Example 7c: Generalized Mixed Models for Binomial Longitudinal Outcomes (% Correct) using SAS GLIMMIX and STATA MELOGIT

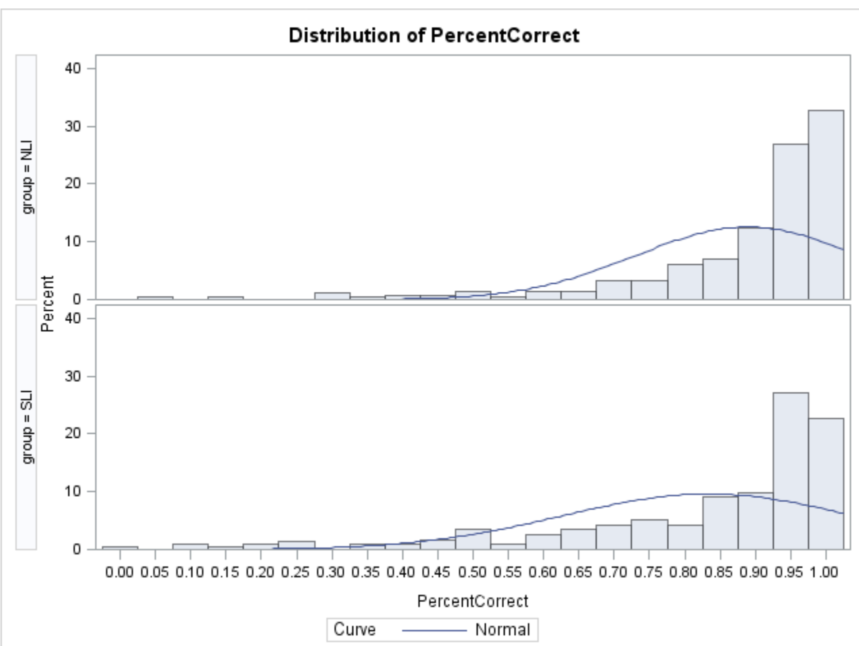
The data for this example are inspired by the publication below, which examined annual growth in a test of grammatical understanding from Kindergarten through 4th grade in children with non-specific language impairment (NLI) or specific language impairment (SLI). Given that percent correct is bounded between 0 and 1, we will use a logit link and a binomial conditional response distribution. In order to use the binomial (which is a discrete distribution), we will need to parameterize the model to predict the number of correct responses out of the number of trials instead of percent correct.

Rice, M. L., Tomblin, J. B., **Hoffman, L.**, Richman, W. A., & Marquis, J. (2004). Grammatical tense deficits in children with SLI and nonspecific language impairment: Relationships with nonverbal IQ over time. *Journal of Speech-Language-Hearing Research, 47*(4), 816-834.

SAS Data Manipulation:

```
DATA work.growthdata; SET work.growthdata;
* Code level-1 ID variable and predictors;
      time=wave-1;
      IF class=2 THEN group="NLI"; * N=56;
      IF class=3 THEN group="SLI"; * N=48;
* Create number of events for binomial model;
      Ntrials=100; PercentCorrect=GrammarOutcome;
      Ncorrect=ROUND(PercentCorrect*Ntrials,1); RUN;

TITLE "Distribution of Percent Correct by NLI and SLI Groups";
PROC UNIVARIATE NOPRINT DATA=growthdata; CLASS group; VAR PercentCorrect;
      HISTOGRAM PercentCorrect / MIDPOINTS= 0 TO 1 BY .05 NORMAL(MU=EST SIGMA=EST); RUN; QUIT;
PROC MEANS DATA=growthdata; VAR PercentCorrect; RUN; TITLE;
```



Individual mean % correct:
M=.86, SD=.13, Min=.24, Max=.99

Occasion-level distribution:
Even though our distributional assumptions will be about the conditional outcome, not the original outcome, odds aren't good it will be normal!

But it may not be strictly binomial, either. The long tails to the left indicate possible over-dispersion (i.e., more variance leftover in the conditional distribution than the binomial distribution would predict). We'll need to model it.

STATA Data Manipulation:

```
* Import data
use "$filesave\growthdata.dta", clear
drop group
gen group=class
label define fgroup 2 "NLI" 3 "SLI"
label values group fgroup
* Distribution of Percent Correct
hist percentcorrect
```

1) Empty Means, Single-Level Model for % correct using DV = Events/Trials

```
* STATA Empty Means, Single-Level Binomial Model
melogit ncorrect , binomial(ntrials),
estat ic, n(104),
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability
estimates store Fit1            // save fit stats
```

Note: I am using Laplace in SAS, which ran faster than Quad. The opposite was the case in STATA. Results are similar.

```
TITLE "Empty Means Single-Level Model";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      ESTIMATE "Intercept" intercept 1 / ILINK; * ILINK un-logits prediction;
RUN;
```

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	**14444.36**
AIC (smaller is better)	14446.36
AICC (smaller is better)	14446.36
BIC (smaller is better)	14450.59
CAIC (smaller is better)	14451.59
HQIC (smaller is better)	14448.01
Pearson Chi-Square	14699.20
Pearson Chi-Square / DF	**28.94**

To inverse link from logits to predicted % correct:

$$\text{Prob}(y=1) = \frac{\exp(1.8475)}{1+\exp(1.8475)} = .8638$$

The sample average probability of getting each item correct is .8638. But Chi-Square/DF > 1, indicating that this model has over-dispersion (too much variance, likely because we haven't incorporated time or the correlation among occasions from the same person).

Parameter Estimates

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|	Gradient
Intercept	1.8475	0.01294	507	142.81	<.0001	-1.04E-8

Estimates

Label	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
Intercept	1.8475	0.01294	507	142.81	<.0001	**0.8638**	0.001522

So even though we are modeling number of correct trials as the DV, the model is phrased to predict percent correct directly (as the conditional mean *p*, the probability that each trial = 1).

2) Empty Means, Two-Level Random Intercept Model for % correct using DV = Events/Trials

```
* STATA Empty Means, Two-Level Random Intercept Binomial Model
melogit ncorrect , || id: , covariance(unstructured) binomial(ntrials),
estat ic, n(104),
estat icc,                      // get ICC
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability
estimates store Fit2            // save fit stats
lrtest Fit2 Fit1                // LRT for random intercept variance

TITLE "SAS Empty Means, Two-Level Random Intercept Binomial Model";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=ID;      * Level-2 subject random intercept;
      ESTIMATE "Intercept" intercept 1 / ILINK;   * ILINK un-logits prediction;
      COVTEST "Need Random Intercept?" 0;         * Test if random intercept is needed;
RUN;
```

COVTEST tests the change in fit if the **G** matrix parameters labeled as 0 were removed from the model.

Convergence criterion (GCONV=1E-8) satisfied.

Fit Statistics

-2 Log Likelihood	**9304.03**
AIC (smaller is better)	9308.03

```
AICC (smaller is better)      9308.06
BIC  (smaller is better)      9313.32
CAIC (smaller is better)      9315.32
HQIC (smaller is better)      9310.18

Fit Statistics for Conditional Distribution
-2 log L(Ncorrect | r. effects)      8819.39
Pearson Chi-Square                   7083.43
Pearson Chi-Square / DF                13.94
```

The ICC for the correlation of occasions within a person:

$$\text{ICC} = \frac{0.9647}{0.9647 + 3.29} = .2268$$

ChiSquare/DF > 1, so this model still has over-dispersion (too much variance, which is to be expected given that no time information has yet been included).

```
             Covariance Parameter Estimates
Cov                                 Standard
Parm       Subject    Estimate        Error     Gradient
UN(1,1)    ID           0.9647       0.1394       -0.001
```

The fixed intercept is not the same as in the previous single-level model because it is now conditional on the random intercept (i.e., expected % correct for someone with $U_{0i} = 0$).

```
                          Solutions for Fixed Effects
                          Standard
Effect        Estimate       Error       DF    t Value    Pr > |t|     Gradient
Intercept       2.1572     0.09787      103      22.04      <.0001     -0.00015

                                       Estimates                           Standard
                          Standard                                            Error
Label         Estimate       Error       DF    t Value    Pr > |t|        Mean         Mean
Intercept       2.1572     0.09787      103      22.04      <.0001      0.8963     0.009093

                         Tests of Covariance Parameters
                            Based on the Likelihood
Label                        DF     -2 Log Like       ChiSq     Pr > ChiSq     Note
Need Random Intercept?        1           14444     5140.32         <.0001     MI
MI: P-value based on a mixture of chi-squares.
```

Random intercept 95% confidence interval in logits = 2.1572 ± 1.96*SQRT(0.9647) = 0.232 to 4.0823, which translates to predicted individual mean probabilities of getting an item correct of .558 to .983.

3) Empty Means, Two-Level Random Intercept Model with Additive Level-1 Over-Dispersion

```
* STATA Empty Means, Two-Level Random Intercept Binomial Model + Over-dispersion
melogit ncorrect , || id: , covariance(unstructured) || case: ,  binomial(ntrials),
estat ic, n(104),
estat icc,                       // get ICC
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability
estimates store Fit3             // save fit stats
lrtest Fit3 Fit2                 // LRT for over-dispersion

TITLE "SAS Empty Means, Two-Level Random Intercept Binomial Model + Over-dispersion";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=ID;        * Level-2 subject random intercept;
      RANDOM INTERCEPT / TYPE=VC SUBJECT=wave*ID;   * Additive level-1 over-dispersion;
      ESTIMATE "Intercept" intercept 1 / ILINK;     * ILINK un-logits prediction;
      COVTEST "Need Overdispersion?" . 0;           * Test if over-dispersion is needed;
RUN;
```

```
        Convergence criterion (GCONV=1E-8) satisfied.
From the log: NOTE: At least one element of the gradient is greater than 1e-3.
          Fit Statistics
-2 Log Likelihood             3547.54
AIC  (smaller is better)      3553.54
AICC (smaller is better)      3553.59
BIC  (smaller is better)      3561.47
CAIC (smaller is better)      3564.47
HQIC (smaller is better)      3556.75
```

```
Fit Statistics for Conditional Distribution
-2 log L(Ncorrect | r. effects)      1643.23
Pearson Chi-Square                     81.02
Pearson Chi-Square / DF                 0.16
```

Over-dispersion has successfully been modeled (< 1)! Now we should be able to have more faith in our results.

```
               Covariance Parameter Estimates
                                        Standard
Cov Parm      Subject      Estimate       Error      Gradient
UN(1,1)       ID             1.0104      0.2642      0.002304
Intercept     ID*wave        3.6173      0.3547      -0.00004
```

The new ICC for the correlation of occasions within a person after accounting for over-dispersion:

$$ICC = \frac{1.0104}{1.0104 + 3.6173 + 3.29} = .1276$$

```
                            Solutions for Fixed Effects
                            Standard
Effect        Estimate         Error       DF    t Value    Pr > |t|     Gradient
Intercept       2.9811        0.1379       91      21.61      <.0001     -0.00321

                                           Estimates
                                                                                     Standard
                            Standard                                                    Error
Label         Estimate         Error       DF    t Value    Pr > |t|         Mean        Mean
Intercept       2.9811        0.1379       91      21.61      <.0001       0.9517    0.006339

                          Tests of Covariance Parameters
                            Based on the Likelihood
Label                       DF    -2 Log Like       ChiSq    Pr > ChiSq    Note
Need Over-dispersion?        1        9304.43     5756.89        <.0001    MI
```

Random intercept 95% confidence interval in logits = 2.9811 ± 1.96*SQRT(1.0104) = 1.011 to 4.951, which translates to predicted individual mean probabilities of getting an item correct of .733 to .993.

3) Add Saturated Means by Wave

```
* STATA Empty Means, Two-Level Random Intercept Binomial Model with Over-dispersion
melogit ncorrect , || id: , covariance(unstructured) || case: ,  binomial(ntrials),
estat ic, n(104),
estat icc,                      // get ICC
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability
estimates store Fit3            // save fit stats
lrtest Fit3 Fit2                // LRT for over-dispersion

TITLE "SAS Saturated Wave Means, Two-Level Random Intercept + Over-Dispersion";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = wave / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      RANDOM INTERCEPT / TYPE=UN SUBJECT=ID;         * Level-2 subject random intercept;
      RANDOM INTERCEPT / TYPE=VC SUBJECT=wave*ID;    * Additive level-1 over-dispersion;
      LSMEANS wave / DIFF=ALL PLOT=MEANPLOT(CLBAND JOIN);     * Plot in logits;
      LSMEANS wave / ILINK PLOT=MEANPLOT(CLBAND JOIN ILINK); * Plot in prob;
RUN;
        Convergence criterion (GCONV=1E-8) satisfied.
From the log: NOTE: At least one element of the gradient is greater than 1e-3.

           Fit Statistics
-2 Log Likelihood            3245.53
AIC  (smaller is better)     3259.53
AICC (smaller is better)     3259.75
BIC  (smaller is better)     3278.04
CAIC (smaller is better)     3285.04
HQIC (smaller is better)     3267.03
```

```
Fit Statistics for Conditional Distribution
-2 log L(Ncorrect | r. effects)      1676.91
Pearson Chi-Square                    105.48
Pearson Chi-Square / DF                 0.21
                 Covariance Parameter Estimates
                                        Standard
Cov Parm      Subject     Estimate         Error      Gradient
UN(1,1)       ID            1.3640        0.2442      0.016008
Intercept     ID*wave       1.3802        0.1472      0.042191
```

Some of the over-dispersion has been reduced by including mean differences over time. This is analogous to explanation of level-1 residual variance (which makes the level-2 random intercept variance increase).

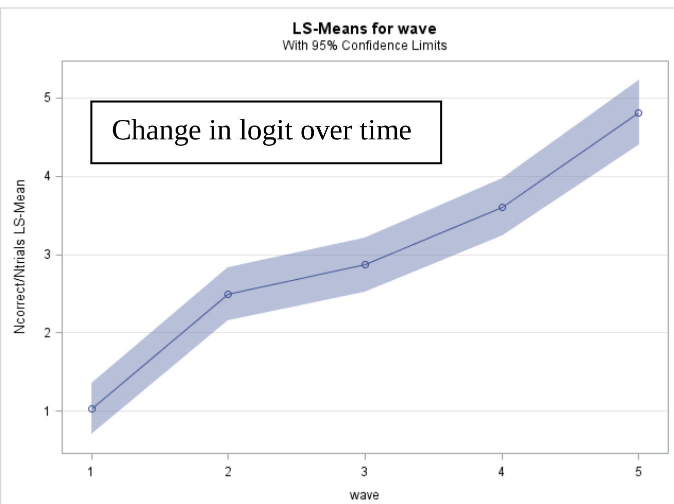
```
          Type III Tests of Fixed Effects
                Num       Den
Effect           DF        DF    F Value    Pr > F
wave              4       412      92.49    <.0001
```

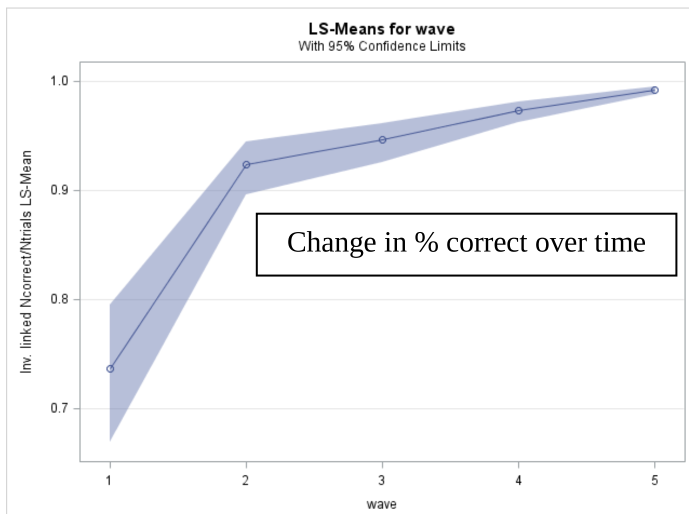
Without the over-dispersion, it's very different!

```
          Type III Tests of Fixed Effects
                Num       Den
Effect           DF        DF    F Value    Pr > F
wave              4       400     877.78    <.0001
```

```
                              wave Least Squares Means
                                                                    Standard
                      Standard                                         Error
wave   Estimate          Error      DF   t Value   Pr > |t|       Mean      Mean
1        1.0269         0.1649     412      6.23     <.0001     0.7363   0.03201
2        2.4917         0.1711     412     14.56     <.0001     0.9236   0.01208
3        2.8673         0.1739     412     16.49     <.0001     0.9462  0.008850
4        3.5998         0.1857     412     19.39     <.0001     0.9734  0.004808
5        4.8145         0.2111     412     22.80     <.0001     0.9920  0.001685
```

Note that the standard error increases slightly with the predicted logit, but decreases with the predicted probability.





3) Add Random Linear Time (to allow for changing variance over time beyond influence of mean)

```
* STATA Empty Means, Two-Level Random Intercept Binomial Model with Over-dispersion
melogit ncorrect , || id: time, covariance(unstructured) || case: ,  binomial(ntrials),
estat ic, n(104),
estat icc,                       // get ICC
nlcom 1/(1+exp(-1*(_b[_cons]))) // intercept in probability
estimates store Fit3             // save fit stats
lrtest Fit3 Fit2                 // LRT for over-dispersion

TITLE "SAS Add Random Linear Time";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = wave / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      RANDOM INTERCEPT time / TYPE=UN SUBJECT=ID;   * Add level-2 subject random time slope;
      RANDOM INTERCEPT / TYPE=VC SUBJECT=wave*ID;   * Additive level-1 over-dispersion;
      COVTEST "Need Random Linear Time?" . 0 0 .;   * Keep intercept and over-dispersion;
RUN;
        Convergence criterion (GCONV=1E-8) satisfied.
From the log: NOTE: At least one element of the gradient is greater than 1e-3.
```

```
           Fit Statistics
-2 Log Likelihood            3227.57
AIC  (smaller is better)     3245.57
AICC (smaller is better)     3245.93
BIC  (smaller is better)     3269.37
CAIC (smaller is better)     3278.37
HQIC (smaller is better)     3255.21
Fit Statistics for Conditional Distribution
-2 log L(Ncorrect | r. effects)     1672.53
Pearson Chi-Square                   103.98
Pearson Chi-Square / DF                0.20

             Covariance Parameter Estimates
                                      Standard
Cov Parm      Subject     Estimate       Error     Gradient
UN(1,1)       ID            0.9883      0.2645     -0.00042
UN(2,1)       ID           0.03870     0.08193     -0.00251
UN(2,2)       ID            0.1212     0.05044      0.00142
Intercept     ID*wave       1.1855      0.1498       0.0006

        Type III Tests of Fixed Effects
              Num      Den
Effect         DF       DF    F Value    Pr > F
wave            4      412      65.13    <.0001
                              wave Least Squares Means
                                                                             Standard
                    Standard                                                    Error
wave    Estimate       Error         DF     t Value    Pr > |t|        Mean        Mean
1         1.0243      0.1473        412        6.95      <.0001      0.7358     0.02863
2         2.4653      0.1598        412       15.43      <.0001      0.9217     0.01154
3         2.8662      0.1758        412       16.31      <.0001      0.9461    0.008955
4         3.6794      0.2084        412       17.66      <.0001      0.9754    0.005003
5         5.1063      0.2675        412       19.09      <.0001      0.9940    0.001601

                        Tests of Covariance Parameters
                           Based on the Likelihood

Label                          DF     -2 Log Like      ChiSq
Need Random Linear Time?        2               .          .            .
```

Relative to previous fixed time model, $-2\Delta LL(2) = 17.96$, $p < .0001$, even if COVTEST couldn't figure it out

4) Finally, the Model of Actual Interest: Group Differences in Change over Time

```
* STATA Add Group and Group*Wave
melogit ncorrect , || id: time, covariance(unstructured) || case: ,  binomial(ntrials),
estat ic, n(104),
contrast i.wave#i.group                              // multivariate Wald tests
margins  i.wave#i.group, predict(xb fixedonly) pwcompare(pveffects) // pairwise comparisons
margins  i.wave#i.group, predict(xb fixedonly)       // predicted logits
marginsplot, name(predicted_logit, replace)          // plot predicted logits
margins  i.wave#i.group, predict(   fixedonly)       // predicted probabilities
marginsplot, name(predicted_probability, replace)    // plot predicted probabilities

TITLE "SAS Add Group and Group*Wave";
PROC GLIMMIX DATA=work.growthdata NOCLPRINT NAMELEN=100 METHOD=LAPLACE GRADIENT;
      CLASS ID group wave;
      MODEL Ncorrect/Ntrials = wave|group / SOLUTION DDFM=BW LINK=LOGIT DIST=BINOMIAL;
      RANDOM INTERCEPT time / TYPE=UN SUBJECT=ID;    * Level-2 subject random int and time slope;
      RANDOM INTERCEPT / TYPE=VC SUBJECT=wave*ID;    * Additive level-1 over-dispersion;
      LSMEANS wave*group / ILINK PLOT=MEANPLOT(CLBAND SLICEBY=group JOIN);       * Plot in logits;
      LSMEANS wave*group / ILINK PLOT=MEANPLOT(CLBAND SLICEBY=group JOIN ILINK); * Plot in prob;
      LSMEANS wave*group / SLICE=group SLICEDIFF=group;     * Simple effects of wave per group;
      LSMEANS wave*group / SLICE=wave;                      * Simple effects of group per wave;
RUN;
```

```
        Convergence criterion (GCONV=1E-8) satisfied.
```

From the log: NOTE: At least one element of the gradient is greater than 1e-3.

Fit Statistics

-2 Log Likelihood	3208.74
AIC (smaller is better)	3236.74
AICC (smaller is better)	3237.59
BIC (smaller is better)	3273.76
CAIC (smaller is better)	3287.76
HQIC (smaller is better)	3251.74

Fit Statistics for Conditional Distribution

-2 log L(Ncorrect | r. effects)	1674.11
Pearson Chi-Square	105.31
Pearson Chi-Square / DF	0.21

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error	Gradient
UN(1,1)	ID	0.9950	0.2612	0.004742
UN(2,1)	ID	0.02209	0.07824	-0.01346
UN(2,2)	ID	0.1008	0.04510	0.017915
Intercept	ID*wave	1.1457	0.1449	0.003692

In theory we could compute pseudo-R^2 reductions for the effect of group on the intercept and linear slope, but the mis-match between the model for the means and variance prevents doing so here. An alternative would be to convert the F-values into standardized mean differences (e.g., Cohen's d). d = 2*SQRT(F/Den DF)

Type III Tests of Fixed Effects

Effect	Num DF	Den DF	F Value	Pr > F
wave	4	408	68.42	<.0001
group	1	90	8.32	0.0049
group*wave	4	408	3.82	0.0047

This is the default ANOVA-style output, of which the only term that is actually useful is group*wave.

Tests of Effect Slices for group*wave Sliced By group

group	Num DF	Den DF	F Value	Pr > F
NLI	4	408	47.57	<.0001
SLI	4	408	27.96	<.0001

These are the simple effects of each predictor at each level of the other, as requested by SLICE.

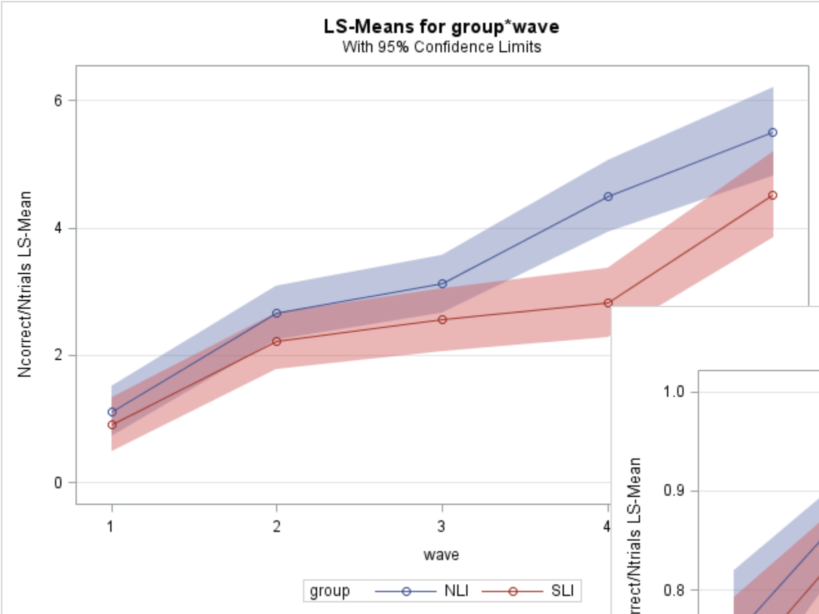
Also given (but not shown for brevity) by SLICEDIFF are all the pairwise differences within group between each occasion.

Tests of Effect Slices for group*wave Sliced By wave

wave	Num DF	Den DF	F Value	Pr > F
1	1	408	0.48	0.4887
2	1	408	2.05	0.1526
3	1	408	2.77	0.0967
4	1	408	18.10	<.0001
5	1	408	4.32	0.0383

group*wave Least Squares Means

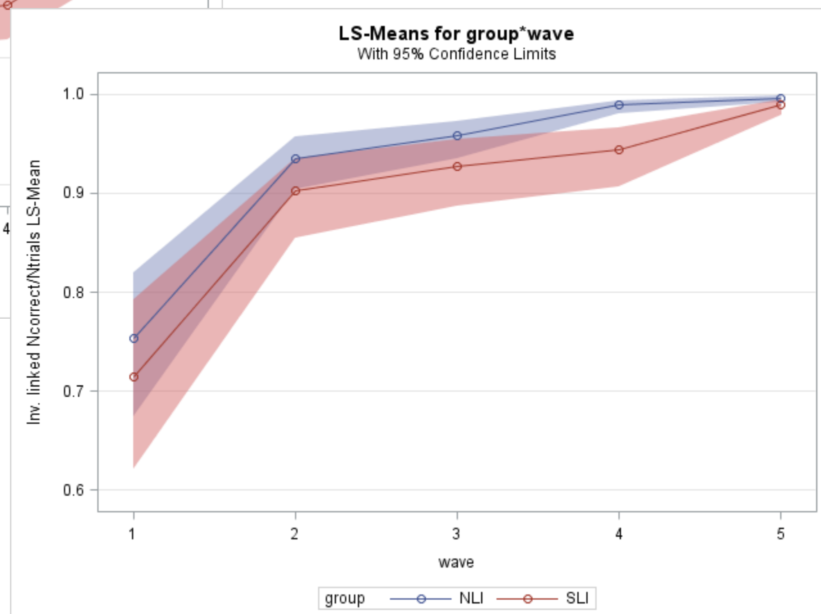
group	wave	Estimate	Standard Error	DF	t Value	Pr > |t|	Mean	Standard Error Mean
NLI	1	1.1180	0.1990	408	5.62	<.0001	0.7536	0.03696
NLI	2	2.6682	0.2144	408	12.44	<.0001	0.9351	0.01301
NLI	3	3.1193	0.2336	408	13.35	<.0001	0.9577	0.009468
NLI	4	4.4938	0.2859	408	15.72	<.0001	0.9889	0.003126
NLI	5	5.5116	0.3557	408	15.50	<.0001	0.9960	0.001425
SLI	1	0.9148	0.2154	408	4.25	<.0001	0.7140	0.04398
SLI	2	2.2193	0.2295	408	9.67	<.0001	0.9020	0.02029
SLI	3	2.5536	0.2480	408	10.30	<.0001	0.9278	0.01661
SLI	4	2.8170	0.2770	408	10.17	<.0001	0.9436	0.01475
SLI	5	4.5201	0.3436	408	13.15	<.0001	0.9892	0.003661



To the left is the growth trajectory for each group in **logits**.

An alternative approach to modeling group differences in these means directly would be polynomial functions. The NLI group looks quadratic, but the SLI group looks cubic.

To the right is the growth trajectory for each group in probability (or **percent correct** here). Because of the ceiling effect at 1,the NLI quadratic trend appears more pronounced, whereas the SLI cubic trend appears less pronounced.



Sample Write-up:

The extent of group differences in change in grammatical understanding (measured as percent correct) across five annual occasions was examined was examined in a series of multilevel models. The sample included 508 occasion-specific outcomes, which were modeled as nested at level 1 within their 104 children at level 2, such that child differences were captured via random effects. In the full sample, the mean percent correct was .86, with a large percentage of observations at or near the ceiling. Accordingly, we predicted the number of correct trials out of the number of possible trials using a logit link function and a binomial conditional outcome distribution, which together keep the predicted percent correct outcomes at or below 1. Thus, while the model predicts the logit (log-odds) of a correct answer for any trial, that prediction can be translated into percent correct via an inverse link function (as requested via ILINK, which provides model-predicted proportions and their standard errors. Model parameters were estimated via full-information maximum likelihood (ML) through the Laplace approximation in SAS GLIMMIX. Accordingly, all fixed effects should be interpreted as unit-specific (i.e., as the fixed effect specifically for patients in which the corresponding random effect = 0). The significance of fixed effects was evaluated with Wald tests (i.e., *t*-test or *F*-tests using between–within denominator degrees of freedom), whereas the significance of random effects was evaluated via likelihood ratio tests (i.e., $-2\Delta LL$ with degrees of freedom equal to the number of new random effects variances and covariances).

Initial models indicated a need for both a random intercept variance for children, $-2\Delta LL(1) = 5{,}140$, $p < .001$, and an additive offset to the level-1 binomial-predicted variance, and $-2\Delta LL(1) = 5{,}757$, $p < .001$. In adding four fixed effects to differentiate the five occasions, significant mean differences over time were found, $F(4, 412) = 92.49$, $p < .001$, such that grammatical understanding appeared to grow over time as expected. We then added a level-2 random linear slope for time in order to capture any additional change in child-level variance over time, which was significant, $-2\Delta LL(2) = 17.96$, $p < .001$. We were then able to test our primary hypotheses regaring differences in growth over time between children with non-specific language impairment (NLI; $n = 56$) and children with specific language impairment (SLI; $n = 48$) by adding group as a main effect and as an interaction with occasion. There was a significant group*occasion interaction, $F(4, 408) = 3.82$, $p < .005$, such that group differences diminished over time. As shown in Figure X, while there were no group differences at occasions 1, 2, or 3, the NLI children significantly outperformed the SLI children at occasions 4 and 5.