Alternative General Linear Models for the Variances in SPSS, SAS, and STATA

Let's consider data from a hypothetical experiment involving control and treatment groups (n = 50 in each). Each group included kids ages 10 or 11 (n = 25 each), for a total of 100 people.

Model #1: Between-Person Analysis

Our first order of business is to create centered versions of the predictors to provide comparability across the results from these models. What ANOVA does is mean-center all predictors, so that's what we'll do ourselves. These examples use the "Example1_Stacked" data files.

```
* Define locations of files used in examples -- CHANGE THIS.
FILE HANDLE example /NAME = "F:\12_ICPSR\ICPSR_2012_Download\SPSS".

* Open SPSS stacked version of example data.
GET FILE = "example/Example1_Stacked.sav".
DATASET NAME Stacked WINDOW=FRONT.

* SPSS code for mean-centering predictors for analysis.
DATASET ACTIVATE Stacked WINDOW=FRONT.
COMPUTE ageM = age - 10.5.
COMPUTE treatM = treat - .5.
COMPUTE ageXtreat = ageM * treatM.
VARIABLE LABELS   ageM      "Age (0=10.5 mean)"
                  treatM    "Treatment Group (0=.5 mean)"
                  ageXtreat "Age by Treatment Interaction (0=means)".
```

```
* SAS code for mean-centering predictors for analysis;
DATA work.Stacked; SET example.Example1_Stacked;
      ageM = age - 10.5;
      treatM = treat - .5;
      ageXtreat = ageM * treatM;
      LABEL   ageM =      "Age (0=10.5 mean)"
              treatM =    "Treatment Group (0=.5 mean)"
              ageXtreat = "Age by Treatment Interaction (0=means)";
RUN;
```

```
 * STATA code for mean-centering predictors for analysis
gen ageM = age - 10.5
gen treatM = treat - .5
gen agetreat = ageM * treatM
```

Summary of Model #1: Between-Person Analysis (Between-Groups ANOVA, Regression)

Full Model equation:

$$y_i = \beta_0 + \beta_1 AgeM_i + \beta_2 TreatM_i + \beta_3 AgeM_i * TreatM_i + e_i$$

Model for the Means:

$$y_i = 53.33 + 7.06 AgeM_i + 2.97 TreatM_i + 2.83 AgeM_i * TreatM_i$$

Model for the Variances:

ONE error term, the residual e_i, with a mean of 0 and a variance of σ_e^2 =16.08

Fitting a Between-Person Model via Univariate GLM in SPSS:

```
* SPSS example of Between-Person analysis via General Linear Model.
DATASET ACTIVATE Stacked WINDOW=FRONT.
UNIANOVA y BY ageM treatM
      /EMMEANS = TABLES(ageM*treatM)
      /DESIGN = ageM treatM ageM*treatM.
```

The BY statement in SPSS indicates the predictors to be treated as categorical.

Tests of Between-Subjects Effects

Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1517.007(a)	3	505.669	**31.440**	.000
Intercept	284494.953	1	284494.953	**17688.613**	**.000**
ageM	1246.006	1	1246.006	**77.471**	**.000**
treatM	220.920	1	220.920	**13.736**	**.000**
ageM * treatM	50.080	1	50.080	**3.114**	**.081**
Error	1544.017	96	**16.084**		
Total	287555.976	100			
Corrected Total	3061.023	99			

a R Squared = .496 (Adjusted R Squared = .480)

Note that the tests of the three effects above are all in the same box, because they are each tested against the same (and only) error term (i.e., the one residual variance).

Fitting a Between-Person Model via Univariate GLM in SAS:

```
* SAS example of Between-Person model via General Linear Model;
PROC GLM DATA=work.Stacked;
      CLASS ageM treatM;
      MODEL y = ageM treatM ageM*treatM;
      LSMEANS ageM*treatM;
RUN; QUIT;
```

The CLASS statement in SAS indicates the predictors to be treated as categorical.

```
The GLM Procedure      Dependent Variable: y    Hypothetical Outcome Y

                                        Sum of
Source                      DF         Squares     Mean Square    F Value    Pr > F

Model                        3     1517.006542      505.668847      31.44    <.0001
Error                       96     1544.016829       16.083509
Corrected Total             99     3061.023371

R-Square      Coeff Var      Root MSE        y Mean
0.495588       7.518879      4.010425      53.33807

Source                      DF     Type III SS     Mean Square    F Value    Pr > F

ageM                         1     1246.006121     1246.006121      77.47    <.0001
treatM                       1      220.920016      220.920016      13.74    0.0004
ageM*treatM                  1       50.080405       50.080405       3.11    0.0808
```

Fitting a Between-Person Model via Regression in SPSS:

```
* SPSS example of Between-Person analysis via Regression.
DATASET ACTIVATE Stacked WINDOW=FRONT.
REGRESSION
      /DEPENDENT y
      /METHOD=ENTER ageM treatM ageXtreat.
```

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.704(a)	**.496**	**.480**	4.01042

a Predictors: (Constant), ageXtreat, treatM, ageM

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1517.007	3	505.669	**31.440**	.000(a)
	Residual	1544.017	96	**16.084**		
	Total	3061.023	99			

a Predictors: (Constant), ageXtreat, treatM, ageM
b Dependent Variable: y

Note that the F-test of the overall model significance matches that of the "corrected model" in GLM.

Note that the residual variance also matches that of GLM.

Coefficients(a)

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	**(Constant)**	**53.338**	.401		132.999	**.000**
	ageM	**7.060**	.802	.638	8.802	**.000**
	treatM	**2.973**	.802	.269	3.706	**.000**
	ageXtreat	**2.831**	1.604	.128	1.765	**.081**

a Dependent Variable: y

Note that the t-values for the fixed effects when squared match the F-tests given for each in GLM.

Fitting a Between-Person Model via Regression in SAS:

```
* SAS example of Between-Person model via Regression;
PROC REG DATA=work.Stacked;
      MODEL y = ageM treatM ageXtreat;
RUN; QUIT;
```

The REG Procedure   Dependent Variable: y Hypothetical Outcome Y

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1517.00654	505.66885	**31.44**	<.0001
Error	96	1544.01683	16.08351		
Corrected Total	99	3061.02337			

Root MSE	4.01042	R-Square	0.4956
Dependent Mean	53.33807	Adj R-Sq	0.4798
Coeff Var	7.51888		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	53.33807	0.40104	133.00	<.0001
ageM	Age (0=10.5 mean)	1	7.05976	0.80208	8.80	<.0001
treatM	Treatment Group (0=.5 mean)	1	2.97268	0.80208	3.71	0.0004
ageXtreat	Age by Treatment Interaction (0=means)	1	2.83070	1.60417	1.76	0.0808

Fitting Between-Person Model via MIXED in SPSS:

```
* SPSS example of Between-Person model via Mixed.
DATASET ACTIVATE Stacked WINDOW=FRONT.
MIXED y WITH ageM treatM
      /METHOD = REML
      /PRINT  = SOLUTION
      /FIXED  = ageM treatM ageM*treatM.
```

MIXED dv WITH continuous predictors
/METHOD = REML or ML
/PRINT = regression solution
/FIXED = predictors for means model

Model Dimension

		Number of Levels	Number of Parameters
Fixed Effects	Intercept	1	1
	ageM	1	1
	treatM	1	1
	ageM * treatM	1	1
Residual			1
Total		4	5

This table tells you how many parameters are in your model for the means (the fixed effects, or 4 here) and in your model for the variances (the residual, or 1 here).

Information Criteria[a]

-2 Restricted Log Likelihood	551.980
Akaike's Information Criterion (AIC)	553.980
Hurvich and Tsai's Criterion (AICC)	554.023
Bozdogan's Criterion (CAIC)	557.544
Schwarz's Bayesian Criterion (BIC)	556.544

The information criteria are displayed in smaller-is-better forms.
a. Dependent Variable: y Hypothetical Outcome Y.

This table provides evidence about the fit of the model to the data (more on this shortly).

Fixed Effects

Type III Tests of Fixed Effects

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	96	**17688.613**	**.000**
ageM	1	96	**77.471**	**.000**
treatM	1	96	**13.736**	**.000**
ageM * treatM	1	96	**3.114**	**.081**

Note these F-values for the tests of our predictors exactly match those from GLM.

Estimates of Fixed Effects

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	**53.338**	.401	96	132.999	**.000**	52.542	54.134
ageM	**7.060**	.802	96	8.802	**.000**	5.468	8.652
treatM	**2.973**	.802	96	3.706	**.000**	1.381	4.565
ageM * treatM	**2.831**	1.604	96	1.765	**.081**	-0.354	6.015

Covariance Parameters

Estimates of Covariance Parameters

Parameter	Estimate	Std. Error
Residual	**16.084**	2.321

Note these fixed effects exactly match the unstandardized weights from regression.

The residual variance exactly matches that from both GLM and regression.

Fitting Between-Person Model via MIXED in SAS:

```
* SAS example of Between-Person model via MIXED;
PROC MIXED DATA=work.Stacked NOITPRINT COVTEST METHOD=REML;
      MODEL y = ageM treatM ageM*treatM / SOLUTION DDFM=BW;
RUN;
```

On MIXED line, METHOD = REML or ML
MODEL dv = fixed effects
/SOLUTION = regression solution
/DDFM = Between-Within denominator degrees of freedom (SPSS does not give you choices)

The Mixed Procedure

Model Information

Data Set	WORK.ANOVA_STACKED
Dependent Variable	y
Covariance Structure	Diagonal
Estimation Method	REML
Residual Variance Method	Profile
Fixed Effects SE Method	Model-Based
Degrees of Freedom Method	Residual

Dimensions

Covariance Parameters	1
Columns in X	4
Columns in Z	0
Subjects	1
Max Obs Per Subject	100

This table tells you how many parameters are in your model for the means ("columns in x", the fixed effects including the intercept, or 4 here) and in your model for the variances ("covariance parameters", the residual, or 1 here).
It also tells you how many observations were read per subject.

Number of Observations

Number of Observations Read	100
Number of Observations Used	100
Number of Observations Not Used	0

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z
Residual	16.0835	2.3215	6.93	<.0001

The residual variance exactly matches that from both GLM and regression.

Fit Statistics

-2 Res Log Likelihood	552.0
AIC (smaller is better)	554.0
AICC (smaller is better)	554.0
BIC (smaller is better)	556.5

This table provides evidence about the fit of the model to the data (more on this shortly).

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	53.3381	0.4010	96	133.00	<.0001
ageM	7.0598	0.8021	96	8.80	<.0001
treatM	2.9727	0.8021	96	3.71	0.0004
ageM*treatM	2.8307	1.6042	96	1.76	0.0808

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
ageM	1	96	77.47	<.0001
treatM	1	96	13.74	0.0004
ageM*treatM	1	96	3.11	0.0808

Note these fixed effects exactly match the unstandardized weights from regression.

Note these F-values for the tests of our predictors exactly match those from GLM.

Fitting Between-Person Model via XTMIXED in STATA:

```
 * STATA example of Between-Person model via XTMIXED
xtmixed y c.ageM c.treatM c.ageM#c.treatM, || PersonID: , noconstant ///
      variance reml residuals(independent),
      estat ic, n(100),
      estimates store Eonly
```

DV = y, Categorical predictors of age, treat, and their interaction
Level 2 ID is PersonID, noconstant = no random intercept
Print variances instead of SD, use reml
residuals → R matrix is diagonal
estat ic → Print IC given N = 100 (as single-level model)
estimates store → Save model results as "Eonly"

```
. * STATA example of Between-Person model via XTMIXED

. xtmixed y c.ageM c.treatM c.ageM#c.treatM, || PersonID: , noconstant ///
>         variance reml residuals(independent),

Note: all random-effects equations are empty; model is linear regression

Mixed-effects REML regression                   Number of obs      =       100

                                                Wald chi2(3)       =     94.32
Log restricted-likelihood = -275.98998          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ageM |   7.059762    .802085     8.80   0.000     5.487705     8.63182
      treatM |   2.972676    .802085     3.71   0.000     1.400618    4.544733
             |
     c.ageM#|
    c.treatM |     2.8307    1.60417     1.76   0.078    -.313415    5.974816
             |
       _cons |   53.33807   .4010425   133.00   0.000     52.55204     54.1241
------------------------------------------------------------------------------
```

Note that STATA provides "_cons" as the fixed intercept

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
               var(Residual) |   16.08351   2.274552      12.18997    21.22066
------------------------------------------------------------------------------

.         estat ic, n(100),

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    100           .     -275.99      5       561.98    575.0058
-----------------------------------------------------------------------------
               Note:  N=100 used in calculating BIC
```

Note that STATA provides LL rather than -2LL (deviance)

Model #2: Within-Person Analysis

What if instead of having 4 groups (age 10 vs. 11, control vs. treatment), we had people measured at age 10 AND age 11 in each group (a longitudinal study)? We would need a different model for the variances that takes into account the fact that (a) multiple data points come from the same person (i.e., the residuals from the same person are likely to be correlated, or *dependent*), or stated another way (b) that people differ systematically from each other in their level of outcome y, regardless of time. So we'll fit a model with a separate error term for within-person variation, too.

In order to conduct a repeated measures ANOVA, the data have to be re-arranged from this type of "stacked" (i.e., long, univariate) format (as in the "Example1_Stacked" data files):

CaseID	PersonID	age	treat	y
1	1	10	0	54.40
2	1	11	0	56.50
3	2	10	0	52.71
4	2	11	0	55.97
5	3	10	0	52.24
6	3	11	0	57.93

……………………………………………………

to a "multivariate" (i.e., wide) format (as in the "Example1_Multiv" data files):

PersonID	Treat	y10	y11
1	0	54.40	56.50
2	0	52.71	55.97
3	0	52.24	57.93

…………………………………………

There is a wizard called "Restructure" under the "Data" menu in SPSS that will do this for you! I've also included instructions on how to restructure data like this in SPSS or SAS.

Fitting a Within-Person Model via Repeated Measures GLM in SPSS:

```
* Open SPSS multivariate version of example data.
GET FILE = "example/Example1_Multiv.sav".
DATASET NAME Multiv WINDOW=FRONT.

* Mean-centering predictor for analysis.
DATASET ACTIVATE Multiv WINDOW=FRONT.
COMPUTE treatM = treat - .5.
VARIABLE LABELS treatM "Treatment Group (0=.5 mean)".

* SPSS example of Within-Person model via RM GLM.
DATASET ACTIVATE Multiv WINDOW=FRONT.
GLM y10 y11 BY treatM
      /WSFACTOR = age 2 Polynomial
      /EMMEANS = TABLES(treatM*age)
      /WSDESIGN = age
      /DESIGN = treatM.
```

Tests of Within-Subjects Effects

Measure: MEASURE_1

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
age	Sphericity Assumed	1246.006	1	1246.006	**177.936**	**.000**
	Greenhouse-Geisser	1246.006	1.000	1246.006	177.936	.000
	Huynh-Feldt	1246.006	1.000	1246.006	177.936	.000
	Lower-bound	1246.006	1.000	1246.006	177.936	.000
age * treatM	Sphericity Assumed	50.080	1	50.080	**7.152**	**.010**
	Greenhouse-Geisser	50.080	1.000	50.080	7.152	.010
	Huynh-Feldt	50.080	1.000	50.080	7.152	.010
	Lower-bound	50.080	1.000	50.080	7.152	.010
Error(age)	Sphericity Assumed	336.122	48	**7.003**		
	Greenhouse-Geisser	336.122	48.000	7.003		
	Huynh-Feldt	336.122	48.000	7.003		
	Lower-bound	336.122	48.000	7.003		

Tests of Between-Subjects Effects

Measure: MEASURE_1
Transformed Variable: Average

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	284494.953	1	284494.953	11305.424	.000
treatM	220.920	1	220.920	**8.779**	**.005**
Error	1207.894	48	**25.164**		

Note the tests of the tests for age and age*treatM are in a different box than the test for treatM. That's because they are tested against a smaller (just within-person) residual error variance term. As a result, the age*treatM interaction is now significant!

The treatM effect is tested against a different error term (all the residual error in the model, the between and within parts of the residual variance).

Previous Between-Person Model SPSS results to compare against:

Tests of Between-Subjects Effects (a)

Dependent Variable: y

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	1517.007(a)	3	505.669	31.440	.000
Intercept	284494.953	1	284494.953	**17688.613**	**.000**
ageM	1246.006	1	1246.006	**77.471**	**.000**
treatM	220.920	1	220.920	**13.736**	**.000**
ageM * treatM	50.080	1	50.080	**3.114**	**.081**
Error	1544.017	96	**16.084**		
Total	287555.976	100			
Corrected Total	3061.023	99			

a R Squared = .496 (Adjusted R Squared = .480)

Fitting a Within-Person Model via Repeated Measures GLM in SAS:

```
* Open SAS multivariate version of anova data into work library;
* Mean-centering predictor for analysis;
DATA work.Multiv; SET example.Example1_Multiv;
      treatM = treat - .5;
      LABEL treatM = "Treatment Group (0=.5 mean)";
RUN;

* SAS example of Within-Person model via RM GLM;
PROC GLM DATA=work.Multiv;
      CLASS treatM;
      MODEL y10 y11 = treatM / NOUNI;
      REPEATED age 2 / ;
      LSMEANS treatM;
RUN; QUIT;
```

Note the tests of the tests for age and age*treatM are in a different box than the test for treatM. That's because they are tested against a smaller (just within-person) residual error variance term. As a result, the age*treatM interaction is now significant!

The treatM effect is tested against a different error term (all the residual error in the model, the between and within parts of the residual variance).

The GLM Procedure
Repeated Measures Analysis of Variance
Univariate Tests of Hypotheses for Within Subject Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
age	1	1246.006121	1246.006121	177.94	<.0001
age*treatM	1	50.080405	50.080405	7.15	0.0102
Error(age)	48	336.122471	7.002551		

The GLM Procedure
Repeated Measures Analysis of Variance
Tests of Hypotheses for Between Subjects Effects

Source	DF	Type III SS	Mean Square	F Value	Pr > F
treatM	1	220.920016	220.920016	8.78	0.0047
Error	48	1207.894358	25.164466		

Previous Between-Person Model SAS results to compare against:

The GLM Procedure   Dependent Variable: y   Hypothetical Outcome Y

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1517.006542	505.668847	31.44	<.0001
Error	96	1544.016829	16.083509		
Corrected Total	99	3061.023371			

R-Square	Coeff Var	Root MSE	y Mean
0.495588	7.518879	4.010425	53.33807

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ageM	1	1246.006121	1246.006121	77.47	<.0001
treatM	1	220.920016	220.920016	13.74	0.0004
ageM*treatM	1	50.080405	50.080405	3.11	0.0808

Fitting a Within-Person Model via MIXED in SPSS:

MIXED requires the original “Example1_Stacked” data file, but now it is important to tell it that the rows that belong to the same person are identified with an ID variable, here “PersonID”:

```
* SPSS example of Within-Person model via Mixed.
DATASET ACTIVATE Stacked WINDOW=FRONT.
MIXED y WITH ageM treatM
      /METHOD = REML
      /PRINT  = SOLUTION
      /FIXED  = ageM treatM ageM*treatM
      /RANDOM = INTERCEPT | SUBJECT(PersonID) COVTYPE(UN).
```

MIXED dv WITH continuous predictors
/METHOD = REML or ML
/PRINT = regression solution
/FIXED = predictors for means model
/RANDOM = person effects

Model Dimension(a)

		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables
Fixed Effects	Intercept	1		1	
	ageM	1		1	
	treatM	1		1	
	ageM * treatM	1		1	
Random Effects	Intercept	1	Identity	1	PersonID
Residual				1	
Total		5		6	

a Dependent Variable: y.

Now we have the same model for the means (4 fixed effects), but a 2-parameter model for the variances (U_0 + e) instead of just one residual variance (e).

Information Criteria[a]

-2 Restricted Log Likelihood	533.554
Akaike's Information Criterion (AIC)	537.554
Hurvich and Tsai's Criterion (AICC)	537.683
Bozdogan's Criterion (CAIC)	544.682
Schwarz's Bayesian Criterion (BIC)	542.682

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: y Hypothetical Outcome Y.

This table provides evidence about the fit of the model to the data (more on this shortly).

Fixed Effects

Type III Tests of Fixed Effects(a)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	48	**11305.424**	**.000**
ageM	1	48	**177.936**	**.000**
treatM	1	48	**8.779**	**.005**
ageM * treatM	1	48	**7.152**	**.010**

a Dependent Variable: y.

Note these F-values exactly match those from the Repeated Measures GLM.

Estimates of Fixed Effects

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Intercept	**53.338**	.502	48	106.327	**.000**	52.329449	54.346687
ageM	**7.060**	.529	48	13.339	**.000**	5.995641	8.123884
treatM	**2.973**	1.003	48	2.963	**.005**	.955438	4.989913
ageM * treatM	**2.831**	1.058	48	2.674	**.010**	.702457	4.958944

Covariance Parameters

Parameter	Estimate	Std. Error
Residual	**7.002**	1.429
Intercept [subject = PersonID] Variance	**9.080**	2.666

Residual Variance for e_{ti}
Random Intercept Variance for U_{0i}

Fitting a Within-Person Model via MIXED in SAS:

```
* SAS example of Within-Person model via MIXED;
PROC MIXED DATA=work.Stacked NOCLPRINT NOITPRINT COVTEST METHOD=REML;
      CLASS PersonID;
      MODEL y = ageM treatM ageM*treatM / SOLUTION DDFM=BW;
      RANDOM INTERCEPT / SUBJECT=PersonID TYPE=UN;
RUN;
```

Dimensions

Covariance Parameters	2
Columns in X	4
Columns in Z Per Subject	1
Subjects	50
Max Obs Per Subject	2

Now we have the same model for the means (4 columns in X), but a 2-parameter model for the variances instead of just one residual variance (covariance parameters = 2, U_{0i} + e).

In addition, it recognizes that each subject has two observations.

Number of Observations

Number of Observations Read	100
Number of Observations Used	100
Number of Observations Not Used	0

Fit Statistics

-2 Res Log Likelihood	533.6
AIC (smaller is better)	537.6
AICC (smaller is better)	537.7
BIC (smaller is better)	541.4

This table provides evidence about the fit of the model to the data (more on this shortly).

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	PersonID	9.0810	2.6659	3.41	0.0003	Random Intercept Variance for U_{0i}
Residual		7.0026	1.4294	4.90	<.0001	Residual Variance for e_{ti}

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	53.3381	0.5016	48	106.33	<.0001
ageM	7.0598	0.5292	48	13.34	<.0001
treatM	2.9727	1.0033	48	2.96	0.0047
ageM*treatM	2.8307	1.0585	48	2.67	0.0102

Type 3 Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
ageM	1	48	177.94	<.0001
treatM	1	48	8.78	0.0047
ageM*treatM	1	48	7.15	0.0102

Note these F-values exactly match those from the Repeated Measures GLM.

Fitting a Within-Person Model via MIXED in STATA:

```
 * STATA example of Within-Person model via XTMIXED
xtmixed y c.ageM c.treatM c.ageM#c.treatM,  || PersonID: ,  ///
      variance reml covariance(unstructured) residuals(independent),
      estat ic, n(50),
      estimates store UandE,
      lrtest UandE Eonly
```

DV = y, Categorical predictors of age, treat, & interaction
Level 2 ID is PersonID, random intercept by default
Print variances instead of SD, use reml
residuals → R matrix is diagonal
estat ic → Print IC given N = 50 persons
estimates store → Save model results as “UandE”
lrtest → deviance test of UandE vs. Eonly

```
. * STATA example of Within-Person model via XTMIXED
. xtmixed y c.ageM c.treatM c.ageM#c.treatM,  || PersonID: ,  ///
>         variance reml covariance(unstructured) residuals(independent),

Note: single-variable random-effects specification; covariance structure set to
identity
Mixed-effects REML regression                   Number of obs      =       100
Group variable: PersonID                        Number of groups   =        50
                                                Obs per group: min =         2
                                                               avg =       2.0
                                                               max =         2

                                                Wald chi2(3)       =    193.87
Log restricted-likelihood = -266.77683          Prob > chi2        =    0.0000

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        ageM |   7.059762   .5292467    13.34   0.000     6.022458    8.097067
      treatM |   2.972676   1.003284     2.96   0.003     1.006276    4.939076
             |
     c.ageM#|
    c.treatM |     2.8307   1.058493     2.67   0.007     .7560913     4.90531
             |
       _cons |   53.33807   .5016419   106.33   0.000     52.35487    54.32127
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
PersonID: Identity           |
                  var(_cons) |   9.080954   2.665923      5.107892    16.14438
-----------------------------+------------------------------------------------
               var(Residual) |   7.002553    1.42939      4.693595    10.44738
------------------------------------------------------------------------------
LR test vs. linear regression: chibar2(01) =    18.43 Prob >= chibar2 = 0.0000

.         estat ic, n(50),
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |     50           .   -266.7768      6     545.5537    557.0258
-----------------------------------------------------------------------------
Note:  N=50 used in calculating BIC

.         lrtest UandE Eonly
Likelihood-ratio test                                 LR chi2(1)  =     18.43
(Assumption: Eonly nested in UandE)                   Prob > chi2 =    0.0000
```

This is the deviance difference test of whether the U + E WP model fits better than the E-only BP model…

```
Note: The reported degrees of freedom assumes the null hypothesis is not on the
      boundary of the parameter space.  If this is not true, then the reported test is
      conservative.
Note: LR tests based on REML are valid only when the fixed-effects specification is
      identical for both models.
```

Results from previous Between-Person Model:

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	53.3381	0.4010	96	133.00	<.0001
ageM	7.0598	0.8021	96	8.80	<.0001
treatM	2.9727	0.8021	96	3.71	0.0004
ageM*treatM	2.8307	1.6042	96	1.76	0.0808

Covariance Parameter Estimates

Cov Parm	Estimate	Standard Error	Z Value	Pr Z	
Residual	16.0835	2.3215	6.93	<.0001	Residual Variance for e_{ti}

Results from Within-Person Model:

Solution for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	53.3381	0.5016	48	106.33	<.0001
ageM	7.0598	0.5292	48	13.34	<.0001
treatM	2.9727	1.0033	48	2.96	0.0047
ageM*treatM	2.8307	1.0585	48	2.67	0.0102

Note that although the estimates for the fixed effects are the same, their standard errors and associated significance tests are not. This is due to both the change in sample size and/or the change in the error terms on which they are based.

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Z Value	Pr Z	
UN(1,1)	PersonID	9.0810	2.6659	3.41	0.0003	Random Intercept Variance for U_{0i}
Residual		7.0026	1.4294	4.90	<.0001	Residual Variance for e_{ti}

Summary of Model #2: Within-Person Analysis

Full Model equation:

$$y_{ti} = \beta_0 + \beta_1 AgeM_{ti} + \beta_2 TreatM_i + \beta_3 AgeM_{ti} * TreatM_i + U_{0i} + e_{ti}$$

SAME Model for the Means:

$$y_{ti} = 53.33 + 7.06 AgeM_{ti} + 2.97 TreatM_i + 2.83 AgeM_{ti} * TreatM_i$$

Different Model for the Variances:

TWO error terms: the e_{ti}, with mean=0 and variance of $\sigma_e^2 = 7.002$, and the U_{0i} with mean=0 and variance of $\tau_{U0}^2 = 9.080$ (sum to 16.084, the original amount of residual variance)

Here's how we get back to the Total Variance from Repeated Measures ANOVA:

Intercept Variance = $(MS_{SUBJECT.ERROR} - MS_{AGE.ERROR})$ / # ages

Intercept Variance = (25.164 – 7.002) / 2 = 9.080

What other terms that could possibly be included are missing? Are they *really* missing?