

**Example 2 Part 2: Models for Accelerated Longitudinal Designs using
Multilevel and Single-Level Structural Equation Modeling (M-SEM and SEM)**
(complete syntax, data, and output available for Mplus electronically)

These data come from Hoffman (2015) chapter 10, which examined prediction of a memory outcome (prose recall) from two metrics of time: years since birth (centered at 84 years) and years in study (centered at the first occasion) in a sample of 557 observations from 207 older adults. This handout shows how to estimate a subset of the models from Part 1 (with time as the level-1 predictor) using multilevel and single-level SEM in Mplus.

Model 0. Empty Means, Random Intercept for Recall:

Level-1: $recall_{ti} = \beta_{0i} + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + U_{0i}$

Model 0 STATA Univariate MLM:

```
display "Model 0: Empty Means, Random Intercept Model for Recall Outcome"
mixed recall , || personid: , mle nolog
```

Model 0 R Univariate MLM:

```
print("Model 0: Empty Means, Random Intercept Model for Recall Outcome")
Empty = lmer(data=Example2, REML=FALSE, formula=recall~1+(1|PersonID))
llikAIC(Empty); summary(Empty); icc(Empty); print("Does the random intercept improve model fit?")
ranova(Empty, reduce.term=TRUE) # LRT for removing random intercept
```

```
$AICtab
      AIC      BIC    logLik  deviance  df.resid
2863.3550 2876.3227 -1428.6775  2857.3550   554.0000  deviance = -2LL (for homework)
```

Random effects:

Groups	Name	Variance	Std.Dev.	
PersonID	(Intercept)	10.4579	3.2339	Level-2 variance of U_{0i}
	Residual	5.1646	2.2726	Level-1 variance of e_{ti}

Number of obs: 557, groups: PersonID, 207

$$ICC = \frac{10.4579}{10.4579 + 5.1646} = .669$$

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)
(Intercept)	9.73491	0.25058	197.01209	38.85	< 2.2e-16

So 67% of the variance in recall is initially due to person mean differences.

Intraclass Correlation Coefficient

Adjusted ICC: 0.669
Unadjusted ICC: 0.669

Model 0 Mplus M-SEM:

```
TITLE: M-SEM Model 0: Empty Means, Random Intercept
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
      TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults
```

VARIABLE:

```
! List of variables in long data file IN ORDER
  NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
  USEVARIABLE = recall;
! Missing data identifier
  MISSING ARE ALL (-999);
! MSEM options
  CLUSTER = PersonID; ! Level-2 ID
  BETWEEN = ; ! Observed ONLY level-2 predictors
  WITHIN = ; ! Observed ONLY level-1 predictors
```

ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR = ML;

MODEL: ! LEVEL-1 = WITHIN, LEVEL-2 = BETWEEN

%WITHIN%

recall; ! L1 R: Residual variance

%BETWEEN%

[recall]; ! Fixed intercept

recall; ! L2 G: Random intercept variance only

Number of Free Parameters 3
 Loglikelihood
 H0 Value -1428.678

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
Variances				
RECALL	5.170	0.394	13.127	0.000
Between Level				
Means				
RECALL	9.736	0.251	38.815	0.000
Variances				
RECALL	10.438	1.306	7.991	0.000

Model 0 Mplus SEM:

TITLE: SEM Model 0: Empty Means, Random Intercept

DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
 TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

! Unstacking to multivariate format

DATA LONGTOWIDE:

! Names of old stacked former variables (without numbers)

LONG = recall|time;

! Names of new multivariate variables (that use numbers)

WIDE = recall10 recall12 recall14 recall16 recall18 |
 time0 time2 time4 time6 time8;

! Variable with level-2 ID info

IDVARIABLE = PersonID;

! Old level-1 identifier

REPETITION = occasion (0 2 4 6 8);

VARIABLE:

! List of variables in long data file IN ORDER

NAMES = PersonID occasion ageT0 tvage recall time;

! Variables to be analyzed in this model (new defined at end)

USEVARIABLE = recall10 recall12 recall14 recall16 recall18;

! Missing data identifier

MISSING ARE ALL (-999);

ANALYSIS: TYPE = GENERAL; ESTIMATOR = ML;

OUTPUT: STDYX; ! Standardized solution in SEM

MODEL:

[recall10-recall18@0]; ! All variable intercepts fixed to 0

recall10-recall18 (Resvar); ! L1 R residual variances held equal

! Recall intercept-only model

Int BY recall10-recall18@1;

! Level-2 model

[Int]; ! Fixed intercept

Int; ! L2 G: Random intercept variance

Number of Free Parameters 3
 Loglikelihood
 H0 Value -1428.678

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT BY				
RECALL0	1.000	0.000	999.000	999.000
RECALL2	1.000	0.000	999.000	999.000
RECALL4	1.000	0.000	999.000	999.000
RECALL6	1.000	0.000	999.000	999.000
RECALL8	1.000	0.000	999.000	999.000
Means				
INT	9.735	0.251	38.786	0.000
Intercepts				
RECALL0	0.000	0.000	999.000	999.000
RECALL2	0.000	0.000	999.000	999.000
RECALL4	0.000	0.000	999.000	999.000
RECALL6	0.000	0.000	999.000	999.000
RECALL8	0.000	0.000	999.000	999.000
Variances				
INT	10.458	1.310	7.986	0.000
Residual Variances				
RECALL0	5.165	0.393	13.141	0.000
RECALL2	5.165	0.393	13.141	0.000
RECALL4	5.165	0.393	13.141	0.000
RECALL6	5.165	0.393	13.141	0.000
RECALL8	5.165	0.393	13.141	0.000

STANDARDIZED MODEL RESULTS

STDYX Standardization

R-SQUARE

Observed Variable	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
RECALL0	0.669	0.035	19.285	0.000
RECALL2	0.669	0.035	19.285	0.000
RECALL4	0.669	0.035	19.285	0.000
RECALL6	0.669	0.035	19.285	0.000
RECALL8	0.669	0.035	19.285	0.000

Model 1b. Syntax and Partial Output for Fixed Quadratic Time, Random Intercept for Recall:

Level-1 Time: $recall_{ti} = \beta_{0i} + \beta_{1i}(age_{ti} - ageT0_i) + \beta_{2i}(age_{ti} - ageT0_i)^2 + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + U_{0i}$, $\beta_{1i} = \gamma_{10}$, $\beta_{2i} = \gamma_{20}$

Where $age_{ti} - ageT0_i = time_{ti}$ (as years-in-study rather than years-since-birth)

Model 1b STATA Univariate MLM:

```
display "Model 1b Time: Fixed Quadratic, Random Intercept Model"
mixed recall c.time c.timesq, || personid: , mle nolog
```

Model 1b R Univariate MLM:

```
print("Model 1b Time: Fixed Quadratic, Random Intercept Model")
RITim = lmer(data=Example2, REML=FALSE, formula=recall~1+time+timesq+(1|PersonID))
llikAIC(RITim); summary(RITim)
```

```
$AICtab
      AIC      BIC    logLik  deviance  df.resid
2856.0088 2877.6216 -1423.0044  2846.0088   552.0000  deviance = -2LL
```

Random effects:

Groups	Name	Variance	Std.Dev.	
PersonID	(Intercept)	10.622	3.2591	Level-2 variance of U_{0i}
Residual		4.983	2.2323	Level-1 variance of e_{ti}

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	9.660987	0.274986	282.359753	35.1326	< 2.2e-16	gamma00
time	0.261331	0.119243	377.995056	2.1916	0.029019	gamma10
timesq	-0.046907	0.015826	366.791758	-2.9640	0.003235	gamma20

Model 1b Mplus M-SEM:

TITLE: M-SEM Model 1b: Fixed Quadratic Time, Random Intercept
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
 TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

VARIABLE:

! List of variables in long data file IN ORDER
 NAMES = PersonID occasion ageT0 tvage recall time;
 ! Variables to be analyzed in this model (new defined at end)
 USEVARIABLE = recall time timesq;
 ! Missing data identifier
 MISSING ARE ALL (-999);
 ! MSEM options
 CLUSTER = PersonID; ! Level-2 ID
 BETWEEN = ; ! Observed ONLY level-2 predictors
 WITHIN = time timesq; ! Observed ONLY level-1 predictors

DEFINE: ! Make squared version of time
 timesq = time*time;

ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR = ML;

MODEL: ! LEVEL-1 = WITHIN, LEVEL-2 = BETWEEN

%WITHIN%
 recall; ! L1 R: Residual variance
 recall ON time; ! No B1i placeholder yet because fixed linear only
 recall ON timesq; ! No B2i placeholder because fixed quad only

%BETWEEN%
 [recall]; ! Fixed intercept
 recall; ! L2 G: Random intercept variance only

Number of Free Parameters 5
 Loglikelihood
 H0 Value -1423.004

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
RECALL ON				
TIME	0.261	0.119	2.187	0.029
TIMESQ	-0.047	0.016	-2.963	0.003
Residual Variances				
RECALL	4.983	0.380	13.122	0.000
Between Level				
Means				
RECALL	9.661	0.275	35.129	0.000
Variances				
RECALL	10.622	1.320	8.044	0.000

Model 1b Mplus SEM:

```

TITLE: SEM Model 1b: Fixed Quadratic Time, Random Intercept
DATA: FILE = Chapter10.csv;           ! Syntax in same folder as data
        TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

! Unstacking to multivariate format
DATA LONGTOWIDE:
! Names of old stacked former variables (without numbers)
LONG = recall|time;
! Names of new multivariate variables (that use numbers)
WIDE = recall0 recall2 recall4 recall6 recall8 |
      time0 time2 time4 time6 time8;
! Variable with level-2 ID info
IDVARIABLE = PersonID;
! Old level-1 identifier
REPETITION = occasion (0 2 4 6 8);

VARIABLE:
! List of variables in long data file IN ORDER
NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall0 recall2 recall4 recall6 recall8
            time0 time2 time4 time6 time8;
! Missing data identifier
MISSING ARE ALL (-999);
! Exact time definition variables
TSCORES = time0 time2 time4 time6 time8;

ANALYSIS: TYPE = RANDOM; ESTIMATOR = ML;

MODEL:
[recall0-recall8@0];           ! All variable intercepts fixed to 0
recall0-recall8 (Resvar);     ! L1 R residual variances held equal

! Recall quadratic growth model using exact time as loadings
Int Lin Qua | recall0-recall8 AT time0-time8;

! Level-2 model
[Int Lin Qua];               ! Fixed intercept, linear quad time slopes
Int Lin@0 Qua@0;            ! L2 G: Random intercept variance (Lin=0 & Quad=0)

Number of Free Parameters          5
Loglikelihood
  H0 Value                        -1422.043
MODEL RESULTS

                Estimate      S.E.  Est./S.E.  Two-Tailed
                P-Value

Means
  INT           9.660      0.275    35.168     0.000
  LIN            0.261      0.118     2.205     0.027
  QUA           -0.047      0.016    -2.983     0.003

Intercepts
  RECALL0        0.000      0.000    999.000   999.000
  RECALL2        0.000      0.000    999.000   999.000
  RECALL4        0.000      0.000    999.000   999.000
  RECALL6        0.000      0.000    999.000   999.000
  RECALL8        0.000      0.000    999.000   999.000

Variances
  INT           10.687      1.324     8.071     0.000
  LIN            0.000      0.000    999.000   999.000
  QUA            0.000      0.000    999.000   999.000

Residual Variances
  RECALL0        4.879      0.377    12.929     0.000
  RECALL2        4.879      0.377    12.929     0.000
  RECALL4        4.879      0.377    12.929     0.000
  RECALL6        4.879      0.377    12.929     0.000
  RECALL8        4.879      0.377    12.929     0.000

```

Model 2b. Syntax and Partial Output for Fixed Quadratic Time, Random Intercept for Recall, adding **Age at Baseline** to Introduce Total Cross-Sectional Birth Cohort Effects:

$$\text{Level-1 Time: } recall_{ti} = \beta_{0i} + \beta_{1i}(age_{ti} - ageT0_i) + \beta_{2i}(age_{ti} - ageT0_i)^2 + e_{ti}$$

$$\text{Level-2: } \beta_{0i} = \gamma_{00} + \gamma_{01}(ageT0_i - 84) + \gamma_{02}(ageT0_i - 84)^2 + U_{0i}$$

$$\beta_{1i} = \gamma_{10} + \gamma_{11}(ageT0_i - 84), \beta_{2i} = \gamma_{20}$$

Model 2b STATA Univariate MLM:

```
display "Model 2b Time: Fixed Quadratic, Random Intercept Model"
display "Controlling for Birth Cohort as Total Effects"
mixed recall time c.timesq c.aget084 c.aget084sq c.time#c.aget084, ///
|| personid: , mle nolog
```

Model 2b R Univariate MLM:

```
print("Model 2b Time: Fixed Quadratic, Random Intercept Model")
print("Controlling for Birth Cohort as Total Effects")
RICohTim = lmer(data=Example2, REMI=FALSE, formula=recall~1+time+timesq
+ageT084+ageT084sq+time:ageT084+(1|PersonID))
l1kAIC(RICohTim); summary(RICohTim);
```

```
$AICtab
      AIC      BIC    logLik  deviance  df.resid
2851.3846 2885.9651 -1417.6923  2835.3846   549.0000
```

Random effects:

Groups	Name	Variance	Std.Dev.	
PersonID	(Intercept)	10.2325	3.1988	→ reduced by another 3.66%
Residual		4.9275	2.2198	→ reduced by another 1.12%

Fixed effects:

	Estimate	Std. Error	df	t value	Pr(> t)	
(Intercept)	9.3880788	0.3414122	263.1359999	27.4978	< 2.2e-16	gamma00
time	0.2876175	0.1192557	380.6947769	2.4118	0.016348	gamma10
timesq	-0.0435466	0.0158152	366.6031559	-2.7535	0.006190	gamma20
ageT084	-0.2879274	0.1000388	250.7639296	-2.8782	0.004345	gamma01
ageT084sq	0.0068459	0.0184967	235.7300477	0.3701	0.711631	gamma02
time:ageT084	0.0389327	0.0177798	400.3882883	2.1897	0.029122	gamma11

Model 2b Mplus M-SEM:

```
TITLE: M-SEM Model 2b: Add Birth Cohort to Fixed Quad Time, Random Intercept
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults
```

VARIABLE:

```
! List of variables in long data file IN ORDER
NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall time timesq at084 at084sq;
! Missing data identifier
MISSING ARE ALL (-999);
! MSEM options
CLUSTER = PersonID; ! Level-2 ID
BETWEEN = at084 at084sq; ! Observed ONLY level-2 predictors
WITHIN = time timesq; ! Observed ONLY level-1 predictors
```

DEFINE: ! Center predictors and make squared versions

```
timesq = time*time;
at084 = ageT0-84;
at084sq = at084*at084;
```

ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR = ML;

MODEL: ! LEVEL-1 = WITHIN, LEVEL-2 = BETWEEN

```
%WITHIN%
recall;           ! L1 R: Residual variance
lin | recall ON time; ! B1i placeholder for linear time slope
recall ON timesq; ! No B2i placeholder because fixed quad only

%BETWEEN%
[recall lin];     ! Fixed intercept and linear time slope
recall lin@0;     ! L2 G: Random intercept variance only
recall lin ON aT084; ! Linear Cohort -> recall int and linear time slope
recall ON aT084sq; ! Quadratic Cohort -> recall int
```

```
Number of Free Parameters      8
Loglikelihood
  H0 Value                      -1417.679
```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
RECALL ON				
TIMESQ	-0.044	0.016	-2.753	0.006
Residual Variances				
RECALL	4.926	0.375	13.126	0.000
Between Level				
LIN ON				
AT084	0.039	0.018	2.189	0.029
RECALL ON				
AT084	-0.288	0.100	-2.878	0.004
AT084SQ	0.007	0.018	0.370	0.711
Intercepts				
RECALL	9.388	0.341	27.497	0.000
LIN	0.288	0.119	2.408	0.016
Residual Variances				
RECALL	10.233	1.277	8.012	0.000
LIN	0.000	0.000	999.000	999.000

Model 2b Mplus SEM:

TITLE: SEM Model 2b: Add Birth Cohort to Fixed Quad Time, Random Intercept
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
 TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

```
! Unstacking to multivariate format
DATA LONGTOWIDE:
! Names of old stacked former variables (without numbers)
LONG = recall|time;
! Names of new multivariate variables (that use numbers)
WIDE = recall0 recall2 recall4 recall6 recall8 |
      time0 time2 time4 time6 time8;
! Variable with level-2 ID info
IDVARIABLE = PersonID;
! Old level-1 identifier
REPETITION = occasion (0 2 4 6 8);
```

```
VARIABLE:
! List of variables in long data file IN ORDER
NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall0 recall2 recall4 recall6 recall8
            time0 time2 time4 time6 time8 aT084 aT084sq;
! Missing data identifier
MISSING ARE ALL (-999);
! Exact time definition variables
TSCORES = time0 time2 time4 time6 time8;
```

```

DEFINE:  ! Center predictors and make squared versions
         aT084 = ageT0-84;
         aT084sq = aT084*aT084;

ANALYSIS:  TYPE = RANDOM; ESTIMATOR = ML;

MODEL:
[recall0-recall8@0];      ! All variable intercepts fixed to 0
recall0-recall8 (Resvar); ! L1 R residual variances held equal

! Recall quadratic growth model using exact time as loadings
Int Lin Qua | recall0-recall8 AT time0-time8;

! Level-2 model
[Int Lin Qua];          ! Fixed intercept, linear quad time slopes
Int Lin@0 Qua@0;       ! L2 G: Random intercept variance (Lin=0 & Quad=0)
Int Lin ON aT084;      ! Linear Cohort -> Int and linear time slope
Int ON aT084sq;        ! Quadratic Cohort -> Int

```

Number of Free Parameters 8

Loglikelihood
H0 Value -1416.747

Information Criteria
Akaike (AIC) 2849.494
Bayesian (BIC) 2876.155
Sample-Size Adjusted BIC 2850.808
(n* = (n + 2) / 24)

MODEL RESULTS

		Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT	ON				
	AT084	-0.288	0.100	-2.875	0.004
	AT084SQ	0.007	0.018	0.364	0.716
LIN	ON				
	AT084	0.039	0.018	2.170	0.030
Means					
	QUA	-0.044	0.016	-2.775	0.006
Intercepts					
	RECALL0	0.000	0.000	999.000	999.000
	RECALL2	0.000	0.000	999.000	999.000
	RECALL4	0.000	0.000	999.000	999.000
	RECALL6	0.000	0.000	999.000	999.000
	RECALL8	0.000	0.000	999.000	999.000
	INT	9.388	0.341	27.520	0.000
	LIN	0.288	0.119	2.428	0.015
Variances					
	QUA	0.000	0.000	999.000	999.000
Residual Variances					
	RECALL0	4.825	0.373	12.935	0.000
	RECALL2	4.825	0.373	12.935	0.000
	RECALL4	4.825	0.373	12.935	0.000
	RECALL6	4.825	0.373	12.935	0.000
	RECALL8	4.825	0.373	12.935	0.000
	INT	10.295	1.281	8.037	0.000
	LIN	0.000	0.000	999.000	999.000

The Mplus M-SEM and SEM results are very close, but they are not quite the same as the univariate MLM results. This is (at least partly) because they are using different definitions of the fixed linear time slope (which is the conditional mean of the person mean slopes in the M-SEM framework because it is predicted by cohort, but not in the univariate MLM version in which level-2 units are not distinguished without an accompanying random slope). The Mplus SEM results differ more, though (perhaps because of optimizing on wide data?).

Model 3b. Syntax and Partial Output to add Random Linear Time to Model 2b:

Level-1 Time: $recall_{ti} = \beta_{0i} + \beta_{1i}(age_{ti} - ageT0_i) + \beta_{2i}(age_{ti} - ageT0_i)^2 + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(ageT0_i - 84) + \gamma_{01}(ageT0_i - 84)^2 + U_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(ageT0_i - 84) + U_{1i}$, $\beta_{2i} = \gamma_{20}$

Model 3b STATA Univariate MLM:

```
display "Model 3b Time: Add Random Linear Time to Model 2b"
mixed recall c.time c.timesq c.aget084 c.aget084sq c.tvage84#c.aget084, ///
|| personid: time, mle nolog covariance(unstructured)
```

Model 3b R Univariate MLM:

```
print("Model 3b Time: Add Random Linear Time to Model 2b")
RLCohTim = lmer(data=Example2, REML=FALSE, formula=recall~1+time+timesq
+ageT084+ageT084sq+time:ageT084+(1+time|PersonID))
llikaIC(RLCohTim); summary(RLCohTim)

$AICtab
      AIC      BIC    logLik  deviance  df.resid
2838.5453 2881.7709 -1409.2726 2818.5453  547.0000
Random effects:
Groups   Name      Variance Std.Dev.  Corr
PersonID (Intercept) 12.4835  3.53320
        time         0.1272  0.35665  -0.473 → new random time slope variance
Residual              3.9405  1.98508

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept)  9.3402104  0.3515637 230.2831758 26.5676 < 2.2e-16
time         0.3132277  0.1123665 366.1427566  2.7876  0.005588
timesq      -0.0455538  0.0149692 340.1904601 -3.0432  0.002523
ageT084     -0.2972341  0.1050941 205.1743603 -2.8283  0.005144
ageT084sq    0.0091296  0.0183207 229.8606122  0.4983  0.618735
time:ageT084 0.0442743  0.0207981 126.2247447  2.1288  0.035217

print("LRT for random linear TVage slope"); anova(RLCohTim, RICohTim)

      npar      AIC      BIC    logLik deviance  Chisq Df Pr(>Chisq)
RICohTim  8 2851.39 2885.97 -1417.69 2835.39
RLCohTim 10 2838.55 2881.77 -1409.27 2818.55 16.8393  2 0.00022049
```

Model 3b Mplus M-SEM:

```
TITLE: M-SEM Model 3b: Add Random Linear Time to Model 2b (Fixed Time + Cohort)
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
        TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

VARIABLE:
! List of variables in long data file IN ORDER
NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall time timesq aT084 aT084sq;
! Missing data identifier
MISSING ARE ALL (-999);
! MSEM options
CLUSTER = PersonID; ! Level-2 ID
BETWEEN = aT084 aT084sq; ! Observed ONLY level-2 predictors
WITHIN = time timesq; ! Observed ONLY level-1 predictors

DEFINE: ! Center predictors and make squared versions
timesq = time*time;
aT084 = ageT0-84;
aT084sq = aT084*aT084;
```

```

ANALYSIS:   TYPE = TWOLEVEL RANDOM; ESTIMATOR = ML;

MODEL:   ! LEVEL-1 = WITHIN, LEVEL-2 = BETWEEN
%WITHIN%
recall;           ! L1 R: Residual variance
lin | recall ON time; ! B1i placeholder for linear time slope
recall ON timesq; ! No B2i placeholder because fixed quad only

%BETWEEN%
[recall lin];           ! Fixed intercept and linear time slope
recall lin;           ! L2 G: Random effects variances
recall WITH lin;       ! L2 G: Random effects covariance
recall lin ON aT084;   ! Linear Cohort -> recall int and linear time slope
recall ON aT084sq;     ! Quadratic Cohort -> recall int

Number of Free Parameters           10
Loglikelihood
  H0 Value                           -1409.272

```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
RECALL ON				
TIMESQ	-0.046	0.015	-3.026	0.002
Residual Variances				
RECALL	3.941	0.360	10.957	0.000
Between Level				
LIN ON				
aT084	0.044	0.021	2.122	0.034
RECALL ON				
aT084	-0.297	0.105	-2.826	0.005
aT084sq	0.009	0.018	0.496	0.620
RECALL WITH				
LIN	-0.596	0.226	-2.644	0.008
Intercepts				
RECALL	9.340	0.352	26.501	0.000
LIN	0.313	0.113	2.778	0.005
Residual Variances				
RECALL	12.483	1.596	7.822	0.000
LIN	0.127	0.043	2.944	0.003

Model 3b Mplus SEM:

```

TITLE: SEM Model 3b: Add Random Linear Time to Model 2b (Fixed Time + Cohort)
DATA: FILE = Chapter10.csv;           ! Syntax in same folder as data
        TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults

! Unstacking to multivariate format
DATA LONGTOWIDE:
! Names of old stacked former variables (without numbers)
LONG = recall|time;
! Names of new multivariate variables (that use numbers)
WIDE = recall0 recall2 recall4 recall6 recall8 |
      time0 time2 time4 time6 time8;
! Variable with level-2 ID info
IDVARIABLE = PersonID;
! Old level-1 identifier
REPETITION = occasion (0 2 4 6 8);

VARIABLE:
! List of variables in long data file IN ORDER
NAMES = PersonID occasion ageT0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall0 recall2 recall4 recall6 recall8
            time0 time2 time4 time6 time8 aT084 aT084sq;

```

```

! Missing data identifier
MISSING ARE ALL (-999);
! Exact time definition variables
TSCORES = time0 time2 time4 time6 time8;

DEFINE: ! Center predictors and make squared versions
aT084 = ageT0-84;
aT084sq = aT084*aT084;

ANALYSIS: TYPE = RANDOM; ESTIMATOR = ML;

MODEL:
[recall0-recall8@0]; ! All variable intercepts fixed to 0
recall0-recall8 (Resvar); ! L1 R residual variances held equal

! Recall quadratic growth model using exact time as loadings
Int Lin Qua | recall0-recall8 AT time0-time8;

! Level-2 model
[Int Lin Qua]; ! Fixed intercept, linear quad time slopes
Int Lin Qua@0; ! L2 G: Random effects variances (Quad=0)
Int WITH Lin; ! L2 G: Random effects covariance
Int Lin ON aT084; ! Linear Cohort -> Int and linear time slope
Int ON aT084sq; ! Quadratic Cohort -> Int

```

```

Number of Free Parameters          10
Loglikelihood
  H0 Value                        -1409.189

```

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
INT ON				
AT084	-0.297	0.105	-2.825	0.005
AT084SQ	0.009	0.018	0.495	0.620
LIN ON				
AT084	0.044	0.021	2.122	0.034
INT WITH				
LIN	-0.589	0.225	-2.613	0.009
Means				
QUA	-0.046	0.015	-3.011	0.003
Intercepts				
RECALL0	0.000	0.000	999.000	999.000
RECALL2	0.000	0.000	999.000	999.000
RECALL4	0.000	0.000	999.000	999.000
RECALL6	0.000	0.000	999.000	999.000
RECALL8	0.000	0.000	999.000	999.000
INT	9.340	0.352	26.504	0.000
LIN	0.313	0.113	2.779	0.005
Variances				
QUA	0.000	0.000	999.000	999.000
Residual Variances				
RECALL0	3.938	0.360	10.942	0.000
RECALL2	3.938	0.360	10.942	0.000
RECALL4	3.938	0.360	10.942	0.000
RECALL6	3.938	0.360	10.942	0.000
RECALL8	3.938	0.360	10.942	0.000
INT	12.480	1.596	7.821	0.000
LIN	0.121	0.043	2.794	0.005

Model 4b. Syntax and Partial Output to add **Random Linear Baseline Age to Model 3b:**

Level-1 Time: $recall_{ti} = \beta_{0i} + \beta_{1i}(age_{ti} - ageT0_i) + \beta_{2i}(age_{ti} - ageT0_i)^2 + e_{ti}$

Level-2: $\beta_{0i} = \gamma_{00} + \gamma_{01}(ageT0_i - 84) + \gamma_{02}(ageT0_i - 84)^2 + U_{0i} + \text{?(ageT0}_i - 84)$

$\beta_{1i} = \gamma_{10} + \gamma_{11}(ageT0_i - 84) + U_{1i}, \beta_{2i} = \gamma_{20}$

Btw, this punctuation mark **?** is called an interrobang

Model 4b STATA Univariate MLM (did converge):

```
display "Model 4b Time: Add Random Linear AgeCoh to Model 3b -- extra iterations"
mixed recall c.time c.timesq c.aget084 c.aget084sq c.tvage84#c.aget084, ///
|| personid: time aget084, mle nolog emiterate(100) covariance(unstructured)
```

recall	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
time	.3135834	.1123693	2.79	0.005	.0933437	.5338231
timesq	-.0455585	.0149707	-3.04	0.002	-.0749006	-.0162165
aget084	-.2974395	.1053803	-2.82	0.005	-.503981	-.090898
aget084sq	-.0351879	.0268437	-1.31	0.190	-.0878005	.0174248
c.tvage84#c.aget084	.0447104	.020784	2.15	0.031	.0039745	.0854463
_cons	9.33669	.3540657	26.37	0.000	8.642734	10.03065

Random-effects Parameters	Estimate	Std. Err.	[95% Conf. Interval]	
personid: Unstructured				
var(time)	.1269922	.043199	.0651964	.2473607
var(aget084)	.0005876	.0075139	7.65e-15	4.51e+07 → new term!
var(_cons)	12.53216	1.637893	9.700139	16.19101
cov(time, aget084)	-.0085308	.0550702	-.1164663	.0994048
cov(time, _cons)	-.6000931	.2283879	-1.047725	-.1524611
cov(aget084, _cons)	.0520171	.290053	-.5164764	.6205106

```
var(Residual) | 3.941339 .3597755 3.295671 4.713501
LR test vs. linear model: chi2(6) = 238.31 Prob > chi2 = 0.0000
```

Model 4b R Univariate MLM (won't converge):

```
print("Model 4b Time: Add Random Linear Time to Model 3b -- won't run")
RL2CohTim = lmer(data=Example2, REML=FALSE, formula=recall~1+time+timesq
+ageT084+ageT084sq+time:ageT084+(1+time+ageT084|PersonID))
l1kAIC(RL2CohTim); summary(RL2CohTim)
```

Model 4b Mplus M-SEM (sort of converged; SEM version would not run at all):

```
TITLE: M-SEM Model 3b: Add Random Linear Cohort to Model 3b (Random Time + Cohort)
DATA: FILE = Chapter10.csv; ! Syntax in same folder as data
TYPE = INDIVIDUAL; FORMAT = FREE; ! Defaults
```

```
VARIABLE:
! List of variables in long data file IN ORDER
NAMES = PersonID occasion aget0 tvage recall time;
! Variables to be analyzed in this model (new defined at end)
USEVARIABLE = recall time timesq at084 at084sq;
! Missing data identifier
MISSING ARE ALL (-999);
! MSEM options
CLUSTER = PersonID; ! Level-2 ID
BETWEEN = at084 at084sq; ! Observed ONLY level-2 predictors
WITHIN = time timesq; ! Observed ONLY level-1 predictors

DEFINE: ! Center predictors and make squared versions
timesq = time*time;
at084 = ageT0-84; at084sq = at084*at084;
```

ANALYSIS: TYPE = TWOLEVEL RANDOM; ESTIMATOR = ML;
 ALGORITHM = INTEGRATION; ! Required for interobang

MODEL: ! LEVEL-1 = WITHIN, LEVEL-2 = BETWEEN

```
%WITHIN%
recall;                ! L1 R: Residual variance
lin | recall ON time; ! B1i placeholder for linear time slope
recall ON timesq;      ! No B2i placeholder because fixed quad only

%BETWEEN%
[recall lin];          ! Fixed intercept and linear time slope
recall lin;            ! L2 G: Random effects variances
recall WITH lin;       ! L2 G: Random effects covariance
recall lin ON aT084;   ! Linear Cohort -> recall int and linear time slope
recall ON aT084sq;     ! Quadratic Cohort -> recall int

Intero | recall ON aT084; ! Define random slope for level-2 predictor
recall ON Intero;      ! Add L2 predictor random slope to predict recall
```

WARNING: THE MODEL ESTIMATION HAS REACHED A SADDLE POINT OR A POINT WHERE THE OBSERVED AND THE EXPECTED INFORMATION MATRICES DO NOT MATCH. AN ADJUSTMENT TO THE ESTIMATION OF THE INFORMATION MATRIX HAS BEEN MADE. THE CONDITION NUMBER IS -0.138D-05. THE PROBLEM MAY ALSO BE RESOLVED BY DECREASING THE VALUE OF THE MCONVERGENCE OR LOGCRITERION OPTIONS OR BY CHANGING THE STARTING VALUES OR BY INCREASING THE NUMBER OF INTEGRATION POINTS OR BY USING THE MLF ESTIMATOR.

THE STANDARD ERRORS OF THE MODEL PARAMETER ESTIMATES MAY NOT BE TRUSTWORTHY FOR SOME PARAMETERS DUE TO A NON-POSITIVE DEFINITE FIRST-ORDER DERIVATIVE PRODUCT MATRIX. THIS MAY BE DUE TO THE STARTING VALUES BUT MAY ALSO BE AN INDICATION OF MODEL NONIDENTIFICATION. THE CONDITION NUMBER IS 0.590D-18. PROBLEM INVOLVING THE FOLLOWING PARAMETER: Parameter 8, %BETWEEN%: RECALL ON AT084

THE MODEL ESTIMATION TERMINATED NORMALLY
 Number of Free Parameters 13
 Loglikelihood
 H0 Value -1409.293

MODEL RESULTS

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
Within Level				
RECALL ON TIMESQ	-0.043	0.015	-2.842	0.004
Residual Variances				
RECALL	3.945	0.360	10.952	0.000
Between Level				
LIN ON AT084	0.044	0.022	2.022	0.043
RECALL ON INTERO	1.052	4.148	0.254	0.800
RECALL ON AT084	-0.199	3.989	-0.050	0.960
AT084SQ	0.009	0.019	0.490	0.624
RECALL WITH LIN	-0.592	0.227	-2.604	0.009
Means				
INTERO	-0.099	3.963	-0.025	0.980
Intercepts				
RECALL	9.458	4.625	2.045	0.041
LIN	0.293	0.114	2.580	0.010
Variances				
INTERO	0.000	0.364	0.001	0.999
Residual Variances				
RECALL	12.478	2.935	4.252	0.000
LIN	0.126	0.043	2.929	0.003