Multilevel modeling for eye tracking data

Dan Mirman

Moss Rehabilitation Research Institute

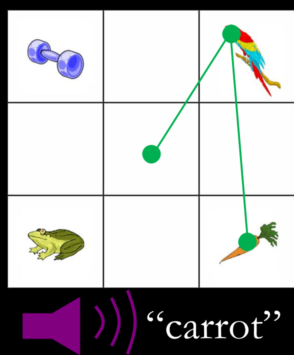
Overview

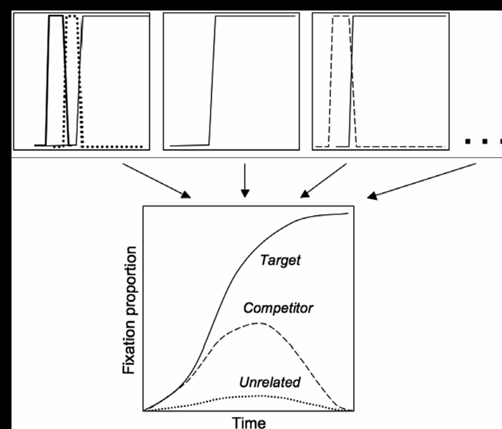
- Conceptual Foundations
  - Eye tracking data, research questions
  - Why use multilevel modeling?
  - What is multilevel modeling?
    - Conceptual description
    - Some MATH

- Walkthrough (afternoon)
- Advanced topic: Individual differences (tomorrow)

Eye tracking data

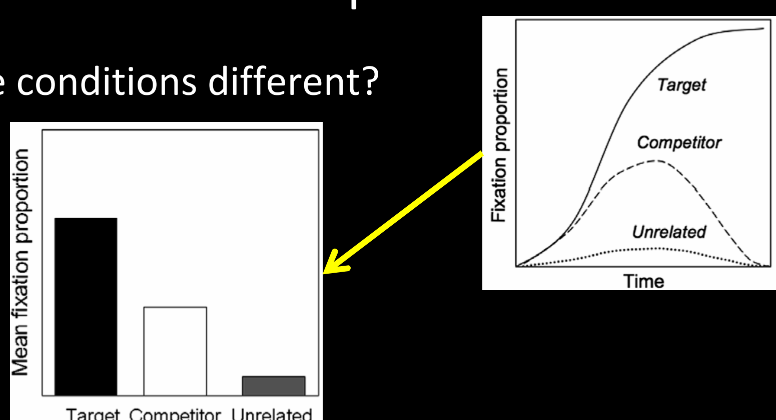
Measuring the time course of language processing





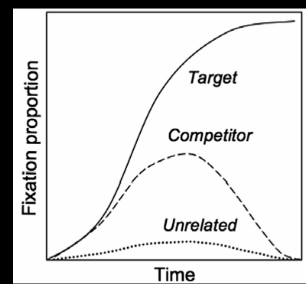
Research questions

1. Are conditions different?

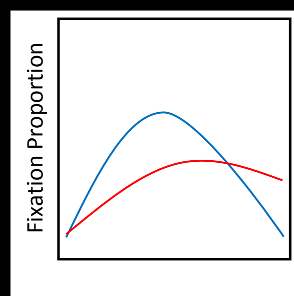


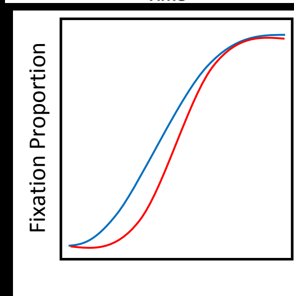
Research questions

1. Are conditions different?
2. When are they different (or not)?
3. How are the time courses different?

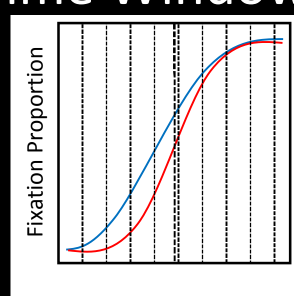


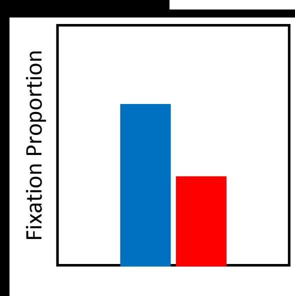
"BLUE is faster than RED"

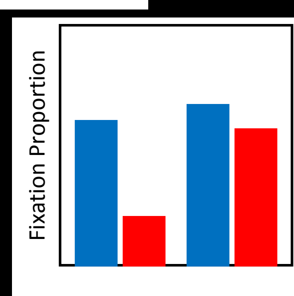




Time Windows





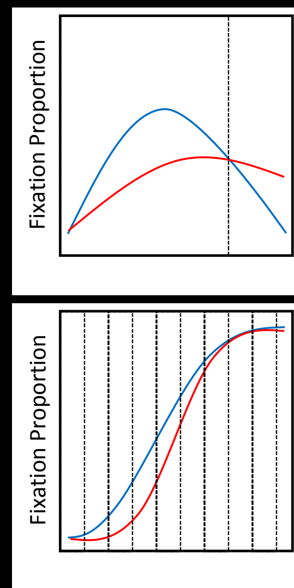


Problems

Å What are the “correct”, experimenter-independent time windows?
  ➢ Especially for cross-over effects

Å Continuity: time bins lose the relationship between time windows

Å Power: each window has relatively little data



The Alternative: Regression

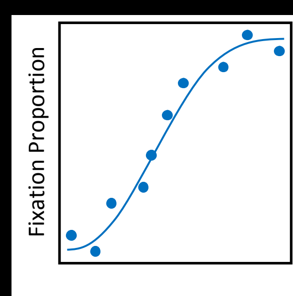
Å Continuous time

Å Effects of covariates

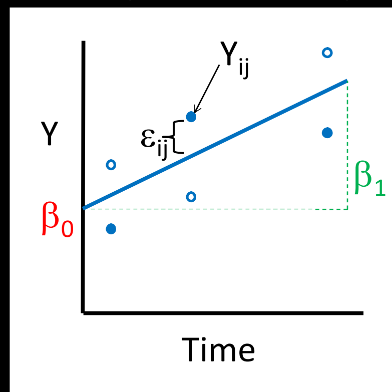
Å Not limited to linear effects

Å Quantify subject/item variability

Å More flexibility about distribution assumptions



A simple example



β_0 : Intercept
β_1 : Slope

"Structural"
"Fixed" Effects

ε_{ij} : residual error from individual observations

"Stochastic"
"Random"
"Residual" Effects

$$Y = \beta_0 + \beta_1 * \text{Time}$$

$$Y_{ij} = \beta_{0i} + \beta_{1i} * \text{Time}_j + \varepsilon_{ij}$$

$$Y_{ij} = \beta_{0i} + \beta_{1i} * \text{Time}_j + \varepsilon_{ij}$$

Fixed effects

- Unique intercept (β_{0i}) and slope (β_{1i}) for each participant
- Multiple values of Time for each participant
- Assuming same functional form for each participant

Random effects

- Unique residual (ε_{ij}) for each observation (Participant x Time)
- Represents measurement error and "unexplained variance"
  - "There are other predictors"
- Distributional assumptions: residuals drawn from normal distribution with mean 0

Multi-level modeling

Level-1

$$Y_{ij} = \beta_{0i} + \beta_{1i}*Time_j + \varepsilon_{ij}$$

Level-2

$$\beta_{0i} = \gamma_{00} + \zeta_{0i}$$

$$\beta_{0i} = \gamma_{00} + \gamma_{0c}*C + \gamma_{0i}*P_i + \zeta_{0i}$$

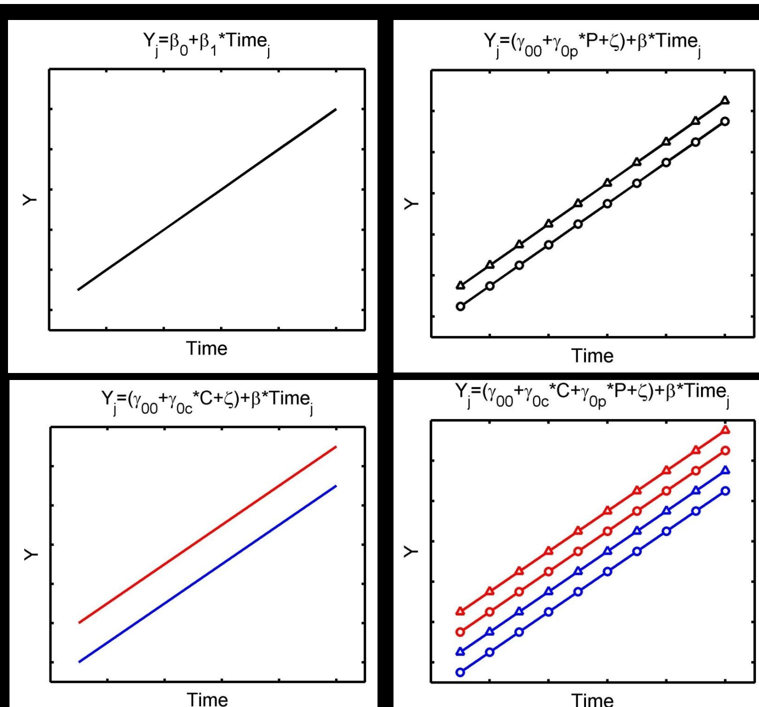
Å Model of the Level-1 parameter (β_{0i})

Å γ_{00} = population mean

Å ζ_{0i} = individual deviation from mean

Å γ_{0c} = fixed effect of condition C on intercept

Å γ_{0i} = fixed effect of Participant i on intercept



Multi-level modeling: Residual error

Level-1 $Y_{ij} = \beta_{0i} + \beta_{1i}*Time_j + \varepsilon_{ij}$

Level-2 $\beta_{0i} = \gamma_{00} + \gamma_{0c}*C + \gamma_{0i}*P_i + \zeta_{0i}$

$\beta_{1i} = \gamma_{10} + \gamma_{1c}*C + \gamma_{1i}*P_i + \zeta_{1i}$

Residual error

Åζ_{0i} = unexplained variance in intercept

Åζ_{1i} = unexplained variance in slope

ÅUnexplained variance ➔ individual differences

ÅRequire a lot of data to estimate

Modeling time

Å Fixation curves are not straight lines

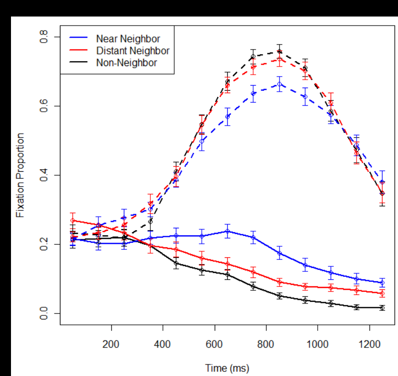
Å Higher-order polynomials

➢ Cons: bad at capturing asymptotic behavior

➢ Pros: (1) can model any curve shape, (2) dynamically consistent

Å How to choose polynomial order?

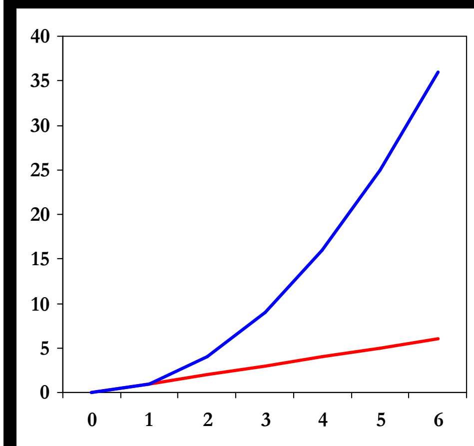
1. Statistical: include only terms that statistically improve model fit
2. Theoretical: include only terms that are predicted to matter
3. Orthodox: use 4^{th} order polynomials because I said so



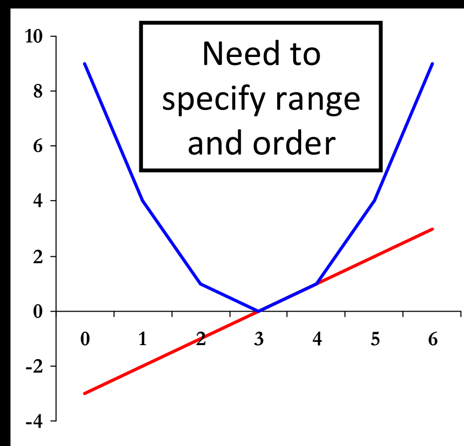
A useful transformation: Orthogonal polynomials

Natural vs. Orthogonal polynomials

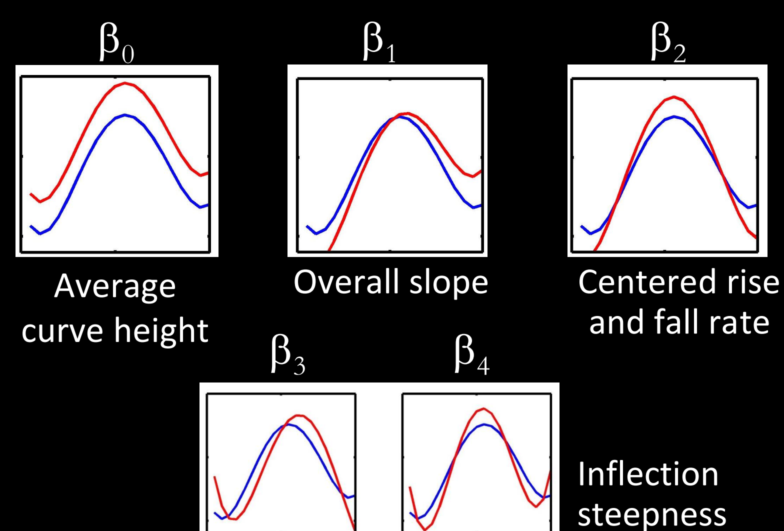
Natural Polynomials
Correlated time terms



Orthogonal Polynomials
Uncorrelated time terms



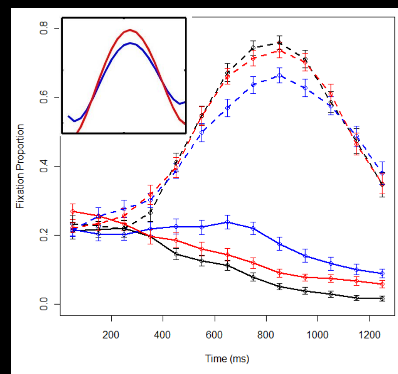
Interpreting orthogonal polynomial terms



Interpreting condition-x-time effects

$Y_{ij} = \beta_{0i} + \beta_{1i}{*}Time_j + \beta_{2i}{*}Time^2_j + \beta_{3i}{*}Time^3_j + \beta_{4i}{*}Time^4_j + \varepsilon_{ij}$

$\beta_{2i} = \gamma_{20} + \gamma_{2c}{*}C + \gamma_{2i}{*}P_i + \zeta_{2i}$



Å γ_{2c} = unique quadratic term for condition C

- ➢Relative steepness for condition C

Å γ_{2i} = unique quadratic term for Participant i

- ➢Relative steepness for Participant i

Maximum Likelihood Estimation

Å Find an estimate of parameters that maximizes the likelihood of observing the actual data

Å Unlike OLS, not a closed-form solution

Å But very powerful and much more flexible

Å Goodness of fit measure: log likelihood (*LL*)

- ï Not inherently meaningful (unlike R^2)
- ï Change in *LL* indicates improvement of the fit of the model
- ï Changes in *-2LL* (aka “Likelihood Ratio”) are distributed as χ^2
  - ÅRequires models be nested (parameters added or removed)
  - ÅDF = number of parameters added

Conceptual summary

- Challenges of longitudinal data
- Multilevel modeling
  - Orthogonal polynomials
  - Fixed effects
    - Evaluate effects of experimental manipulations
  - Random/residual effects
    - Evaluate individual differences
  - Continuous covariates
- Next: gaze, attention, and the visual world paradigm
- Afternoon: Conceptual and hands-on walkthroughs