

Path analysis with AMOS

A real dataset and a medical path model to be tested

For this example we use the same (real) social psychological data but we have fabricated a medical context for it, so the results for this example do not derive from real medical research. Data on five variables were obtained from 211 patients with chronic fatigue syndrome (CFS) in an attempt to identify direct and indirect influences contributing to a reduction in occupational, educational and social ACTIVITY. This is the principal DV, measured on a scale from 1 - 5. The remaining variables are as follows: MONTHS since the onset of the illness; MANAGE, the extent to which the illness is managed by diet, graded exercise, education etc.; FUNCTION, loss of function, based on a check list of physical and cognitive functions; PAIN, a rating of amount and severity of myalgia (muscle pain) experienced. It is proposed that MONTHS since onset will affect MANAGE, which in turn will affect PAIN and FUNCTION, that PAIN will affect FUNCTION and ACTIVITY, and that FUNCTION will also affect ACTIVITY. All of the effects, with the exception of MONTHS → MANAGE and PAIN → FUNCTION are predicted to be negative. The path model is tested using the AMOS package. Three plausible alternative models were also generated, but they will not be described here.

Using AMOS to test the model: getting the data file

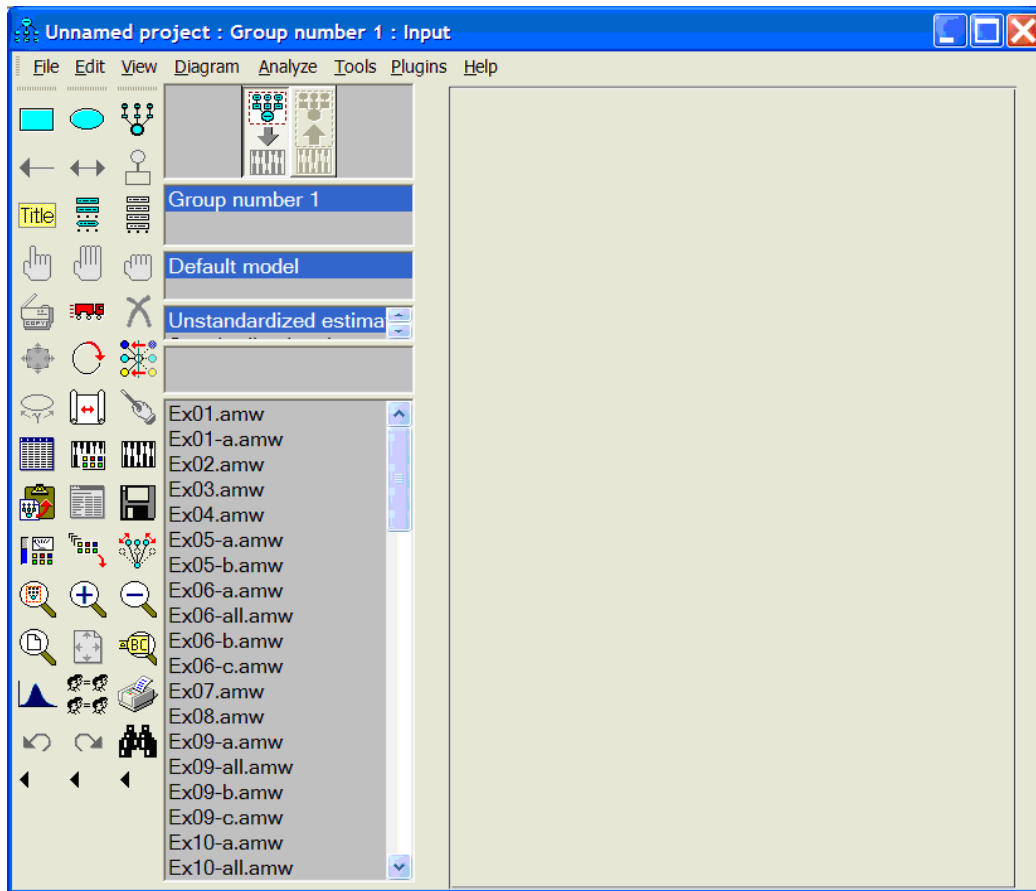
We will begin with the data in five columns in an SPSS datasheet. The first few cases can be seen in Table 7.2 and the full data file, med.path2.sav, is available on the book website.

Table 7.2

First few cases of CFS data (the full dataset can be found as med.path2.sav on the website) (adapted from Sani & Todman, 2002)

case	months	manage	pain	entitati	activity
1	12	19	15	7	1
2	12	25	22	10	2
3	12	26	24	6	3
4	14	29	14	4	2
5	10	22	13	9	2
6	12	26	15	missing	2

Note the presence of missing data that will need to be dealt with. To begin the path analysis, open AMOS Graphics (you have the option of using AMOS basic - an equation-based format, but we think most people will like the graphics version best). The screen will look as in AMOS Screen 7.1. The long list of files to the right of the icons is the contents of the example folder in Amos. First you need to identify your own SPSS datafile to use, so go to **File** on the menu bar, then **Data Files**. Click the **File Name** button and browse to find the file you want (med.path2.sav). Click **OK** to close the dialog box.



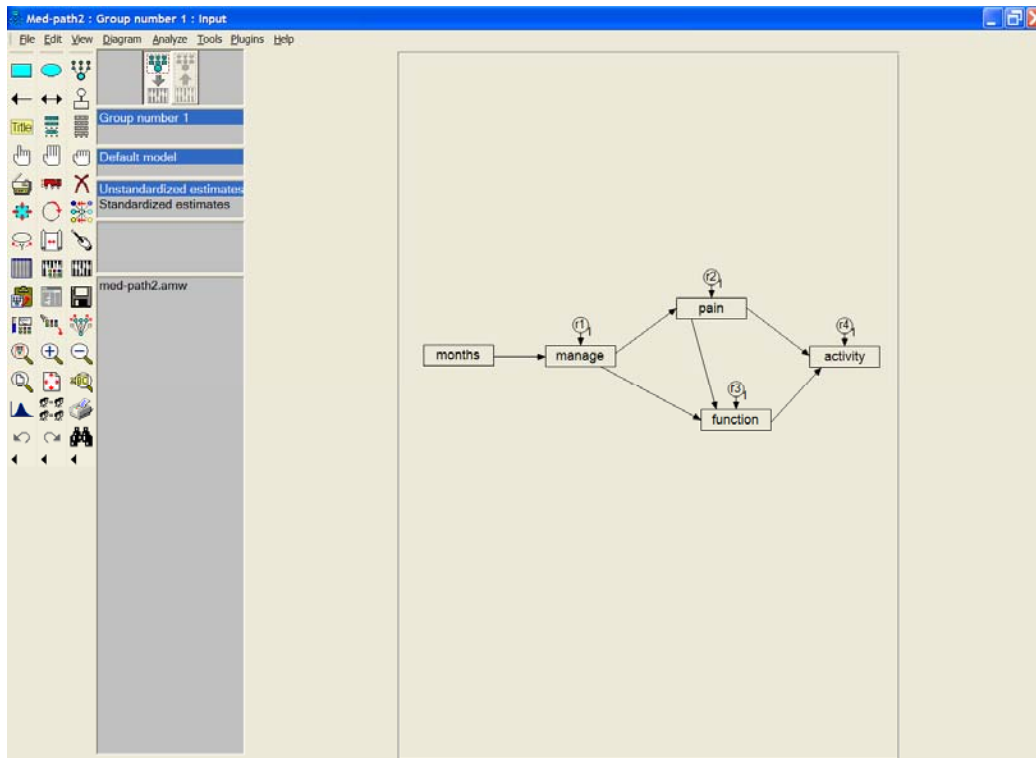
AMOS Screen 7.1. The AMOS graphic screen at the start

Using AMOS to test the model: constructing the input path diagram

You can begin drawing your input path diagram immediately. Go to the **Rectangle** icon at the top left of the icon box and click on it. It will then be highlighted to show that it is active. All of the icons toggle on and off, so you could deselect the rectangle icon by clicking on it again. While the rectangle icon is active, move to the diagram space on the right of the screen. When you click and hold down the mouse button, you can drag the mouse to draw a rectangle. We have five variables, so you need to draw five rectangles. The easiest way is to use the **Copy** icon (5th down in the left column). While it is active, click on the rectangle you have drawn and drag to where you want a copy. Repeat this until you have the five variable boxes positioned as you want them. You can change their size and shape at anytime by activating the icon with arrows

pointing in four directions (6th down in the left column) and you can move objects anytime by activating the **Truck** icon (5th down in the centre column). You can also delete any object by activating the **Cross-out** icon (5th down in the right column). You add arrows to the diagram by activating the appropriate **Arrow** icon, a single arrow in our case (2nd down in the left column). Click where you want the arrow to start and drag it to where you want it to point. You need to realise that all of the icon locations we refer to are as shown in AMOS Screen 7.1. Your columns may be arranged differently depending on the width that is set for the icon display.

Next, you can give the rectangles variable names (they must match the names used in your SPSS data file). Right click on a rectangle and select **Object Properties** from the menu that opens, select the **Text** tab and type the name. When you have entered text in the dialog box, just close the box - there is no **Continue** or **OK** button. Next, activate the **Add a unique variable to an existing variable** icon (2nd down in the right column). This will enable you to attach R^2 value boxes to the four endogenous variables by clicking on the variables. Recall that we mentioned earlier that the values entered in an output path diagram may be R^2 , $1 - R^2$ or $\sqrt{1 - R^2}$; AMOS shows R^2 . You can then name these variables. We will call them r1, r2, r3 and r4, remembering that each of these labels stands for an R^2 value. You should end up with an input path diagram looking something like that shown in AMOS Screen 7.2. This would be a good time to save your work so far: use **File** and **Save As** from the menu bar in the usual way (browse to find the folder where you want to save your work). The name you choose appears in the lowest box to the right of the tools.



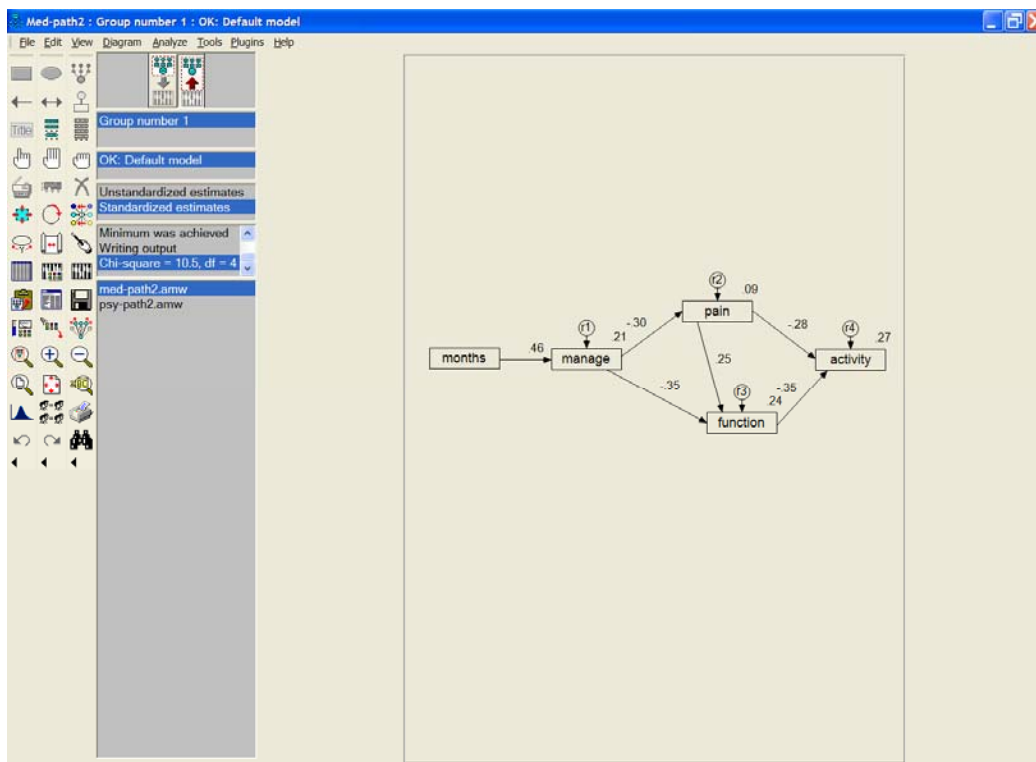
AMOS Screen 7.2. The input path diagram specifying the proposed model

Using AMOS to test the model: requesting output

Now select **Standardized estimates** to the right of the icon columns. Next, click on **View** in the menu bar at the top left, then **Analysis Properties** and the **Estimation** tab. Leave the defaults on and put a tick in **Estimate means and intercepts**. This deals with missing data (recall that we do have data missing in our file (see Table 7.2). Now click on the **Output** tab and put ticks against **Standardized estimates** and **Squared multiple correlations**.

Now you need to locate the **Calculate estimates** icon. It has thick and thin vertical stripes and is 8th down in the right hand column in the icon box. Click on it and you should see a Chi Square value appear in the box below **Standardized estimates**. The analysis has been done and you can look at the output path diagram, with beta coefficients and R^2 values entered, by clicking on the large icon at the top left of the

screen with the red arrow pointing upwards. If you click on the magnifying glass icon with the plus sign on it, your output diagram will be enlarged. Should that result in the diagram becoming too big for the screen, click on the large icon at the top left of the screen with the down arrow to return to your input path diagram, and use the **resizing** and/or the **truck** icon to make the diagram fit. You can then click again on the up arrow icon to recover your output path diagram. The screen should now look something like that shown in AMOS Screen 7.3.



AMOS Screen 7.3. The output path diagram for the CFI data in med.path2.sav

Using AMOS to test the model: understanding the output

If it is unclear to you which values in the output diagram are beta coefficients and which are R^2 values, move the mouse pointer over an arrow and it, and its beta coefficient, will show up red. Likewise, if you move the arrow over a variable box, it and its associated R^2 (labelled r1 etc) value, will show up red.

We see (on the left of AMOS Screen 7.3) that Chi Square for goodness of fit is 10.5 with 4 *dfs*. This was the first statistic to be developed to give information about how well the data fit a proposed model. Our value in this case is significant at $p < 0.05$ (the probability, $p = 0.033$ is given in the **Text Output**, which we will come to next). Just before we do, we may note that, although the fit of the model to the covariance matrix is not 'good' (i.e., Chi Square just reaches significance), this is not a serious setback because, as noted in the section on interpreting the SPSS output, Chi Square very readily reaches significance with large sample sizes even when all other indices indicate a good fit, and in this case Chi Square is only just significant and its value is only slightly more than double the *dfs*.

Using AMOS to test the model: goodness of fit indices

Turning to those other indices, we can look at their values by clicking on the **View text output** icon, which is supposed to represent columns of text (9th down in the centre column). The goodness of fit indices come near the end of the file (select **Model Fit** from the list on the left). The beginning of the goodness of fit summary of models is shown in AMOS Screen 7.4.

The screenshot shows the Amos Output window for a model named 'med-path2.amw'. The 'Model Fit' section is selected in the left-hand navigation pane. The main content area displays four tables of fit indices for three models: Default model, Saturated model, and Independence model.

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	16	10.504	4	.033	2.626
Saturated model	20	.000	0		
Independence model	5	196.382	15	.000	13.092

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.947	.799	.966	.866	.964
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.267	.252	.257
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	6.504	.442	20.154
Saturated model	.000	.000	.000
Independence model	196.382	15.000	13.092

AMOS Screen 7.4. Goodness of fit indices for the CFI (data in med.path2.sav) model

For each index, goodness of fit is given for the model being tested (the **default model**), for the **saturated model** (the just-identified model) and the **independence model** (all correlations among variables are zero). The fit of the default model will lie somewhere between the extremes represented by the other two models. Evaluation of a model hinges on its position between the extremes relative to competing models. In other words, a comparative interpretation is called for. The first index we refer to is CMIN, which is actually the likelihood ratio Chi Square. We have already noted that

its value of 10.5 with $df = 4$ is not inconsistent with a reasonably fitting model. Other indices have been developed to reduce the dependence on sample size. The next block of indices, include two such, NFI (normed fit index) and CFI (comparative fit index), which are among those most frequently reported. The NFI is an example of a so called *descriptive fit index* and indicates the proportion of improvement of the overall fit of the model relative to the independence model. The CFI is also a descriptive fit index and is interpreted in the same way as the NFI, but may be less affected by sample size. For these indices, values close to 1 are generally considered to indicate a good fit, so our values for NFI (0.95) and CFI (0.96) suggest that our model is quite a good fit. Quite a few more batches of related indices follow. Of these, another that is frequently referred to, the RMSEA (root mean square error of approximation) index is one of a number of measures that evaluate the extent to which a model fails to fit the data per degree of freedom, and tends to favour more complex models. It is considered to indicate a bad fit if it is greater than 0.1. It is not on the first screen of fit indices shown in AMOS Screen 7.4, but scroll down and see that the value we get is 0.09, with lower and upper bounds for the confidence interval of 0.02 to 0.16. This suggests that the fit is not very good. There are detailed discussions of all of the indices in the guide to AMOS text and in Marcoulides and Hershberger (1997), both of which we reference in Further reading at the end of the book.

Before leaving the indices though, there is an index called AIC (Akaike's information criterion) that is a modification of the standard goodness-of-fit χ^2 statistic that includes a penalty for complexity. This index does not tell us anything about the fit of a particular model, but it is extremely useful for making comparisons among two or more models, including non-nested models (a model is nested under another when it

is identical except that one or more paths have been removed). Generally, the model with the lowest AIC value is considered to have the best fit. If it is computed for several models estimated from the same data set, the model with the lowest AIC value is considered the best fit. We tested three alternative, theoretically plausible models and they all yielded higher AIC values (the lowest was 48.85) than the value obtained for our proposed model (42.50).

There are many competing goodness of fit indices to choose from and, as no one index is considered clearly preferable in all respects, it is usual to consider a number. If the indices all lead to the same conclusion, you can be reasonably confident in selecting three or four, representing different types, in a report of the results of a path analysis. If the conclusions diverge, it is likely that there is a problem of some kind with the model. One final caution: the relatively good performance of our model with respect to the goodness of fit criteria does not guarantee the plausibility of the model. In fact, it is possible for a model to fit well yet still be incorrectly specified, but discussion of how this may arise is beyond the scope of this book.