PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Impact of the COVID-19 pandemic on the experience and mental health of university students studying in Canada and the United Kingdom: a cross-sectional study
AUTHORS	Appleby, Jennifer; King, Nathan; Saunders, Kate E; Bast, Anne; Rivera, Daniel; Byun, Jin; Cunningham, Simone; Khera, Charandeep; Duffy, Anne

VERSION 1 – REVIEW

REVIEWER	Lunghi, Carlotta Université du Québec à Rimouski
REVIEW RETURNED	06-Apr-2021

GENERAL COMMENTS	I thank the authors and the editor for the opportunity to comment on the manuscript "Impact of the COVID-19 pandemic on the experience and mental health of university students studying in Canada and the United Kingdom: a cross-sectional study". I think the paper is very interesting and addresses an important topic on the impact of COVID-19 pandemia for undergraduate students. I have only a few comments. 1. I would suggest putting Supplemental Figure 1 as Figure 2 in the text. Add the number of responders in the legend of the figure. 2. Similarly, I would suggest putting supplemental Tables 1 and 2 in the text, if possible. 3. A more significant proportion of Queen's University respondents did not have complete data for the variables of interest (12.7% vs. 5.3%). It would be interesting to explore the reasons and that the authors report if there are differences between responders with complete data and without. 4. The negative perception of online learning is largely higher for Queen's University students vs. Oxford's students. Did the authors explore these results?

REVIEWER	Wickersham, Alice King's College London, Institute of Psychiatry, Psychology and Neuroscience
REVIEW RETURNED	12-Apr-2021

GENERAL COMMENTS	Please note that I was invited to review this manuscript with a particular emphasis on the statistical methods and analyses used, and that a full assessment of the clinical or other non-statistical aspects of the manuscript was not required. I have therefore limited my comments to aspects of the design and analysis, and marked some review checklist items as N/A.

Thank you for the opportunity to review this important and interesting study. The analyses are informative. The below comments are primarily requesting clarification on aspects of the design and presentation.

Overall, it could be clearer why certain design/analysis decisions were made - I expect there are very reasonable practical or theoretical explanations for the below, but without knowing them and explicitly acknowledging them in the methods/limitations sections, the analyses reported appear selective:
- Why were only first year participants sampled in Oxford, but all undergraduates at Queen's?
- Why was Oxford's knowledge about COVID-19, associated experiences, and receipt of mental health support not reported?
- Why weren't gender, age and year of study comparisons also done for the Oxford sample?
- Why weren't age group and year of study consistently investigated as predictors of all Queen's outcomes? E.g. year of study was investigated as a predictor of the variables in supplementary table 2, but not supplementary table 1; age group and year of study were not investigated as predictors of average knowledge levels of COVID-19.

Some other points:
- If the authors are able, it might be useful to report more key statistics in the abstract. At present, a couple of percentages for one outcome variable have been included, but it's not clear why that outcome was picked out.
- p.5, line 4: Can the authors say more about how the missing data mechanism was explored, and why a complete case analysis was appropriate? It was implied in the abstract that there was no missing data, so this came as a surprise. I'd suggest rephrasing the abstract's reference to missing data slightly.
- Great to see the Queen's sample being compared to the university database. Can the same be done for the Oxford sample to check for response bias?
- It's difficult to compare the Queen's and Oxford findings without their figure/table results presented side-by-side - would it be possible to do this? (I.e. providing descriptives for Oxford in Table 1; presenting Figure 1 and Supplemental Figure 1 side-by-side).
- It's not clear from the title or abstract that this is a mixed methods study, so the qualitative component came as a surprise. Given that the qualitative findings form a significant part of the paper, can it be given more emphasis earlier on?
- For the limitations, it's worth emphasising that the descriptive analyses can't identify statistically significant differences between Oxford and Queen's - so findings in this regard are purely exploratory.
- STROBE checklist item 16a - some inferential statistics have been conducted, so it's not true that these are descriptive analyses only.

VERSION 1 – AUTHOR RESPONSE

REVIEWER #1
Dr. Carlotta Lunghi, Université du Québec à Rimouski

COMMENT. I thank the authors and the editor for the opportunity to comment on the manuscript "Impact of the COVID-19 pandemic on the experience and mental health of university students studying in Canada and the United Kingdom: a cross-sectional study". I think the paper is very interesting and addresses an important topic on the impact of COVID-19 pandemic for undergraduate students. I have only a few comments.
RESPONSE. Thank you for this positive recognition.

COMMENT. I would suggest putting Supplemental Figure 1 as Figure 2 in the text. Add the number of responders in the legend of the figure.
RESPONSE: As suggested by the reviewer we have added Supplemental Figure 1 as Figure 2 in the main document. The number of respondents has also been added to the figure legend.

COMMENT. Similarly, I would suggest putting supplemental Tables 1 and 2 in the text, if possible.
RESPONSE. Thank you. Supplemental Tables 1 and 2 have now been added to the text (i.e., As Table 2 and 3).

COMMENT. A more significant proportion of Queen's University respondents did not have complete data for the variables of interest (12.7% vs. 5.3%). It would be interesting to explore the reasons and that the authors report if there are differences between responders with complete data and without.
RESPONSE. We have added a comparison of Queen's students with and without missing data to the Results section. Students with missing data were more likely to be male and international students. The lower missing in the Oxford sample compared to Queen's is likely related to the Oxford sample being much smaller in size. Those students who did elect to participate at Oxford may have been more invested in the study (e.g., because they were all first year students and more likely to see changes implemented as a result of their feedback). We did not include a comparison of those with and without missing data for the Oxford sample, as there were only 17 students with missing data, and the Oxford sample was intended to serve as an exploratory informative compliment to the primary analysis.

COMMENT. The negative perception of online learning is largely higher for Queen's University students vs. Oxford's students. Did the authors explore these results?
RESPONSE. We have added sentences to the discussion highlighting this finding and suggesting possible interpretations. For example, we speculated that online learning was more prevalent at Oxford than Queen's pre-pandemic, so that moving fully online was less of an adjustment.

REVIEWER #2
Miss Alice Wickersham, King's College London

COMMENT. Thank you for the opportunity to review this important and interesting study. The analyses are informative. The below comments are primarily requesting clarification on aspects of the design and presentation.
RESPONSE. Thank you for this positive review.

COMMENT. Overall, it could be clearer why certain design/analysis decisions were made - I expect there are very reasonable practical or theoretical explanations for the below, but without knowing them and explicitly acknowledging them in the methods/limitations sections, the analyses reported appear selective:

RESPONSE. Thank you we have addressed this in the responses below.

COMMENT. Why were only first year participants sampled in Oxford, but all undergraduates at Queen's?
RESPONSE. At Oxford the COVID-related questions were added to an ongoing survey of first year students, while at Queen's given that the winter term had ended earlier than in the UK, a supplemental COVID survey was sent out between the scheduled biannual surveys providing the opportunity to expand to all students. [see methods, page 3]

COMMENT. Why was Oxford's knowledge about COVID-19, associated experiences, and receipt of mental health support not reported?
RESPONSE. Unfortunately, not all items that were collected in the Queen's sample were included in the Oxford survey. This has now been clarified in the text (Methods).

COMMENT. Why weren't gender, age and year of study comparisons also done for the Oxford sample?
RESPONSE. A comparison by gender in the Oxford sample has now been added (new Supplemental Tables 1 and 2) and is reported on in the Methods. Comparisons by age and year of study were not done for the Oxford sample, because only first year students covering a limited age range completed the survey at Oxford (compared to all undergraduate students at Queen's).

COMMENT. Why weren't age group and year of study consistently investigated as predictors of all Queen's outcomes? E.g. year of study was investigated as a predictor of the variables in supplementary table 2, but not supplementary table 1; age group and year of study were not investigated as predictors of average knowledge levels of COVID-19.
RESPONSE. A comparison by year of study has been added to Supplementary Table 1 (now Table 2), so that the comparisons by gender, age, and year of study are consistent for the pandemic anxiety and impact items (now Table 2 and 3). Age comparisons of knowledge levels, following of recommendations, and approaches to isolation are now reported on in the Results. Additional comparisons by year of study were not reported for these variables, because year of study is highly collinear with age and therefore its inclusion they would not add to the findings already reported.

COMMENT. If the authors are able, it might be useful to report more key statistics in the abstract. At present, a couple of percentages for one outcome variable have been included, but it's not clear why that outcome was picked out.
RESPONSE. Thank you, we have reported more statistics in the Abstract.

COMMENT. p.5, line 4: Can the authors say more about how the missing data mechanism was explored, and why a complete case analysis was appropriate? It was implied in the abstract that there was no missing data, so this came as a surprise. I'd suggest rephrasing the abstract's reference to missing data slightly.
RESPONSE. Because there was a relatively small amount of missing data we choose to present a complete case analysis, so that the proportions are directly comparable across items (based on the same underlying sample). We have rephrased the statement in the Abstract and added a sentence to the Results section comparing Queen's students with and without missing data on basic demographics. We did not explore missing in the Oxford sample, because of small numbers of missing (n= 17).

COMMENT. Great to see the Queen's sample being compared to the university database. Can the same be done for the Oxford sample to check for response bias?
RESPONSE. A sentence has been added to the results comparing the Oxford sample to the underlying student population in regard to gender.

COMMENT. It's difficult to compare the Queen's and Oxford findings without their figure/table results presented side-by-side - would it be possible to do this? (I.e. providing descriptive statistics for Oxford in Table 1; presenting Figure 1 and Supplemental Figure 1 side-by-side).
RESPONSE. The focus of the analysis was on the Queen's sample, with the smaller Oxford sample intended to serve an exploratory informative compliment. For this reason, and because we have only a subset of measures available in the Oxford sample we have chosen not to combine the tables/figures. However, based on the reviewers comment we have added to the description of the Oxford sample in the prose (Results section).

COMMENT. It's not clear from the title or abstract that this is a mixed methods study, so the qualitative component came as a surprise. Given that the qualitative findings form a significant part of the paper, can it be given more emphasis earlier on?
RESPONSE. Sentences have been added to the Abstract (Design and Results) to highlight the qualitative data component of this study. We would not consider this a mixed methods study however, given the lack of a formal integration process. For that reason we have chosen not to use the terminology 'mixed methods' in our manuscript, but instead point to the qualitative component as an informative compliment to our primarily quantitative findings.

COMMENT. For the limitations, it's worth emphasising that the descriptive analyses can't identify statistically significant differences between Oxford and Queen's - so findings in this regard are purely exploratory.
RESPONSE. We have added a sentence in the Limitations section of the Discussion highlighting that the Oxford cohort provided a secondary or exploratory descriptive analysis and was not the main focus of the analysis.

COMMENT. STROBE checklist item 16a - some inferential statistics have been conducted, so it's not true that these are descriptive analyses only.
RESPONSE. Thank you. We have revised this item and attached an updated checklist with our revised submission.

Thank you again for the opportunity to address the reviewers comments. We appreciate the time and care that went into these reviews, and believe our manuscript has benefited as a result.

VERSION 2 – REVIEW

REVIEWER	Wickersham, Alice King's College London, Institute of Psychiatry, Psychology and Neuroscience
REVIEW RETURNED	18-Aug-2021

GENERAL COMMENTS	Many thanks to the authors for this revised manuscript. It is very much improved, and my queries from the previous round have been addressed. Following these changes, I have outlined below a few more minor suggestions for clarification. 1) Page 4, "...leading to a framework of specific phenomena that appeared increasingly likely to reflect the participant's experience" - can the authors clarify this? It's not clear what is meant by increasing likelihood in this context.

	2) Page 5, the first paragraph of the results jumps from response rate, to data missingness, then back to comparing responded vs remainder of eligible population, then descriptives for the final sample. I'd suggest re-ordering this paragraph to first focus on response rate and comparing responded vs remainder of eligible population, then moving onto missing data and comparing complete vs missing cases, then descriptives for the final sample (plus making sure the Oxford paragraph matches this ordering). 3) Page 4 & 7, it's not clear why knowledge of Covid-19 was treated as continuous (reporting means and t-tests) while following recommendations was treated as categorical (reporting %s and Chi-squared), when both used seven-point Likert scales (additionally, %s are reported for knowledge of Covid-19 in Table 1, so it isn't consistently treated as continuous). I'd suggest picking one approach and using it consistently, or justifying this variation in the text. 4) In the quantitative results section, the authors sometimes helpfully specify "X2 p-value", but other times just report a p-value without clarifying the type of test that the p-value is for; it would be helpful to make this consistent. 5) For the percentages reported in the qualitative results section, what is the denominator? Is it the percentage of survey respondents overall, or the percentage of survey respondents who left qualitative comments? 6) The limitations could reflect further on the fact that the two universities might not be entirely comparable, and that differing restrictions in the two countries might also account for some of the variation between them.

VERSION 2 – AUTHOR RESPONSE

Reviewer 2:

Miss Alice Wickersham, King's College London

Comment 1. Many thanks to the authors for this revised manuscript. It is very much improved, and my queries from the previous round have been addressed. Following these changes, I have outlined below a few more minor suggestions for clarification.

Response. We are glad to have been able to address your previous queries and appreciate all of your thoughtful feedback. Please find our responses to your additional suggestions below.

Comment 1. Page 4, "...leading to a framework of specific phenomena that appeared increasingly likely to reflect the participant's experience" - can the authors clarify this? It's not clear what is meant by increasing likelihood in this context.

Response. Thank you. We have elaborated on this aspect of the methodology which we hope will provide greater clarity to the reader.

Comment 2. Page 5, the first paragraph of the results jumps from response rate, to data missingness, then back to comparing responded vs remainder of eligible population, then descriptives for the final sample. I'd suggest re-ordering this paragraph to first focus on response rate and comparing responded vs remainder of eligible population, then moving onto missing data and comparing complete vs missing cases, then descriptives for the final sample (plus making sure the Oxford paragraph matches this ordering).

Response. The first paragraph of the Results section has been re-ordered as suggested.

Comment 3. Page 4 & 7, it's not clear why knowledge of Covid-19 was treated as continuous (reporting means and t-tests) while following recommendations was treated as categorical (reporting %s and Chi-squared), when both used seven-point Likert scales (additionally, %s are reported for knowledge of Covid-19 in Table 1, so it isn't consistently treated as continuous). I'd suggest picking one approach and using it consistently, or justifying this variation in the text.

Response. The measure of COVID-19 knowledge level is now consistently treated as categorical throughout the manuscript.

Comment 4. In the quantitative results section, the authors sometimes helpfully specify "X2 p-value", but other times just report a p-value without clarifying the type of test that the p-value is for; it would be helpful to make this consistent.

Response. The corresponding statistical test for each p-value has been clarified in the text.

Comment 5. For the percentages reported in the qualitative results section, what is the denominator? Is it the percentage of survey respondents overall, or the percentage of survey respondents who left qualitative comments?

Response. For the proportions reported in the qualitative results section the denominator used was the number of students who provided responses to the open-ended question. The number of responses (denominators) for each university is reported in section 3.5, and a sentence has been added to clarify that this was used as the denominator.

Comment 6. The limitations could reflect further on the fact that the two universities might not be entirely comparable, and that differing restrictions in the two countries might also account for some of the variation between them.

Response. A sentence has been added to the strengths and limitations addressing the comparability of the two universities. Differences between the universities are also discussed in the interpretation of the findings (e.g., perceptions of online learning).

VERSION 3 – REVIEW

REVIEWER	Wickersham, Alice King's College London, Institute of Psychiatry, Psychology and Neuroscience
REVIEW RETURNED	09-Nov-2021

GENERAL COMMENTS	I can confirm that my previous comments have been addressed in this latest revision. Many thanks to the authors for being receptive to the feedback, and congratulations on producing this interesting and important paper!