

REBUTTAL

Breaking the Silos: an online serious game for multi-risk DRR management

REVIEWER 1

Overall assessment

This paper presents a new role-playing game (RPG) 'Breaking the Silos' – designed to replicate the complexities of disaster risk management, the game exposes players to a range of hazards through a series of rounds. Players are asked to make recommendations for spending towards disaster risk reduction measures to the President and are confronted with the consequences of these actions as the game proceeds through the rounds, new hazards are experienced, and the narrative builds.

Overall, this is a very interesting study and I extremely pleased to see some much-needed thinking in this area. Gaming, and particularly RPGs are a novel approach in this field, are a difficult route to investigate, and I think the Breaking the Silos game is a great first step. Often games in this arena fail to provide empirical evidence to support their application, so I am also very happy to see a preliminary evaluation of the games impact. The data collections are fairly shallow, principally comprising quantitative collections and I'd recommend more detailed data collections, including qualitative data during future game tests, which could really shed further light on the games impact and efficacy. For example, an aim of the game is to communicate about the complexities in the decision-making process – consideration for the interaction between players during the game, in terms of collaborations and disagreements, could help to understand if the game is capable of expanding awareness in this area.

We thank the reviewer for their kind words. We fully agree that our study presents a preliminary data analysis. We did collect some qualitative feedback and have updated the manuscript with that. Nonetheless, the reviewer makes a very valid point, and, while out of the scope of this paper we will continue to collect and analyse feedback to keep improving the game into the future. We really like the suggestion of the reviewer to dive deeper into the behavioural aspect and to better understand how players made decisions to (not) interact with other players.

Overall, the paper is pleasant to read, well written and well structured, although a more concise description of the game mechanics would make the paper a little more efficient. Some of the detail on the game mechanics is a little hard to conceptualise, so perhaps a little more precision here might help.

We agree with the reviewer that the description of the game mechanics is quite lengthy, and that a clear and brief explanation of the game mechanics was missing. In line with the comments of the other reviewer we restructured and rewrote the paper to better explain its emphasis on the development of the game. We did add the following at the start of Section 3 How to play the game:

[399] In *Breaking the Silos*, players are assigned their role at the start of the game. Before the start of the game, each player reads their role card, which provides them with detailed information about the DRR measures they can implement (these are unique for each role) and goals they need to meet. The game consists of three rounds. At the start of each round, the moderator selects a disaster and impacted area, and informs the team of the impact that the disaster caused. In each round, the team has a tight budget and time restriction to address the post-disaster situation and to prepare for a next disaster by implementing DRR measures. After each round, the moderator calculates the effects of the implemented DRR measures on the next disaster.

General comments and questions:

- Right at the beginning of the piece, I think the main aims of the game need to be much more strongly stated and with a description of why you think the game, and particularly its mode as an RPG, is able to achieve your aims over other methods of stakeholder education and outreach (e.g. scenarios) and even other game styles.

We agree with the reviewer that the main aim was not stated very clearly. We do not see this particular game version as a replacement to these other forms, but we believe that complementary forms are needed. We have adjusted the paragraph and included an explanation of the use of the RPG-game version in supporting our aim. That paragraph now reads as follows:

[90] To address the aforementioned gaps in current disaster risk serious games, we developed the *Breaking the Silos* game. *Breaking the Silos* is a multiplayer role-playing game (RPG) that aims to raise understanding of the complexities of multi-hazard risk and asynergies of DRR measures among different DRM stakeholders. Solinksa-Nowak et al. (2018) argue that RPGs, more than other game types, allow players to directly experience the uncertainty, chaos, and stress of a DRM situation. Especially multi-stakeholder negotiation RPGs have been shown to be very promising in enhancing a player's knowledge of content and process (Rumore et al., 2016). In *Breaking the Silos*, the players are a team of different decision makers and stakeholders in the DRM process who advise the president of a fictional country on the implementation of DRR measures after different disasters, while considering potential (a)synergies of these DRR measures. The DRM process is mimicked by spreading knowledge and objectives throughout participants and by including randomness to the storyline. The game was developed to help various stakeholders (including policy makers, risk managers, researchers) better understand the complexities of multi-hazard risk and the potential (a)synergies of DRR measures. Unlike past games, this game includes multiple hazards and their spatiotemporal interactions. It also explicitly includes both the response and planning phase of the disaster risk cycle and promotes the examination of (a)synergies between different DRR measures. In doing so, we aim to create a game with a more realistic representation of the growing complexities of risk.

- What is the time duration of each of the rounds in terms of both, how long it physically takes to play but also what is the temporal dimension of each round and between round 1 and round 2?

We thank the reviewer for this question. We realise we did not clearly state the actual playing time in Section 2.3. We have adjusted this accordingly. The actual playing time during the first round is 30 minutes, while the second and third round are 20 minutes each.

The temporal dimension of the game varies per scenario. Both for the time between events as well as the duration of the implementation of DRR measures we account for three different time windows: weeks, months, and years. We realized that this information was scattered over Sections 2.2 and 2.3. To clarify this and in line with a later comment of the reviewer, we have included a timeline showing both the physical time as well as the temporal dimensions.

- The figures are really low resolution and make it hard to see the detail on the game boards. This might be due to the draft document in the review process, but would recommend these being much higher resolution where possible.

The low-quality figures were indeed part of the initial draft document, and we have adjusted this in the revised manuscript.

- In your literature review, be careful to be distinctive where games are table top, digital or role playing games. Different styles and platforms of gaming have varying benefits and drawbacks and can be difficult to compare. Hence, some are more effective than others.

We agree with the suggestion and have made adjustments accordingly. We have used the term “RPG” to refer to our own game and included a sentence in the introduction to explain the existence of different game formats, which makes it difficult to compare them.

[47] In recent years, a large number of serious games relating to DRM have been developed. Solinska-Nowak et al. (2018) conducted a meta-analysis of DRM serious games and found that the majority focus on floods (27 out of 45 reviewed games), earthquakes (10 out 45), and droughts (7 out of 45), while storms (including cyclones, hurricanes, etc) are rarely the main hazards in a game (Solinska-Nowak et al., 2018). **It should be noted that these games encompass a wide variety of game formats, including single- or multi-player video games, single- or multi-player tabletop games, and role-playing games (RPGs), making it difficult to compare them.**

- You mention frequently throughout about randomness in the storyline and game design (e.g. L73 and L348) – can you expand on what you mean by randomness in these situations, how ‘randomness’ is brought about (i.e. is there a game mechanic to facilitate it), and what randomness in the narrative actually looks like?

The randomness of selecting a hazard, hazard intensity, the size of the impacted area (number of cells), and duration between disasters was obtained by creating a random generator. In practice, this means that for each of these categories a random number is selected between the number of possible options. For example, for the hazard type a number between 1, 2 or 3 is generated (using the integrated random generator function RANDBETWEEN) that corresponds to a tropical cyclone, drought or flood, respectively. For the hazard intensity, three levels are possible (1=low; 2=medium; 3=high), for the number of affected cells, it is an integer number between 5 and 15 cells. However, the exact location however is selected by the moderator. Finally for the durations between disasters and recovery time, three levels are possible (1=weeks; 2=months; 3=years) We have tried to explain this better by adding the following sentence:

[442] Before the start of the game, the moderator distributes the roles among the players, randomly selects a series of three hazard types (tropical cyclone, drought, flood), intensity (low, medium, and high) and time between disasters (weeks, months, years), and decides which parts of the board are impacted by the disasters. These selections are not shared a priori with the players and will define the storyline of the game. **Each hazard type, level, and timescale have an equal chance of being selected (uniform distribution). Theoretically, it is possible that three of the same hazard types, levels, and timescales are selected. However, the chance of having three consecutive high intensity hazards of the same type for a storyline is 0.13%, but the number of cells impacted (5 to 15) and the location can still be different.**

- How do the players know what actions they can take per round? Do you provide a ‘cheat sheet’ or do you leave this to their imagination?

All players have a role card (an example of which is shown in Figure 1C) that describes in detail what DRR measures they can implement, what the conditions are (e.g., costs, restrictions in terms of location), etc. Aside from DRR-measure specific restrictions, there are no differences between the rounds in DRR measures that can be implemented. The only difference is that DRR measures implemented in round 1 are marked with a 1 and similar for rounds 2 and 3. This helps the Game Master identify implemented measures and see whether the construction of a measure is finished at the time of a new round. We try not to steer the players too much in their DRM behaviour, so we refrained from providing more detailed guidance on the decision-making process. We have added a brief clarification of this to Sections 2.2 and 2.3.

[412] Each of the role cards has the same structure (Fig 1c). It explains to the player the characteristics of their role including some background information about their position in the team, their relationship with some of the other members of the team, **and the specific DRR measure(s) that their role can implement**. It also provides some detailed knowledge on these DRR measures, including information such as their costs, their advantages, limitations and potential synergies with other hazard types, the time it takes to implement them (discretized between weeks, months or years) and, depending on the role, information about particular hazards, demographic information, etc. Each DRR measure has a different symbol, and a numerical subscript is used to indicate the round during which these DRR measures can be implemented (Fig. 1c). **The types of measures that a role can implement do not change per round, but** some of the DRR measures cannot be built in particular cells of the map; for example, because they cannot be built together with another DRR measure or because they are invalid (e.g., a seawall can only be built in coastal cells), and some of the descriptions of DRR measures warn the player of potential (a)synergies (Fig. 1c). **For example, the agricultural representative can plant both normal and drought-resistant crops (Fig. 1c). They have the following information: the costs of the different crop types (drought-resistant crops are more expensive than regular crops), drought-resistant crops are more vulnerable to floods, to meet the country’s needs they need at least 15 cells of crops, and neither of these crops can be planted in densely populated cells or in cells where Nature Based Solutions (NBS) have been built. The president, the national housing and urban development director, and the citizen representative roles have information about the population per cell. Conflicts can arise when, for example, the citizen representative wants to implement NBS in the same cell where the agricultural representative wants to plant crops or if crops are in the downstream area from where the engineer wants to build a dam to decrease flood risk or upstream droughts.**

[446] The moderator introduces the players to the overall game set-up and leaves some time for the players **to read their role card to learn about their role and possible DRR measures that they can take during each round**. They then read the background story to give all players’ general information on the setting of the game.

- Can you provide some examples of the DRM actions that could be taken by players?

The reviewer raises a very good point; we have included some detailed examples of DRR measures that players can implement as well as their possible conflicts.

[422] **For example, the agricultural representative can plant both normal and drought-resistant crops (Fig. 1c). They have the following information: the costs of the different crop**

types (drought-resistant crops are more expensive than regular crops), drought-resistant crops are more vulnerable to floods, to meet the country's needs they need at least 15 cells of crops, and neither of these crops can be planted in densely populated cells or in cells where Nature Based Solutions (NBS) have been built. The president, the national housing and urban development director, and the citizen representative roles have information about the population per cell. Conflicts can arise when, for example, the citizen representative wants to implement NBS in the same cell where the agricultural representative wants to plant crops or if crops are in the downstream area from where the engineer wants to build a dam to decrease flood risk or upstream droughts.

- You include the game boards from the game sessions, but it's hard to understand the temporal aspect of the story – can you demonstrate this with a timeline of actions perhaps?

We really like this suggestion and have included the following timeline in Section 3.2 Game play.

	Preparation	Round 1	Break 1	Round 2	Break 2	Round 3	Discussion
ACTIONS PLAYERS	<ul style="list-style-type: none"> Read rolecards Familiarise with actions 	<ul style="list-style-type: none"> President to lead the meeting Exchange thoughts on measures to be taken Implement measures on the board 		See round 1	See break 1	See round 1	
ACTIONS MODERATOR	<ul style="list-style-type: none"> Determine disaster Read storyline Set timer for round 1 		<ul style="list-style-type: none"> Determine new disaster Calculate new impacts based on measures 	See round 1	See break 1	See round 1	<ul style="list-style-type: none"> Facilitate discussion General reflections Reflections on debriefings between rounds and comparison of summary tables
TIME	10 min	20 min Days	10 min 5 min Weeks, months, years	15 + 5 min Days	10 min 5 min Weeks, months, years	15 + 5 min Days	30 min

 Disaster (flood, drought, or tropical cyclone) occurs
  Actual playing time
  Temporal dimension

Specific Comments

- P2 Line 46 – “Several studies have demonstrated the successful use of these serious games in increasing risk awareness” – This is a little vague, can you expand on exactly what you mean by ‘successful’, and explain what success would look like for this game.

We thank the reviewer for their feedback and have rewritten the section as follows:

[47] In recent years, a large number of serious games relating to DRM have been developed. Solinka-Nowak et al. (2018) conducted a meta-analysis of DRM serious games and found that the majority focus on floods (27 out of 45 reviewed games), earthquakes (10 out of 45), and droughts (7 out of 45), while storms (including cyclones, hurricanes, etc) are rarely the main hazards in a game (Solinka-Nowak et al., 2018). It should be noted that these games encompass a wide variety of game formats, including single- or multi-player video games, single- or multi-player tabletop games, and role-playing games (RPGs), making it difficult to compare them. Djaouti et al. (2011) created a serious games evaluation system: the G/P/S model (gameplay, purpose, and scope). This model can be used to assess a serious game's main objective (i.e., training, message broadcasting, and knowledge exchange). Several studies have demonstrated the use of serious games in increasing risk awareness (e.g.,

Cremers et al., 2015; Mani et al., 2016; Mossoux et al., 2016; Pereira et al., 2014; Rumore et al., 2016; Solinska-Nowak et al., 2018; Taillandier & Adam, 2018). Rumore et al. (2016) quantified the effectiveness of serious games, and Role Playing Games (RPGs) in particular, in increasing risk awareness, where risk awareness includes risk literacy, an enhanced collaborative capacity to address risk, and social learning.

[92] To address the aforementioned gaps in current disaster risk serious games, we developed the *Breaking the Silos* game. *Breaking the Silos* is a multiplayer RPG that aims to raise understanding of the complexities of multi-hazard risk and asynergies of DRR measures among different DRM stakeholders.

- Citations L 47: hate to be one of those reviewers who push their own work, but Mani, Cole and Stewart, 2016 (using video games for volcanic hazard education and communication with at-risk communities) is also highly applicable here.

We sincerely apologise for missing this paper earlier and we included it in the revised version.

- Minor point of preference on language – Game Master would be more appropriate than moderator, perhaps.

We liked this suggestion; however, we also agree with a comment made by reviewer 2 stating that while Game Master would be correct terminology, “moderator” may be better understood by a wider audience. Therefore, we have added the following to the new Section 2.3 on the role of the moderator:

[199] The game requires a moderator, who can be thought of as a game master.

- L210 – you mention that players are asked to reflect on the game – are their reflections directed by the moderator, or how does this happen? This is important because it is a step in the experiential learning cycle, so I wonder about the scaffolding to support this?

First, the players are asked to share their general thoughts on the game and then the game master asks the players to reflect on their decision-making process, and finally they are asked to compare the different rounds and discuss how their behaviour changed between rounds. Next, the survey is used as a more structured means of obtaining reflections that can be compared between different games.

We did not clearly explain this in the manuscript and have therefore included a new section on the reflections.

[212] **2.4 Learning through debriefing**

While often lacking, a debriefing element in serious games is of utmost importance to support the learning process (Crookall, 2010; Kolb et al., 2014). It is even argued that real learning comes not from playing serious games but from the debriefing element (Crookall, 2010). Several more recent studies have addressed this by including feedback on actions within the game, so-called “learning by doing”, which can increase learning (Solinska-Nowak et al., 2018; Terti et al., 2019). Therefore, we decided to create three rounds, which demonstrate disaster and DRR interactions and allow players to change their approach to DRM in each round. Each round starts after a disaster and the team is asked to agree on the implementation of (a set of) DRR measures. We expect to see the teams responding to the particular hazard type that just caused a disaster rather than to also anticipate future risk of other hazards. Each round

begins after a new disaster, and with the moderator explaining the impacts of that disaster as well as highlighting the impacts of DRR measures that were taken in the previous round. This intermediate debriefing that follows each round and is led by the moderator, was designed as such to enable a reflection on the effects of the actions taken, to allow players to adjust their behaviour in subsequent rounds, and to experience the effects of changing one's behaviour. We expect that this influences the team's behaviour during the next round. Finally, it was decided to include a discussion that takes place at the end of the game to enable players to reflect on the effects of the debriefings. The discussion can be supported by looking at the overview tables that summarize the actions taken after each round and the effects of these actions on subsequent disaster impacts.

- L275 – already mentioned about COVID changes to gameplay, remove redundancy.

We agree and have removed the sentence.

- P11 L323-329 – you use the phrase 'highly realistic', 'highly enjoyed' – this is a bit sensationalist, yes the results are positive but I think you can remove the 'highly' from both (particularly as high relates to scores of both 4.5 and 3.8).

We have removed the mentions of "highly" here and in the conclusions.

- The way the players distributed their spending is of particular interest – does the variation in approach relate to the expertise of the game players?

The expertise of the players certainly plays a role, but it appeared that the group dynamics also play an important role. One of the main differences between the games played at UR versus the ones played with ETH Zurich is the difference in group dynamics: at UR most players did not know each other while at ETH Zurich, the players knew each other very well. Nonetheless, in all games, it appeared that during the second round, much fewer different types of DRR measures were implemented than during the first round. Moreover, the type of measures implemented also changed between rounds. We have expanded the analysis of the spending in Section 4.1, which now reads as follows:

[534] Even though the sequence of disasters and storyline were similar, the teams adopted different DRR strategies as shown in Fig. 3. This is also reflected in Fig. 4, which shows the different investments in DRR strategies between round 1 and 2, demonstrating the many possible choices and outcomes of the game, underscoring its high degrees of freedom. It appears that during all games, teams narrowed down the number of implemented DRR measures between round 1 and 2. While in round 1, on average teams implemented 7 different DRR measures, they implemented on average only 3 measures in round 2. Moreover, while in round 1 only one team invested in early warning systems for tropical cyclones and no one invested in dams, in the second round three teams invested in the tropical cyclone early warning system and two teams spent half of their investments on dams. We refer to supplementary material, Section IV, for a detailed overview of the coins spent per round, per DRR measure and per team.

- To what extent do the players own knowledge play into the scenarios developed? i.e. the ability of players to embody themselves in their characters and take relevant actions.

This is a very important point. All players work in a field related to risk, with the vast majority working on DRR measures related topics. The players of the UR game came from a more diverse background compared to the players of the ETH Zurich games, but in both cases, roles were assigned randomly. We expected the expertise to be very influential in players' decisions. However, our findings from our post-game survey show players working in DRR felt slightly better equipped to make DRR related decision in the game compared to players who work in risk but not in DRR. We therefore think a general background in DRR is of more importance to the game play than the specific expertise of a player. We have added a sentence to Section 4.2 to explain this better.

[666] During the post-game discussions, players' different expertise within the field of DRR appeared to be of less influence on the behaviour of the player.

This piece is very much needed and adds to the growing evidence base for the use of games within the DRR sphere. The authors have taken on a challenging route adopting an RPG, and have demonstrated an approach that has great potential. The paper and analysis is an important first step in understanding its efficacy and I look forward to following this work as it continues to evolve.

We thank the reviewer for the very kind words!

REVIEWER 2

This paper details the development of a roleplay-style game designed to help different stakeholders understand the complexities of disaster risk management (DRM) when considering multiple hazards. Although similar games exist with a DRM theme, Breaking the Silos is novel in the way it forces participants to understand and deal with the multiple hazards, appreciating that actions to mitigate one may have negative consequences with another. The paper is well written, well researched, a pleasure to read and will be of interest to the readers of Geoscience Communication.

We thank the reviewer for these very kind words.

I do not consider that the manuscript requires any major edits to be suitable for publication. However, there are some changes I suggest the authors should consider to improve it further.

1. The structure and focus of the manuscript could be altered to better sell its strengths. Nearly a third of the abstract covers the use of the game and participants' feedback yet the section covering this (Section 4) was disappointingly short. This is not a problem as the main strength of the paper is the design of the game and the comprehensive research that went into it. To reflect this, the manuscript could be restructured as a practise-based research* project where the game itself is the main result. In this form, Section 3 would come before the description of the final game design and include a little more discussion about how your design choices were informed by past research, and how they serve to meet the objectives. The abstract and conclusions should also be edited to make a bigger deal of how the game design was research-led.

We like the reviewer's suggestion and have therefore changed the structure of the manuscript accordingly. We have also made some adjustments in the abstract and the conclusions to reflect this.

2. To increase the significance and citation rate of the final manuscript, greater analysis of the game in action could be included. In particular, the involvement of social scientists specialising

in participatory methods and observations could provide fascinating insight into how participants play the game, the conversations they have, and how they make decisions in different circumstances.

This is a very useful suggestion, and we will take it into consideration when playing and analysing future games. However, at this point we are unable to include such analysis based on the past plays. Nevertheless, we add this suggestion to the discussion section:

[686] **When playing the game in the future, it can be considered to involve social scientists, specialised in participatory methods and behaviour, to further analyse the participants' game play and how behaviour changes over the course of the game through the learning-by-doing process.**

Whether the authors choose to act on either of the above or not, there are a number of minor edits and inclusions that should be made in a revised manuscript.

1. The moderator role seems absolutely vital to the running of the game and ought to be given its own dedicated heading somewhere. Under the heading the role of the moderator should be discussed, including how they are recruited and trained (there's some information on lines 278-280 but it is not clear). Are the moderators independent or one of the research team? What happens if the moderator isn't very good, can the game still succeed? In Lines 358-360, the authors mentions removing the moderator in future versions – will this be possible and what are the implications? As Lara Mani suggests, the role of moderator here would be traditionally called a Games Master in roleplay games – I agree this should be mentioned, but think moderator is more suitable for a wider audience.

We agree that the moderator deserves its own dedicated heading. We have therefore restructured the game design and set up section (previously section 2.2) and expanded the moderator section. Initially we had addresses Lara Mani's comment and changed all mentions of moderator to game master, but we agree that the term moderator is probably more suitable for a wider audience, and we have therefore re-adjusted it and explained that the moderator can be perceived as a game master. We also included a brief explanation of how the moderators were recruited and trained. The role of the moderator is important in the debriefing process and can therefore not be omitted (this was removed from the manuscript). The moderator guides lead the moderator step by step through the game versionation process which should minimize the negative impact of less good moderators. The section now reads as follows:

[198] **2.3 Role of the moderator**

The game requires a moderator, who can be thought of as a game master. The moderator is not actively participating in the game but narrates the storylines, runs the impact and DRR calculations in the background, and keeps track of time. All participants (and the moderator) communicate using an online meeting software such as Teams, Zoom or Skype. The game board and players cards are all shown on the MIRO platform, an online whiteboard for visual collaboration (see Fig. 1). Actions in the gameplay take place on this game board. Additionally, the moderator has access to a calculator tool for keeping track of the score. This is a series of spreadsheets containing the exposure, hazard, and vulnerability relationships between the disasters. As the moderator enters the DRR measures selected by the team, summary tables are automatically updated. These tables indicate whether the team met the round's objectives, the updated budget for the next round, and the difference in risk with the DRR measures selected to a situation in which no DRR measures would have been taken (in terms

of population, building and critical infrastructure impacted). During the games played at UR2020 and ETH Zurich, the moderators were part of the research team. While the moderators' guidelines and storylines (Supplement I) are self-explanatory, we do recommend moderators to first play the game themselves before moderating the game.

2. On Lines 210-211 the authors briefly mention that participants are encouraged to reflect on their experience. The authors should expand on this and consider their choices here as a vital part of their game design. I refer them to work by David Crookall (<https://journals.sagepub.com/doi/abs/10.1177/1046878110390784> from Page 907) discussing the importance of the debrief to achieving the learning outcomes of a game or simulation. This point is somewhat linked to the above as debrief occurs throughout the game too, facilitated by the moderator – again, highlighting the key importance of this role.

We fully agree with the reviewer and have added a new section and expanded another relevant section as follows:

[212] **2.4 Learning through debriefing**

While often lacking, a debriefing element in serious games is of utmost importance to support the learning process (Crookall, 2010; Kolb et al., 2014). It is even argued that real learning comes not from playing serious games but from the debriefing element (Crookall, 2010). Several more recent studies have addressed this by including feedback on actions within the game, so-called “learning by doing”, which can increase learning (Solinska-Nowak et al., 2018; Terti et al., 2019). Therefore, we decided to create three rounds, which demonstrate disaster and DRR interactions and allow players to change their approach to DRM in each round. Each round starts after a disaster and the team is asked to agree on the implementation of (a set of) DRR measures. We expect to see the teams responding to the particular hazard type that just caused a disaster rather than to also anticipate future risk of other hazards. Each round begins after a new disaster, and with the moderator explaining the impacts of that disaster as well as highlighting the impacts of DRR measures that were taken in the previous round. This intermediate debriefing that follows each round and is led by the moderator, was designed as such to enable a reflection on the effects of the actions taken, to allow players to adjust their behaviour in subsequent rounds, and to experience the effects of changing one's behaviour. We expect that this influences the team's behaviour during the next round. Finally, it was decided to include a discussion that takes place at the end of the game to enable players to reflect on the effects of the debriefings. The discussion can be supported by looking at the overview tables that summarize the actions taken after each round and the effects of these actions on subsequent disaster impacts.

[515] At the end, the participants are encouraged to reflect on the game. The moderator first asks the players to share their general thoughts on the game and the moderator asks the players to discuss their decision-making process, to reflect on their behaviour in different rounds, and the effects of the information provided during the debriefing.

Finally, there are some minor things that could be included but not essential. –

1. The authors discuss how the development and implementation of the game were affected by the Covid-19 pandemic however, a global pandemic has not been considered as one of the possible hazards. This seems a glaring gap and missed opportunity and it should at least be mentioned briefly why it was not included.

We agree that it would have been very timely to include COVID-19 or a pandemic in the games played at UR2020 and ETH. As described in the game development Section 2 (previously section 3.1), alternative game versions can be developed and included. As our own expertise is within the field of natural hazards, we would prefer to develop such an alternative game version by involving experts in the field of health-related diseases. This would also allow us to include, for example, a post-flood cholera outbreak, which would be a realistic consecutive disaster type to include. We therefore added the following:

[235] **2.5 Different game versions**

The primary game version reflects a fictional country, with a strong economic dependence on the agricultural sector, facing tropical cyclones, droughts, and floods. Different game versions can be developed, each with their own storyline, socio-economic and political characteristics, hazard types, available financial funds, DRR measures and DRM roles, to adapt the game to players' specific needs or hazard environments. For example, future game versions could include the occurrence of a pandemic, outbreak of a vector-borne disease, or natural-technological disasters. The role cards can also be adjusted such to mimic different government systems.

2. It would be interesting to find out more about how the prior experience of the players influences the decisions made and/or the processes through which decisions are made. Also, in reference to Lines 153-155, how does having different stakeholders missing influence this – could scenarios be set up to represent different governance systems in DRM between nations?

We fully agree that an in-depth analysis of the decision-making process, including the influence of participants' background, would be very valuable to include in a future analysis. Insights obtained from such an analysis could for example also be used to feed into an Agent Based Model.

The role cards allow for the development of game versions that mimic different government systems. We really like this suggestion and have included this (see previous answer).

Missing multiple players would hamper game play. If this is known in advance, the game version and roles can be adjusted. If this happens at the start of the game, it can be decided to add the tasks of the missing player(s) to the other players. I have added the following to better explain this:

[435] In case fewer than eight players are playing, this would mimic the concept that in real life, important stakeholders can also be absent from key meetings. **Alternatively, it can be decided to add the tasks of the absent player to other player(s).**

3. Line 159 – just a clarification here, can the three hazard types, levels, and timescales be the same? i.e., could three high intensity floods be randomly selected?

Theoretically this would indeed be possible. The chance of having three consecutive high intensity floods for a storyline is $(1/3)^6$ so 0.13% chance, but the number of cells impacted (5 to 15) and the location can still be different. We have explained this in the text as follows:

[444] These selections are not shared a priori with the players and will define the storyline of the game. **Each hazard type, level, and timescale have an equal chance of being selected (uniform distribution).** Theoretically, it is possible that three of the same hazard types, levels,

and timescales are selected. However, the chance of having three consecutive high intensity hazards of the same type for a storyline is 0.13%, but the number of cells impacted (5 to 15) and the location can still be different.

4. Lines 165-166 – change “DRR measures they want to implement” to “they want to implement DRR measures” so this reads correctly.

Many thanks for pointing this out, we have adjusted the sentence accordingly.

5. Lines 179-180 – During the participants’ 10 minute breaks are they allowed to discuss the game and their strategy? Does this become an unofficial briefing period?

During the online games, players were not explicitly asked to not discuss game strategy outside of the game environment as the online set-up did not allow participants to continue talking during the break, but this is indeed a valid point when playing the game in-person. I have therefore included the following:

[475] The end of the second timer signals the end of round 1 and a break of 10 minutes for participants to relax. **During the game, participants are asked not to discuss game play outside of the game (for example during breaks) to ensure that reflections on game play and on the debriefing take place within the game’s environment.**

6. Line 220 – I’m not sure “two different game versions” is correct. It would be better to say that the game can be adapted to suit different audiences or for bespoke scenarios.

We agree and have adjusted the sentence as follows and changed all other instances of “game version” to “game version”:

[237] **Different game versions can be developed**, each with their own storyline, socio-economic and political characteristics, hazard types, available financial funds, DRR measures and DRM roles, **to adapt the game to players’ specific needs or hazard environments.**

Thank you to the authors for this interesting and important manuscript. I look forward to seeing the revised version.

We would like to thank the reviewer for his thorough review of our manuscript and the very valuable feedback, which we believe has improved the quality of the manuscript.

Chris Skinner

*A note to the editors – I think the guidance for Research articles could be edited to explicitly state that practice-based research approaches are acceptable.