

# Development of GreenDealz: A public engagement toolkit addressing critical raw materials and the EU Green Deal at informal education settings

## Responses to Editor and Reviewer Comments

Many thanks to the editor and reviewers for providing highly constructive feedback on our manuscript that we have no doubt will help improve the writing and presentation. We have copied, pasted and arranged all editor and reviewer comments in the tables below. Each comment has a response and indication of changes made in the revised manuscript. We include a comment and response # column for additional ease of reference within the tables. Cited manuscript line numbers refer to the revised manuscript PDF **without tracked changes**.

The comments are arranged into Editor, RC1, RC2 and RC3 across broad manuscript sections.

We hope that the editor and referees find our responses and changes to be satisfactory and subsequently deem the revised manuscript fit to be published.

## General comments/overall assessments

#	Editor	Responses and changes made
1	<p>The topic of CRM and the design of the engagement tool, GreenDealz, are great contributions to the GC community and science engagement practitioners in general. The final GreenDealz kit seems very well made. I congratulate the authors on their work.</p> <p>However, all three reviewers (RC1, RC2, RC3) raised concerns regarding the clarity and completeness of the methodology description and the validity of statistical interpretations. This needs to be addressed comprehensively. There is also a need to improve the readability of the manuscript, as mentioned by all reviewers. I concur with all these conclusions.</p> <p>My overall recommendation to the authors is to make the methodology section more robust, particularly describing the decision-making process between iterations for transparency and to help conceptualize the study's aim to design a toolkit and embed an evaluation method. I also suggest considering including the feedback from the audience as suggested by Reviewer 2 (RC2) and Reviewer 3 (RC3). I further recommend presenting the results in a more qualitative nature, given the limitations and representativeness of the sample pools. This is in line with RC3 comment #2.</p>	<p>We thank the editor for their constructive comments on our manuscript.</p> <p>The specific comments and responses outlined in the table sections below highlight the exact changes made.</p> <p>The editions in response to this assessment have aimed to:</p> <ol style="list-style-type: none"> <li>1) Improve readability</li> <li>2) Tighten the theoretical background and study aim</li> <li>3) Enhance clarity of the development process and recruitment</li> <li>4) Re-frame the study as an exploratory pilot focusing on key developmental results only</li> <li>5) Remove the student study due to low validity as suggested, focusing on the development in festival spaces</li> <li>6) Make more explicit the statistical tests used and removed others where unnecessary upon review</li> <li>7) Expand on qualitative items where possible</li> <li>8) Provide adequate descriptions of limitations, strengths and subsequent recommendations</li> </ol> <p>We hope the changes as outlined below address the editor's, and thus, the referee comments and concerns.</p>
	<b>RC1</b>	<b>Responses and changes made</b>

2	<p>It is of utmost importance to our society that the public gains an understanding and knowledge of the role of CRM and its connection to achieving the goals of the Green Deal to be achieved in a variety of ways. Any activities to gain public awareness and to improve the knowledge of the society on key challenges of the society in particular are desperately needed in most societies. The paper picks this need up and provides some simple solutions that can general be easily adopted elsewhere.</p> <p>The title “Development of GreenDealz: A public engagement toolkit addressing critical raw materials and the EU Green Deal at informal education settings “ indicates very clearly the content of the paper. The concept of addressing festival visitors is relatively new and has the potential to reach a very diverse and inexperienced group of people. This is particularly true given that a wide variety of different types of festivals were included in the study.</p> <p>The paper is written in a complex style that is rather atypical for scientific papers. The long and convoluted sentences make it even more difficult to understand it. This distracts from the actual message.</p> <p>The paper seems to address the most of the aspects that need to be considered. A clear description of the method used, the statistical approach and the concept for achieving the overall objective is only available in broad terms.</p> <p>The paper aims to stimulate thinking about CRM in relation to the EU Green Deal. Therefor it focuses on the tools to gain this stimulation rather than the description on the scientific content used for the stimulation, which could be reviewed by the RC. The study seems to be designed rather as a subordinated experiment in a context of a scientific project than as a stand-alone study. It provides a snap-shot of given festival situations, not tailor-made solutions as the many researchers in the topic suggest. The authors fail to explain the selection criteria for the festivals to the reader. A more detailed and structured elaboration would improve the readability of the text. Thus, it could be better structured.</p> <p>The authors discuss their findings in the context of existing research and made an effort to consider transdisciplinary findings as well. The</p>	<p>We thank RC1 for their constructive commentary.</p> <p>We hope our editions to the manuscript’s structure, readability, theoretical framing, method development and result reporting as outlined in the specific sections below have addressed these overall concerns.</p>
---	---	---

	<p>data compiled to obtain the results are not always provided in a comprehensible manner. The methods used are not clearly described in terms of inclusion/exclusion criteria, procedures or tests used, measurements utilized, primary and secondary outcomes or independent and dependent variables, statistical analysis utilized.</p> <p>The article does not test any hypotheses, but rather describes the development of a tool for knowledge transfer and discusses random examples from experience. In this context, it is an interesting opinion piece that encourages further testing and experimentation elsewhere. However, there is no solid evidence that the proposed GreenDealz toolbox actually enhanced the publics' engagement in comparison with other approaches or at all. Yet, the article is a useful source of inspiration for those that view communication with lay audiences as part of their responsibilities. Some revisions required prior publication.</p>	
	<b>RC2</b>	<b>Responses and changes made</b>
3	<p>This paper is an exceptionally novel approach to understanding public knowledge and perceptions of critical minerals. Overall, the paper is well-written with thoughtful consideration given to the evaluative component.</p> <p>Does the paper address relevant scientific questions within the scope of GC? The study's objective is within the scope of GC.</p> <p>Does the paper present novel concepts, ideas, tools, or data? The ideas and data are novel in a geoscience context. The concepts and tools are well-established in social science literature.</p> <p>Are the scientific methods and assumptions valid and clearly outlined? The methods used are valid, but some further clarification is required. A thematic codebook is needed, along with more in-depth descriptions of statistical tests used and the rationale for their use.</p> <p>Are the results sufficient to support the interpretations and conclusions? For the most part results are sufficient. However, there are some issues with sample size. It would be useful to have a statistics expert look at the phrasing used in the discussion/conclusions to make sure the claims are valid.</p>	<p>We thank RC2 for their commentary and helpful suggestions.</p> <p>We hope our editions to the manuscript's structure, readability, theoretical framing, method development and result reporting as outlined in the specific sections below have addressed these overall concerns.</p>

	<p>Do the authors give proper credit to related work and clearly indicate their own new/original contribution? Yes</p> <p>Does the title clearly reflect the contents of the paper? Yes</p> <p>Does the abstract provide a concise and complete summary? Yes</p> <p>Is the overall presentation well structured and clear? Yes, just a few changes will help with flow (specifically, Section 2)</p> <p>Is the language fluent and precise? Yes, language is fluent and very well written.</p> <p>Are the number and quality of references appropriate? Yes</p> <p>*Please note that I've qualified [in-text] comments with "Consideration" and "Necessity". Consideration is something for you to consider (not a necessary change), while necessity likely requires a change.</p> <p>*I've reviewed Reviewer 1's comments as well to ensure I'm not repeating.</p>	
	<b>RC3</b>	<b>Responses and changes made</b>
4	<p>The paper presents a public engagement toolkit aimed at raising public awareness of importance of critical raw materials (CRMs) and, specifically, their role and key contribution to delivery of the EU Green Deal. The paper presents how the toolkit, called GreenDealz, was deployed within a wide range of different types of informal education settings principally festivals. The CRM topic is clearly highly relevant and of clear societal importance in terms of a net zero transition and the authors rightly argue that focus on CRMs as an overlooked component of the energy transition as is the public awareness of the key elements of the issues. As such the work is timely. The methods employed uses festivals as an outreach mechanism/location, although other examples exist of engaging audiences in such locales, this has some rareness in the context of the topic areas and has thus some innovative elements. Overall, the authors' commitment to broadening participation in sustainability discussions with diverse audiences is commendable.</p> <p>However, while the study's intent is valuable, the paper in its current form requires very significant</p>	<p>We thank RC3 for their helpful commentary.</p> <p>We hope our editions to the manuscript's structure, readability, theoretical framing, method development and result reporting as outlined in the specific sections below have addressed these overall concerns.</p>

	<p>revision to ensure that methodological transparency and conceptual clarity is communicated as well as it could be. Overall, the contribution is primarily conceptual and experiential rather than empirical and as such is very descriptive in nature. Whilst this would be fine in my view as there is value in sharing the journey taken in undertaking the work, I do think that a stronger theoretical grounding and clearer structure would help the paper better serve as an important reference for others to make use of. At present the novel elements are rather hidden in a descriptive narrative of what was done rather than any insights gleaned from the work. There are some creative elements in the approaches that could be teased out more fully and benefit the paper as a result.</p> <p>The GreenDealz concept is original and of clear societal relevance, but the paper requires substantial revision in my view. Strengthening the theoretical framing, clarifying the details of the methodology, improving data presentation and the way it is interpreted and discussed, and tightening the text would significantly enhance the paper's potential impact. The work has good potential in showcasing a case study and pilot for informal sustainability communication in CRM once these improvements are made. I hope the authors find my thoughts and review helpful in the refinement of the paper.</p>	
--	---	--

### Technical comments/corrections

#	Editor	Responses and changes made
5	Readability: The text is at times hard to follow. Please remove redundancies to enhance the clarity of the text.	The text has been carefully edited to enhance clarity and readability. We hope these changes are satisfactory.
6	Technical comment. Order of in-text citations: if they are placed in order of relevance, please leave them as they are. Otherwise, please arrange them chronologically.	All in text citations have been revised for order as suggested.
7	<p><b>RC1</b></p> <p>Some of the references are unusually cited in the text, in particular those that are taken from websites (e.g. lines 527-528; 546 and 559). The reference list contains errors, including incomplete and unusual quoting's (e.g. websites) as well as incorrect listed authors that seem not to be in line with common format requirements. If all important relevant references are included the list of cannot be judged by the RC. Thus, the cited</p>	<p><b>Responses and changes made</b></p> <p>All specific citations listed have now been amended (now lines 519-520, 558, 571). All other in text citations have been revised for formatting as per editorial requirements (comment &amp; response #6).</p> <p>Errors on the reference list (e.g., websites, reports) have been corrected as per the guidelines for Geoscience Communication.</p>

	ones are accurately quoted with in the limitations mentioned, e.g.:  Schweingruber and Fenichel (2010) instead of Fenichel and Schweingruber (2010) Gronlund, Norman E. Assessment of student achievement. Allyn & Bacon Publishing, Longwood Division, 160 Gould Street, Needham Heights, MA 02194-2310; tele, 1998.	
8	Some minor typos have been detected such as in line 28 and 32 “Parliament” and in line 831 <a href="https://periodic-table.com/dysprosium/">https://periodic-table.com/dysprosium/</a> .	Typos have now been addressed.
9	Figures and tables missing some descriptions such as N/A in tab 1 (inconsistent with n/a tab 3), Fig 2 label of age range (I assume).	Figures and tables (new and old) have been revised for consistency.
10	Some descriptions are sloppy such as “pure image” instead of image of “pure metal” or “chemical element” (e.g. Fig. A1)	Thank you. Descriptions have been revised. Please see comment & response #65 which addresses these concerns.
11	Some further comments are provided in the draft text.	Please see all specific comment and responses below.
	<b>RC2</b>	<b>Responses and changes made</b>
12	Make sure you are consistent with capitalizations, hyphenations/non-hyphenations throughout	Edits have been made accordingly.
13	See edits in attached document	Thank you, we have included all suggested edits detailed in the manuscript.
	<b>RC3</b>	<b>Responses and changes made</b>
14	4. The manuscript’s readability could be improved by reducing the sometimes long and complex sentences, so I recommend some careful editing for concision and clarity. Figures and tables need clearer labelling, and the reference list should be reviewed for consistency in author order and formatting.	Thank you. We have carefully edited text and captions to improve the manuscripts clarity. We hope you find the revisions acceptable.  Reference list has also been revised as outlined in comment & response #7.

## Specific comments (from EGU sphere report and in-text comments where provided)

### 1. Abstract & Introduction

#	RC1	Responses and changes made
15	Line 14: shouldn’t it be “a” festival? An information on the selected festivals would be good.	The study pertains to several festivals, as such we have kept it as “the” festival environment (line 14). However, we now make broad mention of festival types in the abstract, as suggested (lines 15-16).
16	Line 24: Shortfall in argumentation - one pathway is the green transition for energy supply and storage and GHG-emission poor industry. It requires low-carbon technologies. Those technologies require substantial amounts of CRM often used in battery technologies for various applications.	Thank you. We have now changed “electrification” to “battery technologies”, however sentence remains short to keep in simple terms (lines 24-25).

17	Line 26: More recently e.g. IEA 2025, Global Critical Minerals Outlook 2025	Thank you. This is now included in citation (lines 26-27).
18	Line 27: "are mined from earth Misleading wording – all minerals a first mined from earth (so far no extraterrestrial ones in place). Either you grow or you mine them.	Thank you. We are aware that minerals are first mined from the Earth or grown. This use of language was tailored to readers who may be non-specialists/non-geoscientists to make the connection between mining and the ground beneath our feet. However, this phrasing has now been removed (line 27).
19	Line 28: macroeconomic importance	Thank you. However, as above we wish to keep the language suitable for interdisciplinary audiences and as such will not be including this term (line 28).
20	Line 28: unclear CRMA?	This citation refers to the Critical Raw Materials Act (CRMA). This has now been made clearer in text (lines 27-29)
21	Line 28: AND low sustainability, Reference CRMA <a href="http://data.europa.eu/eli/reg/2024/1252/oj">http://data.europa.eu/eli/reg/2024/1252/oj</a>	Thank you. The reference to sustainability has now been included (lines 28-29).
22	Line 30: Should provide some arguments, e.g. not enough material on stock (urban mining), need for long durability of goods and hence not available for new products, for many CRM there is no sufficient recycling infrastructure (no efficient collection, sorting and recovery in place )..... See also Cristiu et al 2025, <a href="https://doi.org/10.1016/j.resconrec.2025.108250">https://doi.org/10.1016/j.resconrec.2025.108250</a> ; Månberger 2023, 10.2478/ie-2023-0016	Thank you. We appreciate your outlook on the specifics of CRM recycling. However, in the interest of brevity and accessibility, we have intentionally left expansive passages about CRM supply, demand and recycling detail out of the introduction. Instead, we have focused on the key points outlined in the Critical Raw Materials Act and how public engagement is needed.
23	Line 32: AND non of those listed materials at any stage of processing shall be sourced by more than 65% from any third country.	Thank you. The text has been revised to include this clause (lines 31-32).
24	Line 35: suggest to consider also the European perspective where the NIMBY plays an important role as well as the low knowledge on resource availability, geology and modern mining along side with the disconnection of the majority of Europeans to mining and physical production. See also e..g. Mavroudi et al 2025 <a href="https://doi.org/10.1016/j.exis.2025.101689">https://doi.org/10.1016/j.exis.2025.101689</a> “The majority of Europeans have lost trust in mining companies due to the negative legacy of their past and inconsistent engagement in environmental and social practices (Badera and Koco´n, 2015; Barakos and Mischo, 2021; Bice and Moffat, 2014; Fonseca, 2010).”Negrete et al 2024, <a href="https://doi.org/10.1016/j.resourpol.2024.105135">https://doi.org/10.1016/j.resourpol.2024.105135</a>	Thank you. The focus of this introduction is not on why social acceptance to mining is low, rather why this might feed into the lack of public dialogue about CRMs and the energy transition. Therefore, in the interest of brevity, we mention a key aspect of low social acceptance to mining, and hence engagement on raw materials – environmental degradation legacies.  While we accept the point about NIMBYism, we have refrained from including it within this argument because its use as an assessment of people’s concerns is contested in the literature (e.g., works by Patrick Devine-Wright:  <a href="https://onlinelibrary.wiley.com/doi/abs/10.1002/we.124">https://onlinelibrary.wiley.com/doi/abs/10.1002/we.124</a>  <a href="https://onlinelibrary.wiley.com/doi/abs/10.1002/casp.1004">https://onlinelibrary.wiley.com/doi/abs/10.1002/casp.1004</a> ).

25	Line 42: i order to achiev what - see statement in conclusion.	Thank you. This has been re-framed to make clearer our goal (lines 56-58).
26	Line 47: rather describes methods and approaches in my view than coming up with some evidence as suggested by the text.	This paper is cited as it describes what public engagement is. However, we have re-phrased to make this clearer (lines 48-50).
27	Line 47: Seems quotings Rowe & Frewer are mixed with Pidgeon & Fischhoff	These quotes are not mixed in the text. The former refers to a description of public engagement (as above) while the latter refers to the types of beneficial public engagement/communication when dealing with societally relevant topics (e.g., climate change etc). However, we have also re-phrased the sentence linked to the latter citation to make this clearer (lines 50-51).
28	Line 63: I would assume that festival participants are somewhat likeminded and rather open minded towards some aspects but not to all that need to be addressed. Results might be biased.	Thank you. Please see all comments and responses on representativeness/limitations below.
29	Line 69: Focus on learning methosed - suggest to provide reference and quoting above	Thank you. However, it is unclear what is meant by this comment which seems to refer to the citation of Jensen and Buckley (2014). This reference outlines a study on why people attend science festivals and what they value in engagement. Hence, we maintain its citation in the argument on lines 68-71.
30	Line 77: stages of what?	Thank you for flagging this. Pre and post stages refer to the pre and post evaluation stages linked to an engagement or activity. Phrasing has been changed for clarity (lines 76-77).
31	Line 101: a target group within the group of festival visitors?	Target group meaning festival visitors generally. This line has been removed in re-framing and re-writing our aims (section 1.3).
32	Line 107: quit unambitious with low statistical prove.	Apologies. We hope that our revisions and further framing of this research as a developmental process/pilot study addresses this concern.
	<b>RC2</b>	<b>Responses and changes made</b>
33	Line 30: check these references may be worth including a more recent reference	New reference now included (lines 33-34).
34	Line 43: really like this paragraph, I think it provides ample context and shows the magnitude of the issue and eliquitly introduces your project's impact	Thank you.
35	Line 47: make sure this reference works for both components of your definition (the knowledge exchange component and the public participaton component	Thank you. This sentence has been re-worded to better reflect the meaning of public engagement as per this reference (lines 48-50).
36	Line 59: Necessity: adding or changing to one of the more fundamental papers on the dialogue model (see Metcalfe thesis for in-depth review of fundamental scicomm research)	Thank you for flagging this. We had originally included Rodrigues et al., 2023 as we specifically speak about geoscience and dialogue, however we have now included an additional fundamental reference (Bucchi, 2008) as suggested (line 62).

37	Line 60: Consideration: Current phrasing makes me think you are saying that learning gained through informal educational experience is always unexpected, which is not the case. Audiences can go into an informal educational experience expecting to learn something, if you see what i mean.	Thank you for flagging. Sentence is now changed to better highlight that it is not “always” (line 65-66).
38	Line 62: Broad statement, needs references to back it up	Sentence is now removed as the point is already made (with citation) in the following sentence (lines 62-65).
39	Line 63: Necessity: in line with my last comment, again, incidental audiences are definitely one audience, but there are also audiences who likely intentionally seek out these experiences	To better reflect this point, incidental now changed to “informal” (line 67).
	<b>RC3</b>	<b>Responses and changes made</b>
40	1. At present the paper lacks a clear theoretical basis that effectively links informal engagement to learning or to the behavioural outcomes in the engagements undertaken. Perhaps including a defined theory of change or a learning framework (e.g., from informal science education or participatory communication) would strengthen the rationale for the work and the way it was executed. Such a framing would also help to specify which dimensions of the EU Green Deal the toolkit addresses and how well it does that. For example mapping against resource circularity, supply-chain ethics, or awareness of material criticality. At present this is all lumped into a rather broad and bland narrative.	<p>Apologies, we had hoped that our section 1.2 on ‘public engagement as a form of informal education’ had outlined the rationale for why we chose informal learning spaces to develop our toolkit. This text has now been revised to better highlight this theoretical basis (e.g., lines 60-68). We hope this addresses these concerns.</p> <p>Please note that the objective was not to create a toolkit that evaluated specific aspects of knowledge on CRMs and the EU Green Deal as RC3 suggests, but indeed to create a toolkit that facilitates engagement with this topic <i>broadly</i> in informal settings and thus evaluates this knowledge <i>broadly</i>. We have attempted to make this clearer in text (section 1.3).</p> <p>As detailed in the responses below, we hope our further framing of this study as a developmental pilot and further clarity on our process (e.g., Fig. 1, Fig. 2) has addressed these concerns.</p>

## 2. Development of the toolkit (methods & results)

#	Editor	Responses and changes made
41	<p>Structure: The reason to present methods and results together is clear and well founded. However, it is difficult to follow the iteration process the way it is currently written, particularly as the flow is interrupted by lengthy paragraphs of statistical analysis.</p> <p>I propose some specific suggestions below.</p>	<p>Thank you. We hope the editions as highlighted within each comment and response below address this concern.</p> <p>Please note that along with the below changes, we have attempted to make the presentation clearer by dividing this section into two new sections: ‘section 2. Process of toolkit</p>

		development’ which aims to provide an overview of the development structure and general methods and ‘Section 3. Toolkit development’ which outlines methods and results for each individual phase.
42	There could be more transparency and clarity in the description of the iteration phases, including the decision-making process.	Thank you. We hope the editions as highlighted within each comment and response below address this concern.
43	For example, it is not clear how the prompts were classified from easy to difficult, nor how the prompts were chosen. It would be relevant to make explicit how the scientific knowledge presented in GreenDealz was selected from the list of literature sources.	Thank you for flagging this downfall. Please now see new sections 3.4.1. (lines 218-233) and 3.5.1 (lines 296-322) which describe in better detail the process behind creating prompts and content. We also refer to the sources from which we gleaned this information in these sections. These sources are also first reported in the new development process section (see sub-section 2.2).
44	As another example, it is not clear why there was a change in questionnaire wording from the Dingle to the Cork festivals. Namely from, <ul style="list-style-type: none"> <li>• “A3: The EU is self-sufficient when it comes ...” (Dingle) to</li> <li>• “A3: I am surprised that the EU is not self-sufficient when it comes...” (Cork),</li> </ul> since this changed from a neutral statement to an opinion (I am surprised). I would assume this could affect the response in a survey? This is not discussed in the text.	Thank you. It has now been made clearer in text as to why this item (and all other items) were re-phrased (lines 244-249). However, upon reflection we do agree that the use of an opinion-based statement rather than a neutral statement could affect the result presented. Therefore, we now first mention the issue of phrasing differences during presentation of the result, where we have now also removed any statistical comparison (lines 254-255). We then more directly mention this discrepancy in the new sub-section 3.4.3 which covers limitations (lines 278-279).
45	Figure 1 introduces the workflow from scoping to designing phase, as well as characteristics of the design (first under Overview, Line: 110).  The information of the sample size (n) could be added to Figure 1 to make the methods more transparent.  I would further recommend the figure be split into two: (1) a Fig. 1 visualising the phases of the iterative workflow, the corresponding festivals (with festival details as suggested by RC1 and RC3) and participant numbers, and (2) a Table X with the results/conclusions from the iteration process, including the main changes that were made in each step from scoping phase to refinement phase.  This table could help summarise the results from the iteration process, which is not yet clear, as commented by the referees. This would conceptualize study’s aim to design a toolkit and embed an evaluation method.	Thank you. We have made editions accordingly.  Please now see Fig. 1, which has been revised to provide an overview of the iteration process. This figure now includes more detailed information about <ol style="list-style-type: none"> <li>1. the rationale for each phase</li> <li>2. Each of the festivals</li> <li>3. Descriptions of the study samples from each phase.</li> </ol> Please also see a new Fig. 2, which further documents the iteration process. This figure now highlights: <ol style="list-style-type: none"> <li>1. The evolution of the PE toolkit (activity &amp; evaluation methods), as before</li> <li>2. Key findings and observations from each phase</li> <li>3. A summary of the main changes between iterations</li> </ol>
46	Reviewer3, similarly, asks for “qualitative insights into audience engagement”, to which the	Thank you. We have now included an overview of the open feedback responses from

	<p>authors respond that it falls outside the scope of the paper.</p> <p>However, from the statement:</p> <ul style="list-style-type: none"> <li>• Line 521: “From the scoping to the refinement phase, development of the PE toolkit GreenDealz has been driven by audience feedback, responses and field observations.”</li> </ul> <p>and in alignment with the comments by RC2 and RC3, I also strongly suggest considering providing an overview if not an analysis of the audience feedback.</p> <p>This will also serve to make the method section more transparent, and additionally, it will offer the readers a learning process on engagement tool design (not as a final conclusion of effectiveness, but rather as a best practises method).</p>	<p>Dingle and Cork, namely by inclusion of Table 2 and Table 3 (while removing former Fig. 3) which includes anchor samples as suggested by RC2. These tables aim to provide clarity on the response themes we deduced, which we also highlight in text (lines 194-200; lines 261-266).</p> <p>We have made more explicit how we used this feedback to iterate (lines 217-227; lines 281-286). There is also greater reference to these responses in the discussion (lines 534-536; lines 573-577).</p> <p>We hope these changes are deemed beneficial to the manuscript and address all related referee and editor concerns.</p>
47	<p>Line 128-129: “For the data collection carried out at live events, sampling and recruitment was representative of and realistic to the free-flowing nature of the festival environment”</p> <p>a) That recruitment was representative is not evident as we are presented limited information on the different festivals. Do the festivals provide demographic information to which your samples could be compared? For example, most of the Dingle festival participants (88%) had a university degree (compared to ~50% as mentioned in the Irish Central Statistics Office report you refer to). How representative is this sample to (1) assess the change in engagement of an educational toolkit? and (b) overall representative of the “general publics” in festivals?</p>	<p>Thank you. Please now see the revised section 2.3 which re-phrases this claim to better communicate our process and sampling strategy/limitations. We intended the word “representative” to mean reflective of the recruitment capacity in such an environment, however we acknowledge this was misleading and we appreciate the reviewer and editor comments. We hope this section now provides a better foundation for understanding our samples across each phase and that we did not intend for them to be representative of the wider population or publics.</p> <p>Please also see the discussion which now reflects in more detail on the education within samples (lines 517-524).</p>
48	<p>In alignment with RC1 and RC3, I agree the study would benefit to expand the table section to include the festival characteristics, the festival setup and the summary of participants in each setup (as I mention in comment 3c). Indeed, as also RC2 notes, it would help clarify and improve the methods section to include the decision-making process at each iteration and a discussion on the limitations.</p>	<p>Thank you. As suggested, we have now split the original Fig. 1 into two figures to document this as above (see comment and response #45).</p>
49	<p>I agree with RC2 comment #4. I do not see the validity of the university student validation (n= 9, 89% female). The response given to RC2 explains the student and festival groups represent the “control” and “in-festival” settings, respectively; however, these two groups represent different segments. I question whether this small student sample is relevant, particularly in the argument in Lines 363-364, which observes an increase in subjective knowledge in the student sample,</p>	<p>Thank you. In line with all comments, the student study has now been removed from the manuscript. Result presentation, content and discussion throughout have been adjusted to account for this change.</p> <p>Please note, the control “match up” game was carried out at the Science Week events/festivals during the refinement phase, not at the National Ploughing Championships during the testing phase. If this was unclear in</p>

	<p>which is not seen in the festival group (National Ploughing Championships).</p> <p>On the other hand, the control “match up” game in National Ploughing Championships was very clear. Thus, I would suggest reconsidering if the student study is relevant in this context, as it may add unnecessary complexity to the text.</p>	<p>the text, we apologise and we hope that the edits have improved the clarity of the process.</p>
50	<p>I agree with Reviewer2 comment #5. I also question the extent to which the statistics can be used to make the conclusions in the text. I thank the authors for their reply to R2 indicates a statistician may be called in during the revision.</p>	<p>Thank you. Please note that some statistical tests are now removed (e.g., comparison of Dingle and Cork survey items; correlation analysis of shopping score vs. objective knowledge-EA during the refinement phase). Phrasing around the remaining statistical tests have been revised accordingly in text (e.g., lines 450-458; lines 497-502). We also first flag decisions around test choices in the analysis strategy as before (lines 158-162).</p> <p>As we have removed the student study, claims based on statistical tests used within this sample, to which much of this commentary referred, are no longer in the manuscript.</p>
	<b>RC1</b>	<b>Responses and changes made</b>
51	<p>Unfortunately, the number of participants in all experimental setups is very low, so it must be assumed that the experimental setup is not very attractive in comparison to other festival events. The statistical relevance of the test results is very low and corresponds more to a non-representative sample. The results derived from the experiments are therefore only of limited reliability. However, the idea and set-up are of relevance and worth to be developed further.</p>	<p>Thank you. The participants recruited during each phase represent attendees to our exhibits at each festival who consented to being part of our study. The sample sizes are also reflective of the time it took participants to complete our activities and evaluation and thus the sample sizes we deemed sufficient to gain enough information to iterate. We have attempted to explain this better in text (e.g., lines 133-142)</p> <p>We hope that the editions in text in section 2.3, our revised/new figures (Fig.1, Fig. 2), and editions in each phase (lines 178-179; lines 239-240; lines 345-346; line 444) now make the recruitment process clearer.</p> <p>We have also strived to make clearer that results are not representative of wider populations nor conclusive but simply document the iteration process and hint at the potential of our PE toolkit (e.g. lines 140-142; lines 146-150; lines 275-278; lines 512-517).</p>
52	<p>Although some of the information is provided in the text, it is difficult for the reader to compare the different points. The paper would benefit when key figures provided somewhere in the text would be summarised in a tabular overview of the various parameters of the festivals, such as number of participants, social structure and duration of the experiment, would have been helpful as well as description of the elements of</p>	<p>Thank you. Please now see a revised Fig. 1 and a new Fig. 2 which address these concerns regarding presentation and reporting of the iteration process, associated events and study sample sizes, as per similar suggestions from the editor and other reviewers (see comment and response #45).</p>

	the GreenDealz tool and when what part is suitable. Parts such as aspects of the objectives and the experimental setup are scattered throughout the text, while the framework is described inadequately.	We hope these figures allow better conceptualisation of our process and framework, in addition to the revised text.
53	The statistical results are based on various statistical methods without any explanation as to why these different methods had to be used. The stated methods should does not allow another researcher to precisely reproduce the study. The authors point to the low statistical significance but nevertheless draw conclusions about the transferability to other social groups and the applicability of the tool. Based on the available data, this generalisation does not appear to be valid. However, it can still serve as a basis for further studies and experiments. The description of the experimental setup allows replicated by another researcher. The low response to the experiment at all events indicates that the attractiveness of the experimental design still needs to be refined. As outlined by Ford (2019) points out that "tailored content and experiences can resonate with a person at an emotional and intellectual level, they will ultimately become devoted to the" issue.	Thank you. We have now made clearer within each section why certain statistical tests were used. See comment and response #50.  Regarding response rates please see comment and response #51.
54	Line 115: Dias et al 2020 not Carrara	Thank you. However, please note this point is incorrect – the official citation can be found on the front of the PDF report and lists Carrara as first author <a href="https://op.europa.eu/en/publication-detail/-/publication/19aae047-7f88-11ea-aea8-01aa75ed71a1/language-en">https://op.europa.eu/en/publication-detail/-/publication/19aae047-7f88-11ea-aea8-01aa75ed71a1/language-en</a>
55	Line 121: Statistical Package for the Social Sciences? mark it as landmark	Text changed accordingly (line 159).
56	Line 121: which ones?	‘Key assumptions’ as referred to in the analysis strategy (section 2.4) are now better highlighted within each section where statistical tests are used. See comment and response #50.
57	Line 128: how is a sample defined on this context?	Sample is defined as the number of people who engaged with us at each event that consented to being part of our study. Please now see the revised section 2.3 which aims to make this clearer in text. Fig. 1 also aims to provide further clarity on this.
58	Line 131: meaning numbers of participant on the survey?	“maximum possible sample size” refers to the maximum number of participants (as above) that can be recruited to engage in our study at any stage of the process, based on the estimated time taken to complete engagement and the time we spent at the event recruiting. See comment and response #51.

59	Line 136: Figures/Text is lacking any description on Likert or Dichotomous tests. If the readers background is in geology/ economic geology those terms are most likely unknown.	Thank you. Likert type scales are already first mentioned in Section 1.2.: ‘Public engagement as a form of informal education: impact and evaluation’. However, we now include a more detailed description in this section (lines 79-80).  Dichotomous tests are already described in section 3.5.1. as we explain why we used them in the National Ploughing Championships pre- and post- surveys (lines 330-336).
60	Line 151: Readers will not necessarily be familiar with the festivals. A brief assessment / comparison on the main difference of the festival participants and conditions would be useful. It would help to understand the findings. Could be done in an Table, loction map Information on size of the festival (visitors), duration, competing/supplementary activities, cluster of activities etc would help the readers. Seems there is a significant diffrence in the experimental settings that makes comparison and compiling conclusions/recommondaions tricky.	As above, please see the revised Fig. 1, which now provides an overview of the rationale for each iteration phase and describes the festival locations, set-ups, general attendee numbers, and overall duration. The study samples and recruitment set-up are now also described.  We hope this figure addresses your concerns.  We also include further details on audience types and set-up in each section (lines 169-176; lines 290-294; lines 391-396).
61	Line 156: Out of how many? Is this a big number for the festival - statistical wise its low. The figure is rather low for statistical purpose, but may be high for that event, Any further information on e.g. the size of the event, how much time was needed to get those participants interested would be good. - mall number of participants. Percentage of visitors? How attractive was the excercies?	See all prior comments and responses regarding recruitment and sample size.
62	Line 160 re “survey”: Remains unclear, if the participants where ask without any further interactions (status of their knowledge prior to any further gaming)	Thank you. It is unclear what RC1 means by this comment. However, we interpret it to ask whether Dingle participants engaged in anything other than the survey.  Dingle participants only completed a survey. This was used to scope out audience awareness of the topic, interest and engagement preferences as highlighted in Fig. 1 and section 3.4.1. We apologise if this was not clear. This is now made clearer in text (e.g., lines 178-179; lines 257-258) and with revised Fig. 1 and new Fig. 2.
63	Line 160 re “Likert-type”: Needs description or/and at least a reference, e.g. <a href="https://doi.org/10.3390/encyclopedia5010018">https://doi.org/10.3390/encyclopedia5010018</a> or <a href="https://doi.org/10.1016/B0-12-369398-5/00124-9">https://doi.org/10.1016/B0-12-369398-5/00124-9</a>	Thank you. See above comment and response #59 on elaboration of Likert type scales.
64	Line 163 re Cronbach’s alpha: how is that calculated?	Please see lines 158-159, which highlight that the reliability of all scales was conventionally analysed by Cronbach’s alpha (Cronbach, 1951).

65	<p>Line 185: native metal? elemental metals? pure chemical element?</p>	<p>Thank you for flagging this. By “pure” metal, we mean elemental. However, as this is a public facing toolkit, we use the word “pure” on our designs for accessibility and understandability.</p> <p>The text has been revised to reflect this meaning (lines 226-227). See also the revised Fig. A1 caption.</p>
66	<p>Figure 2 (Line 190): The text neither the discription describes whats ment with gaming, creative, workshop and visual.</p> <p>hand-on is mentioned but it remains unclear where are the diffrences, challanges for the particionares. Is it related to the time required, social profile, public presentation, competition issues, privacy.....</p>	<p>Thank you. Please see the revised figure (Fig. 3), which now includes clarification about each style of engagement preference in the caption as highlighted in your comments.</p>
67	<p>Line 196: small number of participants. Percentage of visitors? How attractive was the excercies?</p>	<p>See prior comments and responses regarding recruitment and sample size.</p>
68	<p>Line 206: Needs short description why that is the right choice for the statistical test</p>	<p>Thank you. See comment and response #50.</p>
69	<p>Line 207: BUT Cork has a scientific focus, so it can be assumed that visitors are interested in science, while Dingel deals with Celtic culture. Knowledge of CRM in connection with the energy transition and the circular economy cannot be assumed.</p>	<p>Thank you. Please see lines 275-276 which now mentions this as a limitation.</p>
70	<p>Line 229: poor result in attartivness of the tool, far bejond 1% (0,016%_&gt; 0,005% per day!) and hence not significant.</p>	<p>Thank you. We hope that the revised Fig. 1 and new Fig. 2 along with clarifications in text (e.g., section 2.3) make it clearer to the reviewer that study participants largely reflect the sample sizes possible based on total interaction time, recruiting time spent at festivals/events and informed consent. See comment and response #51.</p> <p>As also outlined in the public author responses, our goal was never to recruit all 250,000 people at a large festival over the course of 2.5 days, nor would this be possible.</p>
71	<p>Line 240: Evidently unsuccessful; see proportion of participants.</p>	<p>See comment and response #51 and all others above regarding sample sizes.</p>
72	<p>Table 2 (Rows 10-11 column 2): Negative phrasing requires the reader to perform a mental "negation" process that should be avoid. see e.g. <a href="https://www.noslangues-ourlanguages.gc.ca/en/writing-tips-plus/clear-communication-choose-positive-over-negative-phrasing">https://www.noslangues-ourlanguages.gc.ca/en/writing-tips-plus/clear-communication-choose-positive-over-negative-phrasing</a></p>	<p>Thank you. While we appreciate this comment, it is conventional to use negatively phrased items (reverse scored items) among positively phrased items to create ‘balanced’ Likert scales to combat acquiescence bias (e.g., <a href="https://doi.org/10.1590/1413-82712023280401">https://doi.org/10.1590/1413-82712023280401</a>).</p> <p>Nonetheless, please note that the scale referred to is not included in the revised</p>

		manuscript as it relates to the student study, which has now been removed.
73	Line 306: unclear - agree with preservation or with utilization or even both?	Thank you. The student sample has now been removed and as such these results are no longer presented.
74	Line 307: very low number not representative at all.	See prior comments and responses regarding recruitment and sample size.
75	Line 339: is the knowledge of the utilizers already high prior to the Green Dealz tool exercise?	Thank you. The student sample has now been removed and as such these results are no longer presented.
76	Line 366: The statistical base is not representative, neither the experimental set up of both tests.	Thank you. The student sample has now been removed and as such these results are no longer presented.
77	Line 372: not necessarily: also raising different items might be useful too - raising different question and providing more or more complex answers.	<p>Thank you. We appreciate this point made by RC1. This paragraph has now been rephrased in line with the removal of the student study which influenced this claim (lines 360-365).</p> <p>However, we wish to make clear that our testing phase evaluation demonstrates the incongruence of even 'quick' 3-item, dichotomously scored, pre- and post- surveys. Therefore, expanding these surveys as the reviewer suggests conflicts with our findings of what is possible and engaging in the festival setting (lines 351-353; lines 382-384). Although, we agree, expanding the surveys to include different or more complex items would be ideal. As such we include this as a recommendation for more controlled settings to validate the toolkit (lines 544-545).</p>
78	<p>Line 390: right reference? which section? term EA is not listed in their index</p> <p>Dori, Y.J., Avargil, S. (2015). Embedded Assessment. In: Gunstone, R. (eds) Encyclopedia of Science Education. Springer, Dordrecht. <a href="https://doi.org/10.1007/978-94-007-2150-0_46">https://doi.org/10.1007/978-94-007-2150-0_46</a> and therein</p>	Thank you. Please note the use of "(EA)" is simply an abbreviation used in this manuscript to refer to embedded assessment. It does not refer to an official abbreviation. We have now made this clearer in text (line 398).
79	Figure 5 (Line 455): why is the post result worse?	<p>In reference to the control group pre- and post- results during the refinement phase, to which this comment refers, please note that the post result is not statistically worse than the pre result as demonstrated by paired t-test (Fig. 5 and lines 450-458).</p> <p>The average pre-score for the control group is <math>0.54 \pm 0.17</math> and the average post-score for the control group is <math>0.54 \pm 0.19</math>, now listed in Fig. 5 caption (formerly in a table but this is now removed).</p>
	<b>RC2</b>	<b>Responses and changes made</b>
80	Line 109: Necessity: This section could benefit from a few clear sentences reviewing the iterative	Thank you. Please see lines 108-116 where we have now added more text as suggested, to provide an overview of each of the phases

	phases and, perhaps, how they fit in your broader research goals. E.g.	before going deeper in the following sections. This, combined with the revised Fig. 1, new Fig. 2, and the process and development split over two sections hopefully makes the process clearer to the reader.
81	Line 113: Necessity: This feedback was used to .... I realize you discuss this in section 2.4 but i think either introducing it here or moving section 2.4 before 2.2 helps the reader understand why you're doing what you're doing	Please see the above comment and response #80.
82	Line 117: Necessity: wondering if this section should be moved lower (after introducing how you evaluated participants), then it would help clarify why you chose particular analyses	Analyses used per section have now been expanded on. See comment and response #50. However, we have kept this analysis strategy sub-section in the new section 2 since this is generally more conventional for method-centric papers. We believe it provides a foundation for analytical methods rather than repeating steps taken within each section.
83	Line 119: Necessity: sentence not making sense. New and existing scales on what? And what survey items? I thought surveys were not done? If you mean scales that you used to measure changes in attitudes/knowledge, etc, specify that.	Thank you. Sentence has been revised for clarity (see lines 156-158).
84	Line 121: I would also add a section here on the various statistical tests used and why each one was warranted	Please see comment and response #50.
85	Line 125: Necessity: no mention of a codebook here. If you have one, identify here. If you don't I would recommend making one - this would clarify some of reviewer 1's concerns on the definitions associated with particular terms you use.	Please see comment and response #46 in relation to qualitative responses. Please see comment and response #66 regarding engagement preference terms in the multiple-choice survey question.
86	Line 131: Necessity: not really understanding this. Why does interaction time have anything to do with sample size? Also, should your target sample sizes be included here?	Please see comment and response #51 regarding sample sizes and recruitment. Please note lines 131-137 now explain target sample calculations more clearly.
87	Line 145: Consideration: is this to use as a basis for comparison for the results of your final implementation of green dealz? If so, specify that here	No. This is not for comparison of results between scoping and refinement. The text has been re-phrased to avoid confusion (lines 167-168).
88	167: Necessity: check these numbers as you cant have 102%	Thank you for flagging this key typo. We have amended the text to reflect the correct % (line 192-193).
89	Table 1 column 2: Necessity: Are the values you present for awareness rounded? Wouldn't they be decimal numbers? I believe the norm is to keep two decimals	Thank you. The values presented for awareness are medians due to the non-normality of the data. Please see line 161-162 where this is specified and Table 1 where the results are reported.
90	Line 199: Necessity: I assume not having pre-knowledge surveys for Dingle will be noted as a limitation?	Thank you. As Dingle was concerned only with a scoping survey (and not any form of intervention i.e., PE activity), we interpret this comment to refer instead to the Cork sample.  Key limitations of each phase are now better accounted for within each section and overall limitations are given in the discussion.

		However, we now attempt to better explain why a pre-post survey method wasn't included in Cork (lines 244-246; lines 274-275).
91	Line 205: Necessity: not clear what this value represents. Before, awareness was presented on a 5-point scale...	Apologies, this value is a mean rank. Mean ranks are compared (rather than medians) in a Mann-Whitney <i>U</i> -test when distribution shapes are dissimilar. This statistical test has now been removed in revision. However, we hope we have made this clearer where we have used this test elsewhere (lines 498-499).
92	Line 207: Consideration: just a note of caution. Sentences like this would typically be in the discussion. However, I understand your paper structure is different in that you are using findings from each phase to provide rationale for changes to the next phase. May be worth looking into other papers like this and ensuring that this is appropriate.	Thank you. We deem this finding/observation to be useful within the iteration process (even with limitations) and as such, a revised version of it remains in this section (lines 257-259).
93	Line 223: Consideration: I recognize hands-on activities was noted by participants, but was there something specific about the act of shopping that made you choose that. I assume it's the connective thinking of "to make a stew, I'm going to need these ingredients". You infer it in section 2.5.1 but I think including an explicit sentence reasoning for this could be beneficial	See lines 221-222 which now explicitly state why the theme of shopping was decided on.
94	Line 228: Necessity: Is there any data on who typically attend this festival? As part of the pre-survey, did you ask anything to do with perceptions on mining? Possibly an interesting mix of people, but perhaps, more pro-industry and thus pro-mining individuals? Potentially worth a brief comment in limitations section.	<p>We did not ask respondents about their perceptions of mining in the pre-survey nor the post-survey. While this indeed would be an interesting study, as specified in the text, our surveying was restricted by time in this kind of setting (e.g., lines 336-340). As we have not surveyed these values, we have not speculated on pro or anti mining attitudes among different audiences.</p> <p>There are no published data on who attends the National Ploughing Championships, however common audience groups are now mentioned in text (lines 293-294).</p> <p>However, we now make a brief, speculative comment about the relatively high perceived need for CRMs recorded at this event and its potential links to the audience (based on the limited surveying used). See lines 368-370.</p>
95	Line 232: Necessity: not clear why validity is needed and how its measured. What does validity refer to in this case? Did you want to see the same impact of green dealz on the results in two different settings? One setting is very uncontrolled (festival) while the other was very controlled? If so, specify this. However, the main issue I'm seeing with this is that it is two completely different audiences which would very	Thank you. As per RC2 suggestions and those from the editor and other reviewers we have now removed the student sample from our study and have reported on limitations of the existing evaluation at the National Ploughing Championships (section 3.5.3).

	much affect differences. I think i understand the intent of this, but the small sample size of the control group and the different population are significant limitations.	
96	Lines 310-318: would recommend having statistician review phrasing	Thank you. This paragraph is now removed in line with the removal of the student sample.
97	Line 385: every time a new festival is mentioned I'm wanting to know more about the audiences who attend these. Wondering if there are other studies out there on who attends festivals that you could draw comparisons to?	Thank you for flagging. For each section, we have now included common audience groups at these festivals (e.g., lines 171-176; lines 291-294; lines 391-396).  In the discussion, we now also include literature on the types of audiences expected at different festivals in response to editor comments on education levels (lines 517-524).
	<b>RC3</b>	<b>Responses and changes made</b>
98	2. The methodology is described only in broad terms in the paper at present. A summary table showing each festival's characteristics (location, participant numbers, activity duration, main observations) would aid comparison. The criteria for selecting festivals, participant demographics, data-collection instruments, and analytical techniques are not clearly stated nor justified as choices – the why is missing. The small and uneven sample sizes limit statistical reliability, and this is not discussed. The application of multiple statistical methods without explanation of why these were chosen also weakens reproducibility and thus the conclusions that can be drawn from the work. As a result I suggest that the work might be better positioned as an exploratory pilot rather than an evaluative trial and that, although semantics, might help the authors in not stretching the data too far in their discussion.	Thank you. We hope that the changes made as outlined in all above comments and responses have addressed these concerns.  In summary, we have included two new figures as outlined in comment and response #45 which aims to better frame and detail the iteration process and the festival settings. In each section we have aimed to make clearer our decision-making process to compliment these figures (e.g., lines 108-116; 146-150; 167-168; 281-286; 289-290; 381-387; 390; 473-475). We have revised our phrasing around statistics and representativeness, removed some unnecessary statistical tests and removed the study which caused the most concern in this regard. We have made more explicit why we used certain tests in each section/phase. We have re-framed the work as an exploratory pilot as suggested.
99	3. The findings are reported narratively, with limited data presentation nor analysis – either of the findings or the limitations. Claims about transferability to other contexts are not supported by the available data and this needs significant rewriting to catch what can and cannot be justified from the data presented. There are really interesting data to discuss that could be expanded upon in my view – such as the qualitative insights into audience engagement and the public misconceptions about CRMs. These really are valuable and merit deeper analysis – this can be speculative given the limited data, so could be part of a discursive analysis.	Thank you. We hope that these concerns are addressed through the revisions outlined in all comments and responses above.  In summary, we have revised our language around results and transferability, removed the student study completely, included limitations within each phase and in the discussion, made more explicit our decision making process between phases including expansion on the reporting and discussion of open feedback responses.

### 3. Discussion & conclusion

#	RC1	Responses and changes made
100	Line 530: no not significant proof	Thank you. Our comments around final conclusions and effectiveness have been reframed/rephrased (e.g., lines 532-533; 540-542). We have also included more recommendations for future research to test the toolkit's effectiveness (e.g., lines 544-545; 565-567).
101	Line 547: how to ensure that the exercise is taken seriously?	Thank you. It is not clear what is meant by this comment, but we interpret to ask how we can ensure that group participants at a festival take the activity seriously. This is a hard variable to measure and our results are based on good faith that consent based participation is indeed taken seriously.
102	Line 561: very speculative - there is no indication in the set up of the exercise that the interpretation is based on findings. However, it seems trivial that issues that are more often/more recently in the news are more commonly remembered (even wrongly) than those that are less public.	Thank you. We appreciate this is speculative, we simply include it as an interesting thought or reflection. We have now re-phrased the text to better highlight this (lines 572-574).
103	Line 568: Ford 2019 seems to be more relevant as she argue that if our tailored content and experiences can resonate with a person at an emotional and intellectual level, they will ultimately become devoted to the topic. Through empowering a community of people who understand and appreciate the topic, they can become advocates and champions for its protection and conservation in the future.	Many thanks for this suggestion. We include Ford (2019) in the paragraph that discusses the importance of meaning and relevance in public engagement (line 574).  Since we cite the idea of the 'residual' public, we include Miller (1983), and now Miller et al., 2024 as a more updated reference (line 563).
104	Line 578: overstreting the findings. As recognised the number of test persons are by far to little and not diverse enough. suggest to delete this or to add further set ups	Thank you. This has now been removed.
105	Line 599: statistics too low to speak about keen interest - see figures of 0,016% of all festival visitors attending the exercise.	Please see prior comments on sampling and recruitment in festival settings. Our claims that people show a keen interest are based on observations at our exhibits and feedback in surveys.
106	Line 610: Doubt that the achiev,ment is proven.	Thank you. We have re-framed our language to be more exploratory (lines 593-596).
	<b>RC2</b>	<b>Responses and changes made</b>
107	Line 530: as noted by the other reviewer, be careful with claims like this, and this is where i would suggest having a statistician consult on the phrasing. Perhaps its just a matter of further specifying the audience here and losing less absolute language?	Thank you. Please see comment and response #100. This sentence has been revised to remove absolute language and instead make clearer that conclusions are tentative due to the pilot/exploratory nature of the study.
108	Line 553: consideration: reiterate that this score measured correctness (right or wrong). Wasn't exactly clear what this sentence meant when i read it at first	Thank you. Revised in text (line 556).

109	<p>Line 561: Consideration: Potentially add acknowledgement that the national ploughing audience would still likely be pro-industry and that potentially has an impact on their knowledge. Though with that said, and this is another potential limitation, even the name of the game greendealz could be considered polarizing (maybe not as much in Ireland?). As in, right-leaning audiences may not even approach a booth with a name including green in it.</p>	<p>Thank you. As we do not see any significant change in knowledge within this sample, nor in their perceived need for CRMs, we do not discuss potential impacts on knowledge.</p> <p>This is an interesting outlook regarding polarisation and certainly if explored further would provide great insight into audience applicability. However, in the Irish context this is not so much of a factor. Nonetheless, we have included a line on this limitation in the discussion (lines 526-527).</p>
110	<p>Line 570: Necessary: get this paragraph reviewed; small sample size and narrow audience may not warrant any comment</p>	<p>This paragraph is now removed.</p>
111	<p>Line 604: I think this paper would really benefit from an additional section on limitations e.g.,</p> <ul style="list-style-type: none"> <li>- representativeness of the sample</li> <li>- is it truly a "incidental audience"</li> <li>- potential polarizing nature of the activity title etc</li> </ul>	<p>Thank you for this recommendation. We have followed this advice, however due to the iterative nature of the study we have now included limitations within each phase (see sections 3.4.3; 3.5.3 and 3.6.3).</p> <p>We now also re-frame our discussion to include key/overall limitations during toolkit development in section 4.1.</p> <p>We hope these editions provide clarity and address these concerns.</p>