**Documentation for PPVT Person Scoring Program**

The scoring process uses three files that each need to be saved to the same file location:

* ***PPVT Person Scoring SAS Input.sas*** 🡪 This SAS syntax file will be edited by the user to describe the characteristics of the dataset to be scored and to select analysis options.
* ***PPVT MCMC Person Scoring.sas*** 🡪 This SAS syntax file actually does the MCMC scoring, and should not be edited by the user. It will be invoked automatically when running the *PPVT Person Scoring SAS Input.sas* syntax file.
* ***Itemparms459.sas7bdat*** 🡪 This SAS data file contains the item parameters needed to conduct the MCMC scoring, and should not be edited by the user.

Below is a description of the data-related options that should be specified by the user, as given in the *PPVT Person Scoring SAS Input.sas* syntax file. For each option, the user should change the value after the “=” on the %LET statement for that option to describe the data to be scored.

* **FILESAVE** 🡪 This is the path to the file directory in which these three files and the data to be scored are stored. The value of the path (e.g., “*F:\PPVT Analysis*” as given in the example) is case-sensitive and space-sensitive and must be specified exactly to be read correctly.
* **RESPFILE** 🡪 This is the name of the SAS dataset that holds the responses to each item for each person, as named by the user. It must be in SAS format (with the extension .sas7bdat). It must contain a single variable to uniquely identify each case to be scored. In addition, it must contain responses to the PPVT items from one or more of the three forms (3a, 3b, and R-M), using these naming conventions, in which the item number is given after the underscore:
  + **For PPVT-3 Form A 🡪** ppvt3a\_1 ppvt3a\_2 …. ppvt3a\_204
  + **For PPVT-3 Form B 🡪** ppvt3b\_1 ppvt3b\_2 …. ppvt3b\_204
  + **For PPVT-R Form M 🡪** ppvtRm\_1 ppvtRm\_2 …. ppvtRm\_175
* **THETAFILE 🡪** This is the name of the SAS dataset to be created by the scoring program, as named by the user. This data file will be saved to the same directory as the other three files. It will contain three variables: the original identification variable, the estimated theta score for PPVT vocabulary ability for each case (theta\_est), and the standard error (SE) for the estimated theta score for the PPVT vocabulary estimate (theta\_se). One can construct 95% confidence intervals that describe the variability in the individual ability estimates as: theta ± 1.96\*SE.
* **FORMUSE 🡪** If a case has responses to more than one form, responses for any common items across forms can only be included once. In the event of such overlap, the user can select which form should contain the responses for any common items (choose 3a, 3b, or Rm as values). If there is no such overlap, Rm can be selected as a default.
* **PERSONID 🡪** This is the name of the variable that uniquely identifies each case in the data to be scored. It will be included in the THETAFILE containing the resulting scores for each case so that the scores can be merged back into the user’s other data for each case.
* **NSUBJECTS 🡪** This is the total number of cases to be scored in the analysis (i.e., rows of data).

Below is a description of the estimation-related options for the MCMC scoring algorithm. Users are encouraged to use the default values unless they have informed reasons to change them.

* **ITEMFILE 🡪** This is the name of the SAS data file included with this program that contains the item parameters that will be used in the MCMC scoring. The current name is *itemparms459*.
* **MCMCSEED 🡪** This is the initial value for the MCMC seed. The default is 0, which sets the value of the seed to that of the system clock when the execute button is pressed, essentially making the draws a completely random process. Replication of analyses with the same (non-0) seed will produce identical results whereas replication of analyses with a 0 seed will produce results that differ slightly (typically in the last decimal place) due to the random nature of the process.
* **NBURNIN 🡪** This is the number of burn-in iterations for the MCMC chain. The default is 1000. This dictates how many iterations are discarded, meant to remove any non-converged values from the result. Larger values take longer, but produce more stable results.
* **NTUNE 🡪** This is the number of tuning iterations for the MCMC chain. The default is 500. This is an internal SAS PROC MCMC specification that helps dictate how fast the program will converge. We do not recommend changing this value.
* **NMCMC 🡪** This is the number of MCMC iterations, excluding the burn-in. The default is 2000. Larger values will produce more stable results.
* **NTHIN 🡪** This is the thinning interval for the MCMC iterations. The default is 10. Larger values will produce results with more stable standard error estimates.

From this point forward in the *PPVT Person Scoring SAS Input.sas* syntax file, nothing else needs to be modified. When this syntax file is executed, it will send this input information to the *PPVT MCMC Person Scoring.sas* syntax file and execute that file automatically via the %INCLUDE statement. Each case to be scored could take several seconds, and so the total time to execute the program depends on the number of cases to be scored. When the program is finished, a SAS data file with the name specified on the THETAFILE option will appear in the directory location specified by the FILESAVE option. Those estimated theta scores can then be used in place of PPVT raw scores in future analyses.

Also included with the program is a SAS data file called “exampledata.sas7bdat” that can serve as an example of how users should specify the options in the *PPVT Person Scoring SAS Input.sas* syntax file. There are 8 cases to be scored in the example file. Cases 1 and 2 have item responses only for the PPVT-3 Form A. Cases 3 and 4 have item responses for both the PPVT-3 Form A and PPVT-R Form M. Cases 5 and 6 only have responses for the PPVT-R Form M. Finally, cases 7 and 8 have item responses for both the PPVT-3 Form B and the PPVT-R Form M. Thus, this example illustrates that this scoring program can provide comparable PPVT vocabulary ability estimates regardless of which of the three forms are used.

Summary of steps for this process:

1. **Prepare the item response data file 🡪** The file should be constructed as described in the RESPFILE option, ensuring that all item responses are either 0=incorrect, 1=correct, or missing.
2. **Prepare the input syntax file 🡪** Edit the options in the *PPVT Person Scoring SAS Input.sas* syntax file to describe the data to be scored by the program.
3. **Run the input syntax file 🡪** Once all options have been specified, the *PPVT Person Scoring SAS Input.sas* syntax file should be executed by selecting the “running man” icon from the SAS editor toolbar. In doing so , the *PPVT MCMC Person Scoring.sas* syntax file will be executed as well, and a data file of PPVT scores will be saved the directory specified in FILESAVE.