Example 2: Unconditional Polynomial Models for Change in Number Match 3 Response Time
(complete data, syntax, and output available for SAS, SPSS, and STATA electronically)

These data (in “Example23” data files) come from a short-term longitudinal study of 6 observations over 2 weeks for 101 adults age 65–80. The goal is to see how performance on this processing speed task (“number match 3”), as measured by response time in milliseconds, declines over the 6 practice sessions.

SAS Code for Data Manipulation:

```
* SAS code to import data, center time for polynomial models;
DATA work.example23; SET filepath.example23;
      c1sess = session - 1; LABEL c1sess = "c1sess: Session Centered at 1";
RUN;
```

SPSS Code for Data Manipulation:

```
* SPSS code to import data, center time for polynomial models.
GET FILE = "example/Example23.sav".
DATASET NAME example23 WINDOW=FRONT.
COMPUTE c1sess = session - 1.
VARIABLE LABELS c1sess "c1sess: Session Centered at 1".
```

STATA Code for Data Manipulation:

```
 * STATA code to center time for polynomial models (and make quadratic version)
gen c1sess = session - 1
gen c1sess2 = c1sess * c1sess
label variable c1sess  "c1sess: Session Centered at 1"
label variable c1sess2 "c1sess2: Quadratic Session Centered at 1"
```

Model 1a. Most Conservative Baseline:
Empty Means, Random Intercept

Level 1: $y_{ti} = \beta_{0i} + e_{ti}$
Level 2: Intercept: $\beta_{0i} = \gamma_{00} + U_{0i}$

```
TITLE1 "SAS Model 1a: Empty Means, Random Intercept Only";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT
  COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / G V VCORR TYPE=UN SUBJECT=ID;
      REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
```

METHOD = ML or REML (default)
CLASS = categorical predictors, nesting
MODEL dv = fixed effects / print solution
RANDOM = person variances in **G**
REPEATED = residuals in **R** matrix

```
TITLE "SPSS Model 1a: Empty Means, Random Intercept".
MIXED nm3rt BY ID session
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV G R
      /FIXED  =
      /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
      /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

MIXED dv BY categorical predictors
WITH continuous predictors
/METHOD = REML or ML
/PRINT = regression solution
/FIXED = predictors for means model
/RANDOM = person variances in **G**

```
* STATA Model 1a: Empty Means, Random Intercept
xtmixed nm3rt ,  || id:  ,  ///
      variance reml covariance(unstructured) residuals(independent,t(session)),
      estat ic, n(101)
      estat recovariance, level(id)
```

DV = nm3rt, random part after ||
Level 2 ID is PersonID, random intercept by default
Print variances instead of SD, use reml
covariance(unstructured) refers to G matrix
residuals(independent) → refers to R matrix by session
estat ic → Print IC given N = 101 persons

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101

                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6
                                                Wald chi2(0)       =         .
Log restricted-likelihood = -4268.4304          Prob > chi2        =         .
------------------------------------------------------------------------------
       nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |   1770.701   45.42063    38.98   0.000     1681.679    1859.724
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |     200883   29471.23      150683.2    267806.8
-----------------------------+------------------------------------------------
               var(Residual) |   44899.96    2825.63      39689.76    50794.13
------------------------------------------------------------------------------
LR test vs. linear regression: chibar2(01) =   691.74 Prob >= chibar2 = 0.0000

.         estat ic, n(101)
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .    -4268.43      3     8542.861    8550.706
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC

.         estat recovariance, level(id)
Random-effects covariance matrix for level id
             |     _cons
-------------+-----------
       _cons |    200883
```

NOTE: LL is given rather than −2LL

This is the fixed intercept (just grand mean so far).

Calculate the ICC for the Number Match 3 outcome:

$$\text{ICC} = \frac{200883}{200883 + 44900} = .82$$

This LR test tells us that the random intercept variance is significantly greater than 0, and thus so is the ICC.

REML-based AIC and BIC are calculated differently in STATA (they count fixed effects), so they won't match the values in other programs.

This is the level-2 **G** matrix, just a random intercept variance so far.

Extra SAS output not provided by STATA:

Estimated R Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	44900					
2		44900				
3			44900			
4				44900		
5					44900	
6						44900

This level-1 **R matrix** (with equal variance over time, no covariance of any kind, known as VC or independence) will be used repeatedly as we add fixed and random effects.

Estimated G Matrix

Row	Effect	Participant ID	Col1
1	Intercept	101	200883

This is the level-2 **G** matrix, just a random intercept variance so far.

Estimated V Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	**245783**	200883	200883	200883	200883	200883
2	200883	**245783**	200883	200883	200883	200883
3	200883	200883	**245783**	200883	200883	200883
4	200883	200883	200883	**245783**	200883	200883
5	200883	200883	200883	200883	**245783**	200883
6	200883	200883	200883	200883	200883	**245783**

The **V** matrix is the total variance-covariance matrix after combining the level-2 **G** and level-1 **R** matrices.

Model 1b. Most Liberal Baseline – Saturated Means, Unstructured Variances (Model Answer Key)

```
TITLE1 "SAS Model 1b: Saturated Means, Unstructured Variances";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = session / SOLUTION DDFM=Satterthwaite;
      REPEATED session / R RCORR TYPE=UN SUBJECT=ID;
      LSMEANS session /; RUN;

TITLE "SPSS Model 1b: Saturated Means, Unstructured Variances".
MIXED nm3rt BY ID session
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV R
      /FIXED  = session
      /REPEATED = session | SUBJECT(ID) COVTYPE(UN)
      /EMMEANS  = TABLES(session).

 * STATA Model 1b: Saturated Means, Unstructured Variances
xtmixed nm3rt ib(last).session, || id: , noconstant ///
      variance reml residuals(unstructured, t(session)),
      estat ic, n(101),
      contrast session,       // omnibus test of mean differences
      margins i.session,      // observed means per session
      marginsplot name(observed_means, replace) // plot observed means
```

Placing *session* on the CLASS/BY statements and in the FIXED/MODEL statements treats it as a categorical predictor. So this is an ANOVA means model. No RANDOM statements mean no random effects.

i. indicates categorical predictor of *session* (ref=last to match others) noconstant = no random intercept (just **R** matrix)

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101
                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6

                                                Wald chi2(5)       =     83.60
Log restricted-likelihood = -4114.8942          Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     session |
          1  |   289.7574   32.69997     8.86   0.000     225.6666    353.8481
          2  |   143.0364   26.20308     5.46   0.000     91.67927    194.3935
          3  |   77.89864    22.8842     3.40   0.001     33.04642    122.7509
          4  |   45.66045   20.78533     2.20   0.028     4.921952    86.39894
          5  |   35.03972   18.11681     1.93   0.053     -.468579    70.54802
             |
       _cons |   1672.136   44.13439    37.89   0.000     1585.634    1758.638
------------------------------------------------------------------------------

------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id:                  (empty) |
-----------------------------+------------------------------------------------
Residual: Unstructured       |
                     var(e1) |   301983.1   42696.65      228893.4    398411.6
                     var(e2) |   259148.8    36635.7      196433.4    341887.4
                     var(e3) |   233366.9   32990.48      176891.6    307872.9
                     var(e4) |   217542.8   30753.82      164896.4    286997.6
                     var(e5) |   212096.8   29984.63      160767.3    279814.7
                     var(e6) |   196732.3   27812.21      149121.6    259543.9
                  cov(e1,e2) |   235657.1   36563.79      163993.4    307320.8
```

This is the multivariate Wald test for all the fixed effects simultaneously (5 mean differences from the fixed intercept here).

Mean diffs relative to session 6

These are the total variances at each occasion…

```
                 cov(e1,e3) |   217992.5    34336.3      150694.6    285290.4
                 cov(e1,e4) |   202605.4   32657.69      138597.5    266613.2
                 cov(e1,e5) |   192152.4   31762.13      129899.7      254405
                 cov(e1,e6) |   195358.7   31224.07      134160.6    256556.7
                 cov(e2,e3) |   230215.6   33672.52      164218.7    296212.5
                 cov(e2,e4) |   213230.6   31899.27      150709.2      275752
                 cov(e2,e5) |     202091   30938.65      141452.3    262729.6
                 cov(e2,e6) |   193267.2   29707.66      135041.2    251493.1
                 cov(e3,e4) |     205208   30462.92      145501.8    264914.2
                 cov(e3,e5) |   196917.7   29697.89      138710.9    255124.5
                 cov(e3,e6) |   188603.5   28532.39        132681    244525.9
                 cov(e4,e5) |   193674.7   28910.48      137011.2    250338.2
                 cov(e4,e6) |     185320   27762.64      130906.2    239733.8
                 cov(e5,e6) |   187839.5   27739.82      133470.5    242208.6
------------------------------------------------------------------------------
LR test vs. linear regression:      chi2(20) =   925.64   Prob > chi2 = 0.0000

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.
```

And these are the total covariances across occasions…

This is the LRT of whether the unstructured R model fits better than the e-only R model…

```
.         estat ic, n(101),
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .   -4114.894     27     8283.788    8354.397
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC

.         contrast session,   // omnibus test of mean differences

Contrasts of marginal linear predictions

Margins      : asbalanced

------------------------------------------------
             |         df        chi2     P>chi2
-------------+----------------------------------
nm3rt        |
     session |          5       83.60     0.0000
------------------------------------------------
```

This is the omnibus test of mean differences across 6 sessions.

```
.         margins i.session,  // observed means per session

Adjusted predictions                            Number of obs   =        606

Expression   : Linear prediction, fixed portion, predict()
THESE ARE THE SATURATED MEANS THE FIXED EFFECTS WILL BE TRYING TO REPRODUCE.
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     session |
          1  |   1961.893   54.68027    35.88   0.000     1854.722    2069.065
          2  |   1815.172   50.65402    35.83   0.000     1715.892    1914.452
          3  |   1750.035   48.06832    36.41   0.000     1655.822    1844.247
          4  |   1717.796   46.41001    37.01   0.000     1626.835    1808.758
          5  |   1707.176   45.82541    37.25   0.000      1617.36    1796.992
          6  |   1672.136   44.13439    37.89   0.000     1585.634    1758.638
------------------------------------------------------------------------------
```

Extra SAS output not provided by STATA:

Estimated R Matrix for ID 101

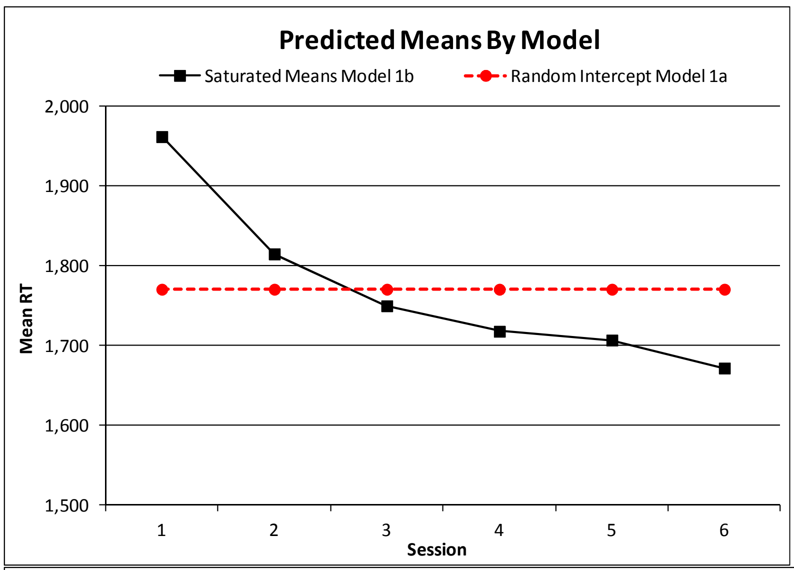
Row	Col1	Col2	Col3	Col4	Col5	Col6
1	**301985**	235659	217994	202607	192154	195360
2	235659	**259150**	230217	213232	202092	193268
3	217994	230217	**233368**	205209	196919	188604
4	202607	213232	205209	**217544**	193676	185321
5	192154	202092	196919	193676	**212098**	187840
6	195360	193268	188604	185321	187840	**196733**

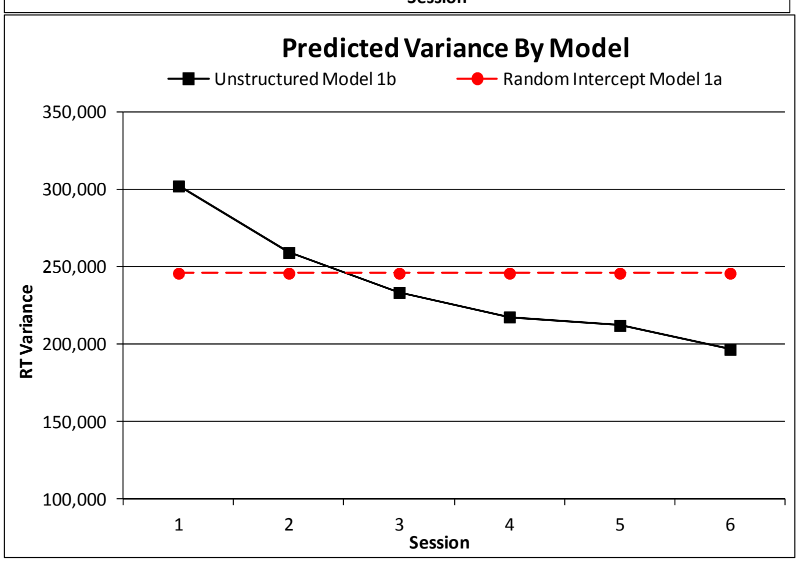
This Unstructured **R matrix** estimates all variances and covariances separately. THIS IS THE DATA we are trying to duplicate with our model for the variances.

Estimated R Correlation Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.8424	0.8212	0.7905	0.7593	0.8015
2	0.8424	1.0000	0.9361	0.8981	0.8620	0.8559
3	0.8212	0.9361	1.0000	0.9108	0.8851	0.8802
4	0.7905	0.8981	0.9108	1.0000	0.9016	0.8958
5	0.7593	0.8620	0.8851	0.9016	1.0000	0.9196
6	0.8015	0.8559	0.8802	0.8958	0.9196	1.0000

So here is what are we trying to model—means and variances, where model 1b is the data:





Model 2a. Fixed Linear Time, Random Intercept

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}\left(\text{Session}_{ti} - 1\right) + e_{ti}$$
$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$
$$\text{Linear Session: } \beta_{1i} = \gamma_{10}$$

```
TITLE1 "SAS Model 2a: Fixed Linear Time, Random Intercept";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = c1sess / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT / G V VCORR TYPE=UN SUBJECT=ID;
      REPEATED session / R TYPE=VC SUBJECT=ID; RUN;

TITLE "SPSS Model 2a: Fixed Linear Time, Random Intercept".
MIXED nm3rt BY ID session WITH c1sess
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV G R
      /FIXED  = c1sess
      /RANDOM = INTERCEPT | SUBJECT(ID) COVTYPE(UN)
      /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

The predictor of *c1sess* will be treated as continuous given that it is not on the CLASS statement (SAS) and it is on WITH (SPSS).

```
 * STATA Model 2a: Fixed Linear Time, Random Intercept
xtmixed nm3rt c.c1sess,  || id:  ,  ///
      variance reml covariance(un) residuals(independent,t(session)),
      estat ic, n(101),
      estimates store FixLin
```

DV = nm3rt, c. means continuous fixed slope for *c1sess*
Level 2 ID is id, random intercept by default
estimates → save results as “FixLin” for next LRT

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101
                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6
                                                Wald chi2(1)       =    131.82
Log restricted-likelihood =  -4207.344          Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      c1sess |  -51.57185   4.491815   -11.48   0.000    -60.37565   -42.76806
       _cons |   1899.631    46.7882    40.60   0.000     1807.928    1991.334
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Identity                 |
                  var(_cons) |   202422.7   29469.85      152172.6    269266.3
-----------------------------+------------------------------------------------
               var(Residual) |   35661.79   2246.481      31519.73    40348.16
------------------------------------------------------------------------------
LR test vs. linear regression: chibar2(01) =   787.61 Prob >= chibar2 = 0.0000

.         estat ic, n(101),
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .   -4207.344      4     8422.688    8433.149
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC
```

The fixed linear effect of c1sess is significant according to the Wald test (*p*-value for fixed effect).

Relative to the empty means, random intercept model 1a, the fixed linear effect of session explained ~21% of the residual variance (which made the random intercept variance increase due to its residual variance correction factor).

Model 2b. Random Linear Time

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}\left(\text{Session}_{ti} - 1\right) + e_{ti}$$
$$\text{Level 2: Intercept: } \beta_{0i} = \gamma_{00} + U_{0i}$$
$$\text{Linear Session: } \beta_{1i} = \gamma_{10} + U_{1i}$$

```
TITLE1 "SAS Model 2b: Random Linear Time";
PROC MIXED DATA=work.example23 NOCLPRINT
NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = c1sess / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT c1sess / G V VCORR TYPE=UN SUBJECT=ID;
      REPEATED session / R TYPE=VC SUBJECT=ID; RUN;
```

```
TITLE "SPSS Model 2b: Random Linear Time".
MIXED nm3rt BY ID session WITH c1sess
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV G R
      /FIXED  = c1sess
      /RANDOM = INTERCEPT c1sess | SUBJECT(ID) COVTYPE(UN)
      /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

Now there are 2 random effects: intercept and linear slope, given by c1sess on the RANDOM statements.

```
 * STATA Model 2b: Random Linear Time
xtmixed nm3rt c.c1sess,  || id: c1sess,  ///
      variance reml covariance(un) residuals(independent,t(session)),
      estat ic, n(101),
      estat recovariance, level(id),
      estimates store RandLin,
      lrtest RandLin FixLin
```

DV = nm3rt, c. means continuous fixed slope for c1sess
Level 2 ID is id, random intercept and c1sess now estimates → save results as "RandLin" for LRT

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101
                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6

                                                Wald chi2(1)       =     70.17
Log restricted-likelihood = -4186.0512          Prob > chi2        =    0.0000
------------------------------------------------------------------------------
       nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      c1sess |  -51.57185   6.156722    -8.38   0.000    -63.63881    -39.5049
       _cons |   1899.631   51.4998    36.89   0.000     1798.693    2000.569
------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(c1sess) |   2233.833   552.9239      1375.178    3628.626
                  var(_cons) |     253258   37897.26      188881.9    339575.3
            cov(c1sess,_cons) |  -12700.79   3621.977     -19799.74   -5601.848
-----------------------------+------------------------------------------------
               var(Residual) |   27905.42   1963.419      24310.74    32031.62
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(3) =   830.20   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

.         estat ic, n(101),
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .   -4186.051      6     8384.102    8399.793
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC
```

```
.         estat recovariance, level(id),

Random-effects covariance matrix for level id
             |    c1sess      _cons
-------------+---------------------
      c1sess |  2233.833
       _cons | -12700.79     253258

.         estimates store RandLin,
.         lrtest RandLin FixLin

Likelihood-ratio test                                 LR chi2(2)  =     42.59
(Assumption: FixLin nested in RandLin)                Prob > chi2 =    0.0000

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.
```

Is the random linear time model (2b) better than the fixed linear time, random intercept model (2a)?

Yep, −2ΔLL= 43, which is bigger than the critical value of 5.99ish on df =~2ish

Extra SAS output not provided by STATA:

Estimated R Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	27905					
2		27905				
3			27905			
4				27905		
5					27905	
6						27905

Estimated G Matrix

Row	Effect	Participant ID	Col1	Col2
1	Intercept	101	253258	-12701
2	C1sess	101	-12701	2233.83

Estimated V Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	**281163**	240557	227856	215155	202455	189754
2	240557	**257995**	219623	209156	198689	188222
3	227856	219623	**239295**	203157	194924	186691
4	215155	209156	203157	**225063**	191158	185159
5	202455	198689	194924	191158	**215298**	183627
6	189754	188222	186691	185159	183627	**210001**

The **V** matrix is the total variance-covariance matrix after combining the level-2 **G** and level-1 **R** matrices. Now the variances and covariances are predicted to change based on time.

Estimated V Correlation Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	1.0000	0.8932	0.8784	0.8553	0.8229	0.7809
2	0.8932	1.0000	0.8839	0.8680	0.8430	0.8086
3	0.8784	0.8839	1.0000	0.8754	0.8588	0.8328
4	0.8553	0.8680	0.8754	1.0000	0.8684	0.8517
5	0.8229	0.8430	0.8588	0.8684	1.0000	0.8636
6	0.7809	0.8086	0.8328	0.8517	0.8636	1.0000

The **VCORR** matrix is the correlation version. The ICC is now predicted to change over time, too (and conditional on linear time).

How the V matrix variances and covariances get calculated in a random linear time model:

$$\mathbf{V}_{\mathrm{i}} \text{ matrix: Variance}\left[y_{\mathrm{time}}\right] = \tau^2_{U_0} + \left[(\text{Session}-1)^2\,\tau^2_{U_1}\right] + \left[2(\text{Session}-1)\,\tau_{U_{01}}\right] + \sigma^2_e$$

$$\mathbf{V}_{\mathrm{i}} \text{ matrix: Covariance}\left[y_A, y_B\right] = \tau^2_{U_0} + \left[(A+B)\,\tau_{U_{01}}\right] + \left[(AB)\,\tau^2_{U_1}\right]$$

Model 3a. Fixed Quadratic, Random Linear Time

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}\left(\text{Session}_{ti} - 1\right) + \beta_{2i}\left(\text{Session}_{ti} - 1\right)^2 + e_{ti}$$

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

$$\text{Linear Session: } \beta_{1i} = \gamma_{10} + U_{1i}$$

$$\text{Quadratic Session: } \beta_{2i} = \gamma_{20}$$

```
TITLE1 "SAS Model 3a: Fixed Quadratic, Random Linear Time";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = c1sess c1sess*c1sess / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT c1sess / G V VCORR TYPE=UN SUBJECT=ID;
      REPEATED session / R TYPE=VC SUBJECT=ID; RUN;

TITLE "SPSS Model 3a: Fixed Quadratic, Random Linear Time".
MIXED nm3rt BY ID session WITH c1sess
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV G R
      /FIXED  = c1sess c1sess*c1sess
      /RANDOM = INTERCEPT c1sess | SUBJECT(ID) COVTYPE(UN)
      /REPEATED = session | SUBJECT(ID) COVTYPE(ID).
```

Interactions can be defined on the fly in SAS and SPSS using *, or in STATA using # (but only for fixed effects in STATA).

```
 * STATA Model 3a: Fixed Quadratic, Random Linear Time
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess,  || id: c1sess,  ///
      variance reml covariance(un) residuals(independent,t(session)),
      estat ic, n(101),
      estat recovariance, level(id),
      estimates store FixQuad
```

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101
                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6
                                                Wald chi2(2)       =     97.86
Log restricted-likelihood = -4170.7386          Prob > chi2        =    0.0000
------------------------------------------------------------------------------------
            nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
           c1sess |  -120.8999   14.54147    -8.31   0.000    -149.4007   -92.39917
c.c1sess#c.c1sess |   13.86561   2.634761     5.26   0.000     8.701578    19.02965
            _cons |    1945.85    52.2433    37.25   0.000     1843.455    2048.245
------------------------------------------------------------------------------------
```

The fixed quadratic effect of c1sess is significant according to the Wald test (*p*-value for fixed effect).

```
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(c1sess) |   2332.667   551.5799      1467.501    3707.891
                  var(_cons) |     254164   37895.62      189758.3    340429.7
            cov(c1sess,_cons) |  -12947.88   3620.697     -20044.31   -5851.442
-----------------------------+------------------------------------------------
               var(Residual) |   26175.83   1844.008      22800.05    30051.42
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(3) =    851.78   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.
```

Relative to the random linear time model 2b, the fixed quadratic effect of session explained another ~6% of the residual variance (which made the random intercept variance increase due to its residual variance correction factor).

```
.        estat ic, n(101),

-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .   -4170.739      7     8355.477    8373.783
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC
```

Model 3b. Random Quadratic Time (and an example of ESTIMATE/TEST/MARGINS statements)

$$\text{Level 1: } y_{ti} = \beta_{0i} + \beta_{1i}\left(\text{Session}_{ti} - 1\right) + \beta_{2i}\left(\text{Session}_{ti} - 1\right)^2 + e_{ti}$$

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

$$\text{Linear Session: } \beta_{1i} = \gamma_{10} + U_{1i}$$

$$\text{Quadratic Session: } \beta_{2i} = \gamma_{20} + U_{2i}$$

```
TITLE1 "SAS Model 3b: Random Quadratic Time";
PROC MIXED DATA=work.example23 NOCLPRINT NOITPRINT COVTEST NAMELEN=100 METHOD=REML;
      CLASS ID session;
      MODEL nm3rt = c1sess c1sess*c1sess / SOLUTION DDFM=Satterthwaite;
      RANDOM INTERCEPT c1sess c1sess*c1sess / G V VCORR TYPE=UN SUBJECT=ID;
      REPEATED session / R TYPE=VC SUBJECT=ID;
      ESTIMATE "Intercept at Session 1"    intercept 1 c1sess 0   c1sess*c1sess 0;
      ESTIMATE "Intercept at Session 2"    intercept 1 c1sess 1   c1sess*c1sess 1;
      ESTIMATE "Intercept at Session 3"    intercept 1 c1sess 2   c1sess*c1sess 4;
      ESTIMATE "Intercept at Session 4"    intercept 1 c1sess 3   c1sess*c1sess 9;
      ESTIMATE "Intercept at Session 5"    intercept 1 c1sess 4   c1sess*c1sess 16;
      ESTIMATE "Intercept at Session 6"    intercept 1 c1sess 5   c1sess*c1sess 25;
      * Predicting linear rate of change at each session (linear changes by 2*quad);
      ESTIMATE "Linear Slope at Session 1"      c1sess 1   c1sess*c1sess 0;
      ESTIMATE "Linear Slope at Session 2"      c1sess 1   c1sess*c1sess 2;
      ESTIMATE "Linear Slope at Session 3"      c1sess 1   c1sess*c1sess 4;
      ESTIMATE "Linear Slope at Session 4"      c1sess 1   c1sess*c1sess 6;
      ESTIMATE "Linear Slope at Session 5"      c1sess 1   c1sess*c1sess 8;
      ESTIMATE "Linear Slope at Session 6"      c1sess 1   c1sess*c1sess 10;  RUN;

TITLE "SPSS Model 3b: Random Quadratic Time".
MIXED nm3rt BY ID session WITH c1sess
      /METHOD = REML
      /PRINT  = SOLUTION TESTCOV G R
      /FIXED  = c1sess c1sess*c1sess
      /RANDOM = INTERCEPT c1sess c1sess*c1sess | SUBJECT(ID) COVTYPE(UN)
      /REPEATED = session | SUBJECT(ID) COVTYPE(ID)
      /TEST = "Intercept at Session 1"     intercept 1 c1sess 0   c1sess*c1sess 0
      /TEST = "Intercept at Session 2"     intercept 1 c1sess 1   c1sess*c1sess 1
      /TEST = "Intercept at Session 3"     intercept 1 c1sess 2   c1sess*c1sess 4
      /TEST = "Intercept at Session 4"     intercept 1 c1sess 3   c1sess*c1sess 9
      /TEST = "Intercept at Session 5"     intercept 1 c1sess 4   c1sess*c1sess 16
      /TEST = "Intercept at Session 6"     intercept 1 c1sess 5   c1sess*c1sess 25
      /TEST = "Linear Slope at Session 1"      c1sess 1    c1sess*c1sess 0
      /TEST = "Linear Slope at Session 2"      c1sess 1    c1sess*c1sess 2
      /TEST = "Linear Slope at Session 3"      c1sess 1    c1sess*c1sess 4
      /TEST = "Linear Slope at Session 4"      c1sess 1    c1sess*c1sess 6
      /TEST = "Linear Slope at Session 5"      c1sess 1    c1sess*c1sess 8
      /TEST = "Linear Slope at Session 6"      c1sess 1    c1sess*c1sess 10.
```

Because <u>twice</u> the quadratic slope is how the linear slope changes per unit time, the value for *c1sess* used in estimating the linear slope per session gets multiplied by 2.

```
* STATA Model 3b: Random Quadratic Time
xtmixed nm3rt c.c1sess c.c1sess#c.c1sess,  || id: c1sess c1sess2,  ///
      variance reml covariance(un) residuals(independent,t(session)),
      estat ic, n(101),
      estat recovariance, level(id),
      estimates store RandQuad,
      lrtest RandQuad FixQuad,
      margins, at(c.c1sess=(0(1)5)) vsquish              // intercepts per session
      marginsplot, name(predicted_means, replace)        // plot intercepts
      margins, at(c.c1sess=(0(1)5)) dydx(c.c1sess) vsquish // linear slope per session
      marginsplot, name(change_in_linear_slope, replace)   // plot quadratic effect
```

The random statement will not accept interaction terms, so we are using the c1sess2 created manually before.

STATA output:

```
Mixed-effects REML regression                   Number of obs      =       606
Group variable: id                              Number of groups   =       101
                                                Obs per group: min =         6
                                                               avg =       6.0
                                                               max =         6
                                                Wald chi2(2)       =     71.74
Log restricted-likelihood = -4151.3728          Prob > chi2        =    0.0000
------------------------------------------------------------------------------------
           nm3rt |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------------
          c1sess |  -120.8999   20.04752    -6.03   0.000    -160.1923    -81.6075
c.c1sess#c.c1sess |   13.86561    3.41541     4.06   0.000     7.171534    20.55969
           _cons |    1945.85   53.84993    36.13   0.000     1840.306    2051.394
------------------------------------------------------------------------------------
------------------------------------------------------------------------------
  Random-effects Parameters  |   Estimate   Std. Err.     [95% Conf. Interval]
-----------------------------+------------------------------------------------
id: Unstructured             |
                 var(c1sess) |   25839.79   5864.685      16561.42    40316.29
                var(c1sess2) |   634.4659    172.375      372.5198    1080.605
                  var(_cons) |   276207.8   41445.59      205831.2    370647.1
         cov(c1sess,c1sess2) |  -3903.291   982.6248       -5829.2   -1977.381
           cov(c1sess,_cons) |  -35734.05   11947.96     -59151.62   -12316.48
          cov(c1sess2,_cons) |   3901.974   1950.304      79.44722      7724.5
-----------------------------+------------------------------------------------
               var(Residual) |   20298.19   1649.117      17310.19    23801.96
------------------------------------------------------------------------------
LR test vs. linear regression:       chi2(6) =    890.51   Prob > chi2 = 0.0000
Note: LR test is conservative and provided only for reference.

.         estat ic, n(101),
-----------------------------------------------------------------------------
       Model |    Obs    ll(null)   ll(model)     df          AIC         BIC
-------------+---------------------------------------------------------------
           . |    101           .   -4151.373     10     8322.746    8348.897
-----------------------------------------------------------------------------
               Note:  N=101 used in calculating BIC

.         estimates store RandQuad,

.         lrtest RandQuad FixQuad,

Likelihood-ratio test                                 LR chi2(3)  =     38.73
(Assumption: FixQuad nested in RandQuad)              Prob > chi2 =    0.0000

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.
```

Is the random quadratic model (3b) better than the fixed quadratic, random linear model (3a)?

Yep, $-2\Delta LL = 39$, which is bigger than the critical value of 7.82ish on df=~3ish

```
.         margins, at(c1sess=(0(1)5))     vsquish                            // intercepts per session
Adjusted predictions                                  Number of obs    =        606
Expression    : Linear prediction, fixed portion, predict()
1._at         : c1sess         =          0
2._at         : c1sess         =          1
3._at         : c1sess         =          2
4._at         : c1sess         =          3
5._at         : c1sess         =          4
6._at         : c1sess         =          5
------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         _at |
          1  |    1945.85   53.84993    36.13   0.000     1840.306    2051.394
          2  |   1838.815   48.48658    37.92   0.000     1743.784    1933.847
          3  |   1759.512   46.99744    37.44   0.000     1667.399    1851.626
          4  |   1707.941   45.89598    37.21   0.000     1617.986    1797.895
          5  |     1684.1   44.23964    38.07   0.000     1597.392    1770.808
          6  |   1687.991   44.20394    38.19   0.000     1601.352    1774.629
------------------------------------------------------------------------------
```

These are the quadratic-model-predicted means per session.

```
.         marginsplot, name(predicted_means, replace)                // plot intercepts
  Variables that uniquely identify margins: c1sess
.         margins, at(c1sess=(0(1)5)) dydx(c1sess) vsquish   // linear slope per session
Conditional marginal effects                          Number of obs    =        606
Expression    : Linear prediction, fixed portion, predict()
dy/dx w.r.t. : c1sess
1._at         : c1sess         =          0
2._at         : c1sess         =          1
3._at         : c1sess         =          2
4._at         : c1sess         =          3
5._at         : c1sess         =          4
6._at         : c1sess         =          5
------------------------------------------------------------------------------
             |            Delta-method
             |      dy/dx   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
c1sess   _at |
          1  |  -120.8999   20.04752    -6.03   0.000    -160.1923    -81.6075
          2  |   -93.1687   13.64968    -6.83   0.000    -119.9216    -66.4158
          3  |  -65.43747   8.002796    -8.18   0.000    -81.12266   -49.75228
          4  |  -37.70624    5.92417    -6.36   0.000     -49.3174   -26.09508
          5  |  -9.975015   9.973315    -1.00   0.317    -29.52235    9.572324
          6  |   17.75621   16.03616     1.11   0.268    -13.67408    49.18651
```

These are the instantaneous linear slopes at each session. Note how the SEs narrow towards the middle of the data.

How well do the predicted means, variances, and covariances from the random quadratic model (3b) match the original means, variances, and covariances from the saturated means model (1b)?

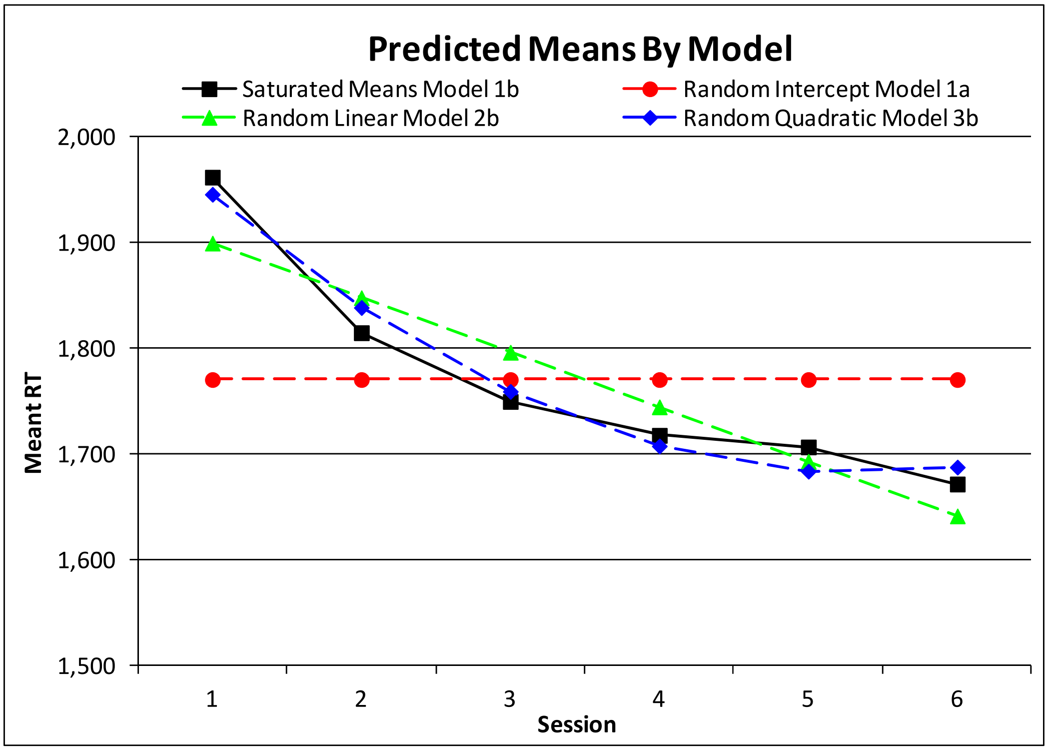
Extra SAS output not provided by STATA:

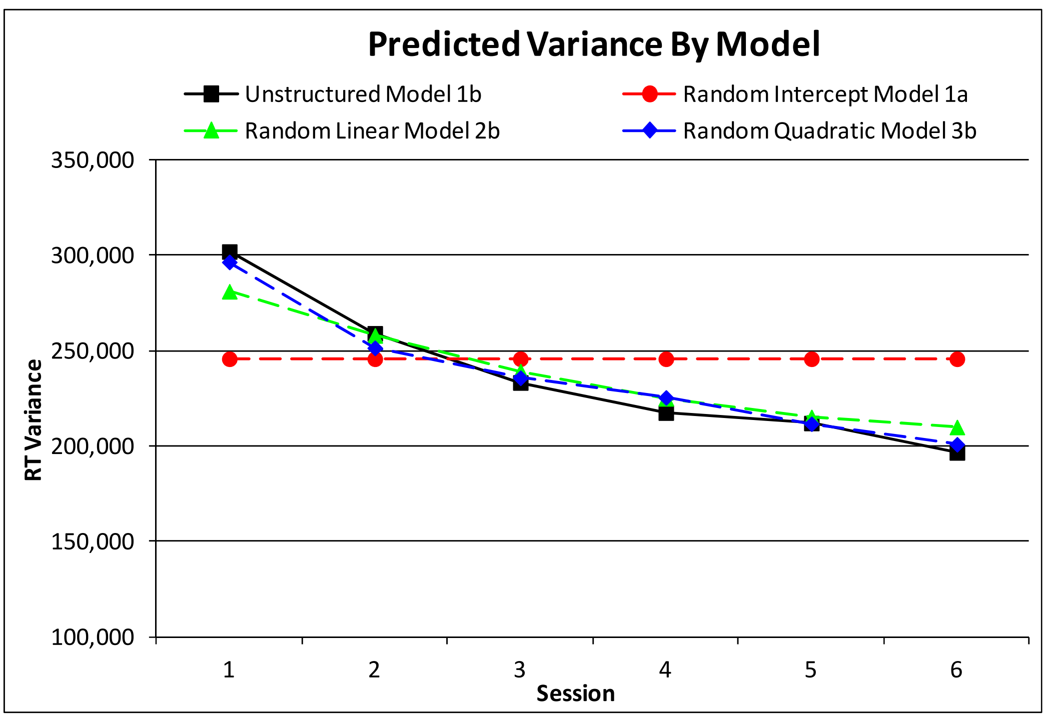
Estimated V Matrix for ID 101

Row	Col1	Col2	Col3	Col4	Col5	Col6
1	**296504**	244374	220346	204122	195702	195085
2	244374	**251508**	219312	208680	199315	191215
3	220346	219312	**235842**	209043	199808	187840
4	204122	208680	209043	**225508**	197182	184958
5	195702	199315	199808	197182	**211735**	182571
6	195085	191215	187840	184958	182571	**200977**

The **V** matrix is the total variance-covariance matrix after combining the level-2 **G** and level-1 **R** matrices. The variances and covariances are predicted to change based on time, but differently.

Figure 1





How the V matrix variances and covariances get calculated in a random quadratic time model:

Predicted Variance at Time T:

$$\text{Var}(y_T) = \sigma_e^2 + \tau_{U_0}^2 + 2*T*\tau_{U_{01}} + T^2*\tau_{U_1}^2 + 2*T^2*\tau_{U_{02}} + 2*T^3*\tau_{U_{12}} + T^4*\tau_{U_2}^2$$

Predicted Covariance between Time A and B:

$$\text{Cov}(y_A, y_B) = \tau_{U_0}^2 + (A+B)*\tau_{U_{01}} + (AB)*\tau_{U_1}^2 + (A^2+B^2)*\tau_{U_{02}} + (AB^2)+(A^2B)*\tau_{U_{12}} + (A^2B^2)*\tau_{U_2}^2$$

Simple Processing Speed – Example Unconditional Models of Change Results

Model Specification

Linear mixed models were estimated using restricted maximum likelihood (REML) in order to examine the overall pattern of and individual differences in response time over six sessions for a simple processing speed test (number match three). The significance of new fixed effects were evaluated using Wald tests, whereas the significance of new random effects was evaluated using likelihood ratio tests (i.e., −2ΔLL), with degrees of freedom equal to the number of new random effects variances and covariances. The 95% confidence interval (CI) for random variation around each fixed effect was calculated as ± 1.96 standard deviations of its accompanying random variance term.

Although the six sessions were held over a period of 6–10 days, given that experience to the test (and not *time* per se) was the most likely reason for changes in response time, session was used as the metric of time (i.e., as opposed to age or day). Session was centered at the first occasion, such that the intercept represented initial status in all models. Observed mean response times (in milliseconds) estimated from a saturated means model (i.e., multivariate analysis of variance) are shown in Figure 1. The intraclass correlation from the unconditional means model (i.e., empty model; random intercept only) was calculated as .82, indicating that over 80% of the variance in number match 3 across sessions occurred between persons in mean RT. Polynomial models were then estimated to approximate the effects of practice across the six sessions, as presented below.

Polynomial Models

Polynomial models were first specified with a random intercept only. A fixed linear effect of session was significant ($p < .001$), such that average response time declined across sessions. The addition of a random linear slope (as well as a covariance between the random intercept and random linear slope) resulted in a significant improvement to the model, −2ΔLL(2) = 43, $p < .001$. However, the magnitude of this linear decline was reduced in later sessions, as indicated by a significant fixed quadratic effect of session (i.e., a decelerating negative trend; $p < .001$). The addition of a random quadratic slope (and its two accompanying covariances with the random intercept and random linear slope) also resulted in a significant improvement in model fit, −2ΔLL(3) = 39, $p < .001$.

The predicted means from the unconditional random quadratic polynomial model for session (i.e., without predictors) are shown in Figure 1, and model parameters using REML estimation are given in Table 1. As shown, the mean predicted response time at session 1 was 1946 ms, with a 95% CI of 916 to 2976 ms. The mean instantaneous linear rate of change at session 1 was −121 ms per session, with a 95% CI of −436 to 194 ms, indicating that not all participants were predicted to improve as evaluated at session 1. Half the mean deceleration in linear rate of change was 14 ms per session, such that the linear rate of change became less negative by 28 ms with each session. The 95% CI for the quadratic effect was of −36 to 63 ms, indicating that not all participants were predicted to decelerate in their rate of improvement across sessions.

<u>Computing random effects confidence intervals for each random effect:</u>

$$\text{Random Effect 95\% CI} = \text{fixed effect} \pm \left(1.96*\sqrt{\text{Random Variance}}\right)$$

$$\text{Intercept 95\% CI} = \gamma_{00} \pm \left(1.96*\sqrt{\tau^2_{U_0}}\right) \rightarrow 1{,}945.9 \pm \left(1.96*\sqrt{276{,}209}\right) = 916 \text{ to } 2{,}976$$

$$\text{Linear Time Slope 95\% CI} = \gamma_{10} \pm \left(1.96*\sqrt{\tau^2_{U_1}}\right) \rightarrow -120.9 \pm \left(1.96*\sqrt{25{,}840}\right) = -436 \text{ to } 194$$

$$\text{Quadratic Time Slope 95\% CI} = \gamma_{20} \pm \left(1.96*\sqrt{\tau^2_{U_2}}\right) \rightarrow 13.9 \pm \left(1.96*\sqrt{634}\right) = -36 \text{ to } 63$$