

Dear Editor and Reviewer,

Thank you again for taking your time to review the revised version of our manuscripts. We try our best to address input and comments from the reviewer point to point. Below is our response to those input. In this response, we use the blue text to indicate our changes from the original text.

Thank you

1	<p>You need to rethink your second objective (Assess if game facilitates knowledge transfer and sharing and supports negotiation and Coordination). The way the second objective is currently addressed is superficial and not transparent. In your response to us you write: ‘the objective of this manuscript is to document the development and adaptation process of the H2Ours game(...)’. If you want to save your data for another manuscript, then drop or revise your second objective.</p> <p>Answer:</p> <p>Thank you for your concern about the objective of this manuscript. Because your first and second comments are very connected, please allow us to response your first comment together with the second comment.</p>
2	<p>I disagree conceptually with ‘In order to achieve the second objective of this paper, to clarify how games can facilitate the sharing and transfer of knowledge, we want to ensure relevance of game situations to the reality (salience), the acceptance that game conditions resemble their conditions (legitimacy), and supported by appropriated data and methods (credibility)’. Salience, legitimacy, and credibility of games are not valid measurements of knowledge transfer and sharing. You could fix many problems in your paper if you formulated the second objective as ‘Assess the salience, legitimacy, and credibility of the H2Ours game.’? You are interested in ensuring salience, legitimacy, and credibility because you assume these are preconditions to allow the game to transfer and share knowledge, and to support negotiation and coordination. You can explain this. But at least the data you present to not allow to assess game outcomes.</p> <p>Answer:</p> <p>Thank you for providing us with very strong and reasonable reasons regarding the second objective and the use of credibility, salience and legitimacy in this study. If we referred to Cash 2002, 2020 and Belcher 2016, the use of credibility, salience and legitimacy in research is to assess the sustainability, practicality and transferability of the research to actions. You have captured our idea in using these criteria in the game is to assess whether the game simulation is useful for the players, especially to lead to the objectives of the game. Now we agree that, to assess whether the game fulfills its objective as a tool that facilitates knowledge transfer and coordination, we need to provide the evidences how the game change their, which we cannot include here because it makes this paper very complex. So we decided to follow your suggestion to change the second objective to assess the quality if the game in term of credibility, salience and legitimacy criteria. Due to the revision of the second objective of this manuscript, we revised several sections related to game evaluation.</p> <p>Introduction line 86:</p>

“Therefore, the objectives of this study are to develop a serious game that is adaptable to different socio-hydrological contexts and issues, and to evaluate the quality of the game in terms of credibility, salience and legitimacy. To achieve our objectives we developed a generic game with two adaptations to two different locations in Indonesia differing largely in hydrological characteristics. First, we developed the H₂Ours game based on the socio-hydrological characteristics of the Rejoso watershed in East Java. Then, we modified the H₂Ours game according to the conditions of the Pawan-Kepulu peatland, West Kalimantan. The qualities of the game were assessed based on several criteria representing credibility, salience and legitimacy which were included in the game development process and post-game assessment. We organized the paper by presenting as method the stages of how we prepared, designed, tested, implemented and evaluated the H₂Ours games. The game itself is the primary ‘result’, illustrated by the game dynamics during test settings and early applications with local stakeholders. Feedback by game participants is presented as an evaluation of the current games. We close by discussing the simplification process from reality to game, effectiveness of the game to achieve the goals set, and the lessons learned”

We simplified the whole section of 2.5 about the explanation of game evaluation to adjust the terminology of credibility, salience and legitimacy:

“The aim of the evaluation stage is to assess the game session process and the quality of the game as the basis for the game's performance to fulfil its objectives.. The game session process was evaluated based on game performances criteria in the form of rules that can be understood, fun and playability over time. While the quality of the game is assessed based on the scientific logic and reliable knowledge used to build the game (credibility), its relevance to the societal issues (salience) and the acceptance by the game participants (legitimacy) (Cash et al., 2002; van Voorn et al., 2016). For the effectiveness of the assessment, we followed input-output assessment process, which evaluated the input used in the game during development process and the output after the game session (Bedwell et al., 2012). We followed the latter approach and carried out the evaluation based on several criteria that refer to credibility, salience, and legitimacy (Table C1 in Appendix C), using some criteria developed by Belcher et al. (2016)

Because Belcher's long list of criteria (Belcher et al., 2016) originally was used to assess the quality of research, for this study we chose several criteria that were relevant to game quality. Each of these criteria were measured during the game design process and after the game implementation. We measured these criteria by how it was associated with the condition and diagnosis of the study area (Section 2.1 and 2.2) game development process (Section 2.3). Please see Table C1 to see the parameters and sections associated with each criteria. A rapid evaluations were conducted after the game session to assess the process and the quality of game session. We converted those game performace criteria and creadibility, salience and legitimay criteria into Likert used questions and asked all game participants to fill in the survey. In the Likert survey, we used five-point scales (strongly disagree, disagree, neutral, agree, and strongly agree) on six statements to ask participants about their feeling during the game, their understanding of the rules of the game, the length of the game simulation, new knowledge that they got from the game, and implementation the game to their reality”

Section 4.2 Discussion we add a paragraph line 495 to explain that in this manuscript we only assessed the quality of the game through credibility, salience and legitimacy as the basis for evaluating the game to meet its objective.

	<p>“We limit the evaluation in this study only to the quality of the game as a product. As a serious game, the H₂Ours carries certain goals that it wants to fulfil (Rodela et al., 2019), namely as a tool that can facilitate the transfer and sharing of knowledge from its players to support the coordination and negotiation process (Section 3.2.1). Evaluating the game in fulfilling its objectives is more complicated than evaluating the game session process. Ideally, the evaluation of the game in achieving its objective can be evaluated after several simulations at various levels of simulation, and should be conducted before, during and after the game sessions (Oprins et al., 2015). The evaluation of the game to meet the objective will be carried out in the next manuscript by providing evidence of changes in participant’s perceptions”.</p>
2	<p>If section 3.3 provides only indicative results, you still need to explain their data basis in the methods section – even if it is a subjective perception of game designers and facilitators.</p> <p>Answer: The data used in Section 3.3 consist of random-walks which are part of the solution space and the results of the game sessions. We have not explained the results that come from the game session, therefore we added a paragraph in Section 2.4 (line 247) to provide explanation regarding the data used in Section 3.3</p> <p>“The game explores the trade-off space between economic and environmental outcomes, with the responses from players during the debriefing adding further insights. The economic and environmental outcomes was calculated based on the average economic and environmental conditions as a result of decision making regarding land use combinations during a game simulation over 10 rounds. We present these results together with the results of the solution space analysis to show the position of players' decisions compared to random decision-making. During the debriefing, we asked participants several questions such as whether they enjoyed the game, what knowledge they gained from the game, how they responded to government regulations of the type included in the game, how they felt seeing other group decisions and (for study case Pawan-Kepulu peatland) their strategies as a member of multi-stakeholder forum”</p>
3	<p>Line 255ff: Please explain in the methods section how measurements during game design and thereafter were done. I assume that these measurements are different from the rapid evaluations.</p> <p>Answer: Thank you for this input. We clarified how we follow these criteria in the game design process by explaining in line 267</p> <p>“Because Belcher's long list of criteria (Belcher et al., 2016) originally was used to assess the quality of research, for this study we chose several criteria that were relevant to game quality. Each of these criteria were measured during the game design process and after the game implementation. We measured these criteria by how it was associated with the condition and diagnosis of the study area (Section 2.1 and 2.2) and game development process (Section 2.3). Please see Table C1 to see the parameters and sections associated with each criteria. A rapid evaluations were conducted after the game session to assess the game session process and the game in achieving its objective. We converted those game performance criteria and credibility, salience and legitimacy criteria into Likert used questions and asked all game participants to fill in the survey. In the Likert survey, we used five-point scales (strongly disagree, disagree,</p>

	<p>neutral, agree, and strongly agree) on six statements to ask participants about their feeling during the game, their understanding of the rules of the game, the length of the game simulation, new knowledge that they got from the game, and implementation the game to their reality”.</p>
4	<p>I did not find a revision in the text related to the clarification that a group represented one role and that a leader was asked to be chosen. This should be included in the paper as it is important to understand the game. Games with group decisions have very different dynamics than games with individual decisions.</p> <p>Answer: Thank you for your input. We clarified the Roles in Section 2.3.2:</p> <p>“According to the ARDI framework (Sect. 2.2), we defined the roles based on the main stakeholders involved in water management in each study area. <i>Most of the players were asked to be a villager, representing the largest stakeholder group, but others had specific roles as agents trying to influence villager decisions.</i> Related to these roles, we designed goals that players must achieve during each simulation based on discussions and interviews with the related stakeholders according to their actual goal. <i>Before the game started, we asked each group to choose a leader to facilitate discussion within the internal team and represent the group in communicating with other groups.</i>”</p>
5	<p>For me as an economist the process of generating the solution space based on 3, 10, 30, 100, 300, and 1000 games with random choice is still not clear. I understand that random choices mean random parameter setting. Using the word ‘choice’ gives the impression that somebody (human) takes a decision. But why 3, 10, 30, 100, 300, and 1000 games? Maybe explain this whole process clearly in an appendix.</p> <p>Answer: Thank you for your input. We decided to replace ‘choice’ with ‘random-walk’ to provide more general meaning because the random parameters refer to the random climate conditions and random land use combinations. Why we chose 3, 10, 30, 100, 300 and 1000, we added explanation in Section 3.2.4.</p> <p>We revised Section 2.3.4 regarding game solution space analysis as follows: “The purpose of game solution space is to define the envelope of possible outcomes within the rules of the game, considering all possible choices made by players in the game (Speelman et al., 2014). In a <i>random-walk</i> any sequence of steps has equal probability, blind to where it may lead. The solution space of the H₂Ours game was explored based on the average of economic and environmental outcomes obtained with a random-generator deciding choices for every step. We mapped the estimated solution space after 3, 10, 30, 100, 300 and 1000 <i>random-walk</i> iterations to obtain a reference for the trajectories observed in a limited number of actual, real-player games. The random-walk conditions were generated in R, then simulated using an Excel spreadsheet representation of the H₂Ours game and its economic and environmental performance indicators. The 1000 <i>random-walk</i> data set was used to assess the probability density of outcomes within the solution space. The economic and environmental performance indicators of actual game implementation refer to player's land use decisions from four</p>

	<p>different game session in Rejoso Watershed which are calculated using the same Excel spreadsheet”</p> <p>And Section 3.2.4, result of solution space analysis: “From the comparison results between 3, 10, 30, 100, 300 and 1000 random-walk iterations, we found that the shape and distribution of economic and environmental outcomes began to stabilize at 300 iterations. Therefore we used 300 games with random conditions as the basis for the solution space of this research. As reference for the player-based game runs, in 300 game runs with a random decision making process, the groundwater distribution varied depending on the location, while the distribution of surface water in the upstream and midstream is almost the same, and in the downstream is wider (Fig. 4A and Fig. 4B). Upstream and midstream had almost the same frequency distribution of surface water flows while runoff from the upstream and midstream areas was dominated by wet years, which then may potentially cause flooding downstream in the same year. Contributions of groundwater from upstream and midstream also responded to wet years, but groundwater utilization by downstream occurs mostly during the dry years. Therefore, the frequency distribution of groundwater contributions were wider than those for surface water”.</p>
6	<p>Line 77: I did not mean to ask you to cite our paper. I was just using it as an example. If you do not make the problematic statement, you do not need to include any citation.</p> <p>Answer: Surface No, we think we have made an inaccurate interpretation regarding the lack of explicit combinations of models and games. Therefore, in the revised version we revised it by providing an explanation of how the model and game can complement each other and including your publication as part of the example.</p>
7	<p>Line 203ff: There is a large body of literature on external validity of experiments and the discrepancy between stated and revealed behavior. The evidence is too ambiguous to support a general assumption that players in a game take the decisions they take in real life. They may experiment, show social desirable behavior, or want to enact a strategy for various reasons. You should share results if you find a correlation between land use decisions in the game and actual land use. Otherwise, I recommend avoiding making such claims also because it is not at all that important for learning games. Knowledge transfer and sharing, negotiation and coordination can all happen even if players do not reveal their real-life behavior in the game. Important is that the choices in the game reflect critical choices people have in real life.</p> <p>Answer:</p> <p>We designed game properties in such a way that it resemble reality in so that participants can imagine and treat the landscape in the H2Ours game as their landscape. I agree that decisions during the game do not all reflect their decisions in real life. But at least they can correlate the impact of their decisions with the impact they experienced during the game simulation with the impact they might experience in real conditions with similar decision. We revised line “Because we expected the decisions made by the participants during the game simulation represented their actual decisions, we developed the game as close to the reality as possible” to prevent ambiguity into (Section 2.3.5, line 210):</p>

	<p>The purpose of game development is to bring the game design into a real form that players can play or touch such as a game board, various required tokens, and other attributes that support the simulation of the game. We developed the game to be close to the perceived reality, so that players can relate their decisions with the consequences obtained during the game session with the impacts that they have experienced or will experience with the similar decisions. The game board, the game’s land-use options, and water simulation miniature are the key elements of recognition for players. Therefore, we adapted these elements to the conditions of each study area.</p>
8	<p>Given the structure of your game, I would still appreciate a more self-critical reflection on possible unintended undesirable learning outcomes. I understand the good intentions of the team. Discussing this risk may be valuable for game designers in future. How could the risk be mitigated?</p> <p>Answer:</p> <p>Thank you for the suggestion. Based on game simulation in Rejoso and Pawan-Kepulu peatland, we realized that lack of social value that encourages continuity of an activity. We include the following paragraph in Section 4.2 (line 532) as a critical lesson learned:</p> <p>“The H₂Ours game clearly shows the trade-off between the economy and the environment by calculating economic and environmental performance indicators in each round after the players change the land use combination and water management. As a result, the relational value between humans and human with nature (e.g. trees and water being inherited from their predecessors and will be a legacy for their descendants, the use of certain woods in religious rituals) sometime becomes blurred. A very clear trade-off between the economic and environmental conditions have led players to make decisions based solely on economic value. Therefore, the cost-benefit calculation of conservation activities needs to be done carefully in this game or include social values as part of the scenario in the game.”</p>
9	<p>I would still appreciate some more concrete thoughts on how this game could be used beyond this project. Who could facilitate the game? Who could be target groups? I understand your higher-level ambitions but how can you imagine a concrete use of the game?</p> <p>Answer:</p> <p>Thank you for your concern to this issue. Further evidence on the potential adaptation of the H₂Ours game to other contexts was recently obtained (Khasanah, Pers. Comm. March 2024) by a World Agroforestry (ICRAF) project in East Nusa Tenggara (NTT) in Indonesia with pastoral land use and shallow groundwater conditions. The primary hydrological issues in that area are groundwater scarcity and drought. In the adaptation process of the H₂Ours game, ICRAF adjusted the ARDi-DPSIR table from this study according to the local issues and recommended water-landscape management in NTT. They played the H₂Ours game with the local communities and multi-stakeholder forum to simulate the land use management in order to conserve springs to secure their water. We did not include this information in the manuscript because we need a consent from ICRAF to include this information in the publications.</p>