Article information: http://dx.doi.org/10.21037/apm-20-2242

Review Comment

Comment 1: It is unclear to me how the current study directly relates to translational research, given the scope of the Annals of Translational Medicine. How can public health officials use the findings of this research? There are no interventions that explain how officials can increase mask wearing among students. **(I)** The constructs chosen (past behavior, country living in) do not provide any tools or intervention strategies in terms of how to increase mask usage rate. **(II)** How would officials go about changing attitudes, subjective norms, and perceived behavioral control (since they can't change the student's residency or past behavior)?

Reply 1(I):

Thank you for your valuable comments! As for the **past behavior**, we removed the construct (past behavior) from the original hypothesis model (Hypothesis 4) and formed the new hypothesis model (Figure 2), because the TPB model is a better fit than the extension of the past behavior according to the view of Ajzen (2002) who is the founder of TPB and the view of Collins and Carey (2007) who compares the TPB model to an extension using past behavior (habitual). And the country variable was regarded as a control variable in the new model (Figure 4). Thus we do not provide any tools or intervention strategies now.

Changes in the text:

We have modified our text as advised in the part of ***Methods*** and ***Results*** respectively:

Methods:

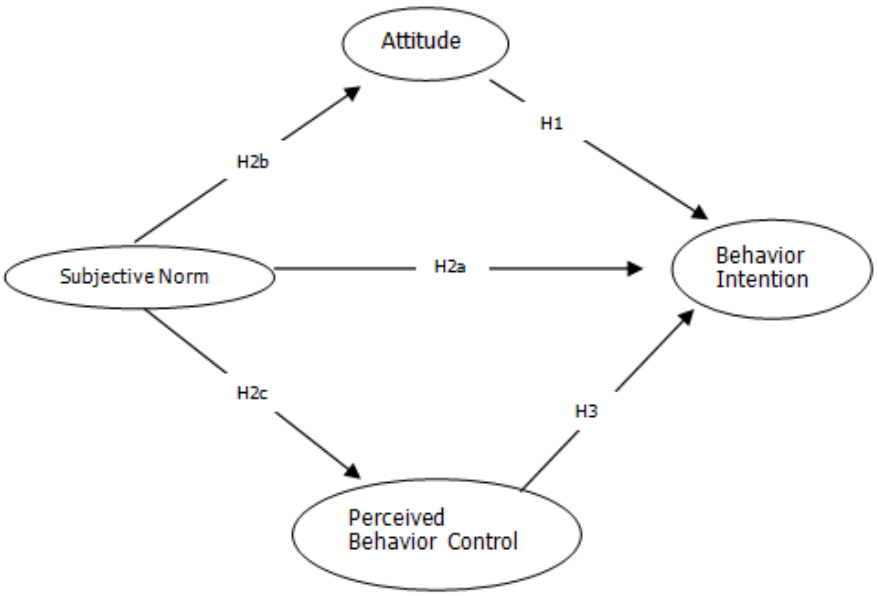

Figure 2 Research model of international students' intention to wear a mask for protection against COVID-19

Results:

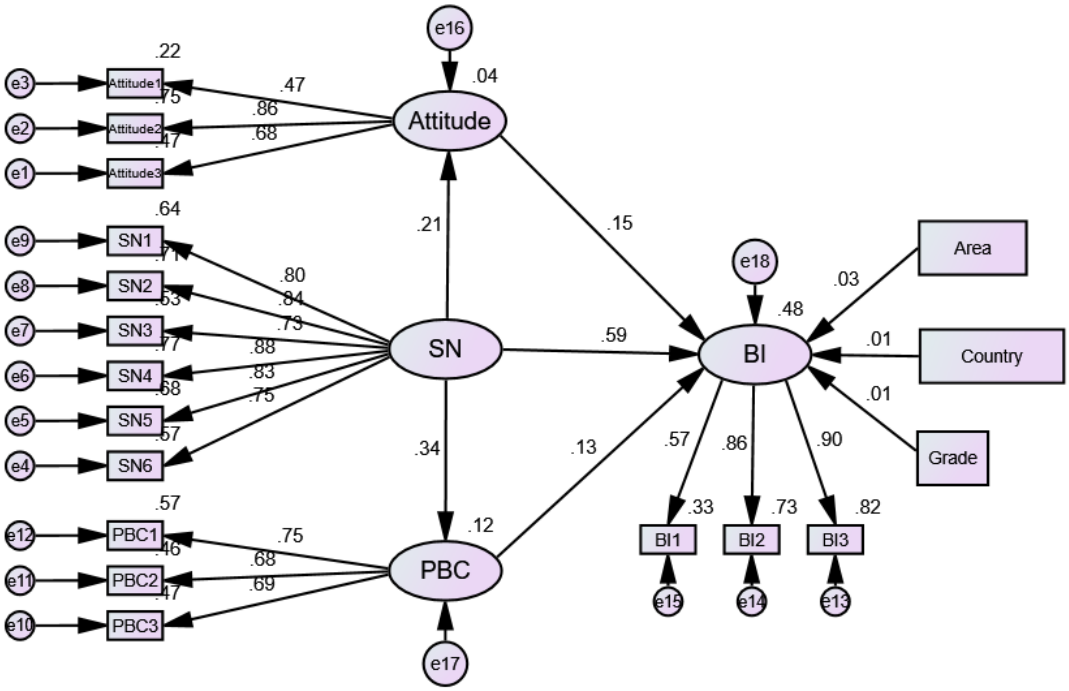

Figure 4 Path diagram for research model (n=477)

Note: BI, behavior intention; PBC, perceived behavioral control; SN, subjective norm; Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents' subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention. Area means the urban area or rural area. Grade included undergraduate and postgraduate. Country means where the international students currently live. The grade ,the area and the country are all performed a dummy variables as a control variable.

(see Page 12, line 249-251) (see Page 25, line 442-450)

Reference:

Ajzen I. Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review*. **2002**;6:107–122.

Collins SE, Carey KB. The theory of planned behavior as a model of heavy episodic drinking among college students. Psychology of addictive behaviors : *Journal of the Society of Psychologists in Addictive Behaviors* **2007**, 21(4), 498–507.

Reply 1(II):

Thank you very much for your comprehensive and incisive advice! We did the following efforts to make the results more orderly and clear. We had deeply analyzed and explained how to improved attitudes, subjective norms, and perceived behavioral control respectively in the part of ***D*iscussion**.

Changes in the text:

SN: Before the international students return, the Chinese government and universities could send some propaganda materials for the importance and usefulness of mask-wearing to their parents and teachers by email or other online tools. Due to the key role of the SN in the model, they could play a significant role in this supervision. Citizens in some countries had been forced to wear a mask otherwise they would get a fine (65). So, after they return to China, the international office of university should develop some appropriate management measures to motivate the students to wear masks to resist the spread of COVID-19. For example, the international office could introduce them to wear a mask as a mandatory requirement whenever they are in or out of the university.

Attitude: Firstly, the Chinese government and universities may public the brochures about the knowledge of COVID-19, which could facilitate the international students' adherence to the infection prevention guidelines. Besides, the officials in the universities should strengthen the education on the preventive role of mask to help the international students reject prejudices and then to enhance their awareness to wear masks.

PBC: Generally speaking, price may be brought down by increasing supply. Therefore, the high price of masks, which was caused by the shortage of supply, could be adjusted by the increasing production. Chinese government and universities could regulate prices and improve production quality, which should eventually provide the convenience of buying masks. Besides, in order to improve the enthusiasm of wearing masks, Chinese government and universities may develop some policies. For instance, the officers in the universities can distribute masks to international students regularly or place a certain amount of masks where necessary, which will reduce the perception difficulty and make them no longer difficult to wear masks.

(see Page 28, line 504-514; Page 29, line 521-527; Page 29, line 535-545)

Reference:

65. Tobol Y, Siniver E, Yaniv G. Dishonesty and mandatory mask wearing in the COVID-19 pandemic. *Economics Letters*,**2020**, 197.

Comment 2: (I) The use of past behavior as an extension in the theory of planned behavior is controversial, and has led to much discussion in the literature. The authors need to address this issue, since Ajzen (2002) himself suggests that past behavior should not be added to the model. See for instance the article by Collins and Carey (2007) that compares the TPB model to an extension using past behavior (habitual), which suggests that the TPB model is a better fit than the extension. **(II)** It is also unclear whether face mask wearing is considered a habitual behavior - if so, you are not directly measuring habitual behavior and would need a good measure for it with latent variables.

Reply 2(I):

Thank you for your valuable comments! In order to improved model fit, we removed the constructs (past behavior) from the original hypothesis model (Hypothesis 4). Because the TPB model is a better fit than the extension of the past behavior according to the view of Ajzen (2002) who is the founder of TPB and the view of Collins and Carey (2007) who compares the TPB model to an extension using past behavior (habitual). So the new hypothesis model was shown in **Figure 2.**

Changes in the text:

We have modified our text as advised in the part of ***Methods.***

(Please see the Figure 2 in reply1(I) or see the Page 12, line 249-251 of the main document)

Reference:

Ajzen I. Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review*. **2002**;6:107–122.

Collins, SE, Carey KB. The theory of planned behavior as a model of heavy episodic drinking among college students. Psychology of addictive behaviors : *Journal of the Society of Psychologists in Addictive Behaviors* **2007**, 21(4), 498–507.

Reply 2(II):

Thank you for your valuable comments! Face mask wearing should not be considered a habitual behavior. Although the death caused by H1N1 virus worldwide, wearing a mask still also was regarded as a physical barrier of freedom and individualism in many countries which indicated that people in many countries didn't have the habit to wear masks according to the article of Wang (2020). So it was not necessary to measure the habit of mask-wearing with latent variables.

Changes in the text:

We have made a clearer explanation in our text as advised in the part of

Introduction:

Although death caused by H1N1 virus worldwide, wearing a mask still also was regarded as a physical barrier of freedom and individualism in many countries (15). (see Page 4, line 90-93)

Reference:

15. Wang J , Pan L , Tang S , et al. Mask use during COVID-19: A risk adjusted strategy. *Environmental Pollution* **2020**, 266(Pt 1):115099.

Comment 3: The authors seem to have developed their own items for the scales. Why not use measures from previous research and adapt them? Otherwise, there are serious questions with the various types of validity, particularly face validity. The measures may be reliable, but I am not convinced that they are valid. More information would be required to ascertain construct, content, face, and criterion validity, including a confirmatory factor analysis. Also, there were 9 measures for subjective norms. Was this definitely a unidimensional measure? Without seeing the measures themselves, it is difficult to determine, and impossible to replicate. I would have recommended using pre-existing scales within the literature, and adapted the items to reflect COVID-19 focus. Also, I don't think it makes sense to report a Cronbach's alpha value of 1.00 for a single measure in Table 3, it would make more sense to leave it blank.

Reply 3:

Thank you so much for your suggestions! We are very sorry for neglecting these issues. Based on the analysis of the current situation of wearing a mask, under the backdrop of COVID-19 and the existing research, combined with TPB, a self-

administered questionnaire was designed. **In order to develop these items**, we reviewed all of the previous research literatures (32,48). In this study, some measures were adapted from previous research. Besides, the usefulness, resource, price etc. of masks in the background of the COVID-19 (as shown in the part of Introduction, Paragraph 1-5) were fully taken into account. Then, the questionnaire was formed.

In order to ensure the reliability and validity, we conducted the Exploratory Factor Analysis **(EFA)** (238 samples) with SPSS V.22.0 and performed Confirmatory Factor Analysis **(CFA)** (239 samples) with AMOS V.24.0 to identify and confirm the factor structure with the 477 samples divided into 2 parts randomly by SPSS V.22.0 . As for the dimension of SN, we removed 3 items because of the similarities with other items to ensure a unidimensional measure before EFA. And we also measured **internal consistency, composite reliability, convergent validity and discriminant validity.** The Final structural model was well reflected by fit indices and Cronbach's α, composite reliability, average variance extracted (AVE), the square roots of the AVE and correlations among variables as shown in Table 2, Table 3, Table 4, Table 5 , Table 6, Figure 3 and Figure 4.

Changes in the text:

We have modified our text as advised in part of ***Methods and Results***:

Methods:

2.2 Instruments and measures

Based on the analysis of the current situation of wearing a mask, under the backdrop of COVID-19 and the existing research, combined with TPB, a self-administered questionnaire was designed. In order to develop these items, we

reviewed all of the previous research literatures (32,48). In this study, some measures were adapted from previous research. Besides, the usefulness, resource, price etc. of masks in the background of the COVID-19 (as shown in the part of Introduction, Paragraph 1-5) were fully taken into account. Then, the questionnaire was formed.

(see Page 12-13, line 252-260)

2.4 Data analysis and statistics

Data were recorded using Microsoft Excel. Descriptive statistics including frequencies and proportion of each variables of demographics, such as gender, age, major, grade, area, number of family members and family income (RMB per month) etc. were calculated using SPSS V.22.0. The hypothesis model was analyzed by the Structural Equation Modeling (SEM) using AMOS V.24.0. Before test the hypothesis model, we performed factor analysis to identified and confirmed the questionnaire's constructs. Exploratory Factor Analysis (EFA) was conducted by SPSS V.22.0. with the index of Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity. Principal component analysis and Varimax were used to obtain the factor structure matrix. Cronbach's $\alpha \geq 0.600$ indicated sufficient internal consistency (49). Confirmatory Factor Analysis (CFA) was performed by AMOS V.24.0. with fit index included Chi-square Value of Minimum Sample/Degree of Freedom (CMID/DF), Root Mean Square Residual (RMR), Standardized RMR (SRMR), Root Mean Square Error of Approximation (RMSER), Goodness of Fit Index (GFI), Normed Fit Index (NFI), Comparative Fit Index (CFI), Tucker-Lewis index (TLI) etc. Composite reliability (CR), convergent validity and the discriminant validity were also assessed by Cronbach's α, average variance extracted (AVE), the AVE square root and the correlations among variables.

(see Page 14-15, line 294-312)

Results:

3.2 Factor Analysis

The data of 477 samples were divided into 2 parts randomly by SPSS V.22.0. One part of the data of 238 samples were conducted to perform the EFA. Another part of data of 239 samples were undertaken to do the CFA to confirm the result of EFA. The reliability and validity of this self-designed questionnaire were also analyzed.

3.2.1 Exploratory Factor Analysis (EFA)

The single sample t-test was conducted by critical ratio (CR) between ranked top 27% of others 16 items scores and the last 27%. The t value of each item were all significant ($p < 0.001$). Pearson correlation coefficient was also significant ($p < 0.01$). The value of KMO was 0.867, and Bartlett's test of sphericity was significant ($\chi^2=2080.587$, $p < 0.001$), which estimated the suitability of EFA (54). The results of EFA on the one part of samples (n=238) were shown in Table 2. Four factors were obtained according to the eigenvalue ≥1, and total variance explained was 69.64 %. We removed the item 20, because it's factor loading was only 0.496 which was far below other values. All factor loadings of other 15 items were above 0.600. And all factors' Cronbach's α were above 0.600 which arrived the minimum criterion and indicated the acceptable internal consistency (49).

Table 2 The results of EFA (n=238)

Factors/Items	Factor loadings	Eigenvalue	Cronbach's alpha

Attitude		1.237	.643
Item1 (Attitude1)	0.731		
Item2 (Attitude2)	0.832		
Item3 (Attitude3)	0.742		
SN		6.360	.929
Item4 (SN1)	0.819		
Item5 (SN2)	0.866		
Item6 (SN3)	0.829		
Item7 (SN4)	0.836		
Item8 (SN5)	0.829		
Item10 (SN6)	0.781		
PBC		1.536	.731
Item13 (PBC1)	0.845		
Item14 (PBC2)	0.789		
Item15 (PBC2)	0.731		
BI		2.009	.824
Item16 (BI1)	0.641		
Item18 (BI2)	0.835		
Item19 (BI3)	0.838		

Note : Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents' subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention.

3.2.2 Confirmatory Factor Analysis (CFA)

The results of CFA on the other part of samples (n=239) were shown in Figure 3 and Table 3. The standardized regression weights of all items were above 0.05. The values of the Critical Ration (C.R.) were all significant ($p<0.001$). The values of the

standard error of estimate were below 2.5, which indicated the model was suitable with the sample. Then, Chi-square Value of Minimum Sample/Degree of Freedom (CMID/DF)= 2.063<5 (55) and fit indices (Root Mean Square Residual (RMR)=0.039<0.05 (56), Standardized RMR (SRMR)= 0.058< 0.08) (57), Root Mean Square Error of Approximation (RMSER)=0.067<0.08 (58), Goodness of Fit Index (GFI)=0.914>0.900 (59), Normed Fit Index (NFI)=0.906>0.900 (60), Comparative Fit Index (CFI)=0.948>0.900 (57),Tucker-Lewis index (TLI)=0.935>0.900 (57)) showed the structure was acceptable.

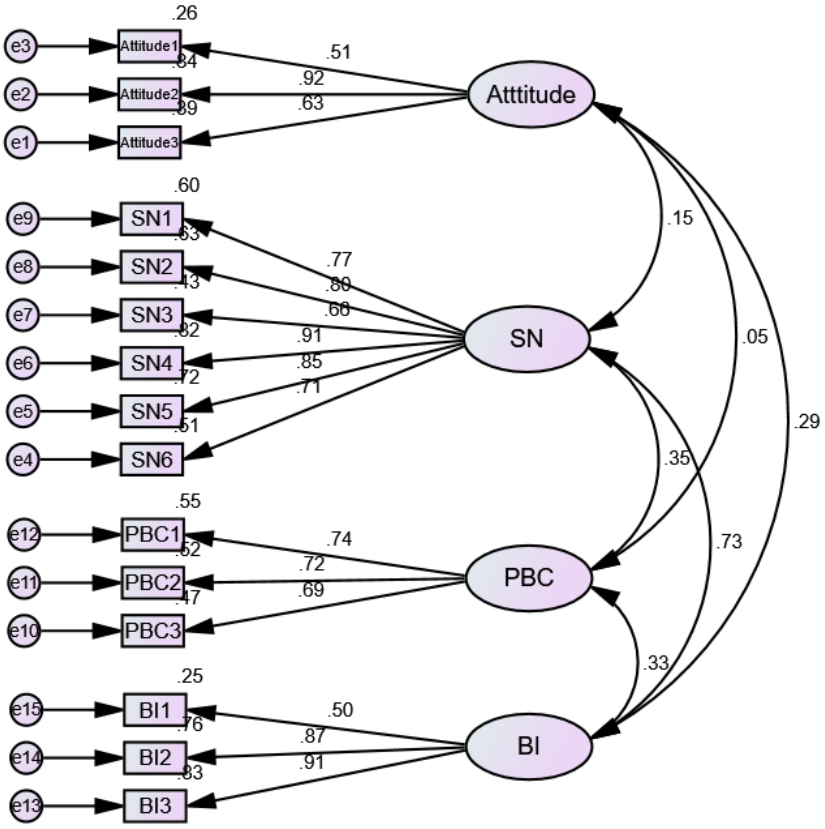

Figure 3 The model of CFA on the Attitude, SN, PBC and BI (n=239)

Note : BI, behavior intention; PBC, perceived behavioral control; SN, subjective norm; Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents'

subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention.

Table 3 The Standardized Regression Weights of the model (n=239)

Path	Standardized Regression Weights	S.E.	C.R.	P
Attitude1<--- Attitude	0.513	0.171	6.804	0.000
Attitude2<--- Attitude	0.918	0.261	6.655	0.000
Attitude3<--- Attitude	0.627			
SN1<--- SN	0.774	0.084	11.561	0.000
SN2<--- SN	0.796	0.084	11.891	0.000
SN3<--- SN	0.657	0.101	9.809	0.000
SN4<--- SN	0.907	0.075	13.502	0.000
SN5<--- SN	0.846	0.086	12.639	0.000
SN6<--- SN	0.715			
PBC1<--- PBC	0.745	0.152	8.267	0.000
PBC2<--- PBC	0.723	0.159	8.237	0.000
PBC3<--- PBC	0.686			
BI1<--- BI	0.502	0.079	8.045	0.000
BI2<--- BI	0.869	0.067	16.285	0.000
BI3<--- BI	0.910			

Note : Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents'

subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention.

As shown in Table 4, the reliability and validity of the questionnaire included Cronbach's α, composite reliability, average variance extracted, and correlations among variables. Cronbach's α≥0.700 indicated sufficient internal consistency or reliability, and that composite reliability was adequate. Meanwhile, the value of average variance extracted (AVE) showed that the convergent validity was acceptable. The discriminant validity of the questionnaire was assessed. It is evident that the AVE square root of each construct is higher than the absolute value of their correlation (61); the cross-loadings show that all items loaded on their respective constructs are higher than on the other constructs and the cross-loadings differences are above the threshold of 0.10 (62).

Table 4 Correlations among variables (n=239)

	R	CR	AVE	BI	PBC	SN	Attitude
BI	0.786	0.817	0.612	0.782			
PBC	0.757	0.762	0.516	0.327	0.718		
SN	0.901	0.906	0.619	0.725	0.348	0.787	
Attitude	0.702	0.738	0.500	0.294	0.053	0.150	0.707

Note: BI, behavior intention; PBC, perceived behavioral control; SN, subjective norm; R, Cronbach's α; CR, composite reliability; AVE, average variance extracted. The square roots of the AVE are shown on the diagonal and italicized elements, below which are the correlations between the construct's values.

3.3 Evaluation of structural model

Chi-square Value of Minimum Sample/Degree of Freedom (CMID/DF)= 3.527 and fit indices (Root Mean Square Residual (RMR)=0.036, Standardized RMR (SRMR)=0.068, Root Mean Square Error of Approximation (RMSER)=0.073, Goodness of Fit Index (GFI)=0.904, Normed Fit Index (NFI)=0.881 , Comparative Fit Index (CFI)=0.911, Tucker-Lewis index (TLI)=0.935) indicated the model was acceptable.

The model's measurement results were shown in Table 5 (n=477), Figure 4 (n=477) and Table 6 (n=477). The values of the Critical Ration (C.R.) were all above 3.25 which indicated the estimates were significant ($p<0.001$). As shown in Figure 4 (n=477), about 48% of variance in the intention to wear a mask was explained: attitude is 4% and PBC is 12%, and the path diagram showed how A, SN and PBC predicted the BI to wear a mask.

Table 5 The Standardized Regression Weights of the model (n=477)

Path	Standardized Regression Weights	S.E.	C.R.	P
Attitude1<--- Attitude	0.468	0.116	8.754	0.000
Attitude2<--- Attitude	0.864	0.161	9.327	0.000
Attitude3<--- Attitude	0.685			
SN1<--- SN	0.801	0.051	18.140	0.000
SN2<--- SN	0.845	0.054	19.279	0.000
SN3<--- SN	0.728	0.062	16.267	0.000
SN4<--- SN	0.879	0.049	20.175	0.000

SN5<--- SN	0.826	0.057	18.770	0.000
SN6<--- SN	0.754			
PBC1<--- PBC	0.753	0.116	11.362	0.000
PBC2<--- PBC	0.682	0.116	11.196	0.000
PBC3<--- PBC	0.689			
BI1<--- BI	0.573	0.056	13.150	0.000
BI2<--- BI	0.856	0.050	21.367	0.000
BI3<--- BI	0.904			

Note：Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents' subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention.

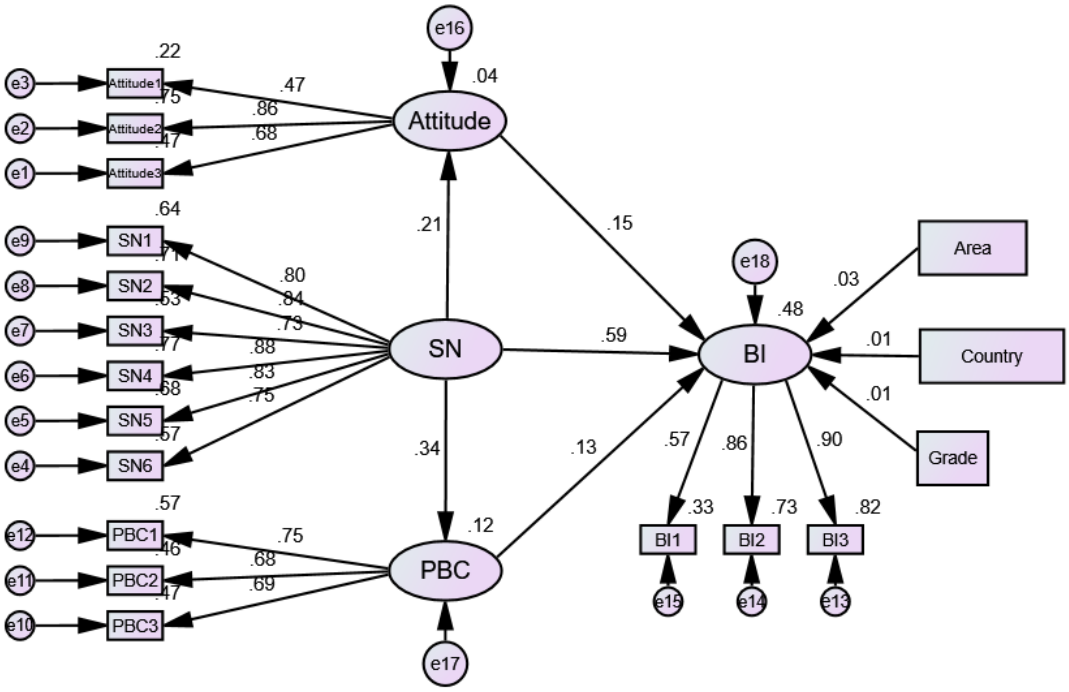

Figure 4 Path diagram for research model (n=477)

Note: BI, behavior intention; PBC, perceived behavioral control; SN, subjective norm; Attitude1–Attitude3 denote the three items used to measure the respondents' attitudes; SN1–SN6, the six items used to measure the respondents' subjective norm; PBC1–PBC3, the three items used to measure the respondents' perceived behavioral control; BI1-BI3, the three items used to measure the respondents' behavior intention. Area means the urban area or rural area. Grade included undergraduate and postgraduate. Country means where the international students currently live. The grade, the area and the country were all performed as control variables.

Table 6 The Path coefficients of the BI to wear a mask (n=477)

Causal Variables	Outcome Variable	Standardized Total Effect	Standardized Direct Effect	Standardized Indirect Effect
Attitude	**BI**	.148	.148	.000
SN	**BI**	.664	.588	.075
	Attitude	.210	.210	.000
	PBC	.340	.340	.000
PBC	**BI**	.131	.131	.000

Note: BI, behavior intention; PBC, perceived behavioral control; SN, subjective norm.

Figure 4 (n=477) and Table 6 (n=477) showed the results of hypothesis testing based the model. Attitude and PBC were respectively directly related to BI positively. SN also had a direct effect on Attitude , PBC and BI. Besides, SN also indirectly promoted the BI through Attitude and PBC. In addition, the effects of the living area, countries where international students living and grade were controlled. Country was regarded as a dummy variable that was set to a value of 1 if the international students

currently lived in China, 0 otherwise, in the SEM. The same code was applied to Grades and Living area. The standardized regression weights were 0.033 ($p>0.05$), 0 .010 ($p>0.05$),0.013 ($p>0.05$) respectively which indicated they were not significantly related to the BI to wear a mask. Thus this hypothesis model was strongly confirmed and all hypotheses were supported.

(see Page 19-26, line 358-465)

References:

32. Hansstein FV, Fabián E. Exploring motivations behind pollution-mask use in asample of young adults in urban China. Globalization and Health 2018, 14, 2-10.
48. Li Z. The Investigate of The Risk Behavior of Driving on The Road in Chiang Mai. Master Degree, Kunming University of Science and Technology, Kunming, 2018.
49. Griethuijsen RALF, Michiel W, Helen H, et al. Global patterns in students' views of science and interest in science. Res. Sci. Educ 2015, 45, 581–603.
54. Suh, Y. The performance of maximum likelihood and weighted least square mean and variance adjusted estimators in testing differential item functioning with nonnormal trait distributions. Struct. Equ. Model. A Multidiscip. J 2015, 22, 568–580.
55. Byrne BM. Structural equation modeling with AMOS: Basic concepts, applications, and programming. Springer: Routledge, New York, 2010.
56. Byrne BM. Structural equation modeling with LISREL, PRELIS, and SIMPLIS: basic concepts, applications, and programming. Springer: Psychology Press, New York, 2013.
57. Hu Li-tze, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling 1999, 6, 1-55.
58. MacCallum RC, Browne MW, Sugawara HM. Power analysis and determination of sample size for covariance structure modeling. Psychol Methods 1996, 1:130-149.

59. Miles J, Shevlin M. A time and a place for incremental fit indices. Personality & Individual Differences 2007, 42:869–874.
60. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin 1980, 88: 588-606.
61. Fornell C, Larcker D. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 1981, 18, 39-50.
62. Gefen D, Straub D. A Practical Guide to Factorial Validity Using PLS-Graph: Tutorial and Annotated Example. Communications of the Association for Information Systems 2005, 16, 91-109.

Comment 4: How did you measure "country living in" and does this mean their nationality or their residency status? How many countries were identified and how was this information coded? Was this measure categorical, and if so was it nominal or ordinal? If nominal, how was the SEM run? It is unclear how the path from "country living in" to behavioral intention should be interpreted, particularly if the data is nominal. And if it is nominal, shouldn't it be a multigroup analysis?

Reply 4:

Thank you very much for your comprehensive and incisive advice! We did the following efforts to make the results more orderly and clear. "Country living in" was measured by the item "Which country do you currently live in ? ". It means the international students' residency status. Their nationalities were reflected by another item: Nationality. 52 nationalities and 37 countries (currently residing) were identified. We removed the H5 of original model, and the variable named "currently living in" was regarded as a dummy variable that was set to a value of 1 if the

international students currently lived in China, 0 otherwise, in the SEM as shown in Figure 4. The same code was applied to Grades and Living area.

Changes in the text:

We have made a clearer explanation in our text as advised in the part of ***Results.***

Results

In the web-based survey, of the 550 questionnaires distributed, 492 were returned. Among the 700 questionnaires, 477 were usable for a response rate of 86.7% . According to the item "Which country do you currently live in ", 87 of them still were in China, and 390 had been in home (the countries currently live in).

3.1Demographics and characteristics

According to the report of international education in China, 59.95% of the international students in mainland of China came from Asia, followed by Africa (16.57%), Europe (14.96%), Americas (7.26%) and Oceania (1.27%) (53). The Asian students were mainly from Korea, Thailand, Pakistan, India, Laos, Kazakhstan, Vietnam, Bangladesh, Malaysia etc. (53). The final database of this study included a total of 477 records. The participating international students came from 52 countries across 6 continents, namely Asia, Africa, Europe, North America, South America, and Oceania, and were currently residing in 37 countries across the globe, which covered the top 13 countries of the international students' nationality or residence (53). The sample framework is generally consistent with the report of international education in China in 2018.

(see Page 16-17, line 329-347)

3.3 Evaluation of structural model

(Please see the Figure 2 and the Figure 4 of **Reply 1(I)** or see the Page 12, line 249-251 and Page 25, line 442-450 of the main document, and the text description behind the Table 6 as following:)

Figure 4 (n=477) and Table 6 (n=477) showed the results of hypothesis testing based the model. Attitude and PBC were respectively directly related to BI positively. SN also had a direct effect on Attitude , PBC and BI. Besides, SN also indirectly promoted the BI through Attitude and PBC. In addition, the effects of living area, countries where international students living and grade were controlled. Country was regarded as a dummy variable that was set to a value of 1 if the international students currently lived in China, 0 otherwise, in the SEM. The same code was applied to Grades and Living area. The standardized regression weights were 0.033 ($p>0.05$), 0 .010 ($p>0.05$),0.013 ($p>0.05$) respectively which indicated they were not significantly related to the BI to wear a mask. Thus this hypothesis model was strongly confirmed and all hypotheses were supported.

(see Page 25-16, line 454-465)

References:

53. Ministry of Education of the People's Republic of China. International students studying in China in 2018, April 12, **2019**.Available online: http:// www.moe.gov.cn/jyb_xwfb/gzdt_gzdt/s5987/201904/t20190412_377692.html (accessed October 8,2020).

Comment 5: Hypothesis 5 is unclear. Why do the authors expect student residency to impact behavior intention? Is this behavior intention while at the university or when at home? The paper seems to suggest that most of the students were not in China when the study was conducted (page 7). Are the authors measuring their intention to wear a mask while in their own country, while in China, or while at the university. For that

matter, when are they supposed to be wearing a mask, since the recommendations vary by country - while outdoors all the time, outdoors when around others and unable to social distance, when indoors in public spaces...? It is also unclear whether their intention to wear a mask would differ depending on where they are (on campus, in China, back in their country). If it is based on universities in China, then are the mask-wearing policies across the universities the same or do they differ?

Reply 5:

Thank you so much for your suggestions. We are very sorry for neglecting these issues. In order to make the results more orderly and clear, we have made the following revisions: Because there was no sufficient evidence, we removed the H5 of the original model, and revised the research model as shown in Figure 2. And now the variable named countries where the international students currently lived in was performed as a control variable in the SEM (Figure 4). According to the item "Which country do you currently live in, "87 of them still were in China, and 390 had been in home. "The home" means the countries where the international students currently live in. We focus on the outdoors and indoors where there is high risk of infection. We measured the international students intention to wear masks in different countries (e.g. their own countries or China) and in different locations (e.g. the subway, the theater, the elevator, a park, theater, library, classroom, square, business or other places where crowds gather etc.)

There are unified mask-wearing policies across the universities in China. In order to effectively prevent and control the epidemic in campus, the Ministry of Education of the People's Republic of China issued the *Guidelines on COVID-19 Prevention and Control in Higher Education Institutes* (7). It required the students to

wear masks in the public transportation, the classroom, and the elevator etc. Besides, the college students must conform to rigorous mask-wearing standards laid down by the Ministry of Education of the People's Republic of China (8).

Changes in the text:

We have modified our text as advised in the part of ***Introduction(Paragraph 3), Motheds and Results.*** Please see the Figure 2 and the Figure 4 of **Reply 1(1)** or see Page 12, line 249-251 and Page 25, line 442-450 of the main document**,** and the text description as following:

Introductio*n (Excerpt from Paragraph 3)*

And there are unified mask-wearing policies across the universities in China. In order to effectively prevent and control the epidemic in campus, the Ministry of Education of the People's Republic of China issued the *Guidelines on COVID-19 Prevention and Control in Higher Education Institutes* (7). It required the students to wear masks in the public transportation, the classroom, and the elevator etc. Besides, the college students must conform to rigorous mask-wearing standards laid down by the Ministry of Education of the People's Republic of China (8).

(see Page 4, line 79-86)

Motheds:

2.2 Instruments and measures

We measured international students intention to wear masks in different regions (e.g. their own countries or in China) and in different locations (e.g. the subway, the theater, the elevator, a park, theater, library, classroom, square, business or other places where crowds gather etc..)

(see Page 13, line 274-278)

Results:

In the web-based survey, of the 550 questionnaires distributed, 492 were returned. Among the 700 questionnaires, 477 were usable for a response rate of 86.7% . According to the item “Which country do you currently live in “, 87 of them still are in China, and 390 of them have been in home (the countries currently live in).

(see Page 16, line 329-334)

References:

7. Guidelines on COVID-19 Prevention and Control in Higher Education Institutes; People’ Medical Publishing House: Beijing, China, 2020.
8. Ministry of Education of the People's Republic of China. The Programme for Prevention and Control of the Epidemic Disease of COVID-19 in Colleges and Universities, April 13, **2020**. Available online: http://www.moe.gov.cn/jyb_xxgk/moe_1777/moe_1779/202010/t20201021_495955.html (accessed December 1,2020).

Comment 6: I don't see an extensive reporting of goodness of fit measures. I did see SRMR, but did not see a report on the model's chi square, RMSEA, and CFI.

Reply 6:

Thank you for your valuable comments! We replace the partial least square structural equation model (PLS-SEM) performed by PLS 3.2.8 with covariance-based structural equation modeling (CB-SEM) conducted by Amos V.24.0. to analyze the hypothesis model. Exploratory Factor Analysis (EFA) (238 samples) and Confirmatory Factor Analysis (CFA) (239 samples) were performed to identify and confirm the factor structure. The structural model was well reflected by Kaiser-

Meyer-Olkin (KMO) , Bartlett's test of sphericity, Cronbach's α, composite reliability, average variance extracted (AVE), the square roots of the AVE, correlations among variables and fit indices (Chi-square Value of Minimum Sample/Degree of Freedom (CMID/DF), Root Mean Square Residual (RMR), Standardized RMR (SRMR), Root Mean Square Error of Approximation (RMSER), Goodness of Fit Index (GFI), Normed Fit Index (NFI), Comparative Fit Index (CFI), Tucker-Lewis index (TLI) etc.) as shown in Table 2, Table 3, Table 4, Table 5 , Table 6, Figure 3 and Figure 4.

Changes in the text:

We have modified our text as advised in part of ***Methods and Results***:

Methods

2.4 Data analysis and statistics

Results

3.2 Factor Analysis

3.2.1 Exploratory Factor Analysis (EFA)

3.2.2 Confirmatory Factor Analysis (CFA)

3.3 Evaluation of structural model

Please see the **Reply 3** or see the Page14-15, line 394-312 and Page 19-26, line 358-465 of main document.

Comment 7: There are numerous grammatical errors in the manuscript.

Reply 7:

We apologized for the grammatical errors and spelling errors in this manuscript. We have checked and modified the grammatical errors. Moreover, to improve the quality of the English, we had the manuscript edited by an English native speaker.