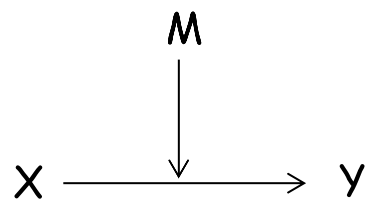
Moderation

A moderator is a variable that specifies conditions under which a given predictor is related to an outcome. The moderator explains **'when'** a DV and IV are related. *Moderation* implied an interaction effect, where introducing a moderating variable changes the direction or magnitude of the relationship between two variables. A moderation effect could be (a) Enhancing, where increasing the moderator would increase the effect of the predictor (IV) on the outcome (DV); (b) Buffering, where increasing the moderator would decrease the effect of the predictor on the outcome; or (c) Antagonistic, where increasing the moderator would reverse the effect of the predictor on the outcome.



Moderation

Hierarchical multiple regression is used to assess the effects of a moderating variable. To test moderation, we will in particular be looking at the *interaction* effect between X and M and whether or not such an effect is significant in predicting Y.

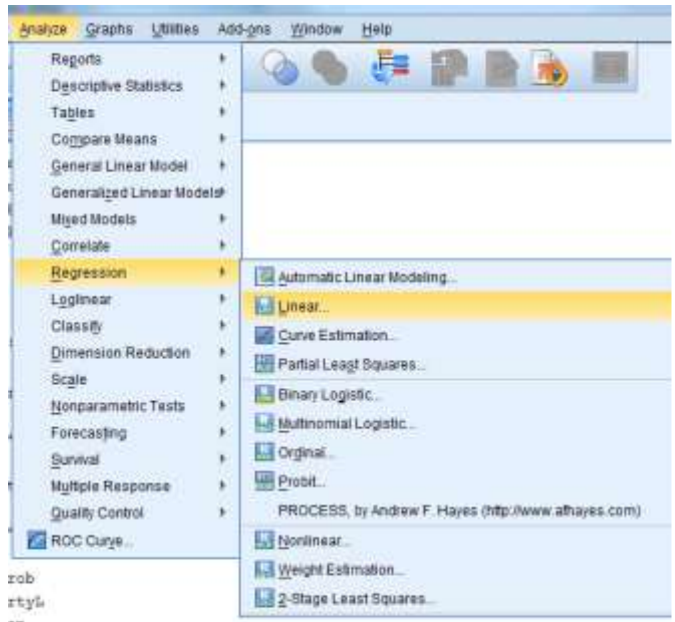
Steps in Testing Moderation

In order to confirm a third variable making a moderation effect on the relationship between the two variables X and Y, we must show that the nature of this relationship changes as the values of the moderating variable M change. This is in turn done by including an interaction effect in the model and checking to see if indeed such an interaction is significant and helps explain the variation in the response variable better than before. In more explicit terms the following steps should be followed:

1. First, you need to standardize all variables to make interpretations easier afterwards and to avoid multicolliearity (the SPSS process described below does this for you automatically).
2. If you are using regular regression menu items in SPSS or similar software, you would also need to dummy code categorical variables and manually create product terms for the predictor and moderator variables (dummy coding is still necessary with the discussed process, however product terms are created automatically).
3. Fit a regression model (block 1) predicting the outcome variable Y from both the predictor variable X and the moderator variable M. Both effects as well as the model in general (R^2) should be significant.
4. Add the interaction effect to the previous model (block 2) and check for a significant R^2 change as well as a significant effect by the new interaction term. If both are significant, then moderation is occurring.
    - If the predictor and moderator are not significant with the interaction term added, then complete moderation has occurred.
    - If the predictor and moderator are significant with the interaction term added, then moderation has occurred, however the main effects are also significant.

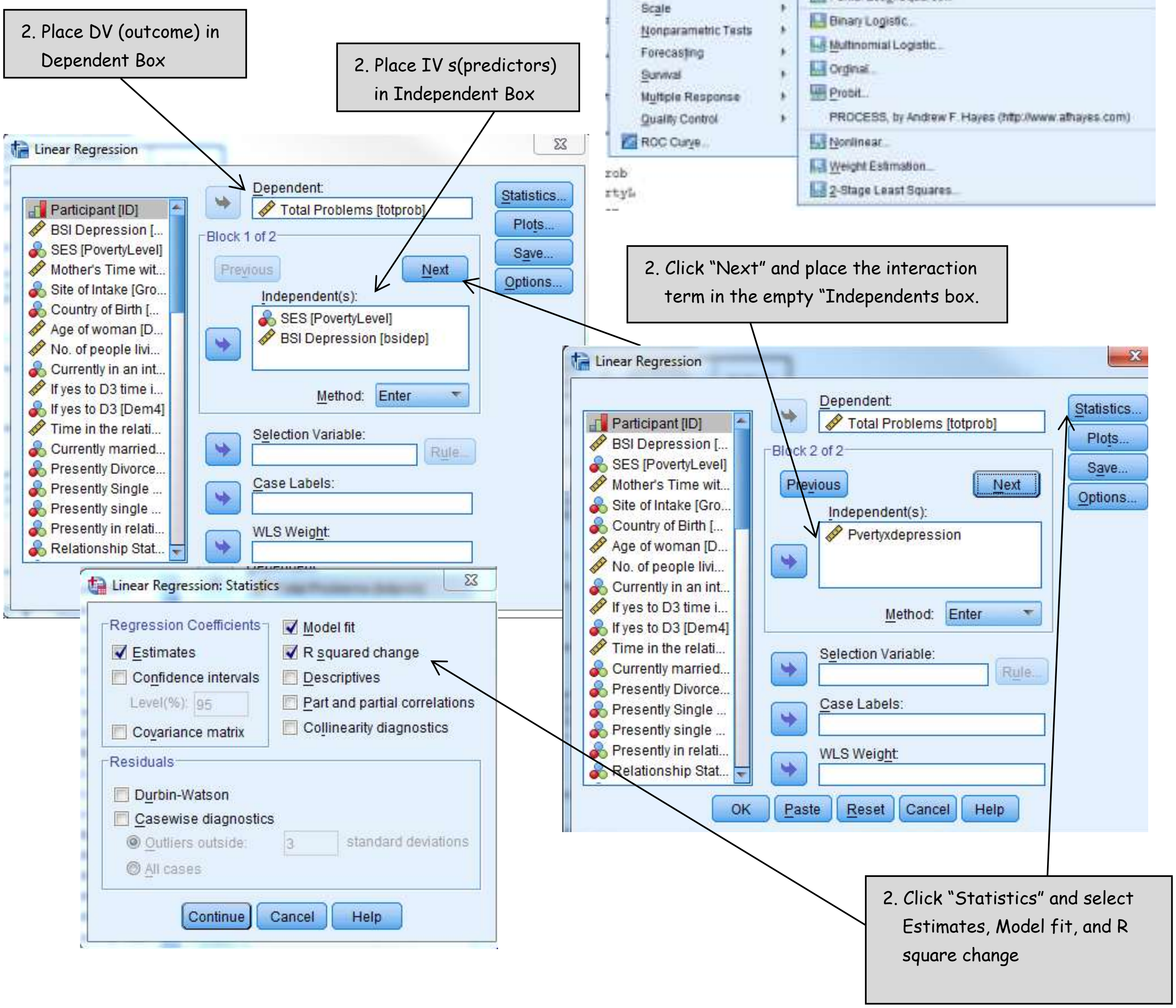
Conducting the Analysis in SPSS

Similar to mediation, moderation can also be checked and tested using the regular linear regression menu item in SPSS. For this purpose you would need to dummy code categorical variables, center the variables as well as create the interaction effect(s) manually. We on the other hand will use the PROCESS developed by Andrew F. Hayes which does the centering and interaction terms automatically. You do however still need to dummy code categorical variables with more than 2 categories before including them in the model.



1. Create the uncentered interaction term. Transform → Compute → Var1*Var2

2. Start by running the model with the uncentered interaction to get the amount of variance accounted for by the predictors with and without the interaction.



Step 1 - At this step, you are only interested in if the models are significant and if the amount of variance accounted for in Model 2 (with the interaction) is significantly more than Model 1 (without the interaction).

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	77437.125	2	38718.563	76.573	.000[a]
	Residual	150175.871	297	505.643		
	Total	227612.997	299			
2	Regression	81995.064	3	27331.688	55.558	.000[b]
	Residual	145617.933	296	491.952		
	Total	227612.997	299			

a. Predictors: (Constant), BSI Depression, SES
b. Predictors: (Constant), BSI Depression, SES, Pvertyxdepression
c. Dependent Variable: Total Problems

Is model 1 (without the interaction term) significant?

→ Yes, $F(2, 297) = 76.57$, $p < .001$

Is model 2 (with the interaction term) significant? → Yes, $F(3, 296) = 55.56$, $p < .001$

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.583[a]	.340	.336	22.48650	.340	76.573	2	297	.000
2	.600[b]	.360	.354	22.18000	.020	9.265	1	296	.003

a. Predictors: (Constant), BSI Depression, SES
b. Predictors: (Constant), BSI Depression, SES, Pvertyxdepression
c. Dependent Variable: Total Problems

Does model 2 account for significantly more variance than model 1?

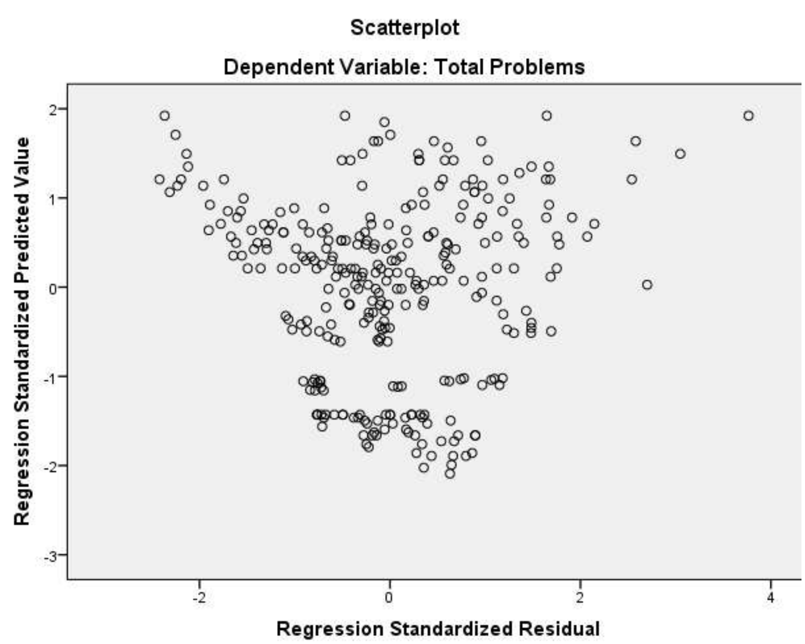
In this example, Model 2 with the interaction between depression and poverty level accounted for significantly more variance than just depression and poverty level by themselves, R^2 change = .020, $p = .003$, indicating that there is potentially significant moderation between depression and poverty level on child's behavior problems.

Syntax for Step 1

```
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA COLLIN TOL CHANGE
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT totprob
  /METHOD=ENTER PovertyLevel bsidep
  /METHOD=ENTER Pvertyxdepression
  /SCATTERPLOT=(*ZPRED ,*ZRESID).
```

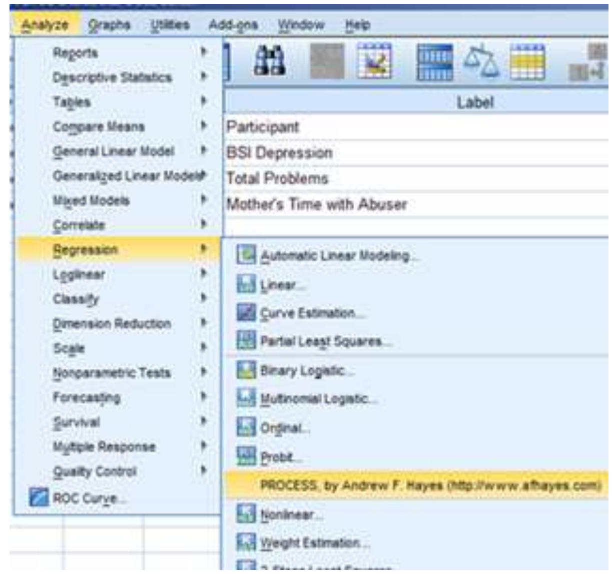
This scatterplot syntax will give you a graph of the residuals so you can examine heteroskedasticity.

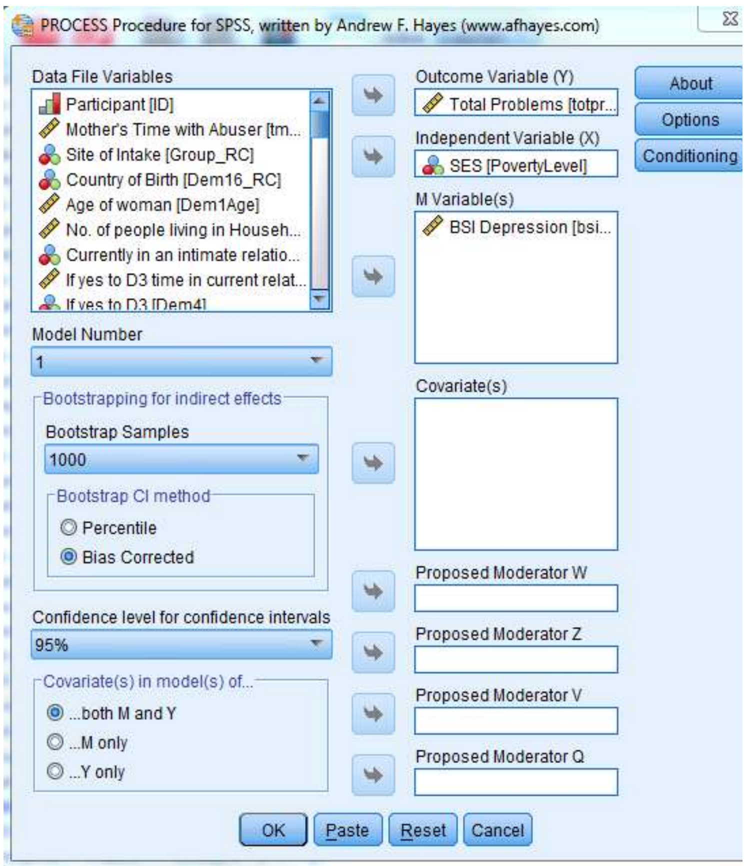
You want the scatter plot to be well distributed.



Step 2 - Since there is a potentially significant moderation effect, we can run the regression on the centered terms to examine the effect. While you can do this by centering the terms yourself and building the regression, this is best done using an add-on process.

3. Your dataset must be open. To run the analysis, click on analyze, then regression, then PROCESS, by Andrew F. Hayes (http://www.afhayes.com). If you don't see this menu item, it means that this process first needs to be installed on your computer.

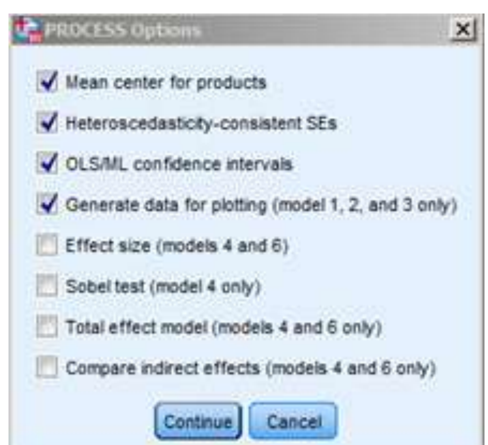




4. The PROCESS Dialog will open. Select and move the initial IV (X), the DV (Y) and the moderator variable (M) into their appropriate boxes as shown in the picture.

5. You can also include any covariates in the appropriate box.

6. In order to test a *moderation* effect, make sure that the Model Number is set to 1.



7. Click on the Options button and select appropriate options. To better examine the effect of a moderating variable, the first four options (Mean center for products, Heteroscedasticity-consistent SEs, OLS/ML confidence intervals, and Generate data for plotting) can be selected.

8. The syntax for this process is very long. You can create a syntax file by clicking on Paste.

Output - After running this process, the output you will see will look similar to what is shown below. Since bootstrapping is used to calculate standard errors and confidence intervals, this might take a little while.

```
Run MATRIX procedure:

************* PROCESS Procedure for SPSS Beta Release 140712 *************

        Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com
**************************************************************************
Model = 1
    Y = totprob
    X = PovertyL
    M = bsidep

Sample size
        300
**************************************************************************
Outcome: totprob
Model Summary
          R       R-sq          F        df1        df2          p
      .6002      .3602    56.6464     3.0000   296.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant    42.4127     1.2801    33.1317      .0000    39.8934    44.9320
bsidep        .5487      .2126     2.5802      .0104      .1302      .9672
PovertyL    10.8893      .9639    11.2975      .0000     8.9924    12.7863
int_1         .4319      .1525     2.8323      .0049      .1318      .7320

Interactions:

 int_1    PovertyL    X     bsidep
*************************************************************************
Conditional effect of X on Y at values of the moderator(s)
     bsidep     Effect         se          t          p       LLCI       ULCI
    -6.6380     8.0223     1.0958     7.3207      .0000     5.8657    10.1789
      .0000    10.8893      .9639    11.2975      .0000     8.9924    12.7863
     6.6380    13.7564     1.6452     8.3617      .0000    10.5187    16.9941

Values for quantitative moderators are the mean and plus/minus one SD from mean
**************************************************************************
Data for visualizing conditional effect of X of Y
   PovertyL     bsidep       yhat
    -1.4453    -6.6380    27.1759
      .0000    -6.6380    38.7707
     1.4453    -6.6380    50.3654
    -1.4453      .0000    26.6742
      .0000      .0000    42.4127
     1.4453      .0000    58.1513
    -1.4453     6.6380    26.1724
      .0000     6.6380    46.0548
     1.4453     6.6380    65.9372

******************** ANALYSIS NOTES AND WARNINGS *************************
Level of confidence for all confidence intervals in output:
    95.00
NOTE: The following variables were mean centered prior to analysis:
 PovertyL bsidep
NOTE: All standard errors for continuous outcome models are based on the HC3 estimator
------ END MATRIX -----
```

Use these values to plot the interaction using the Excel file "Interaction Plot"

The first part of the output lists the variables in the analysis, indicating which is considered as a dependent variable (Y), which an independent variable (X) and which a moderator (M). The total sample size is also displayed. Then the results from a regression model are displayed which includes the interaction effect between the independent variable and the moderator.

Step 3 – Plot the interaction points to interpret the interaction.

Open the Excel file " Interaction Plot" and enter the values from the output in the green cells (B4:D6). Also change the labels in A3 and C2 to reflect your variable names.

Change the variables names in the blue cells only to accurately describe your chart. Keep the "Low Average High".

	A	B	C	D
1				
2			Depression	
3	Poverty	(1 SD Below)	(Mean)	(1 SD Above)
4	(1 SD Below)	27.176	26.6742	26.172
5	(Mean)	38.771	42.4127	46.054
6	(1 SD Above)	50.365	42.4127	65.937
7				
8		Low Depression	Average Depression	High Depression
9	Low Poverty	27.176	26.674	26.172
10	Average Poverty	38.771	42.413	46.055
11	High Poverty	50.365	42.413	65.937
12				

Sample Write up

To test the hypothesis that the child behavior problems are a function of multiple risk factors, and more specifically whether mother's depression moderates the relationship between poverty level and child behavior problems, a hierarchical multiple regression analysis was conducted. In the first step, two variables were included: poverty level and mother's depression. These variables accounted for a significant amount of variance in child's behavior problems, $R^2 = .340$, $F(2, 297) = 76.57$, $p < .001$. To avoid potentially problematic high multicollinearity with the interaction term, the variables were centered and an interaction term between poverty level and mother's depression was created (Aiken & West, 1991).

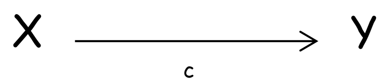
Next, the interaction term between poverty level and mother's depression was added to the regression model, which accounted for a significant proportion of the variance in child behavior problems, $\Delta R^2 = .02$, $\Delta F(1, 296) = 9.27$, $p = .001$, b = .432, $t(296) = 2.83$, $p < .01$. Examination of the interaction plot showed an enhancing effect that as poverty and mother's depression increased, child behavior problems increased. At low poverty, child behavior problems were similar for mother's with low, average, or high depression. Children from high poverty homes with mother's who had high depression had the worst behavior problems.

References

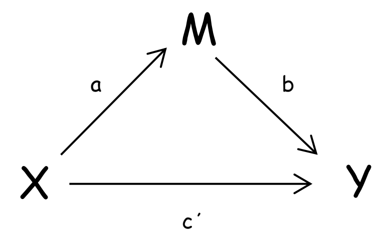
Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage.

Mediation

Mediation implies a situation where the effect of the independent variable on the dependent variable can best be explained using a third *mediator* variable which is caused by the independent variable and is itself a cause for the dependent variable. That is to say instead of X causing Y directly, X is causing the mediator M, and M is in turn causing Y. The causal relationship between X and Y in this case is said to be *indirect*. The relationships between the independent, the mediator and the dependent variables can be depicted in form of a path diagram/model.



Direct Causality



Indirect Causality

Each arrow in a path diagram represents a causal relationship between two variables to which a coefficient or weight is assigned. These coefficients are nothing but the standardized regression coefficients (betas) showing the direction and magnitude of the effect of one variable on the other.

Variables

Instead of using the terms independent and dependent variables, it would make more sense in the context of path models to speak of exogenous and endogenous variables.

Exogenous Variables – Variables which in the context of the model have no explicit causes. That is to say they have no arrows going to them.

Endogenous Variables – Variables which in the context of the model are causally affected by other variables. That is to say they have arrows going to them.

From a regression standpoint, for every endogenous variable in the model a regression model should be fitted.

Assumptions

1. Continuous Measurements. All variables are assumed to be measured on a continuous scale.
2. Normality. All variables are assumed to follow a Normal distribution.
3. Independence. The errors associated with one observation are not correlated with the errors of any other observation.
4. Linearity: relationships among the variables are assumed to be linear.

Steps in Testing Mediation

In order to confirm a mediating variable and its significance in the model, we must show that while the mediator is caused by the initial IV and is a cause of the DV, the initial IV loses its significance when the mediator is included in the model. In more explicit terms, we should follow the following four steps:

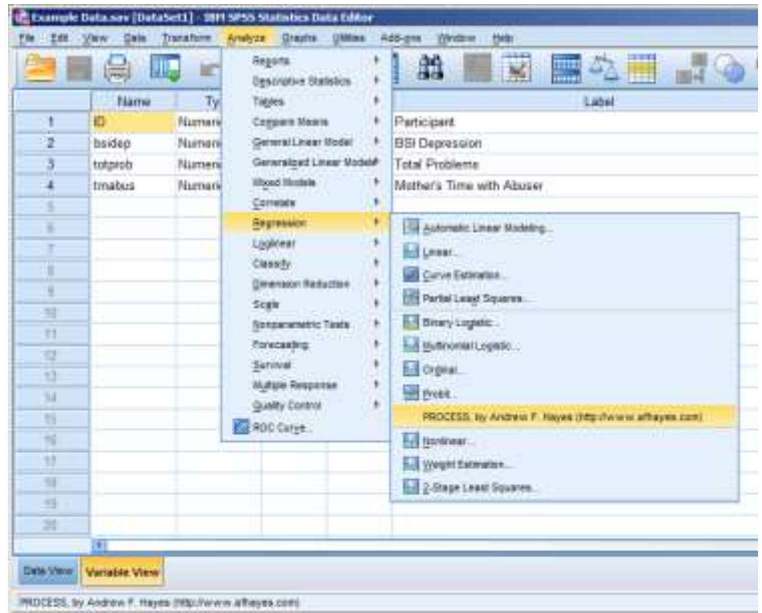
1. Confirm the significance of the relationship between the initial IV and DV (X → Y)
2. Confirm the significance of the relationship between the initial IV and the mediator (X → M)
3. Confirm the significance of relationship between the mediator and the DV in the presence of the IV (M|X → Y)
4. Confirm the <u>insignificance</u> (or the meaningful reduction in effect) of the relationship between the initial IV and the DV in the presence of the mediator (X|M → Y)

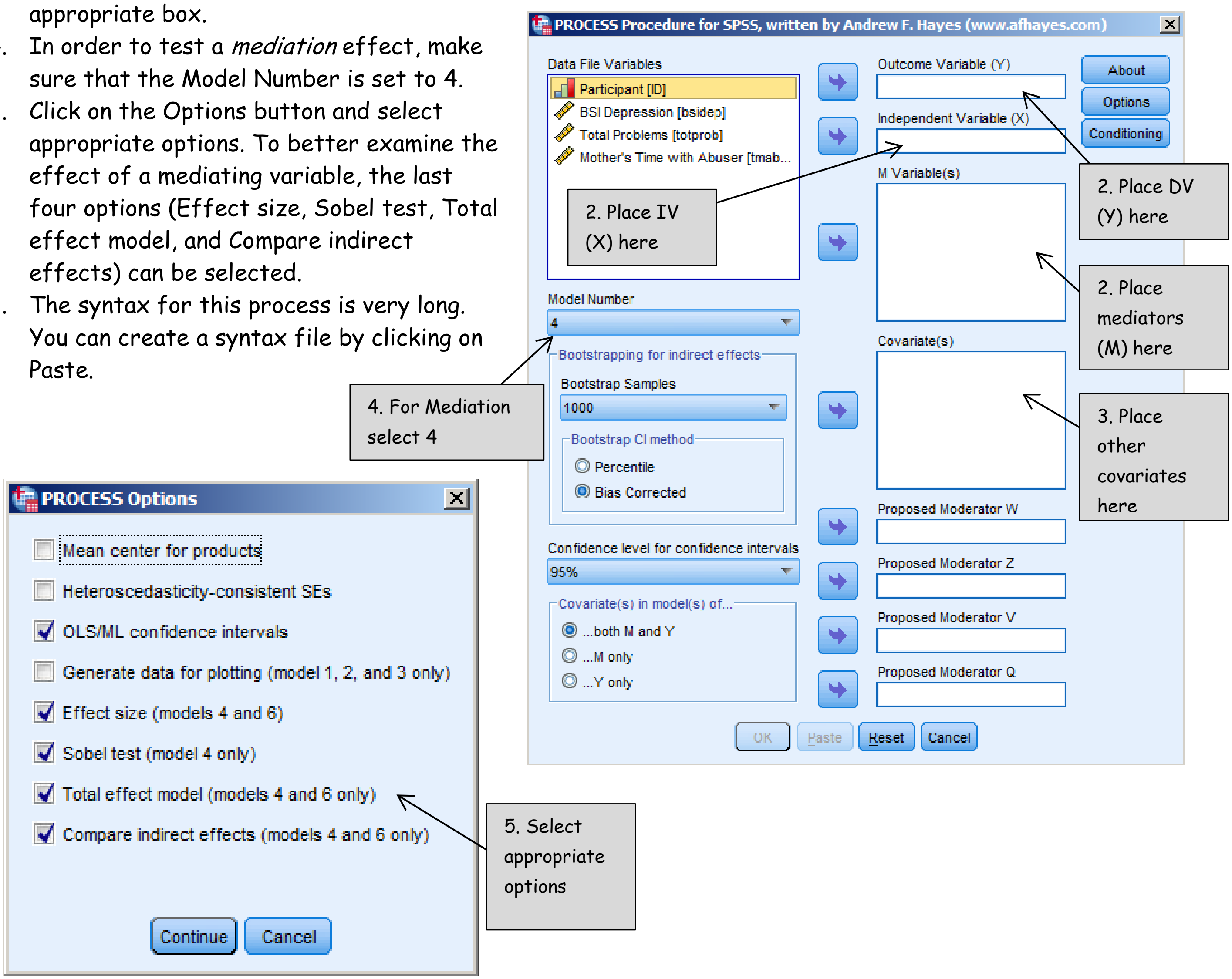
Steps 3 and 4 will involve the same regression model.

Conducting the Analysis in SPSS

Mediation can be tested by following the above steps using the regular linear regression menu item in SPSS, or more conveniently using a special PROCESS developed by Andrew F. Hayes which is described below.

1. Your dataset must be open. To run the analysis, click analyze, then regression, then PROCESS, by Andrew F. Hayes (http://www.afhayes.com)
   If you don't see this menu item, it means that this process first needs to be installed on your computer.
2. The PROCESS Dialog will open. Select and move the initial IV (X), the DV (Y) and the mediator variable (M) into their appropriate boxes as shown in the picture.
3. You can also include any covariates in the appropriate box.
4. In order to test a *mediation* effect, make sure that the Model Number is set to 4.
5. Click on the Options button and select appropriate options. To better examine the effect of a mediating variable, the last four options (Effect size, Sobel test, Total effect model, and Compare indirect effects) can be selected.
6. The syntax for this process is very long. You can create a syntax file by clicking on Paste.





Output After running this process, the output will look similar to what is shown below. Since bootstrapping is used to calculate standard errors and confidence intervals, this might take a little while.

```
Run MATRIX procedure:

************* PROCESS Procedure for SPSS Beta Release 140712 *************

        Written by Andrew F. Hayes, Ph.D.   http://www.afhayes.com

**************************************************************************
Model = 4
    Y = totprob
    X = tmabus
    M = bsidep

Sample size
        300

**************************************************************************
Outcome: bsidep

Model Summary
          R       R-sq          F        df1        df2          p
      .7079      .5011   299.3041     1.0000   298.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant     5.0764      .3748    13.5439      .0000     4.3388     5.8140
tmabus       9.3921      .5429    17.3004      .0000     8.3237    10.4605

**************************************************************************
Outcome: totprob

Model Summary
          R       R-sq          F        df1        df2          p
      .2974      .0884    14.4054     2.0000   297.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant    32.3713     2.6812    12.0735      .0000    27.0948    37.6479
bsidep        .9080      .3260     2.7852      .0057      .2664     1.5496
tmabus       5.4908     4.3256     1.2694      .2053    -3.0220    14.0035

************************** TOTAL EFFECT MODEL ****************************
Outcome: totprob

Model Summary
          R       R-sq          F        df1        df2          p
      .2542      .0646    20.5867     1.0000   298.0000      .0000

Model
              coeff         se          t          p       LLCI       ULCI
constant    36.9809     2.1332    17.3357      .0000    32.7828    41.1790
tmabus      14.0191     3.0898     4.5373      .0000     7.9386    20.0997

***************** TOTAL, DIRECT, AND INDIRECT EFFECTS ********************

Total effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
    14.0191     3.0898     4.5373      .0000     7.9386    20.0997
```

Variables in the analysis

X significant predictor of M

M|X significant predictor of Y
X|M not a significant predictor of Y

X significant predictor of Y

```
Direct effect of X on Y
     Effect         SE          t          p       LLCI       ULCI
     5.4908     4.3256     1.2694      .2053    -3.0220    14.0035

Indirect effect of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
bsidep     8.5283     3.1283     2.7878    14.9841

Partially standardized indirect effect of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
bsidep      .3091      .1078      .1041      .5142

Completely standardized indirect effect of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
bsidep      .1546      .0539      .0520      .2573

Ratio of indirect to total effect of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
bsidep      .6083      .2777      .1872     1.2285

Ratio of indirect to direct effect of X on Y
           Effect    Boot SE   BootLLCI   BootULCI
bsidep     1.5532   250.7987    -4.9630   109.7770

R-squared mediation effect size (R-sq_med)
           Effect    Boot SE   BootLLCI   BootULCI
bsidep      .0597      .0228      .0232      .1196

Preacher and Kelley (2011) Kappa-squared
           Effect    Boot SE   BootLLCI   BootULCI
bsidep      .1130      .0386      .0399      .1910

Normal theory tests for indirect effect
     Effect         se          Z          p
     8.5283     3.1065     2.7453      .0060

******************** ANALYSIS NOTES AND WARNINGS *************************

Number of bootstrap samples for bias corrected bootstrap confidence intervals:
     1000

Level of confidence for all confidence intervals in output:
    95.00

------ END MATRIX -----
```

Indirect effect of X on Y significantly greater than zero

The first part of the output lists all variables in the analysis, indicating which is considered as a dependent variable (Y), which an independent variable (X) and which a mediator (M). The total sample size is also displayed.

Then a series of regression models are fitted, first predicting the mediator variable using the independent variable (step 2); then the dependent variable using both the independent variable and the mediator (steps 3 and 4); and finally the dependent variable using the independent variable (step 1). In this case, while the independent variable was a significant predictor for both the dependent and the mediator variables, it is no longer significant in the presence of the mediator variable; confirming the mediation effect. A measure for the indirect effect of X on Y is also presented after the regression models. In this case the effect size was 8.5283, with a 95% confidence interval which did not include zero; that is to say the effect was significantly greater that zero at α = .05.

Sample Write up

In Step 1 of the mediation model, the regression of mother's time spent with the abuser on child behavior problems, ignoring the mediator, was significant, $b = 14.02$, $t(298) = 14.02$, $p = <.001$. Step 2 showed that the regression of the mother's time spent with the abuser on the mediator, depression, was also significant, $b = 9.39$, $t(298) = 17.30$, $p = <.001$. Step 3 of the mediation process showed that the mediator (depression), controlling for mother's time with the abuser, was significant, $b = .908$, $t(297) = 2.79$, $p = .0057$. Step 4 of the analyses revealed that, controlling for the mediator (depression), mother's time with the abuser scores was not a significant predictor of child behavior problems, $b = 5.49$, $t(297) = 1.23$, $p =.2053$. A Sobel test was conducted and found full mediation in the model ($z = 2.74$, $p = .006$). It was found that depression fully mediated the relationship between mother's time spent with the abuser and child behavior problems.