



Peer-Review Report Form

Manuscript Title:	An Evaluation of Research Colloquium Series in Senior High School Using Kirkpatrick's Four-Level Model
Type of Article:	Original Article

FIRST ROUND

Editor

General Comments and Recommendations

The reviewers acknowledge the significance of your manuscript and recommend several revisions to further improve its clarity, coherence, organization, and overall scholarly quality. In addition, you are advised to thoroughly review and validate all references included in the manuscript. Please ensure that all cited sources are accurate, accessible, and retrievable through Google Scholar to uphold the credibility and reliability of the study. Additionally, make sure to include the DOI or website link for every reference whenever available to facilitate verification and accessibility of sources. Furthermore, only sources that are cited within the manuscript should appear in the References section, and all entries listed in the References section must also be cited in the text. For example, if 20 unique sources are cited in the manuscript, then only those 20 sources should be included in the References section. Please also note that repeated citations of the same source within the text are counted as a single reference entry; therefore, the total number of references may be lower than the total number of in-text citations.

Reviewer 1

General Comments and Recommendations

I read your manuscript with big interest. The study gives a very practical look at how a research colloquium series can help senior high school students improve their research skills. Choosing Kirkpatrick's Four-Level Model is a very good decision because it gives a complete picture of the training program, starting from the students' feelings up to the final results like conferences. The flow of the paper is mostly logical, and the problem you want to solve is clear. However, I have some major worries about your methodology and how you interpreted the data. Your claims about what the students "learned" are too strong because your research design is missing a baseline. Also, there is a very high drop-out rate of the participants that you did not explain, which might make your results look more positive than they really are. I suggest you fix these critical areas before the manuscript is finalized.



<u>Specific Comments and Recommendations</u>	
	In the Introduction, you mentioned that the initiative is made up of seven key events. But later, you stated that the last two episodes were not done in a plenary format. Instead, these were just discussed by the research teachers in their regular classes. It is a bit confusing to call them part of the "colloquium series" if they are just normal classroom lectures. Please clarify why you still included them in the program evaluation as colloquium events.
	In the Methodology, you reported that 494 students were enrolled, but only 214 completed the whole colloquium series. Table 1 also shows this is only 43.32% of the total cohort. Less than half finished the program. Why did so many students stop attending? You must discuss this in your paper. There is a big risk of "survivor bias" - meaning, maybe only the highly motivated students stayed until the end, which is why your Level 1 Satisfaction ratings are very high ("Highly Satisfied"). You should mention this in your Limitations section.
	Under the Instruments section, you explained that you used a structured rubric to check the data analysis accuracy and application of statistical tools for Level 3. But when I checked the Results section, Table 6 only shows the count of groups (45 out of 75) who did the analysis independently. Where are the actual scores from the rubric? If you used a rubric, you need to show the data. If you did not analyze the rubric scores, you should remove the mention of it in the Methodology to avoid confusing the readers.
	You showed that the students got a mean score of 40.63 out of 50 in the assessment. From this, you concluded that there is a "high level of knowledge retention" and successful "knowledge acquisition". But, as you admitted, no pre-tests were conducted. Without a pre-test, you really cannot claim that the colloquium caused this learning. Maybe the students already knew the statistics from their regular Practical Research 2 classes before they even attended your session. I strongly recommend changing your wording. Instead of saying "knowledge acquisition," you can say "high post-training knowledge level." You must also explicitly mention this lack of a baseline when you present Table 5, not just at the very end of the paper.
	You reported that 7 groups (9.33%) had their papers accepted for publication. You honestly noted that these publications have minimal indexing. To make this part stronger, it would be helpful to briefly mention what kind of publications these are. Are they school-based internal publications, or are they external peer-reviewed platforms? Providing this context will give readers a realistic idea of the students' achievements at Level 4.
	The recommendations you gave are very practical, especially the idea to give preparatory materials for difficult technical topics and the need to implement pre- and post-tests next time. Please just make sure that your final concluding sentences match the toned-down claims about Level 2 learning.

Please indicate your recommendation by checking the appropriate box below.

<u>Decision</u>	
	Accept the manuscript for publication.
/	Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.
	Reject the manuscript, as it is not suitable for publication.



Reviewer 2

<u>General Comments and Recommendations</u>	
The manuscript discusses a relevant school-based initiative that may help strengthen research culture among senior high school students. The use of Kirkpatrick's model is appropriate, but the paper needs stronger methodological clarity and more cautious interpretation of the results because most findings are descriptive.	
<u>Specific Comments and Recommendations</u>	
The study makes strong claims about the effectiveness and impact of the colloquium series, but the design does not fully support causal conclusions. Since there was no pre-test, comparison group, or baseline data, it is difficult to say that the improvements were caused mainly by the colloquium series.	
The methodology needs more detail on the participants and sampling. The paper states that 214 students completed the whole series, but the post-assessment involved 494 students. This difference should be clearly explained because it affects how the results should be interpreted.	
The instruments are not sufficiently described. The paper should explain how the survey, assessment test, rubric, and interview guide were developed, validated, and scored. Without this information, it is hard to judge the quality and credibility of the data.	
There is some inconsistency in the treatment of the seven colloquium events. The study says there were seven events, but the satisfaction tables appear to include only five sessions. The two class-based episodes should either be analyzed clearly or separated from the plenary sessions to avoid confusion.	
The qualitative analysis needs stronger reporting. The manuscript presents themes and sample responses, but it does not explain the number of interviewees, coding process, how themes were developed, or whether another coder checked the analysis. This is important to support trustworthiness.	
The Level 4 results are useful, but the interpretation is somewhat overstated. Conference participation and publication acceptance may also be influenced by teacher mentoring, student ability, school support, or selection of stronger groups. These possible factors should be acknowledged.	
The conclusion should be more balanced. Instead of saying that the program significantly strengthened research culture, it may be better to state that the findings suggest positive outcomes and potential contribution, while also recognizing the limits of the evaluation design.	

Please indicate your recommendation by checking the appropriate box below.

<u>Decision</u>	
	Accept the manuscript for publication.
/	Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.
	Reject the manuscript, as it is not suitable for publication.



SECOND ROUND

<u>Decision</u>	<u>Editor</u>	<u>Reviewer 1</u>	<u>Reviewer 2</u>
Accept the manuscript for publication.	/	/	/
Reconsider the manuscript after the authors have satisfactorily addressed and complied with the reviewers' comments and recommendations.			
Reject the manuscript, as it is not suitable for publication.			