

Guidance Manual for Running and Interpreting the NM FAS Syntax

FY 2011



Pacific Institute for Research and Evaluation

Table of Contents

1. Data Cleaning	3
2. Module A: CORE Data Screening.....	3
3. Inconsistency Checks	4
4. Create Baseline Demographic Variables.....	4
5. Baseline Recodes	5
6. Post-Test Recodes	5
7. Baseline Scale Scores	5
8. Post-Test Scale Scores.....	6
9. Differences between Baseline and Post-Test frequencies for ATOD use (Dichotomous variables).....	6
10. Differences between Baseline and Post-Test frequencies for ATOD use (Dichotomous variables) ONLY among those reporting ATOD use at Baseline	6
11. Differences between Baseline and Post-Test frequencies for ATOD use (Ordinal variables).....	7
12. Reliability checks for Module A Baseline Scales.....	7
13. Reliability checks for Module A Post-Test Scales	8
14. GLM Analyses for Module A.....	9
15. Module B Analyses	12
16. Module C Analyses	13
17. Module D Analyses.....	14
18. Module E Analyses	14

If you have questions or encounter problems, contact Mary Choe at mchoe@pire.org or 919-265-2638.

Instructions for Running and Interpreting SFS Syntax

The syntax file is annotated to guide the user through each data step. Data steps are organized in numbered sections in the syntax file. To run a data step, highlight the syntax with the mouse and then press the blue triangle on the toolbar to run the syntax. It is recommended that the user run one step at a time and check the output before beginning the next step. This is especially important following the Data Screening cleaning and screening steps as the data will need to be cleaned if there are any problems indicated in the results of these steps. *Remember that the results are only as good as the data, so take the extra time to thoroughly review and clean the data.*

It is a good idea to record user commands in the log. Click on Edit, then Options, and then Viewer to locate the “Display Commands in Log” option and make sure it is checked. When output is generated, the commands will appear first.

1. Data Cleaning

Begin by making a copy of your data file and placing it somewhere safe. Then make sure that you have entered participant id numbers at post-test for those you have post-test data for. This syntax selects all cases where Baseline and Post-Test Participant ID match and will delete all other cases. In other words it selects matched pairs. **IT IS ESSENTIAL THAT AFTER YOU RUN THIS SYNTAX, YOU RE-NAME YOUR DATA FILE.** Otherwise you may lose your original data set. If you already have only matched pairs, then this is less important.

2. Module A: CORE Data Screening

The purpose of this step is to make sure there are no data entry errors or a great deal of missing data for any variables. Make sure you review your data and clean any errors before creating new variables or scales in subsequent steps. This syntax runs cross-tabs for baseline by post-test Module A variables. This allows you to look at the baseline variable against the post-test variable. It also includes those typically excluded responses such as “legitimate skips”. Make sure you check the following for each cross-tabulation: 1) The # of missing for each measure. A lot of missing data may indicate that the question is sensitive question or a bad question. Alternatively, you may have participant IDs with no data in which case you want to remove them from the data file. 2) Values that are out of range of the responses. If the responses in the codebook are coded 0 to 3 and the crosstabs indicate values from 1 to 4 then you need to recode the data. Otherwise, later you will end up with errors in your estimates.

Step: Run Cross-tabs syntax. Check output to make sure data have been entered correctly. If there is a lot of data missing, there may be a data entry mistake or an issue with those items.

3. Inconsistency Checks

Technically, this step is a continuation of the data cleaning process that started in section one. First flags are created for inconsistent variables at baseline and post-test (e.g., respondent reports that they haven't used non-prescribed prescription medications in general and then in the subsequent questions report using a certain type of non-prescribed prescription medication like Oxycodone). This means the data are inconsistent. We check whether there are inconsistencies within each test (ie., within the baseline test and within the post-test). New variables are created for each check and then a frequency is run to identify the number of inconsistencies that exist for that substance.

Step: Run the syntax as listed in the "Inconsistency Checks" section.

For each record that has a "1" in the new variables created ("adiscrep", "adiscrep_2", ...), clean the data if possible by returning to the original survey data and confirming accurate response. If the discrepancy cannot be resolved, set the data to missing for both variables and then be sure to save the changes in the dataset. The syntax will repeat these steps for the other substance use variables. Repeat the cleaning process. *You don't want to include inconsistent data in analyses if at all possible because it can lead to inaccurate results.*

One suggestion is to create a table to track the inconsistencies. For example:

	Baseline	Post-test
Alcohol use	2	1
Prescription drugs	3	1

For something as minimal as this, it may not be necessary to keep track of these, but it's helpful to see if a particular substance seems to have more discrepancies than another and it helps to decide whether you need to go back and check the data again, because cases with discrepancies will be dropped from substance specific analyses later on.

4. Create Baseline Demographic Variables

This syntax recodes q04-q04h into one Ethnicity variable with four categories: Non-Hispanic White (1), Hispanic (2), non-Hispanic American Indian or Alaska Native (3), and all Others (4).

Step 4A: Run syntax to create the new race/ethnicity variable

The next syntax provides descriptive statistics for the demographic variables at baseline only including age, race/ethnicity, language spoken at home, education level, work status, and Medicaid status (both child and parent). It also will produce the average (or mean), standard deviation and range for the age categories. These will go into Table 1 or just below Table 1 in the FAS reporting template.

Step 4B: Run frequencies and descriptive statistics on baseline measures

Check results to make sure that this corresponds to what you think these data should look like. For example, if you are analyzing data from a predominantly Native American population, but there seem to be more Non-Hispanic whites than Native Americans, check the coding in the data set and make sure the syntax looks correct as well.

5. Baseline Recodes

In this step, the syntax recodes variables that need to be reverse coded in the Family Interaction scale, and collapses several continuous variables and multi-category variables into dichotomous variables (yes/no, agree/disagree) for baseline data.

Steps 5A through 5D: Run the entire “Baseline Recodes” section of syntax at one time. Then run the entire “Posttest Recodes” section of syntax at one time.

The reverse coded Family Interaction scale variables (qA16, qA18, qA19) and the dichotomized efficacy (qA1-11) and substance use (qA33-40) variables will have an “r” for revised after the variable name (ie: qA16r). There will not be any output created, but the user can confirm that the variables were created by clicking on the “Variable View” tab at the bottom of the dataset. Scroll down the variable list; the revised variables will be last.

In addition, a new code called “subuse” will be created that dichotomizes the participants based on their report of substance use at baseline only. Anyone who reports using any tobacco, alcohol, or nonprescribed prescription medication at least 1 day in the previous 30 days will show up in the dataset as a “1”. Those with no reported substance use will have a “0” value. We do not need to worry about missing data in this case because we only use this measure as a filter later on to identify only respondents who report any ATOD use at baseline.

Finally, we create a combined measure of the 4 measures of prescription drug use at baseline and again at post-test that will later be used in the GLM analyses. This way, any inconsistent responses on the prescription drug measures will be corrected for.

6. Post-Test Recodes

Steps 6A-6C: Run the entire “Posttest Recodes” section of syntax at one time.

7. Baseline Scale Scores

This step will create mean scale scores for the Home Environment variables (qA1-qA4), Social Services Utilization variables (qA5-qA11), Family Interaction variables (qA12-qA21), and Social Support variables (qA22-qA31).

Step: Run the entire “Baseline Scale Scores” section of syntax at one time.

Check that the mean scale score variables were created by looking at Variable View. Does the range correspond to what is in the codebook? For example, if the range for the mean scale score is from 1 to 5

and the range in the codebook is from 0 to 4 then either the data are coded incorrectly or the syntax is wrong.

8. Post-Test Scale Scores

Step: Run the entire “Post-Test Scale Scores” section of syntax at one time.

9. Differences between Baseline and Post-Test frequencies for ATOD use (Dichotomous variables)

This syntax calculates frequencies for the dichotomized yes/no ATOD use variable in the entire sample.

Step: Run all syntax. Report “Yes” percentages at baseline and post-test & Percent Change in Table 2.

To calculate the percentage change, do the following:

- 1) Subtract post-test percentage from baseline percentage. $24.5 - 19.6 = 4.9$; $4.9/24.5 = 0.2$
- 2) Divide the difference by the baseline value.

Ex 1. $24.5 - 19.6 = 4.9$; $4.9/24.5 = 0.2$ Ex 2. $16.3 - 18.2 = -1.9$; $-1.9/16.3 = .117$

Substance	Baseline %age of respondents reported past 30 day use	Post-test %age of respondents reported past 30 day use	Percent Change
Past 30 day smoking	24.5%	19.6%	20.0% (Decrease)
Past 30 day drinking	16.3%	18.2%	11.7% (Increase)

10. Differences between Baseline and Post-Test frequencies for ATOD use (Dichotomous variables) ONLY among those reporting ATOD use at Baseline

This syntax does the exact same thing as the previous syntax with the exception of adding a filter that excludes anyone reporting no ATOD use at pretest. It calculates frequencies for the dichotomized yes/no ATOD use variable only among those participants who reported any ATOD use at baseline.

Step: Run all syntax. Report Yes percentages in Table 3. Calculate percentage change and report in Table 2.

11. Differences between Baseline and Post-Test frequencies for ATOD use (Ordinal variables)

This syntax calculates means for the original ATOD use variables at baseline and posttest only among those participants who reported ATOD use at baseline rather than in the whole sample. Since in most cases, at baseline most respondents report no ATOD use, rather than have a mean that is essentially zero and as a result is most likely to increase at post-test, we think a better representation is to show the frequency of use only among those who report any ATOD use. In this case, the sample is far smaller and while there will be some who report no use for some substances, the overall frequency of use for any given substance should be larger and more likely to show change at post-test in the desired direction.

Step: Run all syntax. Report Means in Table 4. Calculate mean differences and report in Table 4.

12. Reliability checks for Module A Baseline Scales

The reliability syntax generates statistics for the components of multiple-item additive scales. It uses Cronbach's alpha which measures how well a set of items (or variables) measures a single unidimensional latent construct. Cronbach's alpha can be written as a function of the number of test items AND the average inter-correlation among the items. Typically, if you increase the number of items, you increase Cronbach's alpha. Additionally, if the average inter-item correlation is low, alpha will be low. As the average inter-item correlation increases, Cronbach's alpha increases as well. Technically speaking, Cronbach's alpha is not a statistical test - it is a coefficient of reliability (or consistency). This makes sense intuitively - if the inter-item correlations are high, then there is evidence that the items are measuring the same underlying construct. This is really what is meant when someone says they have "high" or "good" reliability. They are referring to how well their items measure a single unidimensional latent construct or the internal consistency of the items.

We should note here, however, that since this is the first year of using these scales, we do not have data in existence to predict which measures should be included in a scale. Therefore, based on the "face validity" of items, we have grouped measures accordingly. This first year, we may find that the reliability or internal consistency of a scale is very low. However, using this year's data we will be able to conduct factor analyses to better determine construct validity and in turn hopefully also improve reliability. Cronbach's alpha should ideally be similar at baseline and at post-test if the measure is measuring the same construct the same way at both times.

Step: Run the entire “Reliability Checks for Scales_Baseline ” syntax section and look at the Reliability table.

Example: The reliability statistic in this example will be based on results for your sample. However, because respondents are not *required* to answer all the questions, missing data will be dropped from the calculation of the statistic. According to the table below, 246 records were not included in the calculation of the reliability statistic in Table 16. Before interpreting a reliability statistic, confirm that most of the records are included in the calculation of the reliability statistic.

Scale: ALPHA

Table 15: Case Processing Summary

		N	%
Cases	Valid	2570	91.3
	Excluded ^a	246	8.7
	Total	2816	100.0

a. Listwise deletion based on all variables in the procedure.

Here, the reliability shown in the example below is said to be good using all nine items because alpha is .866. (Note that a reliability coefficient of .70 or higher is considered "acceptable" in most social science research situations, although it is common to see lower cut-off points used).

Table 16: Reliability Statistics

Cronbach's Alpha	N of Items
.866	9

13. Reliability checks for Module A Post-Test Scales

Step: Run the entire “Reliability Checks for Scales Posttest” syntax section and look at the Reliability table. **Only report Post-Test alphas in Table 6.** However, compare baseline and post-test Cronbach’s alpha coefficients to make sure they are similar. If they are not, please let the evaluator know.

14. GLM Analyses for Module A

This analysis is called "Repeated Measures MANOVAs" and it is done through SPSS's GLM procedure. The results will give you an *effect size (partial eta squared)* that may be used instead of the R-squared provided by regression analyses, as well as test for significant difference between pre- and post-test means. The GLM syntax also lets you consider differences within each subject and between each subject (within difference and between differences). In these cases, the within subjects factor of interest to us is time. Specifically the time between when baseline and post-test data were gathered and whether over that time, changes in the measures occurred.

The GLM syntax will run GLMs for the ATOD items and the Module A scales. It will provide you with an F-test value, a significance value and a Partial Eta Square value for TIME. Essentially this answers the question, "Did the time spent receiving prevention programming make a significant difference in the measure?"

Of note, this year we have chosen to run individual GLM analyses by substance rather than all together. We do this because the GLM will remove any observation with any missing data on any of the variables in the model, meaning, if a respondent has missing data for prescription drug use, they are also dropped for the cigarette use analysis. Therefore, so as to keep as many respondents in the analyses as possible for each measure, we will run each individually.

*The examples below were run using a very small fabricated NM FAS data set; your results will more than likely look very different from these.

Step 14-1: Run all of the syntax labeled "GLM".

Step 14-2: Look at the Within-Subjects Factors Table to get the F-test value, the significance value, and the partial eta squared. Enter required information into Tables 5 & 6.

Table 21: Within-Subjects Factors

Within-Subjects Factors

Measure	time	Dependent Variable
g_cig	1	qA33r
	2	qA33r_2
g_alc	1	qA34r
	2	qA34r_2
g_intox	1	qA35r
	2	qA35r_2
g_binge	1	qA36r
	2	qA36r_2
g_pmed	1	qA37r
	2	qA37r_2

g_pain	1	qA38r
	2	qA38r_2
g_reg	1	qA39r
	2	qA39r_2
g_tranq	1	qA40r
	2	qA40r_2

The first table displays the dependent variables that correspond to each combination of within-subjects factors (meaning you are looking at the difference between baseline and posttest for each individual. The *Measure* column in the table indicates this because it shows you that you have a “1” and “2” measurement for every individual on each one of the variables listed. The within-subjects factor, *Time*, designates groups of measurements taken at the same time (in this case 2 time periods: baseline=“1” and posttest=“2”).

Step 14-3: Look at the Descriptive Statistics Table

One of the assumptions for an analysis of variance is that the data in each cell come from populations with the same variance. This table provides an opportunity to check that assumption and it also displays the number of cases for each level of each factor and allows you to see if any of the factor levels do not contain any observations.

Descriptive Statistics

	Mean	Std. Deviation	N
Baseline 30-day cigarette use dichotomized (no/yes)	.5714	.53452	7
Posttest 30-day cigarette use dichotomized (no/yes)	.1429	.37796	7
Baseline 30-day alcohol use dichotomized (no/yes)	1.0000	.00000	7
Posttest 30-day alcohol use dichotomized (no/yes)	.5714	.53452	7
Baseline 30-day alcohol to intoxication use dichotomized (no/yes)	.7143	.48795	7
Posttest 30-day alcohol to intoxication use dichotomized (no/yes)	.2857	.48795	7
Baseline 30-day binge alcohol use dichotomized (no/yes)	.5714	.53452	7
Posttest 30-day binge alcohol use dichotomized (no/yes)	.4286	.53452	7

Baseline 30-day unauthorized prescription med use dichotomized (no/yes)	.5714	.53452	7
Posttest 30-day unauthorized prescription med use dichotomized (no/yes)	.2857	.48795	7
Baseline 30-day unauthorized prescription med use-pain pills dichotomized (no/yes)	.2857	.48795	7
Posttest 30-day unauthorized prescription med use-pain pills dichotomized (no/yes)	.2857	.48795	7
Baseline 30 day unauthorized prescription med use-stimulant or antidepressant dichotomized (no/yes)	.1429	.37796	7
Posttest unauthorized prescription med use- stimulant or antidepressant dichotomized (no/yes)	.2857	.48795	7
Baseline 30-day unauthorized prescription med use-sleep aid or tranquilizer dichotomized (no/yes)	.4286	.53452	7
Posttest 30-day unauthorized prescription med use-sleep aid or tranquilizer dichotomized (no/yes)	.2857	.48795	7

Step 14-4: Look at the Multivariate Tests Table

The multivariate tests table displays four multivariate tests of significance for the effects in the model. Pillai's trace is the first multivariate test listed. Wilks' lambda is sometimes called the U statistic. Lambda ranges between 0 and 1, with values close to 0 indicating the group means are different and values close to 1 indicating the group means are not different (equal to 1 indicates all means are the same). Hotelling's trace is based on the sum of eigenvalues.

Of the four test statistics, Wilks' lambda is convenient to use and related to the likelihood-ratio criterion. The hypothesis and error degrees of freedom of the F distribution are shown. The value of the test statistic is displayed followed by the F statistic, which is a transformed value of the corresponding test statistic and has an approximate F distribution. When the significance level is relatively small (less than 0.05) for the effect being tested, then we conclude that the effect is significant.

Multivariate Tests(c)

Effect			Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power(a)
Between Subjects	Intercept	Pillai's Trace	.996	603.187(b)	2.000	5.000	.919	.996	1206.374	.002
		Wilks' Lambda	.004	603.187(b)	2.000	5.000	.919	.996	1206.374	.002
		Hotelling's Trace	241.275	603.187(b)	2.000	5.000	.919	.996	1206.374	.002
		Roy's Largest Root	241.275	603.187(b)	2.000	5.000	.919	.996	1206.374	.002
Within Subjects	time	Pillai's Trace	.166	.499(b)	2.000	5.000	.634	.166	.998	.006
		Wilks' Lambda	.834	.499(b)	2.000	5.000	.634	.166	.998	.006
		Hotelling's Trace	.200	.499(b)	2.000	5.000	.634	.166	.998	.006
		Roy's Largest Root	.200	.499(b)	2.000	5.000	.634	.166	.998	.006

- a Computed using alpha = .05
- b Exact statistic
- c Design: Intercept
Within Subjects Design: time

In the example above, the effect of time is not significant within subjects ($p=.634$) and the group means are not that different. In this analysis, the value of the Partial Eta Squared statistic is very high meaning, even though the effect of time was not significant, it still accounted for quite a bit of variance in the outcome. (Remember these are made up data and your results should look different. In general, when conducting research involving human beings and social behaviors, rarely does any one variable have a large effect size, because predictors of human behavior are almost infinite and we can never measure them all.)

These are the guidelines for determining how large or small an effect is:

- a small effect: partial $\eta^2 = .01$ to $.059$;
- a medium effect: partial $\eta^2 = .06$ to $.139$; and
- a large effect: partial $\eta^2 \geq .14$

Step 14-5: Report means, F-test statistic and significance values in Tables 5 & 6.

15. Module B Analyses Only run this syntax if you have data for Module B.

We only provide syntax for running prenatal data at baseline. We assume that anyone entering the program pregnant are no longer pregnant at the end of the program. This may be an incorrect

assumption in which case please let PIRE know so we can also provide syntax for analyzing post-test data. You do not have to present pre-test data only. This decision should be made by the prevention provider and local evaluator together.

15A. Frequencies. This syntax runs frequencies for all Module B variables.

Step: Run syntax. Look at output to become familiar with the data, check for any unusual patterns of responding.

15B. Means. This syntax produces means for Intention to use ATOD after birth items (qB18-qB21).

Step: Run syntax. Report means in Table 7.

15C. GLM Analyses. This syntax will run a GLM for the Parenthood Preparedness scale. Only run this syntax if you have both baseline and post-test data for this scale.

Step: Run all of the syntax. Report means, F statistic and significance and effect size for the analysis in Table 8.

15D. Scale reliability. Alpha reliability for the Parenthood Preparedness scale. Only run syntax for post-test reliability if you have post-test data.

Step: Run syntax. Report post-test alpha in Table 8 if you have post-test data. Otherwise report baseline alpha.

16. Module C Analyses Only run this syntax if you have data for Module C.

16A. Data Cleaning Crosstabs. This syntax runs crosstabs for all Module C variables baseline and post-test. Use this step to check for any data entry errors or missing data problems.

Step: Run syntax. Look at output to become familiar with the data, check for any unusual patterns of responding.

16B. Child Safety Items. This syntax runs frequencies for Child Safety Items.

Step: Run syntax. Report frequencies in Table 9.

16C. Items C12 and C13. This syntax runs means for qC12 and qC13.

Step: Run syntax.

16D. and 16E. Scale Scores. This syntax creates mean scale scores for the Parenting Skills items (qC1-qC19 but no including qC12 and qC13) and Child Well-Being items (qC20-qC23) at baseline and post-test.

Step: Run syntax.

16F. Scale reliabilities. Alpha reliabilities for the Parenting Skills and Child Well-being scales.

Step: Run syntax. Report post-test alphas in Table 10.

16G. GLM Analyses. This syntax will run GLMs for Module C scales.

Step: Run all of the syntax. Report means, F statistic and significance and effect size for the Parenting Skills and Child Well-Being scales analyses in Table 10.

17. Module D Analyses Only run this syntax if you have data for Module D.

17A. Data Cleaning Crosstabs. This syntax runs crosstabs for all Module C variables baseline and post-test. Use this step to check for any data entry errors or missing data problems.

Step: Run syntax. Look at output to become familiar with the data, check for any unusual patterns of responding.

17B. Recodes. In this step, the syntax recodes variables that need to be reverse coded in the Handling of Stress scale.

Step: Run syntax. There will not be any output created, but the user can confirm that the variables were created by clicking on the “Variable View” tab at the bottom of the dataset. Scroll down the variable list; the revised variables will be last.

17C. Scale Scores. This syntax creates mean scale scores for the Handling of Stress items (qD1-qD19) at baseline and post-test.

Step: Run syntax.

17D. Scale reliability. Alpha reliabilities for the Handling of Stress scale.

Step: Run syntax. Report post-test alpha in Table 11.

17E. GLM Analyses. This syntax will run GLMs for the Handling of Stress scale.

Step: Run all of the syntax. Report means, F statistic and significance and effect size for the Handling of Stress scales analysis in Table 11.

18. Module E Analyses Only run this syntax if you have data for Module E.

18A. Data Cleaning Crosstabs. This syntax runs crosstabs for all Module C variables baseline and post-test. Use this step to check for any data entry errors or missing data problems.

Step: Run syntax. Look at output to become familiar with the data, check for any unusual patterns of responding.

18B-D. Recodes. In this step, the syntax recodes variables that need to be reverse coded in the Effective Discipline, Parent Child Interaction, and Positive Reinforcement scales.

Step: Run syntax. There will not be any output created, but the user can confirm that the variables were created by clicking on the “Variable View” tab at the bottom of the dataset. Scroll down the variable list; the revised variables will be last.

18E-G. Scale Scores. This syntax creates mean scale scores for the Positive Reinforcement items (qE1-qE6), Parent Child Interaction items (qE7-qE12), and the Effective Discipline items (qE13-qE23) at baseline and post-test.

Step: Run syntax.

18H-J. Scale reliability. Alpha reliabilities for the Positive Reinforcement, Parent Child Interaction, and Effective Discipline scales.

Step: Run syntax. Report post-test alphas in Table 12.

18K. GLM Analyses. This syntax will run GLMs for the Positive Reinforcement, Parent Child Interaction, and Effective Discipline scales.

Step: Run all of the syntax. Report means, F statistic and significance and effect size for the Positive Reinforcement, Parent Child Interaction, and Effective Discipline scales analyses in Table 12.