video views and average view duration, as videos with a high avg. view duration are hypothesized to be promoted within an algorithm and shown to more viewers, thus yielding higher views (particularly on platforms like YouTube). A high average view duration also indicates that the audience is more interested in the video content and that it holds their attention rather than scrolling to the next video, thus providing a useful gauge in how effective the science communication video is. However, because TikTok can support such short videos (~15 s), those videos can very easily yield a high avg. view duration, but that doesn't necessarily mean they will also have a high number of views. We see videos with the highest number of views (>90,000) have avg.

	view durations of >40%, indicating that this is the value creators should aim for. Overall, we do observe a weak positive statistically significant correlation between the number of video views and the average video view duration.
New Text:	See discussion and caption for new Figure 5.
Line:	Figure 8
Reviewer Comment:	See comment on Fig. 7. It's unclear why views are plotted against engagement rate. What would the discovery of a relationship between these metrics help with? What hypothesis or idea would that address?
Author Reply:	The higher the engagement rate, the more people are interacting with a piece of content, thus showing higher interest on the side of the viewer (see Habibi and Salim, 2021, another study that measures user engagement of educational science content on TikTok). As is hypothesized with how algorithms work, content that is interacted with more will be shown to more people, thus yielding higher views. What is interesting is that we observe no correlation between the number of video views and the engagement rate—even videos with lower views received relatively high engagement rates. What we suggest is that a video needs both a high engagement rate and a high avg. view duration to be shown more in the algorithm and receive more views.
New Text:	See discussion and caption for new Figure 6.
Line:	Figure 9
Reviewer Comment:	See comment on previous 2 figures. What's would be the significance of the x-y relationship? Furthermore, the caption does not guide the reader through the figure in any way. It does not explain anything and I am left wondering why it is shown and what the authors would like me to see in it.
Author Reply:	We find that there is no clear relationship between the average view duration and engagement (two factors that are likely to increase video reach in the algorithm). Thus, just because a person views more of a video does not necessarily mean they are more likely to interact with the content. However, the average view duration here is also linked to the length of the video, as ~15 s videos can much more easily yield a high avg. view duration than can a ~150 s video (Fig. 10–the longest videos actually have consistently high engagement rates, especially for shares).
New Text:	See discussion and caption for new Figure 7.

Line:	Section 6.4
Reviewer Comment:	I understand why duration and time posted may be valuable information. These may be useful for communicators in their planning to maximise views. However, I do not see the significance of a relationship between views and average % watched.
Author Reply:	Explained in reply to above comment on Fig. 7.
Line:	Section 7
Reviewer Comment:	Any thorough, literature-informed discussion of the study's limitations (and the limits to recommendations made) is missing. For example, I think it would be important to: - Comment on importance of different metrics for assessing communication effectiveness. - Comment on how sensitive the study's results are to the specific algorithms used by TikTok, how transparent they are and how (often) these may change. How does this affect the take-away messages of this study? - Comment on technology exclusivity and how it limits reach. - Comment on the effect of video content on audience demographics/nationality/location. This may be important in context of the interpretation of the viewing times also.
Author Reply:	There is essentially zero transparency to TikTok's algorithm, hence the importance of this study and others like it. The existing literature related to analyzing content on TikTok and the algorithm is very limited (see Klug et al., 2021, citation added in paper). Our study is thus important because it provides additional data and insight for the very little that has been analyzed of science communication on TikTok. We make note in the text that algorithms can change over time, and thus it will be useful to conduct future surveys over time. However, algorithmic changes that fundamentally impact how the app functions are unlikely to occur over short periods of time. 95% of teens in the United States say that they have access to a smartphone (Pew Research Center). Thus there is almost no exclusion to the technology, demonstrating that TikTok has nearly unlimited reach potential. We have included additional discussion of audience demographics/location in the text.
Added Text:	(Section 6.1) "85% of the account's followers are located within the United States, followed by 3% from Canada. 84% of the total number of video views came from users located within the United States, with each individual video averaging 78% of views from users in the United

States. Canada was the second highest video viewer region on 22 videos, Australia on 17 videos, the United Kingdom on five, and the Philippines on four, with the secondary country yielding on average 5% of a video's views."

(Section 6.4) "Regardless of the global location where the earthquake occurred, 84% of each videos' views on average came from users located within the United States. Therefore, these GMV videos were largely not being shown to users in countries where the earthquake occurred, but rather shown to users in North America (where the visualization is focused)."

(Section 7.2) "The Hunga Tonga video was also shown to a wider international audience, with only 66% of views coming from users located in the United States. Thus, if the focus of the video is on a specific international location/event, the video will likely be served to a broader audience."

(Section 7.5) "As well, we acknowledge that TikTok's algorithm may change over time, although we don't anticipate such significant changes would occur in a short amount of time that would invalidate our results."

Line:	405–412
Reviewer Comment:	I question several of these recommendations (and related conclusions) based on some of the study's flaws: (a) The recommendation to post earlier in the day seems based on the discussed relationship between time posted and video views. More precisely, it seems based on the observation that videos that received the most views were uploaded in the morning or early afternoon. By simply looking at the data plot, I do not make the same observation. In fact, it is difficult to see any significant pattern at all except maybe a general increase in video views later in the (MST) day. This contradicts the authors' findings and highlights the need for rigorous statistics to make sure a signal/observation is real. Furthermore, this recommendation is very US centric and/or vague. To elaborate: The authors use a US time zone to define the time of the day. Does the recommendation pertain to MST morning or morning in whatever time zone the communicator lives? To answer this, the authors need to

establish where the views are from. If they could establish (1) that their viewers are US (or MST) based, and (2) that there is an actual relationship between views and time of day, a recommendation for upload time could be merited (albeit restricted to the US). As it is, however, no such conclusions can be drawn from the TikTok analytics data.

(b) As a communicator, I want to communicate a specific set of problems and topics. Recommendations to choose a topic based on what is newsworthy (or tied to a specific location) is therefore questionable. This recommendation seems focused on increasing views (regardless of intended subject of communication) rather than on effective communication of a specific topic. The recommendation to include gender-related topics is similarly problematic. I would like to know "What do I do to increase the number of female viewers for videos about this specific topic?" and not "What topic should I choose to increase the number of female viewers." The recommendation would be merited if the strategy is to include more gender-related topics to gain more female followers, who then view videos of other topics. However, this strategy is also questionable given that relatively few views are gathered from "following" (Fig. 5).

(c) Given how few views the videos got from hashtags (I cannot even see the percentage of it for most of Fig. 5) I do not see how the authors can list this as a recommendation here.

| Author Reply: | (a) We have removed the recommendation related to post-time from this list, and instead provide new text below in Section 7.1.

(b) We do not intend to suggest that every topic be related to a newsworthy event or place-based geology, rather that these are ways in which to maximize the algorithmic impact of the video content, thus reaching the highest amount of viewers and sharing the science with a broad audience. These are also tools to make the topic communicated relevant and relatable to the audience. As we demonstrated, even videos that had lower views yielded high engagement rates, which indicates that the content was still impactful, even if it was received by a smaller audience. We find that demonstration videos < 30 s are an impactful way to communicate topics, as are lecture-style videos ~40 - 120 s in duration. We include this discussion of gender and this recommendation as we hypothesize that there may be inherent sexism |

	baked into the TikTok algorithm, where accounts that are identified as female are shown less or are not shown science-related content on their "For You" page. "*The recommendation would be merited if the strategy is to include more gender-related topics to gain more female followers, who then view videos of other topics*" is precisely what we suggest and recommend. A nuance to note is that videos from an account a person is following are also often shown on the 'For You' page. Hence, even if the video is viewed by an account follower, the view will still be counted as coming from the 'For You' page. (c) We re-emphasize here that views on TikTok generally do not come directly from hashtags (like one might look up a hashtag on Twitter), rather that hashtags are tools to provide additional algorithmic context for the video (especially beneficial if highlighting a location, e.g., #california).
Added Text:	(Section 7.1) "We also found that there was little relationship between the time a video was posted and how well it performed (Fig. 2, 9d). Post time optimization is the concept of posting social media content at a time when most users/followers are active and online and likely to see the content, thus increasing its potential reach. As the majority of our account's followers are located within the United States, we focused our posting during daytime hours within the U.S. before followers would logoff online for the night. However, it would be interesting, especially if wanting to attract an international audience, to see how videos perform if they are posted during nighttime hours in the U.S. We anticipate that videos should be posted at a time local to when the desired or majority audience is primarily online."

Anonymous Reviewer #2:

Reviewer Comment:	First of all, it's not clear which research gap this paper tries to address. The goal to increase the visibility of geosciences and geophysics on TikTok is achieved to some extent. The 2 million views in 4 months are impressive, but there is almost no description of why would and how the authors created those very popular videos in such ways. Are there any different considerations when the authors made those videos? Is there any hypothesis that led the authors to make those videos? Or the success videos are simply due to good luck? Above all, what has video creators learned from the process? I think that the authors should reflect on this process from a critical standpoint.

Author Reply:	This study is the very first to produce and evaluate geoscience content on TikTok. TikTok and short-form video content has exploded in popularity, and until this point there had been no research on how to capitalize on this success for geoscience communication efforts. We thus fill a crucial research gap related to science communication efforts via short-form video content and provide the first ever analysis of geoscience content on TikTok. There is no way to predict which videos may go 'viral' on TikTok (as is the case for any type of viral content). Virality likely requires some type of luck, but also content resonates with the viewer. Our goal with this study was to produce different types of content (different video length, video topic, presentation style, etc.) and evaluate what trends (if any) we can discern in content that had the highest success in terms of reach and engagement.
Reviewer Comment:	The authors listed two goals in the manuscript: To increase the visibility of geosciences and to determine the best strategies for effective geoscience communication on TikTok. However, the engagement is not discussed thoroughly, and the listed data are insufficient to analyze the geoscience communication's effectiveness on TikTok. Statistics of likes, views, and percentage of watched are useful metrics for showing the engagement of videos on TikTok but more importantly, stemming from the previous point, the authors should give their reasons. Although some design factors have been discussed loosely, like memes and music, the authors may want to give a more clear message about why some of the videos you made are more popular than others (at least your hypotheses and using your data and experiment design to test them). The current analysis and the reported data are not sufficient. I am wondering why the authors did not do any qualitative analysis of comments (considering you have so many comments and some of them are not relevant to learning). Lacking experiment design and the reported data as written hardly can give an insightful answer. I will recommend the authors analyze the contents of the comments and use a mixed method to critically discuss the potential reasons. This leads to another potential weakness: the literature review on how to design geoscience videos to engage the public on social media is insufficient.
Author Reply:	In this paper, we focus primarily on quantitative metrics for analyzing the 'success' and impact of science communication content on TikTok. Assessing the average video view duration and engagement rate provides important insight into whether people are interested in or

	engaged with the content. If we see videos that people are watching for a long/full duration and liking/commenting/sharing, that indicates that the science communication was interesting and engaging (and we aim to share interesting facts and knowledge with the broad public). We did not include memes or music in our videos, and thus those are not factors in our video design. We discuss in the paper reasons for why some videos may be more popular than others (users may be more likely to watch shorter demonstration videos, users are likely to receive location-based video suggestions, etc.) As we discuss in the section on ethical implications, user privacy on TikTok is an important consideration, and thus we do not include any specific comments, as no users gave explicit agreement to be part of this study beyond TikTok's terms of use. Given that this is the first study examining using TikTok for geoscience communication (we demonstrate that the platform offers exceptional reach and low-barriers to entry for science communicators), our plan is for future studies to build upon this initial study and focus more on qualitative data, particularly from user comments. For future studies, we would like to follow Institutional Review Board protocols and receive approval to qualitatively evaluate the effectiveness of the science communication on TikTok. However, in this paper, we have added a number of sentences generalizing user comments and responses to our videos. We find this to be a happy medium, especially since our videos received over 3,500 comments during this period, and it is beyond the scope of this study to compile and evaluate all comments.
Added text:	(Section 7.2) "User comments on GMV videos typically expressed how it was a unique, interesting visualization that they had not seen before." "User comments also often expressed interest in the distribution of seismic instruments in North America." "Comments left on these videos typically expressed that the users enjoyed the overview of information and how it was conveyed." "Based on these video comments, the high engagement was generated in part by users from those geographic locations. Many commented

	that they were from the locations being covered; some recognized features in the town they were familiar with, and others learned something new and interesting about the geology of where they live."
Reviewer Comment:	Also, the authors may want to do a more thorough literature review about the existing geoscience efforts on TikTok and how your efforts differ from others (again, what research gap this paper tries to address). The current literature review gives a valuable introduction to TikTok for geoscientists and communicators unfamiliar with TikTok, but this is not a research gap. Without this, this manuscript is not a rigorous research inquiry. The authors mentioned several possible theories, such as place-based video design (may want to use field-based or location-based design since the place-based can be misleading to place-based educational design) and geohazard event-based video design. However, there is no literature review about these designs or what is the template of the authors' video design. Very little is said about how this effort and results extend generalizable knowledge about how communicators can replicate the success of designs.
Author Reply:	There is no other existing literature related to geoscience communication efforts on TikTok, as ours is the very first study to exist. The majority of geoscience content on TikTok is produced by young individuals who have recently or are currently completing bachelor's or graduate degrees in the geosciences. Content is often split between "meme-style" videos, showing what's inside a rock when you break it open, and explainer videos of geology topics. Our organizations are pioneering an effort to produce concerted geoscience education videos on TikTok and encourage more scientists and science communicators to do the same. Our study provides the first knowledge base of what types of geoscience communications videos may be successful on TikTok. While there is no existing literature on the topic, there are many museums that now have TikTok accounts (e.g., the American Museum of Natural History), and an analysis of museum-related content would be complementary to this study, but outside the scope of the current study. Future studies will be able to build on our work and test the hypotheses that we offer. We agree that the term "place-based" can be misleading, and we have changed this term to "location-based" throughout.

Reviewer Comment:	The authors may want to discuss what their hypothesis works, what may not, and how they predicted the videos' success on TikTok (considering there is no existing formula or clear pattern in the data). How is feedback incorporated into the development of new videos? Although this work shared valuable data to confirm and highlight the possibility of TikTok, especially the unique advantage of FYP and TikTok algorithm for new geoscience communicators, it is not novel enough to consider extending our knowledge boundary about geoscience communication on TikTok.
Author Reply:	There is no singular way to predict the success of a video, and unfortunately there will never be a formula to guarantee success with social media content. However, we have been able to observe certain trends both related to the topic of the video and the 'construction' of a video that may aid in its success and increase the impact of the communication effort. For example, after making a video talking about the geology of a specific location that got a large number of views, we made more videos featuring location-based geology that were also largely successful. As the videos were all made during a pilot period, the pilot was focused on producing various different types of content. After full analysis of the videos (this study), that will allow us to fully assess patterns and trends in video success, which can then be incorporated into the development of new videos.
Reviewer Comment:	For example, what are the teaching goals of each video? why do some of the videos get so many views? While some others are not. Is it the content? The narrative? The visuals? The overall design? Why they were recommended more than others? Is it really because of the geohazard feature? If so, the authors should organize their information better. The GVM could be a good visualization, but the evidence is not enough to be convincing. The paper tested several metrics but didn't show convincing evidence of what factors can increase the visibility of geosciences and geophysics videos. The reported results have no clear pattern to determine what type of videos or design would be more popular. As written, the amount of GVM videos (9 videos) is not enough to discuss the statistical significance of the effectiveness of this design element. This is the same as the geohazard videos. (A minor suggestion: The authors may want to define the 'lecture-style video'. Is it a universal definition or just for TikTok?)
Author Reply:	The teaching goals of each video are essentially to convey geoscience-related information (in our organizations' case seismology, geodesy,

	and topography) in a way that is accessible and engaging to the viewer. There unfortunately is no way to specifically discern the exact reason for why a video is successful, as it is likely a combination of factors (the topic + the design/format + the visuals). Regarding the number of GMV videos, they are restricted to when magnitude 6+ earthquakes occur, so it is generally not possible to get a large sample size of videos over a short period of time. However, even with a smaller sample size, the GMV videos are the best way to isolate variables, as the video design/format, visuals and description are the same, and observe if there are any other factors that may impact the videos' success. We evaluate statistical significance with p-values. As a 'lecture-style' video is our own definition and wording, we included a longer explanation of what we mean with this term. (Essentially, like a teacher would 'lecture' to a class using a PowerPoint presentation, our 'lecture-style' videos feature the host directly talking to the audience like a teacher would, showing visuals and imagery in the background.)
Added Text:	"Our videos featured three rotating front-facing hosts and were typically done in a "lecture-style" format, where the host talks directly to the audience explaining a geoscience topic while using engaging background visuals."
Reviewer Comment:	Moreover, two million views in 4 months with only 48 videos are amazing, but does the views of TikTok comparable to those of YouTube videos? Considering its extremely short durations (5-60sec), a video on the same topic can be much longer. This question also applies to the percentage of watched and the likes. On YouTube, IRIS channels also have many good geohazard videos, but it looks like they didn't get so many views. Is it the timeliness, the platform, the audiences, or the design? I think the authors have the special advantage in giving a good discussion about this.
Author Reply:	Views on TikTok can be difficult to compare to views on YouTube, as TikTok videos are so much shorter than YouTube videos, as noted by the reviewer. For a few of our TikTok videos, we pulled a number of visuals directly from longer videos on the IRIS and UNAVCO YouTube pages, which would be the best comparison of a long-form vs short-form video. However, it is difficult to directly compare the number of views, as some of the YouTube videos have been up for

	seven years accumulating views and likely have variable spikes in views over time. A future study could perform a more specific direct evaluation of short-form vs. long-form educational geoscience video content.

The best direct comparisons we can provide for TikTok vs. YouTube as platforms are the Terra Explore lidar TikTok videos that were uploaded to the OpenTopography YouTube channel as 'YouTube Shorts.' The same exact videos on YouTube received extremely few views, orders of magnitude lower, compared to the videos on TikTok. This observation would suggest that there is a unique opportunity for expanded reach of science communication topics on TikTok. However, that is beyond the scope of this paper. |
Reviewer Comment:	Furthermore, the work provides some valuable data. But, many of the figures have little value and may be removed or revised. More importantly, the results alternate between assertions about the evidence presented, point-of-view statements that are not identified, and overreaching claims. For example, Line 278 to 280, "Videos 40 sec to 2 minutes in length received the highest engagement rates (Fig. 10)", Line 282 to 284 "High engagement rates on videos with lower view counts …a wider audiences..", I cannot find a clear pattern or empirical evidence supporting these claims.
Author Reply:	We have since performed more rigorous statistical analyses per the comments of Reviewer #1, which better aids in the clarification of such statements. New text is added throughout the results and discussion sections.
Line:	30-32, 291-299
Reviewer Comment:	Interestingly, there is almost no literature review or efforts mentioned or discussion about using YouTube to communicate geosciences (the authors even talked about Twitter in the introduction). Shouldn't YouTube be a more comparable video-based social media platform to TikTok? Especially considering IRIS, UNAVCO, and Open Topography all have their YouTube channels. Is there a particular reason?
Author Reply:	There is little existing literature that we are aware of that discusses geoscience communication on YouTube. However, there is a new study by Wang et al. (2022) in *Geoscience Communication* that does evaluate geoscience videos on YouTube that we have now cited in our

	paper. While YouTube and TikTok both are video platforms, the short-form, 'casual', vertical videos of TikTok and its algorithm-driven nature make them less comparable than at first glance (It would only be apt to compare 'YouTube Shorts' with TikTok videos). Our organizations' YouTube channels have content that differs quite greatly from the videos produced for TikTok. For example, the OpenTopography YouTube channel primarily features tutorials and webinars geared towards researchers and scientists, rather than direct public outreach efforts. IRIS on their YouTube have videos that range from public outreach (e.g., 'Women in Geoscience Series') to educational animations and explainers. UNAVCO's YouTube channel largely has videos from webinars and short course videos. A discussion of this material on YouTube is not relevant to or related to our current study, which specifically focuses on TikTok.
Added Text:	(Section 7.2) "Research from educational geoscience videos on YouTube also suggests that short, timely videos about natural hazards is useful for engaging the public, especially those who live near where the hazards occur (Wang et al., 2022)."
Line:	60
Reviewer Comment:	I think the current limitation for TikTok videos is 10 minutes, not the 3mins anymore.
Author Reply:	TikTok now does allow users to directly upload a singular 10 minute video file within the app or via desktop, however you can only film videos on TikTok that are 3 minutes in length (which is the primary mode of creation).
Added Text:	"There is now also the limited ability to upload a single 10-minute video file from your device to TikTok."
Reviewer Comment:	Literature Review is not enough. As mentioned at the beginning, the literature review introduces some interesting concepts and facts about TikTok and science communication on TikTok. Still, these do not carry over to the study's variables and measurement.
Author Reply:	There currently is no other existing literature related to the geosciences on TikTok, and there are very few other studies that evaluate or discuss science communication on TikTok or short-form video. Our study is the very first to create and evaluate geoscience communication on TikTok, and we are pioneering evaluations and investigations rather than drawing upon an existing knowledge base. The analytical toolkit that we use is more generally related to performance of content on

	social media. We have included further discussion of the types of geoscience videos on TikTok (primarily produced by individuals who have recently graduated or are currently in bachelor's or graduate geoscience programs).
Added Text:	(Section 3) "Much of geoscience-related content on TikTok is created by young individuals who have recently or are currently completing undergraduate or graduate degrees in the geosciences. Videos typically feature geology-related "memes" or jokes, showing what's inside a rock when you break it open, or explaining geoscience topics."
Line:	Figure 5
Reviewer Comment:	Regarding figure 5, why do you not report the statistics of shares? Have any of your videos been shared and viewed on Twitter or embedded on other websites? How will these be categorized? Will they be categorized into FYP, Personal Profile, or Unknown? This may affect how videos are being found and watched.
Author Reply:	Shares are not reported as a video view category. There is no way to know how or where a video has been shared (copied the link vs. texted to a friend vs. posted to a social platform, etc.). We hypothesize that video views from shares are counted as a 'Personal Profile' view, as videos shared are the direct link to the video. Additionally, a video being shared does not equate to a video view, and there is no way to know whether the video was actually viewed when it was shared.
Line:	350-355
Reviewer Comment:	If the gender of audiences on TikTok cannot be accessed or the data is unreliable, then a more thorough discussion of the limitation should be added.
Author Reply:	We have no reason to believe that the gender percentages reported on TikTok are unreliable. While users do not self-contribute this data, TikTok works to discern gender from accounts that the user links (Facebook, Instagram) where users can report gender, or from algorithmic assessment of user behavior, users' names, etc. (If anything, TikTok is known for over-mining user data). These are the same demographics that are given to multi-million dollar companies that run targeted advertising campaigns on the app, who would rely on robust demographic data. We do note though that the gender % of followers does not account for non-binary individuals, and what is lacking is a more nuanced categorization of gender. We have added some additional text in this regard.

Added Text:	(Section 7.4) "However, we have no reason to believe that this information would be inaccurate due to TikTok's heavy data mining and the fact that this same demographic information is also provided to million-dollar corporations for targeted advertising campaigns."
Reviewer Comment:	I personally want to see more data about the timing of geohazard events and the release time of videos. Is there any relationship between the speed of release and the views of the video?
Author Reply:	All videos related to geohazards (primarily earthquake and GMV videos) were released the day of or the day after the event. Thus, there is essentially no lag time that can be analyzed.
Reviewer Comment:	The demonstration using food could not be the major reason for different views. One possibility in my mind is the interests of the topic (e.g., more people have motivations in learning the magnitude of an earthquake than the types of faults), is there any useful comments to give insights about this?
Author Reply:	What we suggest in the text is that the difference in the videos' performance is that one video was ~30 s, and the other video was ~1 min. The significantly longer duration of the video is what we hypothesize led to its smaller reach, suggesting that shorter demonstration videos better capture an audience's attention.
Reviewer Comment:	The authors may want to clarify how this work contributes to broader theoretical debates like how geohazards affect or how place-based design affects engagements. Current discussions are not critically reflected on the empirical data (including yours and others).
Author Reply:	We have added additional text and a citation in this regard.
Added Text:	(Section 7.2) "These findings support pedagogical research that suggests that recent, real-world incident examples should be included when teaching geoscience or environmental science topics to increase student engagement and interest (Singh et al., 2022)."