

**All datafiles:**

ID should be capitalized (for example in CBCL files, ID is "id"). This may cause problems in case sensitive programs.

**ACS:**

For when we distribute data (raw and derived) to grad students, make a note or correct the "raw data" file to match the pre/post coding schema (so that someone with this raw data can derive the summary score we included in the data)

**CBCL:**

1. change the variable names that ends with "\_0" which indicates "baseline/pre-treatment" to "2" indicating "FU6"; For FU12, change "\_0" to "\_3" or or use "\_FU6" and "\_FU12"

2. variable names should have the same format, in CBCL FU 6 profile data, these three variables should be "CBCL\_113\_a\_2. CBCL\_113\_b\_2, CBCL\_113\_c\_2"

3. for CBCL, we did control for sex and age in the pre post analysis, and the T score used sex and age norms, so just curious what you meant for accounting for sex and age in one of the emails?

4. small inconsistency issues: for example, for the CBCL subscale depressed/anxious, in FU6 dataset, this variable was called "CBCL\_anxious". In FU12 this subscale was called "CBCL\_anxiety"

5. need to create an additional CBCL subscale, it's the 7-item adolescent sleep wake scale composite, using the following code, for both FU6 and FU12

```
egen CBCL_sleep_composite_FU12 = rowtotal (CBCL_47_3 CBCL_54_3 CBCL_76_3 CBCL_77_3 CBCL_92_3 CBCL_100_3  
CBCL_108_3), missing
```

**CDRS:**

- variables names should end with \_FU6 or \_FU12, this helps with merging..

**CMEP:**

- I find sometimes the scored file contains the raw data, others only has the sum score. might be a good idea to have the raw data + summary score in the same file (or consistent for all measures). that way when we send these data to others, it'll be consistent/not confusing.

- For variable naming, CMEP\_6\_x should actually be CMEP\_x\_6 or CMEP\_x\_FU6. The last bit should indicate the time point.

**MAQ and PSQI**

- need to add \_FU6, FU12

- ID variable change from "StudyID" to "ID"

**Sleep diary**

- ID variable change from id to ID

- ending of variable names should be \_FU6 or FU12

- about 1000 extra rows at the end of both files; this causes merging problem (these are treated as extra data, though all missing). need to correct this.

**PHQ**

- sum variable name should be "PHQ\_sum\_FU6"

**SSS**

- should end with "6" indicating FU6

**YSAS**

- this is great. is the last sum score "YSAS\_SOC\_FU12" a summary score of all?

**Other random qu:**

- Also needs sex and age at FUs. Sex should be the same, but age would be different?

- Does the codebook for FU6 and FU12 need to be updated?

- Could you confirm the following IDs, are not in the study (but these IDs are in the sleep diary data files). My thought is that if they are not in the study, then we can delete these IDs/rows.

ID  
21  
145  
267  
103  
271  
86  
181

189  
100  
254  
220