15-Sep-2019   
  
Dear Mr Turner:   
  
Your manuscript entitled "The Impact of Academically Homogeneous Classrooms for Undergraduate Statistics", which you submitted to PRIMUS, has been reviewed.  The referee comments are included at the bottom of this letter, along with those of the editor who coordinated the review of your paper.   
  
The overall impressions of the current version of the manuscript are favorable.  Subject to some relatively minor revisions to your paper, we expect to ultimately accept the paper following said revisions. Therefore, we invite you to respond to the referees' comments and edit your manuscript accordingly; when you do so, please be sure to follow the directions from the associate editor (which appear last following my signature).   
  
When you revise your manuscript, please indicate the changes you have made in the author response field during the upload process.   
  
Please be aware that while figures will appear in color in the online issues of the journal, they will appear in grey-scale in the print issue.  Please verify that all figures will reproduce satisfactorily in grey-scale, or make any necessary adjustments.   
  
To submit the revision, log into https://mc.manuscriptcentral.com/upri and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. Please enter your responses to the comments made by the referee(s) in the space provided. You can use this space to document any changes you made to the original manuscript. Please be as specific as possible in your response to the referee(s).   
  
Alternatively, once you have revised your paper, it can be resubmitted to PRIMUS by way of the following link:   
  
\*\*\* PLEASE NOTE: This is a two-step process. After clicking on the link, you will be directed to a webpage to confirm. \*\*\*   
  
https://mc.manuscriptcentral.com/upri?URL\_MASK=65f5f82122524ecba0a57bb644f06409   
  
IMPORTANT:  Your original files are available to you when you upload your revised manuscript.  Please delete any redundant files before completing the submission.   
  
Once again, thank you for submitting your manuscript to PRIMUS and I look forward to receiving your revision.   
  
Sincerely,   
Matt   
--   
Dr Matthew Boelkins   
Editor in Chief, PRIMUS   
boelkinm@gvsu.edu   
  
\*\*\*   
  
Referee(s)' Comments to Author:   
  
Referee: 1   
  
Recommendation: Accept   
  
Comments:   
(There are no comments.)   
  
Additional Questions:   
<i>Brief Synopsis</i>: Please include in your synopsis a classification of the type of paper this is – e.g. a reflection on a teaching experience, a discussion of a pedagogical practice, a description of a new idea or technology for teaching, a case study, an assessment study, or …:

*This is a study comparing heterogeneous to homogeneous classrooms.*  
  
<i>Audience</i>: Who is the primary audience that will find the paper valuable?  Is the paper clearly addressed to this audience? Is the idea of the paper specific to a particular cohort (e.g. a specific course such as linear algebra, or a specific institution such as a community college)?  If so, can the ideas be reasonably exported to other cohorts? If not, can the paper be enhanced in some manner to broaden its applicability?:

*The audience that will find this paper valuable is the group of undergraduate math instructors. There is nothing specific to certain courses or types of institutions. There may be applications to other disciplines, but I think their audience is already very broad.*  
<i>Exposition</i>: Is the paper focused and well-organized? Is there a clear theme?  Do the sections flow well?  Is the writing engaging?  Is the writing grammatically correct and fluid?  Is the length of the paper appropriate to its content?  Please comment on how these aspects of the paper might be improved.:

*The paper is well-written. It was very accessible and had a great flow. I thought the length of the paper was appropriate.*  
<i>Mathematical Content</i>: If a specific mathematical concept is the central focus of the article, are there new and interesting insights into the mathematics itself or into how students learn and understand it?  How are these connections between the underlying mathematical content and the pedagogy explored? Is all mathematical content in the paper correct?:

*The focus of the paper is on comparing student success in two different class structures: classes with students of equal ability and classes with students of mixed ability. It does not focus on mathematical content.*  
<i>Professionalism</i>: Is the paper written objectively and with a respectful tone toward students, colleagues, and the profession? Are the claims of the paper substantiated in appropriate depth for PRIMUS?  Is the scholarship responsible, giving sufficient detail and providing careful references with proper credit to sources?  (No significant portion of the work should be published elsewhere and no part of the work may be plagiarized, including self-plagiarization.  While adhering to professional ethics is the author’s responsibility, please do note any concerns you may have in your report.): The tone was very respectful to all, and the scholarship was very responsible. In particular, the paper did not get into the perceived advantages/disadvantages of the two approaches, but merely compared results. 

*This is the first time I have seen a study of these two approaches.*  
<i>Fit</i>: Good PRIMUS articles often have all or most of the following traits:   
(1) a novel idea for teaching and learning (this could be at any level -- individual students, classroom, programmatic, institutional or even multi-institutional, keeping in mind that PRIMUS stands for “Problems, Resources, and Issues in Mathematics Undergraduate Studies”)   
(2) discussion of how that idea impacts student learning (the claims of the article may be substantiated by anecdotal reports or a summary of a formal study, but should be compelling to a typical mathematics instructor)   
(3) practical, transferable advice facilitating implementation by practitioners   
(4) exploration of at least one interesting problem in undergraduate mathematics pedagogy or curriculum, as well as substantive insight about that problem and how it might be addressed   
  
In what ways does this paper have these traits?  In which areas might the paper need to be strengthened in order to do so?  If the paper does not have most of these traits, is there nevertheless a highly compelling reason it would be of interest to PRIMUS readers?:

*I have not seen a study of heterogeneous versus homogeneous classrooms, and many schools are making decisions that are not data-driven. This paper shows that homogeneous classrooms do not produce greater student success.****The only idea that I would like to hear more about are potential strengths of a heterogeneous classroom, but the paper stands on its own merit without it.***  
<i>Overall Merit</i>: In summary, is this paper well-written and substantive? Is there sufficient content, reflection, and novelty to merit publication?  What are its overall strengths?:

*I enjoyed this paper. The design and results of the experiment are clear. As many colleges are now accelerating the pathway through mathematics, this paper could be quite valuable for colleges deciding whether to use a cohort or comingle model for their corequisite courses.*

Referee: 2   
  
Recommendation: Major Revision   
  
Comments:   
(There are no comments.)   
  
Additional Questions:   
<i>Brief Synopsis</i>: Please include in your synopsis a classification of the type of paper this is – e.g. a reflection on a teaching experience, a discussion of a pedagogical practice, a description of a new idea or technology for teaching, a case study, an assessment study, or …:

*This paper describes a study in which the authors compared the performance of students in an introductory statistics course at a military academy based on their perceived ability going into this course.  Using multiple analysis techniques, the authors found that there is no significant difference in student performance based on whether they were randomly assigned to a class vs. grouped by ability.  Furthermore, there was no significant difference in student performance based on the ability level.*  
<i>Audience</i>: Who is the primary audience that will find the paper valuable?  Is the paper clearly addressed to this audience? Is the idea of the paper specific to a particular cohort (e.g. a specific course such as linear algebra, or a specific institution such as a community college)?  If so, can the ideas be reasonably exported to other cohorts? If not, can the paper be enhanced in some manner to broaden its applicability?:

***This paper's main audience appears to be teachers of introductory statistics as well as perhaps academic administrators.  I believe that the scope of this paper is fairly limited to those that teach introductory statistics courses.  The authors do claim that this research could apply to "all undergraduate statistics courses," but I did not find that to be substantiated.***  
<i>Exposition</i>: Is the paper focused and well-organized? Is there a clear theme?  Do the sections flow well?  Is the writing engaging?  Is the writing grammatically correct and fluid?  Is the length of the paper appropriate to its content?  Please comment on how these aspects of the paper might be improved.: The article is straightforward and easy to read.   
  
***I suggest that sections 4 and 5 be combined into one.***  
***Some of section 4.2 appears to contradict what is said in 4.1.  In 4.1, the authors state that "the significance of the statistical results are limited to this type of course," but in 4.2, they state that "these results can be applied with confidence to all undergraduate statistics courses."  I do not believe that the quoted statement in 4.2 is substantiated.  Furthermore, I think that the scope is even more limited than to that which the quoted statement in 4.1 suggests. In 4.1, the authors also state their belief that "other STEM courses would yield similar results" without any justification for this statement.***  
***Grammatical/wording corrections/suggestions:   
1.  page 1, line 38/47:  "ability or/and randomly grouped sections" should be changed to "ability- or/and randomly-grouped sections" (respectively, based on /)   
2.  page 1, line 40:  add comma after "course"   
3.  page 2, line 34:  "students" should be the possessive "student's"   
4.  page 2, lines 41-43:  "students" should be changed to the plural possessive "students'"   
5.  page 2, line 52:  "freedom" might be a better word choice than "ability"   
6.  page 3, line 7:  change "whom" to "who"   
7.  page 3, line 7:  move "at a military academy" to the end of the sentence   
8.  page 3, line 16:  change "draw" to "drawing"   
9.  page 3, line 27:  I suggest putting "as indicated in Table 1" in parentheses   
10.  page 3, line 52:  add a comma after "i.e."   
11.  page 4, line 42 and Table 1 (page 5):  On page 4, courses are referred to by title, but in Table 1, courses are only referred to by number.  This makes it difficult for the reader to know which course is which.   
12.  page 4, line 49:  I suggest adding "fits of the" before "models," since it is the model fit that is summarized in the table, not the models themselves.   
13.  page 5, Table 1:  Non-STEM should be hyphenated   
14.  page 5, lines 22-26 (Table 1 caption): The second sentence does not completely make sense and should be reworded.   
15.  page 5, line 36:  "interpretability" does not need to be hyphenated, and a comma should be added after this word   
16.  page 5, line 38:  the reference should go before the period   
17.  page 5, line 51:  change "assign" to "assigned"   
18.  page 5, Table 2:  For the sake of consistency, I think the 2nd and 3rd descriptions should have "Final Grade" added at the end.   
19.  page 7, line 15 (Table 3 caption):  add "the fits of the" before the word "course"   
20.  page 7, line 28:  change "increase" to "increases"   
21.  page 7, line 26:  "course wide" should be hyphenated   
22.  page 7, line 47:  change "analyze" to "analyzed"   
23.  page 8, line 18 (Table 4 caption):  add a period at the end   
24.  page 8, line 21:  add a comma after "level"   
25.  page 8, line 25:  "end of course" should be hyphenated   
26.  page 9, lines 30-32:  change "Ability 8 Random vs Ability" to "Ability 8 vs Random"   
27.  page 9, line 45:  "side by side" should be hyphenated   
28.  page 10, line 43:  perhaps change "position" to "ability"   
29.  page 11, line 48:  add a comma after "course"   
30.  page 12, line 43:  change "outweighs" to "outweigh"   
31.  page 15, line 17:  insert "likely" before "what"   
32.  page 15, line 21:  "higher ability" should be hyphenated   
33.  page 15, line 21:  perhaps change "rethought" to "reconsidered"   
34.  page 1, line 40:  add a comma after "course"***  
<i>Mathematical Content</i>: If a specific mathematical concept is the central focus of the article, are there new and interesting insights into the mathematics itself or into how students learn and understand it?  How are these connections between the underlying mathematical content and the pedagogy explored? Is all mathematical content in the paper correct?:

*Mathematical/statistical content is not the focus of this article.*  
  
<i>Professionalism</i>: Is the paper written objectively and with a respectful tone toward students, colleagues, and the profession? Are the claims of the paper substantiated in appropriate depth for PRIMUS?  Is the scholarship responsible, giving sufficient detail and providing careful references with proper credit to sources?  (No significant portion of the work should be published elsewhere and no part of the work may be plagiarized, including self-plagiarization.  While adhering to professional ethics is the author’s responsibility, please do note any concerns you may have in your report.):

***In the second paragraph of the Introduction, the authors state that Dr. Tom Loveless is the most referenced author.  While that might be true, the particular source that is cited is from 21 years ago.  Thus, I believe that the statistics quoted are likely fairly outdated.  The authors should consider consulting a more recent source.  In this same paragraph, the authors make reference to those surveyed being "binned by those that taught in middle school," but the meaning of this is unclear.***  
***In the paragraph that begins on line 27 of page 3, more information about the course examined in this study should be included.  What is the prerequisite for this course?  Is this an algebra-based or calculus-based probability/statistics course?***  
***In the last paragraph of section 2.2, the authors state that "the groups were structured in a way that minimized bias." This claim should be justified a bit more.***  
***In section 2.3, the authors mention the 9 different ability groups, but I feel that more information should be given here. Although it is clear that 1 contains the students with the highest ability and 9 those with the weakest, how were these divisions made.  Furthermore, why did the researchers choose to have so many different groups.  I would like to see more details on both the motivations and the methods of the authors here.***  
  
***Also in section 2.3, the authors mention the final exam.  Is this exam comprehensive?***  
***In Tables 4 and 6, the hour indicators of B, C, H, and I are used.  However, there is no explanation of what these mean.  Are they arbitrary?  Also in table 4, it is unclear as to what the parenthetical "2nd" and "3rd" indicators refer.***  
***Based on the instructor allocation shown in Table 6, it appears that there might be some confounding of ability with instructor.  Is there a reason why each instructor had only one of the ability levels?  The authors claim on page 6 that it was to "reduce the impact on instructors," but it is not clear why or how this design does that.***  
***The interpretation of the confidence interval in the last sentence of the second paragraph of section 3.2 is incorrect.  They could state that with 95% confidence they cannot state that there is a difference in the two means.  However, they cannot be confident that there is no difference.  Although these sound similar, they have different implications.***  
***For the pairwise comparisons that are shown in Table 8, it is highly suggested that the authors use Dunnett's multiple comparisons procedure since the 9 ability levels are all being compared to a common control (the randomly assigned group).***  
***The last sentence on page 9 needs more justification/explanation, since I do not see anecdotal evidence of what is stated.***  
***Since the print version of this article will not include color, the figures need to be altered to make the groups more distinguishable.  I suggest that different plotting symbols be used for each of the two groups in Figure 1.  In Figure 2, I suggest that the authors indicate in the caption which group is on the left and which is on the right in each pair of graphics.***  
***Please see my comments regarding the Discussion/Conclusion in the "Exposition" section of this review.  I do not believe that these results are generalizable in the ways that the authors claim.***  
***In the Conclusion, the authors make reference to "qualitative findings."  However, only quantitative findings were presented in this article.***  
<i>Fit</i>: Good PRIMUS articles often have all or most of the following traits:   
(1) a novel idea for teaching and learning (this could be at any level -- individual students, classroom, programmatic, institutional or even multi-institutional, keeping in mind that PRIMUS stands for “Problems, Resources, and Issues in Mathematics Undergraduate Studies”)   
(2) discussion of how that idea impacts student learning (the claims of the article may be substantiated by anecdotal reports or a summary of a formal study, but should be compelling to a typical mathematics instructor)   
(3) practical, transferable advice facilitating implementation by practitioners   
(4) exploration of at least one interesting problem in undergraduate mathematics pedagogy or curriculum, as well as substantive insight about that problem and how it might be addressed   
  
In what ways does this paper have these traits?  In which areas might the paper need to be strengthened in order to do so?  If the paper does not have most of these traits, is there nevertheless a highly compelling reason it would be of interest to PRIMUS readers?:

***I do believe that it is useful for readers to see what appears to be a lack of impact of homogeneous classrooms.  Most institutions would lack the ability to incorporate such a scheme since students mostly have the freedom to choose their own courses.  For this reason, I do think it would be of interest to PRIMUS readers.  However, I think that much work needs to be done (as indicated in other portions of this review) in order to make it ready for publication in PRIMUS.***  
<i>Overall Merit</i>: In summary, is this paper well-written and substantive? Is there sufficient content, reflection, and novelty to merit publication?  What are its overall strengths?:

***This paper is fairly well-written, but it needs a substantial revision.  It's major strength is as I described in the "Fit" section of this review.  Institutions that have less rigid structure and influence in student scheduling would be able to see that such groupings of students by ability were not associated with a difference in performance in this statistics class.  That might ease any pressure they feel to instill such a course structure.***

Referee: 3   
  
Recommendation: Minor Revision   
  
Comments:   
(There are no comments.)   
  
Additional Questions:   
<i>Brief Synopsis</i>: Please include in your synopsis a classification of the type of paper this is – e.g. a reflection on a teaching experience, a discussion of a pedagogical practice, a description of a new idea or technology for teaching, a case study, an assessment study, or …:

*This article describes the results of an experiment designed to test the hypothesis that grouping students by ability in undergraduate introductory statistics course sections leads to higher final average grades in the course when compared to randomly assigning students to sections of the course.*  
<i>Audience</i>: Who is the primary audience that will find the paper valuable?  Is the paper clearly addressed to this audience? Is the idea of the paper specific to a particular cohort (e.g. a specific course such as linear algebra, or a specific institution such as a community college)?  If so, can the ideas be reasonably exported to other cohorts? If not, can the paper be enhanced in some manner to broaden its applicability?:

*Administrators and educators, especially statistics educators, at the high school and university levels will find this information valuable.  Administrators and department chairs at larger colleges and universities will find this extremely interesting and valuable.  The results of the study could reasonable be applied to STEM educators at the high school and university level which would broaden the pool of interested readers.*  
<i>Exposition</i>: Is the paper focused and well-organized? Is there a clear theme?  Do the sections flow well?  Is the writing engaging?  Is the writing grammatically correct and fluid?  Is the length of the paper appropriate to its content?  Please comment on how these aspects of the paper might be improved.:

*The length of the paper is good.  The sections are logically organized and labeled.  There are a number of punctuation errors involving apostrophes towards the beginning of the paper, including the abstract.  The overall writing flows well in most parts of the paper.*  
***It is not clear, on page four, what is meant by "validating" a course.  The last two sentences of section 2.1 were not clearly written.***  
***In the second sentence of section 4.1, I would recommend inserting the word "practical" when referring to significance so readers do not confuse this with statistical significance.***  
***There is a typo in reference 2.***  
  
<i>Mathematical Content</i>: If a specific mathematical concept is the central focus of the article, are there new and interesting insights into the mathematics itself or into how students learn and understand it?  How are these connections between the underlying mathematical content and the pedagogy explored? Is all mathematical content in the paper correct?:

***Most of the mathematical content of the paper is correct.  The conclusion that the authors reach in the second paragraph of section 3.2 is not valid.  It is extremely disconcerting that the authors have not reported the p-values for each test.  Perhaps these values could be included in tables 8 and 9.  The significance level of 5% should be mentioned somewhere as well.***  
***In the third sentence of section 4.2, the authors state "These results can be applied with confidence to all undergraduate ..." however this study only considered introductory statistics courses.  More justification or guarded language would be welcomed.***  
***Specifics of the linear models would be welcomed in an appendix.  These would make it more clear what are the independent variables considered in the second linear model.***  
<i>Professionalism</i>: Is the paper written objectively and with a respectful tone toward students, colleagues, and the profession? Are the claims of the paper substantiated in appropriate depth for PRIMUS?  Is the scholarship responsible, giving sufficient detail and providing careful references with proper credit to sources?  (No significant portion of the work should be published elsewhere and no part of the work may be plagiarized, including self-plagiarization.  While adhering to professional ethics is the author’s responsibility, please do note any concerns you may have in your report.):

***The authors do not explain whether this experiment was blinded.  Did the instructors know whether their sections contained students grouped by ability or at random?  This information should be included and incorporated as appropriate into the analysis and applicability of their results.***  
<i>Fit</i>: Good PRIMUS articles often have all or most of the following traits:   
(1) a novel idea for teaching and learning (this could be at any level -- individual students, classroom, programmatic, institutional or even multi-institutional, keeping in mind that PRIMUS stands for “Problems, Resources, and Issues in Mathematics Undergraduate Studies”)   
(2) discussion of how that idea impacts student learning (the claims of the article may be substantiated by anecdotal reports or a summary of a formal study, but should be compelling to a typical mathematics instructor)   
(3) practical, transferable advice facilitating implementation by practitioners   
(4) exploration of at least one interesting problem in undergraduate mathematics pedagogy or curriculum, as well as substantive insight about that problem and how it might be addressed   
  
In what ways does this paper have these traits?  In which areas might the paper need to be strengthened in order to do so?  If the paper does not have most of these traits, is there nevertheless a highly compelling reason it would be of interest to PRIMUS readers?:

***This paper addresses points (3) and (4).  It is novel and interesting.***  
<i>Overall Merit</i>: In summary, is this paper well-written and substantive? Is there sufficient content, reflection, and novelty to merit publication?  What are its overall strengths?:

*The paper is a valuable addition to the literature on how to group students across multiple sections of the same course.  The experiment is well-designed and the results are interesting and valuable.*  
*The paper's strength derives from its applicability and novelty.  This well-designed experiment quantitatively tests an important hypothesis.*  
\*\*\*   
  
Associate Editor Comments to the Author:   
  
This is an interesting and unique paper as in most cases students take courses as they like. Most institutions couldn't run the experiment you ran. Two of the referees provide thoughtful comments, one of which caught numerous grammatical errors. Most, if not all, of these should be addressed. In particular, **I agree with the referee that suggested adding p-values and possibly an appendix with the models you used for the analysis. Also, consider your language about how this may or may not generalize to other courses and possibly other institutions. It isn't clear that another college, even if it could do what you do, would necessarily get the same results.**  
  
I think there is some work to do addressing referee comments, but it is mostly routine and doesn't impact what you have to say. Hence the recommendation is for minor revisions.