CSV file “Data with pure face base on general score”

The 2nd last column (“Bio\_Score”) is the result of the participants’ culture standard based on general scores calculated by the scaled mean of response in the ‘cultural measure’ question.

The last columns (“Bio\_Birth”) represent the result of the participants’ culture standard base on their FBorn, MBorn, Born place

Number represents different type of culture standard

0 represents pure dignity

1 represents pure face

2 represents bio-culture with dignity and face

NaN (Blank in “Bio\_score”) represents the culture which are excluded from the research

As 6 questions for ‘dignity’, 5 questions for ‘face’, 4 questions for ‘honour’, scale (6+5+4 = 15) is applied to reduce the bias brought by the number of questions of each type of culture.

Calculating the scale mean of dignity: sum(dignity) \* (6/15)

Calculating the scale mean of face: sum(face) \* (5/15)

Calculating the scale mean of honour: sum(honour) \* (4/15)

(Estimated) Initial Threshold for distinguishing whether a general score is high enough:

Dignity Threshold = 10, Face Threshold = 8, Honour Threshold = 7

### From my perspective, the result from the general score might not exactly match what we’ve got from the birth place (Father, Mother, Me) results. Please read the csv file (especially the last 2 columns for classification results) and figure out what might be a better way to classify the participants (based on the current results if possible).

## As we figure out the exact way of classification based on the result we’ve got in the last 2 columns, we could then calculate the average score of informational and emotional unethical tactics for each participant for further research usage.

### If you have any better idea in calculating the general score for each type of culture or a better selection of threshold value, please let me know in the Messenger and I could change the code immediately to generate a new result csv file for checking

# Besides, if anyone is good at doing the reliability analysis, please help me with the uploaded csv file as I am completely unfamiliar with that concept (Coding might not be difficult to me, but the unlearned concept would surely trick me 😭).

Explanation of Data Cleaning code

1. Drop all participant who only preview the survey but does not attend
2. Drop the uninformative columns (unrelated with the research)
3. Convert the string columns with numerical value into a float value for calculation
4. Drop the participants with too much missing value in SINS scale questions
5. Generate a list of collecting dignity culture locations
6. Generate a list of collecting face culture locations
7. Use the element in these 2 list to match with participants’ FBorn, MBorn, Born location
8. If FBorn, MBorn, Born are all in dignity, Bio\_Birth = 0 (Pure dignity birth)
9. If FBorn, MBorn, Born are all in face, Bio\_Birth = 1 (Pure face birth)
10. If FBorn, MBorn, Born contains dignity and face, Bio\_Birth = 2 (Bio birth)
11. Else case, discard directly
12. After the above process, 157 participants are kept
13. Calculating the scale mean of dignity: sum(dignity) \* (6/15)
14. Calculating the scale mean of face: sum(face) \* (5/15)
15. Calculating the scale mean of honour: sum(honour) \* (4/15)
16. Set Dignity Threshold = 10, Face Threshold = 8, Honour Threshold = 7
17. Record Bio\_Score = 0 if only scale mean of dignity > threshold
18. Record Bio\_Score = 1 if only scale mean of face > threshold
19. Record Bio\_Socre = 2 if both and only scale mean of dignity and face > threshold
20. Else case, record as NaN (Blank)
21. After the above process, 157 participants are kept (discarding code is not processed)