

Statistical methodology for modelling the relation between effectiveness of anti-bullying interventions and time

A.-G. Chalamandaris, M. Wilmet-Dramaix, S.K. Ertesvåg, E.
Andreou, D. Piette

Vème Congrès International d Epidémiologie ADELFF/EPITER
Epidémiologie et santé mondialisée

Bruxelles, 13 September 2012

Introduction

- Researchers in several countries have implemented and evaluated the effectiveness of **school-based anti-bullying interventions (SBABI)**.
- Others have worked on summarising and drawing conclusions on the effectiveness of SBABIs through meta-analyses.
- Nevertheless, the **relationship between time and effectiveness** of SBABI has not being fully studied.

Description of SET-Bullying

- Through a literature review **25 references** (i.e. **18 unique studies**) were identified as potentially eligible.
- Owners of the databases were contacted in order to participate to *SET-Bullying*.
- **3 data-bases** received :
 - *RESPEKT*, Norway (Ertesvag and Vaaland, 2007)
 - *DFE-SHEFFIELD*, United Kingdom (Eslea and Smith, 1998)
 - *HELABI*, Greece (Andreou et al., 2007)

Objective: to **understand**, **describe** and **statistically model** the relationship between the time and the effectiveness of a SBABI, through a **secondary analysis** of the included studies.

Similarities between the 3 studies

- As per inclusion criteria for SET-Bullying, their study design was supposed to include study measurements at :
 - **baseline (T1)**
 - **end-of intervention (T2)**
 - **follow-up**: one (**T3**) or two (**T4**)
- **Hierarchical structure** of data: students are nested within school grades or schools (i.e. **clusters**).
- All studies included anonymous data collection procedures
⇒ **longitudinal structure at cluster level**.

Heterogeneity of the 3 studies

Among other things, the 3 studies differ in terms of:

- **Theoretical model** of the anti-bullying intervention.
- **Educational system** of implementation.
- **Length** of intervention.
- **Timing** of study measurements.
- Included population (**country**, **age**, **number** of participants).
- **Outcomes** used to assess the bullying status.
- **Database structure**

Outcomes I

Issue: Each one of the 3 studies has used different scales or questionnaire items.

Action:

- Assumed that similar outcomes measure the **same latent unobserved variable**.
- **Grouping outcomes** in 9 semantic categories.
- Implementing the same analysis strategy (i.e. **similar model**) for all outcomes.
- **Expecting similar pattern** in results for each semantic category.

Outcomes II

The 9 semantic categories:

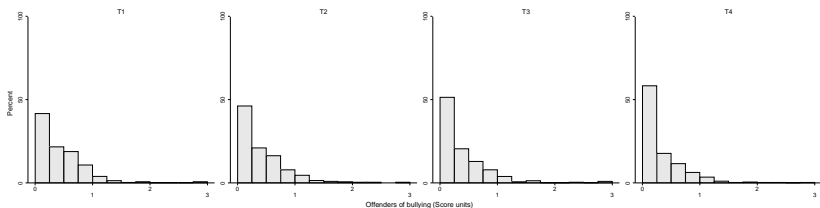
- Self-reported frequencies of being bullied (19)
- Self-reported frequencies of bullying others (10)
- Estimated level of bullying at class (2)
- Participation to bullying related roles (4)
- Opinion towards bullying (1)
- Self-efficacy (3)
- Behaviours towards confrontation of bullying (7)
- General items on socialization at school (4)
- Opinion on efficacy of intervention (2)

Distributional assumptions I

Issue: Analysis of **scales as continuous data**:

- common practice in the literature.
- may deviate from a normal distribution.
- are not genuinely continuous data (i.e. a subject can fall in one of several distinct values).

Database RESPEKT: Histogram of scale *Offenders of bullying (Score units)* at each study measurement (all clusters combined)



Based on students (all clusters combined) with a least one non-missing value in scale *Offenders of bullying (Score units)*.

Distributional assumptions II

Action: when possible analyse both the **scale as continuous** data and the consisting **items as categorical** data, and assess any differences in the results

Outcome

Self-reported frequencies of bullying others

Offenders of bullying

This school year, how often have you **bullied/hassled others** at school?

This school year, how often have you bullied/hassled others at school **by teasing them**?

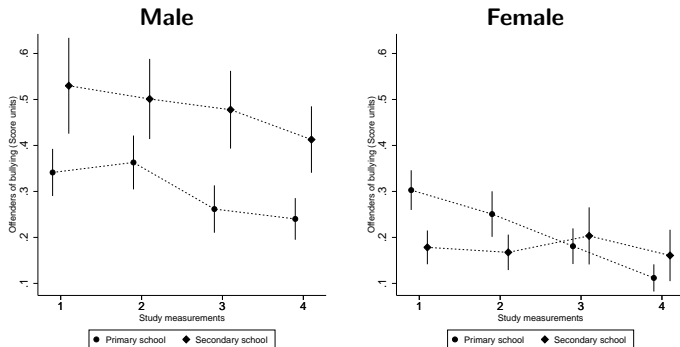
This school year, how often have you bullied/hassled others at school by **isolating/shutting them out**?

This school year, how often have you bullied/hassled others at school **by hitting, kicking, or shoving them**?

Confounding in bullying reports

Issue: Differences by students' age or gender in responses in bullying related outcomes.

Database RESPEKT: Mean and 95% confidence intervals of scale *Offenders of bullying (Score units)* at each study measurement by school type and students' gender



Points correspond to raw means and lines to 95% confidence intervals.

Based on students by school type and gender with a least one non-missing value in scale *Offenders of bullying (Score units)*.

Data structure I

Issue: **Hierarchical** data structure

Responses from students in the same cluster are expected to be more correlated than responses of students from different clusters.

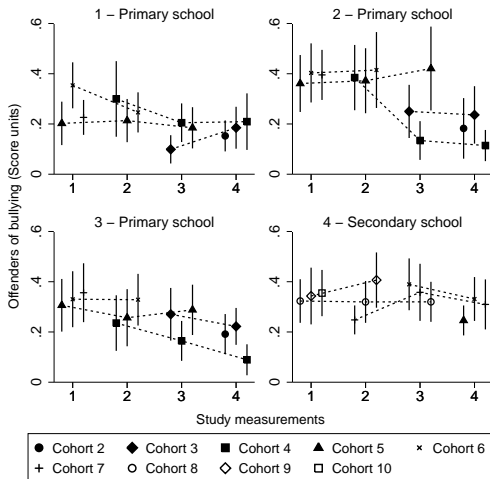
Issue: **Longitudinal** data structure **at cluster level**

Due to the use of anonymous questionnaires, the longitudinal structure of the data can be tracked at the cluster level.

Issue: Effectiveness of the anti-bullying intervention over time may **differ by cluster**.

Data structure II

Database RESPEKT: Mean and 95% confidence intervals of scale *Offenders of bullying* (Score units) at each study measurement by grade clusters



Points correspond to raw means and lines to 95% confidence intervals.

Based on students by grade clusters with a least one non-missing value in scale *Offenders of bullying* (Score units).

Statistical model - General form I

$$y^*_{cti} = \underbrace{\beta_0 + \beta_T T_c + \beta_A A_{cti} + \beta_G G_{cti} + \beta_X X}_{\text{fixed}} + \underbrace{\zeta_{0C} + \zeta_{CT} T_{ct} + \epsilon_{cti}}_{\text{random}}$$

T_c **time** of study measurement (months) $\left\{ \begin{array}{l} \text{categorical} \\ \text{linear} \\ \text{linear + quadratic} \end{array} \right.$

A_{cti} students' **age** (years) \implies **categorical**

G_{cti} students' **gender** \implies **categorical**

X **interaction** terms between T_c, A_{cti}, G_{cti} $\left\{ \begin{array}{l} \text{all 2x and 3x} \\ T \times A \\ T \times G \\ \text{none} \end{array} \right.$

ζ_{0C} **random intercept** for each cluster

$\zeta_{CT} T_{ct}$ **random slope for time** for each cluster

ϵ_{cti} residual

Statistical model - General form II

y^*_{cti} corresponds to

- y_{cti} for **scales** as continuous data
- $\ln \left\{ \frac{\Pr(y_{cti} = 1)}{1 - \Pr(y_{cti} = 1)} \right\}$ for items with **binary responses**
- $\ln \left\{ \frac{\Pr(y_{cti} > s)}{1 - \Pr(y_{cti} > s)} \right\}$ for items with **ordered responses**

Model Selection I

- **52 outcomes** in 9 semantic categories
- **12 models** will be estimated
(3 models on time x 4 models on interactions).
- Model selection based on
BIC_p - Penalized Bayesian Information Criterion

$$BIC_p = - \underbrace{(-2 \cdot LL_0 + 2 \cdot LL)}_{\chi^2} + 2 \cdot \underbrace{(P - 1)}_{df_{\chi^2}} \cdot \ln(N)$$

Model Selection II

- **Base-model** \implies age, gender, time as categorical variables + all interaction terms
- Calculation of the **relative difference** from each model to the base model for every outcome

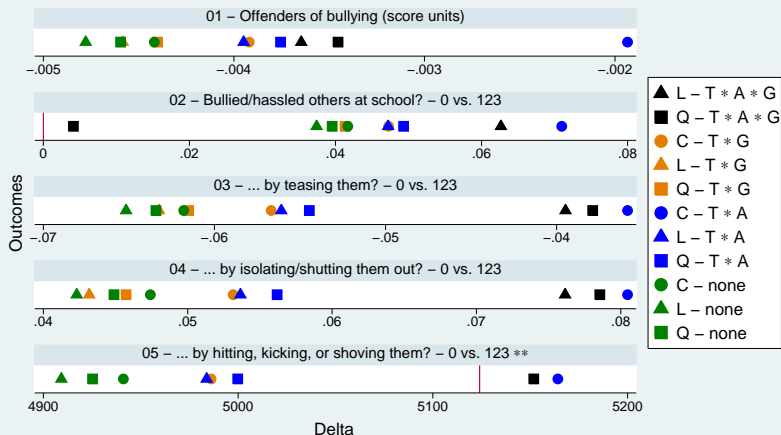
$$\delta = \ln \left(\frac{BIC_{p_{model}}}{BIC_{p_{base}}} \right)$$

- **Smallest the value** of $\delta \implies$ best fit of the data

Model Selection III

Semantic category: **Self-reported frequencies of bullying others**

Database RESPEKT: Preliminary results on model selection



** Base model could not be estimated due to non-convergence. Therefore Delta cannot be computed.

Conclusions

- Heterogeneity and characteristics of the original studies may threaten the validity of the results.
- Extensive modelling of a variety of bullying-related outcomes.
- Analysis methodology imposing minimal assumptions.
- However, *SET-Bullying* is the first effort to study the relationship between time and effectiveness

Acknowledgements

We would like to express their gratitude for the provision of their databases and all their help and communications regarding their studies to:

- Peter K Smith
- Mike J Eslea
- Sigrun K. Ertesvåg
- Eleni Andreou

Without their contributions *SET-Bullying* could have never been realised.

References

- Andreou, E., Didaskalou, E. and Vlachou, A. (2007), 'Evaluating the effectiveness of a curriculum-based anti-bullying intervention program in greek primary schools', *Educational Psychology* **27**(5), 693–711.
- Ertesvag, S. K. and Vaaland, G. S. (2007), 'Prevention and reduction of behavioural problems in school: An evaluation of the respect program', *Educational Psychology* **27**(6), 713–736.
- Eslea, M. and Smith, P. K. (1998), 'The long-term effectiveness of anti-bullying work in primary schools', *Educational Research* **40**(2), 203–18.

thank you