Introduction to Longitudinal Analysis

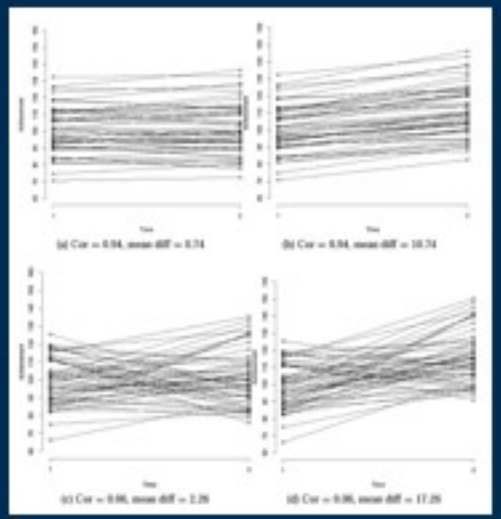
Andrew S. Zieffler

Means vs. Correlations

- Primary focus is examination of mean change of response variable over time
- Correlations are also informative
  - Indicate strength of dependency of response
  - Often more closely spaced observations are more highly correlated
- Focus should depend on research question

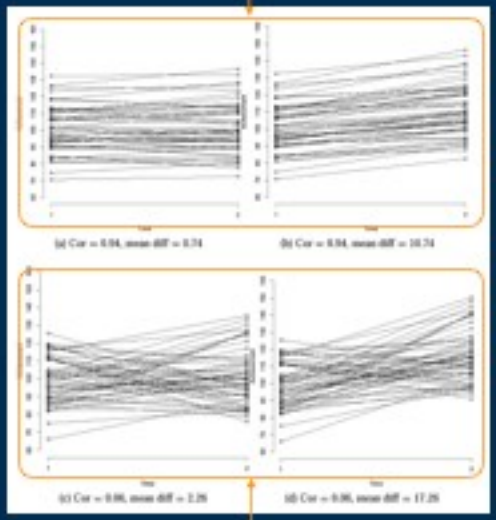
Relative vs. Absolute Change

- Correlations used to make inferences about relative change
  - Based on z-scores and when computed for each time point, have constant mean = 0 across time
  - Precludes meaningful inference about absolute change (i.e., change in quantity of response)
  - Focus in on change in relative position (rank order) over time or *stability* of individuals



Achievement (*y*) measured at two time points for same cohort of individuals

Generally high stability

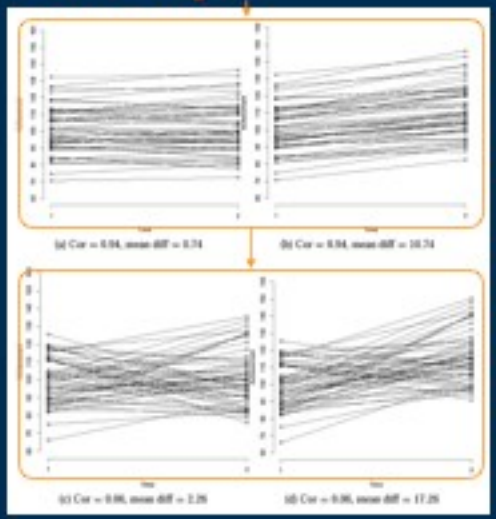


Generally low stability

Achievement (*y*) measured at two time points for same cohort of individuals

Extent of crossing lines is related to strength of stability coefficient (correlation)

Equal stability, different mean change



Achievement (*y*) measured at two time points for same cohort of individuals

Extent of crossing lines is related to strength of stability coefficient (correlation)

Mean difference is indifferent to between-time correlation

What Accounts for Instability

- Variability in individual rate of change
- Random measurement error
  - Scores are not perfectly reliable
  - Statistical model needed to sort out change between these two sources

Measurement Issues

- Same instrument used across time
  - Developmentally appropriate? Construct can change as individuals age–Externalizing behavior (e.g., aggression) measured from age 5 to age 30. Measurement will change (e.g., number of playground bites to ability to get along with boss at work)
  - Empirical check on validity–longitudinal IRT models and examine invariance of IRT parameters

Measurement Issues

- For analysis of mean change, construct must be measured in absolute sense
  - Measurement determined in reference to objective scale (e.g., height)
  - Not influenced by measurement of other subjects
  - Relative measurements do not allow mean change to be studied (e.g., rating of social cooperation)

Response Variable Assumptions

★ Constitutes valid foundation for making inferences about mean change

- Quantitative, continuous (only for this course)

Conceptual Overview of Linear Mixed Effects Regression (LMER)

- Used to analyze mean change, but also allows estimation of individual variation in change
- Incorporates random error term to account for measurement unreliability
- Extension of linear (multiple) regression

Conceptual Overview of Linear Mixed Effects Regression (LMER)

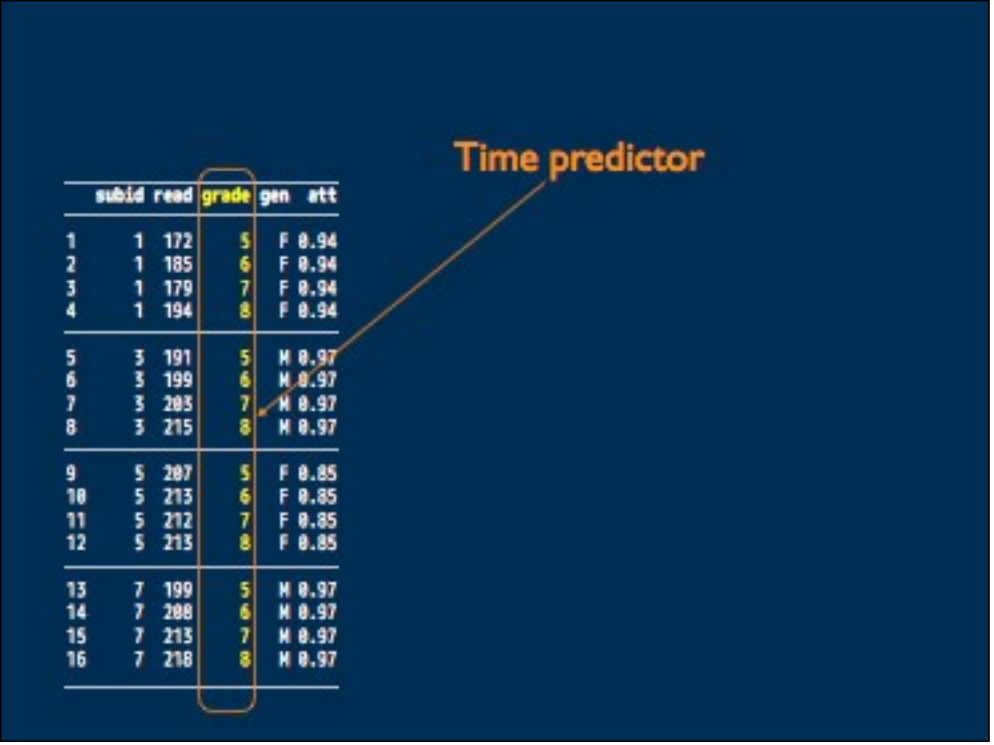
- Goal is to account for variation of the response variable
  - **Within-subjects variation:** row-to-row change for individual subjects
    - Due to changes in response variable over time (i.e., variation in repeated measures)
  - **Between-subjects variation:** block-to-block change in response variable
    - Due to individual differences (i.e., variation in starting values and growth rates between individuals)

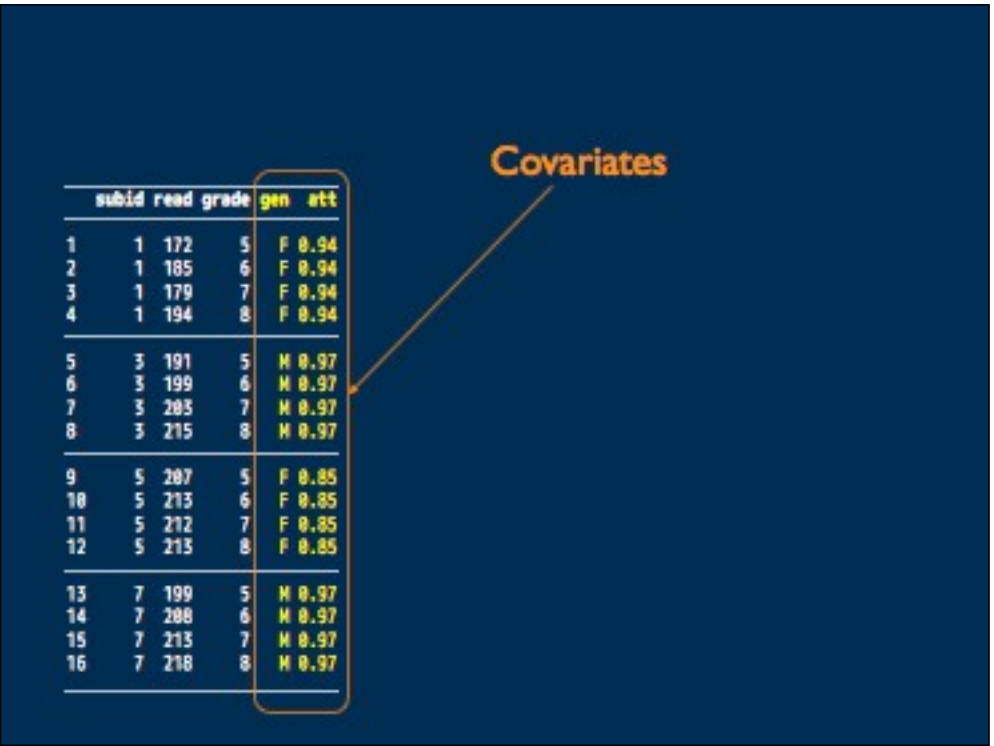
Subject identifier

	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97

Response variable

	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97


Time predictor
subid read grade gen att
1 1 172 5 F 0.94
2 1 185 6 F 0.94
3 1 179 7 F 0.94
4 1 194 8 F 0.94
5 3 191 5 M 0.97
6 3 199 6 M 0.97
7 3 203 7 M 0.97
8 3 215 8 M 0.97
9 5 207 5 F 0.85
10 5 213 6 F 0.85
11 5 212 7 F 0.85
12 5 213 8 F 0.85
13 7 199 5 M 0.97
14 7 208 6 M 0.97
15 7 213 7 M 0.97
16 7 218 8 M 0.97


Covariates
subid read grade gen att
1 1 172 5 F 0.94
2 1 185 6 F 0.94
3 1 179 7 F 0.94
4 1 194 8 F 0.94
5 3 191 5 M 0.97
6 3 199 6 M 0.97
7 3 203 7 M 0.97
8 3 215 8 M 0.97
9 5 207 5 F 0.85
10 5 213 6 F 0.85
11 5 212 7 F 0.85
12 5 213 8 F 0.85
13 7 199 5 M 0.97
14 7 208 6 M 0.97
15 7 213 7 M 0.97
16 7 218 8 M 0.97

	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97

Within-subject variation to account for

	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97

Between-subject variation to account for in start value of response

	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97

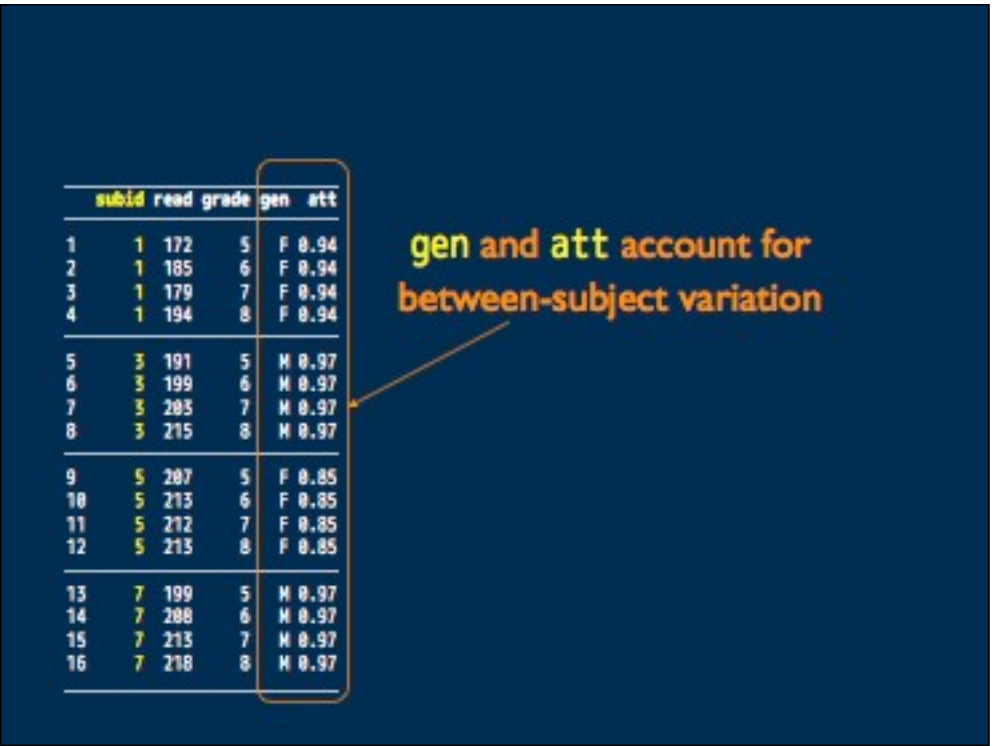
Between-subject variation to account for in growth rates of response

Predictors and Accounting for Variation

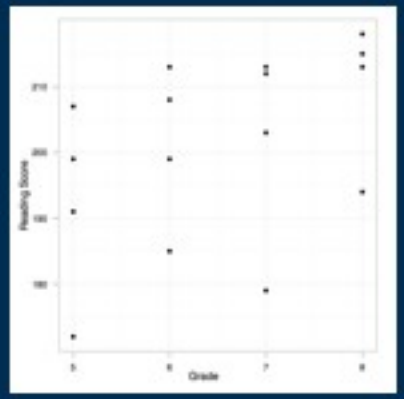
- Predictors that change over time (time predictors, dynamic predictors) account for within-subject variation
- Predictors that do not change over time (static predictors), but vary between individuals, account for between-subject variation

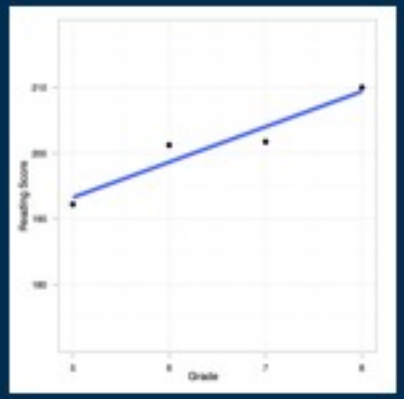
	subid	read	grade	gen	att
1	1	172	5	F	0.94
2	1	185	6	F	0.94
3	1	179	7	F	0.94
4	1	194	8	F	0.94
5	3	191	5	M	0.97
6	3	199	6	M	0.97
7	3	203	7	M	0.97
8	3	215	8	M	0.97
9	5	207	5	F	0.85
10	5	213	6	F	0.85
11	5	212	7	F	0.85
12	5	213	8	F	0.85
13	7	199	5	M	0.97
14	7	208	6	M	0.97
15	7	213	7	M	0.97
16	7	218	8	M	0.97

grade accounts for within-subject variation

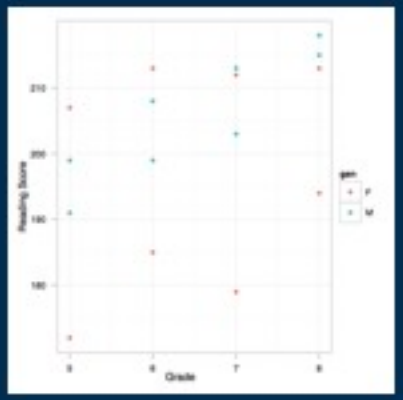


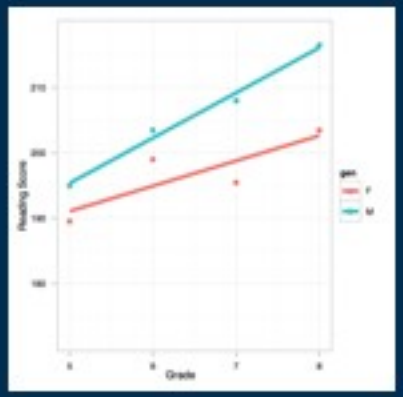
Using grade to Account for Variation in read



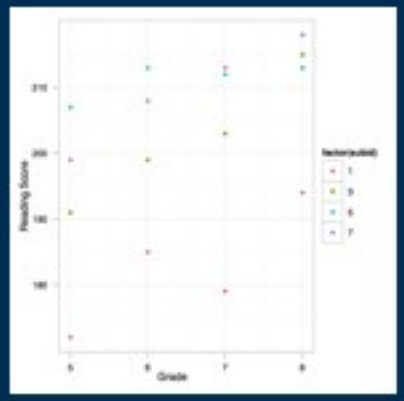


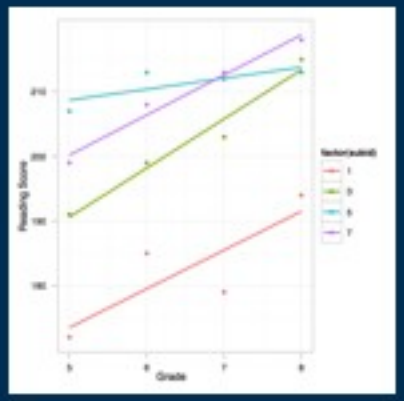
Using grade and gen to Account for Variation in read





Examining Random Effects



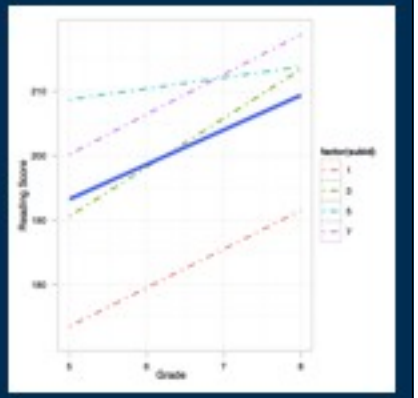


Within-Subject Variation

- Scatter of observed data around single subjects' regression line
- Variation considered to be measurement unreliability
- Typical assumptions of normality, mean = 0, constant variance

Examining Random Effects

subid	intercept	slope
1	143.50	6.00
3	152.60	7.60
5	200.20	1.70
7	169.20	6.20
Mean	166.38	6.20



Between-Subject Variation

- Deviation of individual lines around mean regression line (random effects)
- Variation in random intercepts and random slopes (subject-to-subject variation)
- Correlation between random intercepts and random slopes, $r = -0.87$

Examining Random Effects Conditioning on Gender