Factorial ANCOVA

Here's the means for the factorial analysis of how gender and marital status relate to depression.

Descriptive Statistics

Dependent Variable: DEP

GENDER	MARITAL	Mean	Std. Deviation	N
male	single	5.0492	5.58459	122
	married	6.1064	6.69923	47
	Total	5.3432	5.91411	169
female	single	9.0417	8.10995	120
	married	5.8919	4.57081	74
	Total	7.8402	7.12801	194
Total	single	7.0289	7.22053	242
	married	5.9752	5.47032	121
	Total	6.6777	6.69899	363

Interpretation of these means suggests that…

- Females have higher average depression scores than males.
- Singles have higher average depression scores than married folks (a small effect).
- The simple effect of marital status is different for males and females.

However, since there is to RA or manipulation of these conditions, we might wonder if part of this difference is due to differential

Descriptive Statistics

Dependent Variable: STRESS

GENDER	MARITAL	Mean	Std. Deviation	N
male	single	8.93	6.927	122
	married	7.00	7.863	47
	Total	8.40	7.228	169
female	single	10.08	7.701	120
	married	6.77	7.411	74
	Total	8.82	7.742	194
Total	single	9.50	7.328	242
	married	6.86	7.558	121
	Total	8.62	7.500	363

Some interesting differences…

- A larger gender simple effect in the same direction
- A larger marital status effect
- Interaction pattern is different – simple effect for males is opposite that for depression

We might also consider whether social support should be considered in the analysis…

Descriptive Statistics

Dependent Variable: total social support

GENDER	MARITAL	Mean	Std. Deviation	N
male	single	5.5561	1.05512	122
	married	5.5553	1.21629	47
	Total	5.5559	1.09858	169
female	single	5.8271	1.27056	120
	married	5.7514	1.07645	74
	Total	5.7982	1.19792	194
Total	single	5.6905	1.17239	242
	married	5.6752	1.13190	121
	Total	5.6854	1.15751	363

Not much going on here…

So, based on these analyses, we might decide that out analysis of the relationship between gender and marital status and depression might benefit from the inclusion of stress, but probably not social support.

If we decided to include stress, we might also want to consider the interactions of stress with the other variables in the model. There are three approaches…

1. include every known variable and all possible interactions -- you can see where this would lead…
2. Include the "theoretically interesting" or "hypothesis based" ones – and waste serendipity?
3. test whether these effects contribute to the model – concerns about sufficient power/sample size, and how the model would change with other variables added

Here's the code to make the necessary variables for #3

- dummy codes for marital and gender
- centered stress variable
- interaction codes as products of respective effects

```
If (gender = 1) genc = 1.
if (gender = 2) genc = 0.

if (marital = 1) marc = 1.
if (marital = 2) marc = 0.

compute c_stress = stress - 8.62.

compute gm_int = genc * marc.

compute sg_int = c_stress * genc.

compute sm_int = c_stress * marc.

compute gms_int = genc * marc * c_stress.
```

You're likely to see two approaches to testing the inclusion of these interaction terms…

1. Include them in a full model and see if they contribute – keep them in the model if they do (remembering that you should have all the lower order effects for any higher order effect in the model)
2. Test the R^2 difference between a model that does and doesn't include these variables. Consider that this is a less specific model, in that the R^2 test really tests if the contribution of the added effects is significant on average – so it is possible to have a non-significant R^2? with one or more of the added effects being significant.

Doing the latter also gets you the former… Hang on – mixing depression and stress is always "interesting"

Model Summary

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	.460[a]	.212	.203	5.98053	.212	24.050	4	358	.000
2	.503[b]	.253	.238	5.84705	.041	6.510	3	355	.000

a. Predictors: (Constant), C_STRESS, GENC, MARC, GM_INT

b. Predictors: (Constant), C_STRESS, GENC, MARC, GM_INT, GMS_INT, SM_INT, SG_INT

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3440.800	4	860.200	24.05	.000[a]
	Residual	12804.489	358	35.767		
	Total	16245.289	362			
2	Regression	4108.543	7	586.935	17.17	.000[b]
	Residual	12136.746	355	34.188		
	Total	16245.289	362			

a. Predictors: (Constant), C_STRESS, GENC, MARC, GM_INT

b. Predictors: (Constant), C_STRESS, GENC, MARC, GM_INT, GMS_INT, SM_INT, SG_INT

c. Dependent Variable: DEP

Looks like these stress-related interactions should play a role in the analysis !

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	6.535	.700		9.341	.000
	GENC	.135	1.116	.010	.121	.904
	MARC	1.997	.895	.141	2.231	.026
	GM_INT	-3.727	1.356	-.263	-2.749	.006
	C_STRESS	.348	.043	.389	8.167	.000
2	(Constant)	6.502	.701		9.278	.000
	GENC	.653	1.118	.049	.584	.560
	MARC	1.896	.887	.134	2.138	.033
	GM_INT	-4.034	1.351	-.285	-2.985	.003
	C_STRESS	.330	.092	.369	3.574	.000
	SG_INT	.317	.143	.233	2.213	.028
	SM_INT	.110	.116	.098	.949	.343
	GMS_INT	-.655	.177	-.392	-3.705	.000

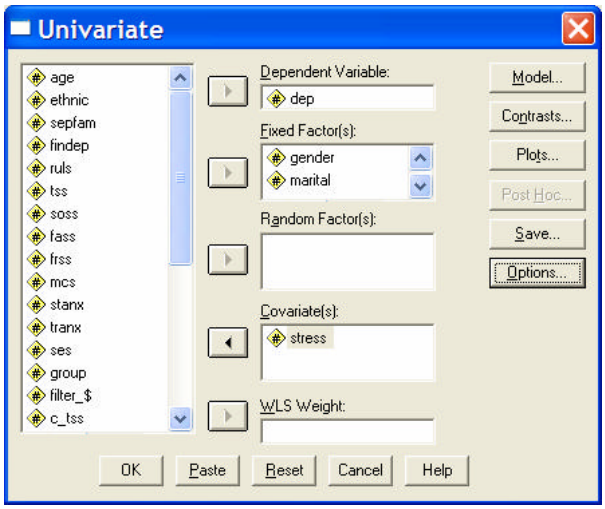
a. Dependent Variable: DEP

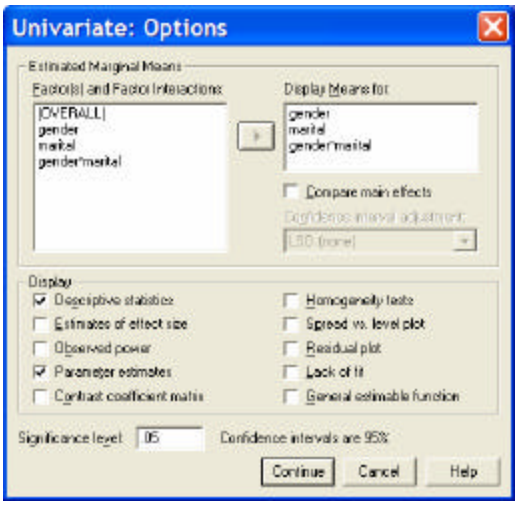
When you don't include the interactions involving the covariate in the model, two things might happen…

1. As in this case, you might miss one or more of these interactions that are significant and interesting – here there are stress*gender and stress*gender*marital interactions that we would miss doing a "regular ANCOVA without the covariate interactions (If you don't look for it you can't find it)
2. There might be a 3-way interaction involving the covariate the pattern of which makes the 2-way involving the 2 IVs misleading – an underspecification effect from not having "all the variables" in the model

Of course, this gets complicated quickly if you have multiple covariates – the number of 3- and 4-way interactions can add up!!

Here's the factorial ANCOVA without the stress-related interactions (just as an example).





Descriptive Statistics

Dependent Variable: DEP

GENDER	MARITAL	Mean	Std. Deviation	N
male	single	5.0492	5.58459	122
	married	6.1064	6.69923	47
	Total	5.3432	5.91411	169
female	single	9.0417	8.10995	120
	married	5.8919	4.57081	74
	Total	7.8402	7.12801	194
Total	single	7.0289	7.22053	242
	married	5.9752	5.47032	121
	Total	6.6777	6.69899	363

Tests of Between-Subjects Effects

Dependent Variable: DEP

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3440.800[a]	4	860.200	24.050	.000
Intercept	2037.009	1	2037.009	56.953	.000
STRESS	2385.611	1	2385.611	66.699	.000
GENDER	232.844	1	232.844	6.510	.011
MARITAL	1.355	1	1.355	.038	.846
GENDER * MARITAL	270.217	1	270.217	7.555	.006
Error	12804.489	358	35.767		
Total	32432.000	363			
Corrected Total	16245.289	362			

a. R Squared = .212 (Adjusted R Squared = .203)

1. GENDER

Dependent Variable: DEP

GENDER	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
male	5.806[a]	.514	4.795	6.817
female	7.535[a]	.442	6.666	8.404

a. Evaluated at covariates appeared in the model: STRESS = 8.62.

2. MARITAL

Dependent Variable: DEP

MARITAL	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
single	6.737[a]	.386	5.977	7.497
married	6.604[a]	.563	5.497	7.710

a. Evaluated at covariates appeared in the model: STRESS = 8.62.

3. GENDER * MARITAL

Dependent Variable: DEP

GENDER	MARITAL	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
male	single	4.941[a]	.542	3.876	6.006
	married	6.671[a]	.875	4.950	8.392
female	single	8.534[a]	.549	7.453	9.614
	married	6.536[a]	.700	5.160	7.912

a. Evaluated at covariates appeared in the model: STRESS = 8.62.

Parameter Estimates

Dependent Variable: DEP

Parameter	B	Std. Error	t	Sig.
Intercept	3.537	.753	4.699	.000
STRESS	.348	.043	8.167	.000
[GENDER=1]	.135	1.116	.121	.904
[GENDER=2]	0[a]	.	.	.
[MARITAL=1]	1.997	.895	2.231	.026
[MARITAL=2]	0[a]	.	.	.
[GENDER=1] * [MARITAL=1]	-3.727	1.356	-2.749	.006
[GENDER=1] * [MARITAL=2]	0[a]	.	.	.
[GENDER=2] * [MARITAL=1]	0[a]	.	.	.
[GENDER=2] * [MARITAL=2]	0[a]	.	.	.

a. This parameter is set to zero because it is redundant.

Recall that the corrected marginal means and their differences (main effects) will be depend upon if you use the ANOVA summary table and tables of correct marginal means (using effect codes) or the parameter estimates -- interpreting bs as corrected marginal mean differences (using dummy coding). The corrected cell mean differences (simple effects) and so the interaction effects will be the same (but not the absolute cell mean values).