

Author's response to reviewer and editor comments

Thank you very much for all of your detailed comments. Please find responses to your comments outlined below.

These comments were all very insightful and made us rethink the structure and the narrative of the article. In revisiting the manuscript, we also built on many of the reviewers' comments and criticisms related to the small sample size and adopted a more conservative approach to our statistical analysis. We have reported these changes in the methodology and the results sections, and believe that these changes improves the quality of the paper substantially, while refining and narrowing its claims.

Reviewer 1:

The manuscript is overly long and reads more like a thesis or report. It needs significant trimming.

The manuscript has been shortened by about 10% in word count and by 3 pages. As a few tables were added, the word count in the body of the article should be reduced by 15-20%. This has been done by shortening the overall article, especially the regional setting and the results and discussions sections. Efforts were made to shorten every section where available.

The introduction should be rewritten to focus on key points relevant to readers, with unnecessary regional details minimized.

We have shortened the introduction by setting a tighter focus. We have reduced the breadth of the points provided with less details of the southeast Arabia region.

The statement of aims could be more concise (e.g., condensed to 1–2 sentences).

The aims of the study have been shortened and reduced to three sentences, in an attempt to follow this recommendation, while maintaining the comprehensive description of the overall aim of the study.

The description of methods needs to be clearer, especially Section 3.3.

The methodology outlined in Section 3 needs refinement to improve clarity and accuracy, especially Section 3.3. For instance, in the first sentence of Section 3.3, it is stated that "An online experiment was conducted using the videos produced in the second phase of the study," which is incorrect. The study started in Phase 1 with a focus group—independent of an online format (or was it?).

We have adjusted the wording to make the methodology clearer and avoid causing any confusion.

The sample size for each phase (focus group, pre-test, post-test, follow-up) should be explicitly stated.

Thank you for the suggestion. We have added Table 2 in the manuscript to make this information more apparent earlier in the paper.

Use of the term “a wave” for surveys is not appropriate; consider standard terminology from similar peer-reviewed studies.

As some terminology used in the study may be unfamiliar to the audience of this journal, we have added Table 1 to outline terms that are developed in the social sciences field.

Ethical concerns are raised regarding the collection of personal information (name, email, WhatsApp, institution); anonymisation should be used instead.

A critical concern is the reason for requesting participants to specify their names in the questionnaires. If the aim is to track individual opinion evolution, participants should be assigned anonymised codes. Furthermore, asking for extensive personal information—such as email addresses, WhatsApp contacts, and study institution—is unnecessary and raises ethical considerations.

Thank you for pointing this out. This was a topic mentioned in the ethics review conducted for the study as well. However, given recommendations from local collaborators for continuous engagement in a longitudinal study. The personal data were considered necessary to reach out to the participants effectively for completion of the study. The manuscript was edited to directly address these concerns, and pointed to the ethics statement section as well for better clarity in the ethical approval.

The sampling was based on convenience and is not representative of the broader population, limiting generalisability.

Another important limitation is that the sample is not representative of the wider population. The assumption underlying the sampling method (as outlined in...) pertains to random sampling within a homogeneous population. However, the sampling was non-random and based on convenience. Consequently, the results are only applicable to a highly educated subgroup within a specific locale.

We are aware of this limitation in our study, and have added multiple references to this limitation in the study. The sampling, however, was based on targeting rather than convenience. We believe we have taken the necessary measures to address this point, by tightening our wording so it is more evident. Furthermore, we have edited the results to include the confidence intervals to highlight the limitations of the smaller sample that we had in our study.

The use of different samples across the three survey phases undermines longitudinal comparisons. Only data from participants who completed all three phases (n = 82) should be used in longitudinal analysis. Discrepancies in sample sizes may explain some unexpected patterns in the results.

We agree that potentially non-random patterns in survey drop-outs are a concern; our longitudinal analysis (DiD) used only the samples from those who completed all three waves (n=82), with the per-phase samples (specifically the larger Wave 2 sample, n=104) being only used for summary statistics and t-test reports, not for the main analyses. We have clarified the paper text to make this more clear, as the presentation of exactly which samples are being used in which analyses was confusing in the original draft.

Reviewer 2:

The paper lacks structure and clear organisation, making it difficult to follow.

We have restructured the paper and refined it to provide a clearer organisation. Also, tables and figures were added to better guide readers through the paper.

However, in its present form, the study lacks a lot of structure and provides few details, making it difficult to follow for communicators interested in replicating this experiment.

We will make the revisions based on your inputs along with the inputs provided by the fellow reviewers.

There is insufficient information on how the focus group was performed and how this informed the video. I was expecting much more regarding the criteria to include the Examples in the two videos. Is it because these sites could be candidates for becoming UNESCO Global Geoparks? Or is it because these sites have some well-documented archaeological research attached? Are they critically endangered? This could have been described in the regional setting.

As you pointed out, the focus group results lacked an explanation of how it fed into the contents of the video. We have added more information in this regard. The different sites used as examples were selected to represent a wide array of Quaternary geoheritage sites in the region. We have provided a clearer explanation of why these sites were selected, to highlight the features that makes them distinctive and interesting to be included in the video.

There is insufficient information on how the focus group informed the video's script. Did they provide the examples, or were they discussed from a short list of

candidates? Did they provide examples of the religious texts to be discussed and how they should have been?

As pointed out in this comment, the focus group results lacked an explanation of how it fed into the contents of the video. We have added this important element into the manuscript. We have provided a better explanation of why the video sites were selected, to highlight the features that makes them distinctive and interesting to be included in the video.

Regarding the influence of the focus groups: did they know what the audience targeted for the videos was?

Yes, there was comprehensive explanation in the beginning of the focus group sessions. It was one of the first information provided in the introduction of the focus group. However, this information was not explicitly mentioned in the manuscript. We have made revisions so that this information is clearly presented.

Were the videos intended to be screened in museums or universities? Attached to a special exhibition?

No, the videos were not screened in museums or universities, or attached to a special exhibition. The video was design specifically for the research conducted. We have added text in the manuscript to clarify this point.

Currently, both videos have a lot of specialised content that is not explained. Worth mentioning are the following: a) the video does not explain what a facies is, but it is used as an important plot line to say these quaternary geosites are important to understand climate changes; b) the video moves too fast in temporal scale, by including information regarding the 4.56 billion years of geological record but without really explaining much of this geology in southeast Arabia; c) the cultural aspect is glossed over and it is not readily clear if they four sites represent the same ages in the Quaternary or if they are all from different sections. It seems that most participants were very familiar with the jargon and chronologies to understand much of the video, but it is not aimed at outreach to the public; d) non-specialised public is introduced to too many topics in 9 minutes, namely deep-time, several geological scales, archaeological evidence, sedimentation and depositional environments, types of erosion, and names of geomorphological structures. Nevertheless, my comments come from assuming that the video is addressed to the general public. I can see how that long video could be more informative if these topics were explained in an exhibition at a site museum or a geopark.

The video was not produced as a material to provide comprehensive information about geoheritage sites or geology in general. We agree that a wide array of specialised knowledge was introduced, but the purpose of this video was not to provide full explanations of each structure or concept, but to provide an overview and see if such an overview can interest viewers and change their knowledge level and/or attitude for

conservation of Quaternary geoheritage sites. The participants were also given a list of resources that they can use to gain further information about the topics introduced. We will add this point into the manuscript for better clarity. Furthermore, a long video for use at a site museum or a geopark could be interesting to understand its effects, but such a consideration is beyond the scope of this current study.

How was the survey translated into Arabic? Was the difference between the responses given in Arabic and English relevant?

The survey was translated by one of the authors Mr Al Rawahi, who is a native speaker of Arabic. This information has been added for clarity in the text. The differences between the languages were considered, but did not find meaningful differences between these groups in our preliminary analysis. This explanation has been added in the text.

How was the quiz compiled? Can the quiz be included as a supplement as well? Currently, it only reports identifying the photos of potential quaternary geoheritage sites. What was the question? How was it produced? Did the focus group also discuss the nature of the quiz? What about questions regarding the archaeological findings?

The details of the quiz were incorporated in the questionnaire. There were two knowledge test questions in each questionnaire (Q11 and Q18.1 in the first wave, Q3 and Q12.1 in the second wave, Q11 and Q19.1 in the third wave) and they are outlined in appendix i. The questions were prepared by the authors, and it was not an element discussed in the focus group, as the focus group was conducted with the objective of identifying the most appropriate elements to include in the video. The answers to the Quaternary landscape question were outlined or explained in the video, but the relationship between Quaternary geology and question was only explained in the video with the cultural components. We have rephrased the term “quiz” to “knowledge test questions” to clarify the nature of the question asked, as “quiz” may give the impression that there were many questions asked. Also, we have added text to reflect the points above.

Also missing is a breakdown of the survey answers. This would enable us to understand the demographics of the respondents and the difference between different groupings (non-nationals v nationals, specialists v non-specialists, geoscientists v others, among other combinations).

The results of the study, especially in the statistical analysis, already provides a breakdown of specialists vs non-specialists, as identified in the text and Table 6. Geoscientists vs others would result in a similar composition to the specialists vs non-specialists analysis. Nationals vs non-nationals was not reported as there were insufficient non-nationals in the study to give statistically robust inferences (this attribute was not a focus in our recruitment of participants as it did not form part of our

original study design). In preparing our response to the reviewer's comments, we did run all of our models with a dummy variable representing the contrast between UAE nationals and non-nationals specifically, but none of these yielded statistically significant results due to the lack of sufficient samples (only about 20 on each side of the contrast, far too few for a fixed effects model such as DiD). We have noted this in the paper text.

Reviewer 3:

I would also expect more detailed information about the production of the videos and the design of quizzes.

As mentioned above, a better description of the video production process and design of the quizzes, including revisions to the Table 4, has been made to respond to this point.

My general impression is that the paper is overlong in the Results sections and the results of different tests are better presented in a synthetic form, with the aid of one composite graph and table rather than split into many. As it is now, it offers a monotonous, not very engaging reading.

We have consolidated the graphs and tables and adjusted the text so that the overall paper is shorter, more varied, and easier to read (e.g. the interest and knowledge sub-sections were combined to make the overall paper less repetitive and more engaging).

The second impression is that not all aims of the study formulated in L227-232 have been finally addressed. This mainly regards the aim #2. This paper does not really offer any recommendations "how effectively the connections between Quaternary geoheritage sites and local culture (i.e. history and archaeology) can be communicated to the public in this region". Perhaps this intended aim can be simply removed.

This is a very important and valid point. We have taken it into account, and we took out aim #2. This has enabled a tighter more targeted aim for the study. Thank you for pointing it out.

Third, two limitations are listed in the penultimate paragraph (L1013-1017), but the third important factor is the nature of Quaternary geoheritage itself which varies from one environment to another (as well as its geocultural dimension). In arid lands such as Oman and UAE the sedimentary and landform record of the Quaternary is well visible; it is not so in humid, vegetated environments, where we only have occasional insights. Further, landforms such as dunes or cave speleothems are much more captivating than many other types of relief, equally

important to understand the Quaternary. I guess that even non-specialists in Arabia are familiar with dunes, whereas the vast majority of general public in central and western Europe is likely unaware of the existence of large dune fields, now under vegetation.

The point that Quaternary geoheritage sites are visible and apparent in Oman and the UAE is a very important point, and we have emphasised this more in the regional setting section. Thank you for your suggestion.

Another issue is that while the paper as a whole is quite extensively referenced, there is very little conceptual background provided about the linkages between geoheritage and cultural heritage, which is now explored in different papers from different geomorphological and cultural settings. Multiple themes are addressed within this broader framework (a few review-type papers and book chapters have been published in the recent years) and it would be interesting to have an overview which ones can be developed in Oman/UAE and which ones have been included into the video for the treatment group.

This is a very good point to improve the opening of this paper, and we have included a short review section to address this point in the introduction.

There are also some terminological issues to solve:

L79 It is incorrect to identify a 'geomorphosite' through solely its aesthetic attributes. I fear that Reynard et al. (2009) was misunderstood. Please see his chapter "Geomorphosites: definition and characteristics" (available from https://www.researchgate.net/publication/266021163_The_assessment_of_geomorphosites) for the thorough discussion of the concept. Another useful reference source is Coratza and Hobléa (2018), The Specificities of Geomorphological Heritage, included into "Geoheritage" book by Elsevier.

Thank you for your comment, we have revised this point accordingly to make sure that this wrong impression is not given.

L97-98 This suggests that Quaternary sites are synonyms of geomorphosites. They are not. Some geomorphosites show landform elements which have roots in the more distant past and were subject to minimal change during the Quaternary, while some Quaternary sites are not really geomorphological sites (e.g., lake sediments).

Thank you for your comment, we have adjusted the text accordingly so this is not the case.

L243 How is the difference between "alluvium and gravel" and "fluvial" understood? Alluvium is a fluvial (waterlaid) deposit, whereas adding "gravel" is unnecessary and confusing, as it introduces a component of lithology missing in other categories.

Fluvial describes processes, actions, or systems related to rivers and streams, while alluvial describes the material (sediment) deposited by rivers and streams or the features formed by this material. We have taken out “gravel” from the manuscript

Finally, I think that in L280, L386 and L407 the correct phrasing should be "focus group MEETINGS (were designed)", whereas in L842 it should be "Relationship between Quaternary Geology and Archaeology" (not "Quaternary geoheritage sites" again).

Thank you, we have revised the manuscript accordingly.

A purely technical comment is that section 1.5 is missing (1.6 should become 1.5).

Thank you, we have revised the manuscript accordingly.

Editor’s comment to authors

Specifically, it is crucial to clearly delineate, in the early portion of the methodology section, the number of participants involved in all phases, with particular emphasis on Group A, Group B, and the pre- and post-video stages. This clarification can be effectively achieved by revising Fig. 2. Alternatively, a summary table may be incorporated following Fig. 2, which would specify the total number of participants, as already detailed in Tables 1 and 3.

Thank you for this comment, with a very good idea to respond to the reviewer’s comments. We have added Table 2 in the manuscript to respond to this suggestion.

Given that some terminology originates from the social sciences, we recommend providing explanations of how these terms are used. This is particularly important as we are publishing in a journal focused on geoscience communication, and the potential readership may encompass individuals from various academic disciplines.

Thank you again for another very constructive comment. We have added Table 1 to respond to this point and to better explain terms used in the study.

To address ethical considerations, authors may include an ethical statement in their manuscript, as outlined in the guidelines for contributors available at <https://www.geoscience-communication.net/submission.html>. For further guidance on structuring this section, authors are encouraged to consult the editorial found at <https://gc.copernicus.org/articles/4/493/2021/>, which also provides insights on organizing a research paper.

An ethical statement had already been included in the manuscript. However, as described in the comments above, we have added a clear discussion to address this point and pointed to the ethical statement and the procedures within the manuscript.

Given that the study presents certain limitations, as identified not only by the reviewers but also by the authors themselves, for instance, in L580 (where it is noted that, although the study initially aimed to target the general public, the participant profile is restricted to a highly educated, nature-loving, and young segment of the population), it would be prudent to include a dedicated paragraph discussing these limitations. This section should also address the extent to which these limitations may impact the results and propose directions for future research to address these issues.

A paragraph dedicated to the limitation to the study and future direction that similar studies can take has been added in the conclusion to address this point. Thank you for pointing out the need for such a section.

In response to Reviewer #2's comments on the video content, please provide a more detailed clarification of paragraph 1.6 (Aims of the study) in accordance with your response to Reviewer #2.

The aim of the study and the content of the video in the manuscript was edited to clarify the scope and the intentions of the study and use of videos.

Following rev. #3, consider the possibility of summarizing the results using graphs and tables for clarity.

The summarisation of results has been conducted by the insertion of Figure 3 and Table 6.