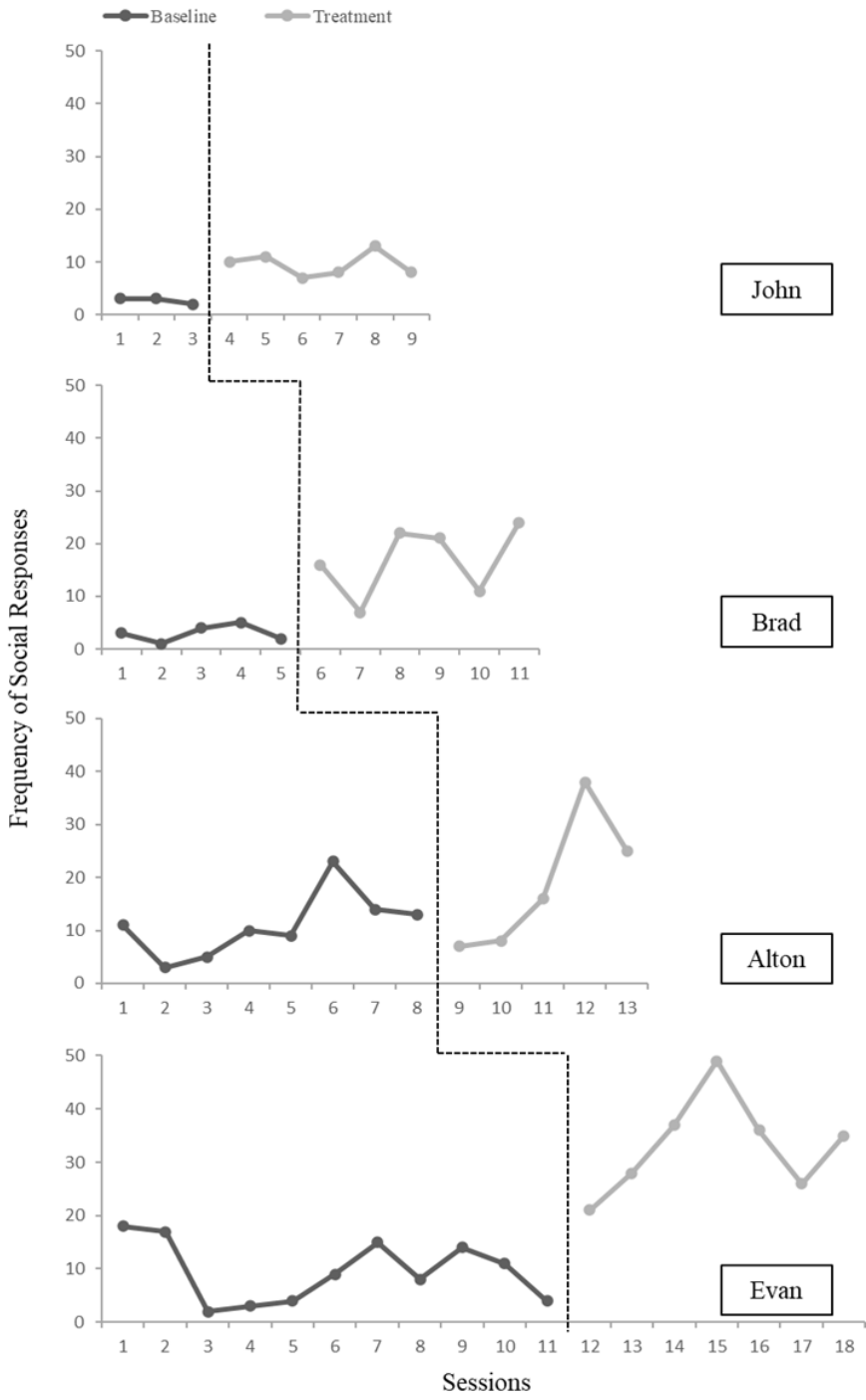
# GLMMs for Overdispersed Count Data in SCED Studies: Does Autocorrelation Matter?

Haoran Li, Wen Luo, Eunkyeng Baek, Kwok Hap Lam, Wenyi Du, Noah Koehler
Texas A&M University
June 28, 2023

# Definition of Single-Case Research

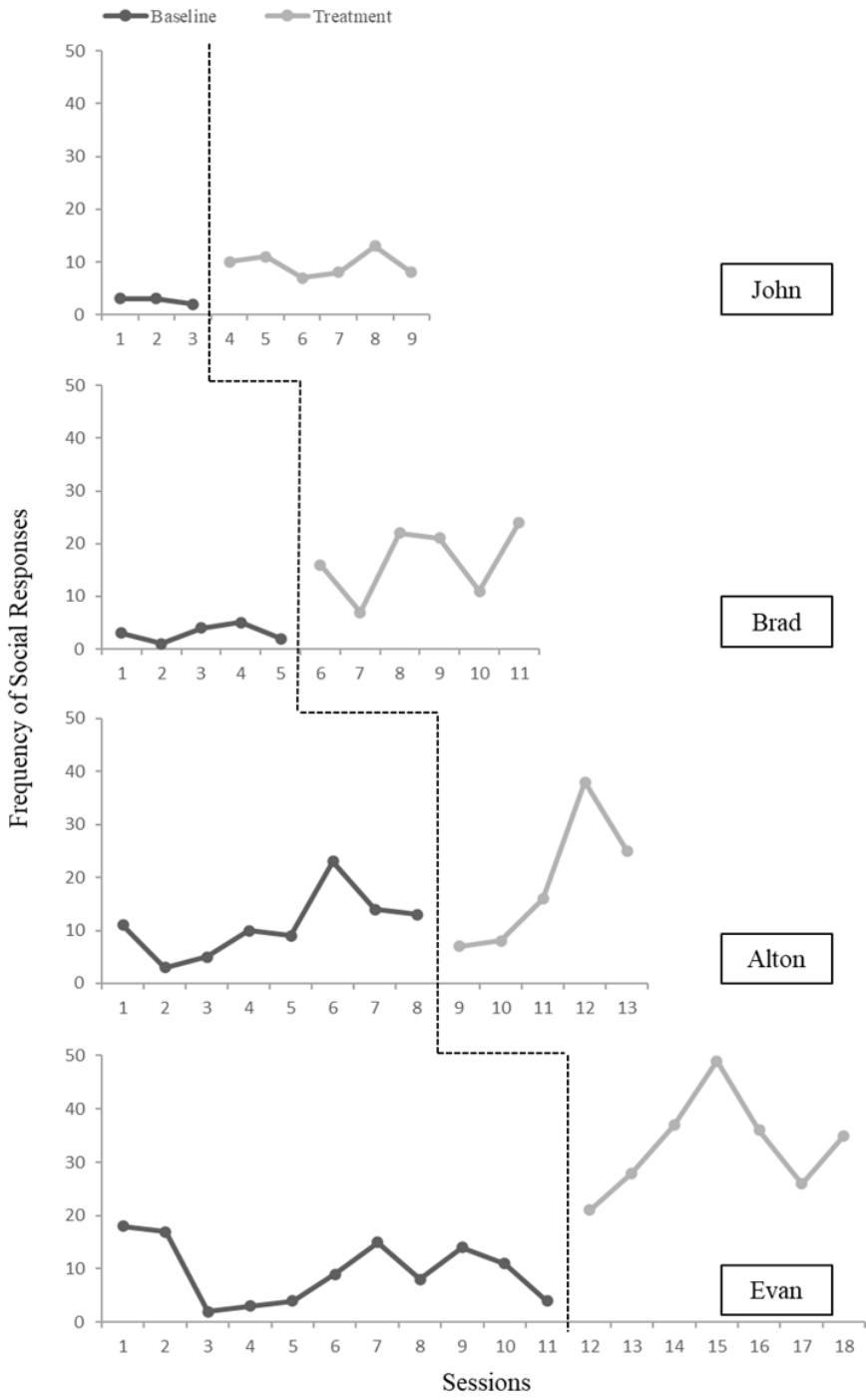
Single-case research is the intensive study of a **single case** or **small samples** by **repeatedly measuring** an outcome before and after an intervention to examine treatment effects.



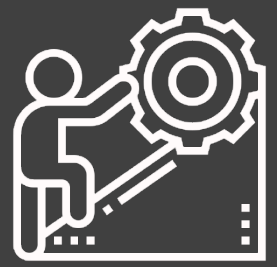
# Multiple Baseline Design

Multiple baseline design (MBD) is comprised of interrupted time series data from multiple cases, settings, or behaviors where an intervention is introduced **sequentially** within different time series (Baek & Ferron, 2013; Ferron et al., 2010).

- The basic interrupted time series in MBD include two phases: baseline and treatment.
- Inferences about the intervention are usually made by comparing different conditions (baseline vs. treatment) presented to cases over time.



# Challenges



Nonnormal outcomes in SCEDs such as count and proportion data

How to deal with autocorrelated count data with trend effects

- A false positive conclusion is likely to be obtained via visual analysis for count and proportion data due to the dependency between mean and variance.
- Visual analysis cannot appropriately deal with autocorrelation, which could lead to inconsistent results and inflated type-I error rates.

# Generalized Linear Mixed Models (GLMMs)

Data consideration:

- Provide **various distributions for count and proportion data** (e.g., Poisson and binomial).

Examples:

- Poisson: number of problematic behaviors observed during a session.
- Binomial: number of intervals of observing social interactions out of a total number of intervals during a session.

# GLMMs

GLMMs have the general form $\boldsymbol{y}|\boldsymbol{u} \sim Distr(\boldsymbol{\mu},\boldsymbol{V}\mu)$, $g(\boldsymbol{\mu})=\boldsymbol{X\beta}+\boldsymbol{Zu}$,

$$Y_{ij} \sim Poisson\left(\lambda_{ij}\right)$$

$$\text{Level 1: } \log\left(\lambda_{ij}\right) = \beta_{0j} + \beta_{1j}Time_{ij} + \beta_{2j}Phase_{ij} + \beta_{3j}Time'_{ij}Phase_{ij}$$

$$\text{Level 2: } \begin{cases} \beta_{0j} = \gamma_{00} + u_{0j} \\ \beta_{1j} = \gamma_{10} + u_{1j} \\ \beTa_{2j} = \gamma_{20} + u_{2j} \\ \beta_{3j} = \gamma_{30} + u_{3j} \end{cases}$$

$$\begin{bmatrix} \mu_{0j} \\ \mu_{1j} \\ \mu_{2j} \\ \mu_{3j} \end{bmatrix} \sim MVN\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{u0} & \sigma_{u0u1} & \sigma_{u0u2} & \sigma_{u0u3} \\ \sigma_{u1u0} & \sigma^2_{u1} & \sigma_{u1u2} & \sigma_{u1u3} \\ \sigma_{u2u0} & \sigma_{u2u1} & \sigma^2_{u2} & \sigma_{u2u3} \\ \sigma_{u3u0} & \sigma_{u3u1} & \sigma_{u3u2} & \sigma^2_{u3} \end{bmatrix}\right),$$

However, the Poisson distribution assumes that the $E(Y) = Var(Y)$, which is often violated due to a data issue called overdispersion.

# Overdispersion

Overdispersion in count data occurs when there is **excessive variance** than what a Poisson can deal with.

- Overdispersed count data: $Var(Y) > E(Y) = \lambda$

Overdispersion source: **correlated measurements**, **extra noise**, and **zero-inflation**.

Overdispersion is not uncommon for count data in SCEDs (Pustejovsky et al., 2019).

Ignoring overdispersion could lead to **biased** standard errors and **inflated** Type I error rates (Hilbe, 2011, 2014; Li et al., 2023).

# Models to Handle Overdispersed Count Data

Negative binomial: $Y_{ij} \sim$ Negative binomial $(\lambda_{ij}, \theta)$

- $E(Y) = \lambda$ and $var(Y) = \lambda + \frac{\lambda^2}{\theta}$, $\theta > 0$

Observation-level random effects (OLRE) model: $Y_{ij} \sim$ Poisson $(\lambda_{ij})$

- $\log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} Time_{ij} + \beta_{2j} Phase_{ij} + \beta_{3j} Time'_{ij} Phase_{ij} + e_{ij}$, $e_{ij} \sim N(0, \sigma_e^2)$
- $E(Y) = \lambda$ and $var(Y) = \lambda + \lambda^2[\exp(\sigma_e^2) - 1]$

# GLMMs with SCED count data

## Performance

**When count data are Not overdispersed:**

| | Data generation | Fitted model | Estimation method | Estimates accurate? | Inferential results reliable? |
|---|---|---|---|---|---|
| Count data | Negative binomial | Poisson | Laplace (Wald test) | ✓ | X |
| | | Negative binomial | | ✓ | X |
| | | OLRE | | ✓ | X |
| | Negative binomial | Poisson | Pseudo likelihood (t test with Kenward-Roger) | ✓ | ✓ |
| | | Negative binomial | | ✓ | ✓ |
| | | OLRE | | ✓ | ✓ |
| Count data | OLRE | Poisson | Laplace (Wald test) | ✓ | X |
| | | Negative binomial | | ✓ | X |
| | | OLRE | | ✓ | X |
| | OLRE | Poisson | Pseudo likelihood (t test with Kenward-Roger) | ✓ | ✓ |
| | | Negative binomial | | ✓ | ✓ |
| | | OLRE | | ✓ | ✓ |

Li, H., Luo, W., Baek, E., Thompson, C. G., & Lam, K. (in press). Multilevel modeling in single-case studies with count and proportion data: A demonstration and evaluation. *Psychological Methods*.

# GLMMs with SCED count data

## Performance

## When count data are **Overdispersed**:

| | Data generation | Fitted model | Estimation method | Estimates accurate? | Inferential results reliable? |
|---|---|---|---|---|---|
| Count data | Negative binomial | Poisson | Laplace (Wald test) | ✓ | X |
| | | Negative binomial | | ✓ | X |
| | | OLRE | | ✓ | X |
| | Negative binomial | Poisson | Pseudo likelihood (t test with Kenward-Roger) | ✓ | X |
| | | Negative binomial | | ✓ | ✓ |
| | | OLRE | | ✓ | ✓ |
| Count data | OLRE | Poisson | Laplace (Wald test) | ✓ | X |
| | | Negative binomial | | ✓ | X |
| | | OLRE | | ✓ | X |
| | OLRE | Poisson | Pseudo likelihood (t test with Kenward-Roger) | ✓ | X |
| | | Negative binomial | | ✓ | ✓ |
| | | OLRE | | ✓ | ✓ |

# A Remaining Issue

- Did not consider autocorrelations in the data generation process. In fact, previous reviews found that the overall magnitude of autocorrelation among SCED data was small (ρ=.20, Shadish & Sullivan, 2011) to moderate (ρ=.46, Barnard-Bark et al., 2021).

- Some researchers argued that GLMMs with distributions for overdispersion (e.g., NB models) can account for autocorrelated errors because one potential source for overdispersion is autocorrelation (Barron, 1992; Hilbe, 2011), this has not been tested in the context of SCEDs.

# A Promising Solution

Observation-level random effect model with an AR(1) error structure (OLRE_AR1): $Y_{ij} \sim$ Poisson $(\lambda_{ij})$

- $\log(\lambda_{ij}) = \beta_{0j} + \beta_{1j} Time_{ij} + \beta_{2j} Phase_{ij} + \beta_{3j} Time'_{ij} Phase_{ij} + e_{ij}$, $e_{ij} \sim AR(1)$
- $E(Y) = \lambda$ and $var(Y) = \lambda + \lambda^2[\exp(\sigma_e^2) - 1]$

# GLMMs with SCED autocorrelated count data

## Purpose: Performance of GLMMs

1) Are the estimates of treatment effects from Poisson, NB, and OLRE models still accurate ?

2) Can we still trust inferential results?

3) Does the newly introduced OLRE_AR model have better performance?

# Simulation Conditions

| Parameter | Value | Rationale |
|---|---|---|
| Series length ($I$) | 10 (starting points of the intervention: 3, 4, 6, 7) or 20 (starting points of the intervention: 6, 8, 12, 14) | Typical SCED setting |
| Number of cases ($J$) | 4 or 8 | Typical SCED setting |
| $\gamma_{10}$ | log (1.00) | |
| $\gamma_{20}$ | log (1.00), log (1.50) or log (3.00) | Zero, medium, and large immediate treatment effects |
| $\gamma_{30}$ | log (1.00), log (1.05) or log (1.10) | Zero, medium, and large treatment effects on the trend |
| $\sigma_{u0}^2$ | 0.1 | |
| $\sigma_{u1}^2$ | 0.0001 | Meta-analytical results of SCEDs with count data |
| $\sigma_{u2}^2$ | 0.1 | |
| $\sigma_{u3}^2$ | 0.0001 | |
| Autocorrelation ($\rho$) | 0.1, 0.3, 0.5, 0.7, or 0.9 | Review of SCEDs |
| $[\gamma_{00}, \sigma_e^2]$ | [log (4.77), 0.095], [log (4.57), 0.182] or [log (4.39), 0.262] | Baseline level: $\exp(\gamma_{00} + \sigma_e^2/2) = 5.00$<br>Dispersion ratio: 1.5, 2.0, 2.5 |

There was a total of 540 conditions, and each had 1000 independent datasets (i.e., replications) generated from the OLRE_AR model.

# Data Analysis for the Primary Simulation

We fitted Poisson, NB, OLRE and OLRE_AR models estimated by pseudo-likelihood using SAS GLIMMIX Procedure.

- Convergence issues for OLRE_AR models. Random effects for the trend effects are removed
- The t test with Kenward Roger adjustment was adopted to conduct statistical inference for treatment effects.

Performance measures:

- Bias, MSE and coverage rate of the immediate treatment effect and treatment effect on the trend
- Type I error rates

# Results of Simulation

- Bias, Coverage Rate, and MSE

**Table 2**

*Bias, Coverage Rate, and MSE of Estimators of Immediate Treatment Effect and Treatment Effect on the Trend*

| Performance | AR | Immediate treatment effect ($\gamma_{20}$) | | | | Treatment effect on the trend ($\gamma_{30}$) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Poisson | NB | OLRE | OLRE_AR | Poisson | NB | OLRE | OLRE_AR |
| Bias | | −0.011 | −0.008 | −0.011 | −0.009 | −0.009 | −0.006 | −0.006 | −0.002 |
| Coverage rate | 0.1 | .908 | .941 | .944 | .954 | .924 | .956 | .959 | .937 |
| | 0.3 | .907 | .934 | .937 | .954 | .915 | .942 | .945 | .931 |
| | 0.5 | .908 | .925 | .928 | .952 | .907 | .927 | .930 | .922 |
| | 0.7 | .918 | .928 | .928 | .952 | .905 | .914 | .915 | .904 |
| | 0.9 | .931 | .938 | .938 | .954 | .915 | .922 | .923 | .882 |
| MSE | | 0.117 | 0.116 | 0.113 | 0.117 | 0.011 | 0.010 | 0.010 | 0.009 |

*Note.* NB = negative binomial; OLRE = observation-level random effects.

# Results of Simulation

- Type I Error Rate

**Table 3**

*Type I Error Rates of Tests for Immediate Treatment Effect/Treatment Effect on the Trend with Count Data*

| Performance | AR | Immediate treatment effect ($\gamma_{20}$) | | | | Treatment effect on the trend ($\gamma_{30}$) | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Poisson | NB | OLRE | OLRE_AR | Poisson | NB | OLRE | OLRE_AR |
| Type I error rate | 0.1 | .090 | .057 | .054 | .053 | .076 | .044 | .041 | .065 |
| | 0.3 | .095 | .064 | .062 | .055 | .085 | .057 | .053 | .072 |
| | 0.5 | .094 | .074 | .071 | .055 | .092 | .073 | .071 | .080 |
| | 0.7 | .084 | .071 | .071 | .057 | .095 | .085 | .083 | .101 |
| | 0.9 | .068 | .060 | .060 | .049 | .083 | .077 | .076 | .121 |

*Note.* NB = negative binomial; OLRE = observation-level random effects.

# Implications and Discussion

- Overall, all models yield accurate estimates for the immediate treatment effect and treatment effect on the trend.

- For immediate treatment effects, OLRE_AR models can successfully control type I error rate at all levels of autocorrelation, while inferential statistics from Poisson models are anticonservative. NB and OLRE models can still control type I error rate when the autocorrelation is small (0.1 or 0.3).

- For treatment effects on the trend, the inferential statistics from OLRE_AR models do not outperform than other models in terms of controlling type I error rates.

# Limitations and Future Research Agenda

- Simulated data based on the MBD across cases, which is the most common design in SCEDs. However, other designs such as ABAB designs are not uncommon in single-case studies.
- Only considered the positive valance outcomes.
- Convergence issues were frequently encountered for OLRE_AR models when adding random effects for trend effects.

# Thank You