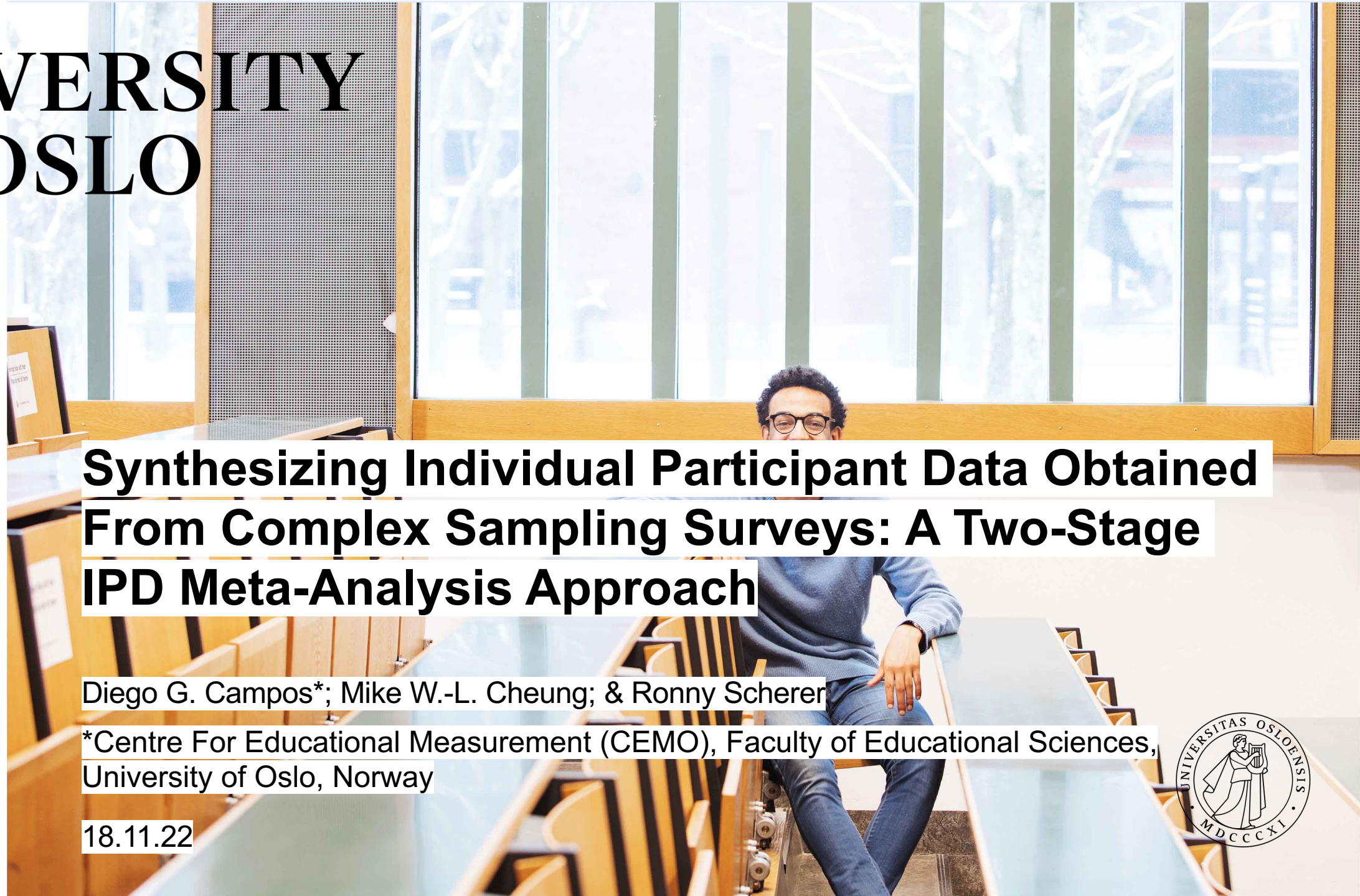
UNIVERSITY OF OSLO



# Synthesizing Individual Participant Data Obtained From Complex Sampling Surveys: A Two-Stage IPD Meta-Analysis Approach

Diego G. Campos*; Mike W.-L. Cheung; & Ronny Scherer

*Centre For Educational Measurement (CEMO), Faculty of Educational Sciences, University of Oslo, Norway

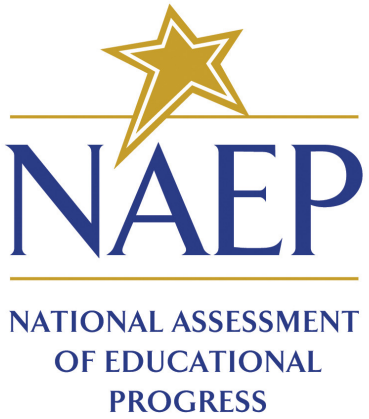
18.11.22

# What Are Complex Surveys?

## Examples

- Program for International Student Assessment (PISA)
- Trends in International Mathematics and Science Study (TIMSS)
- National Assessment of Educational Progress (NAEP)
- National Educational Panel Study (NEPS)





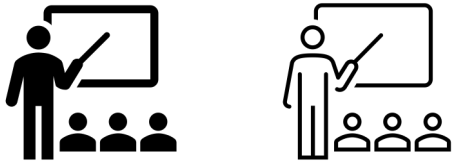

# What Are Complex Surveys?

- Increasing the understanding of critical factors influencing teaching and learning
- Identifying key educational issues: educational inequalities
- Informing national strategies for monitor and improve the educational system



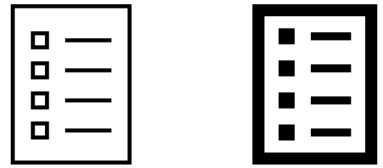
Scherer, Siddiq & Nilsen (2021)

# What Are Complex Surveys?



**Multistage Sampling**

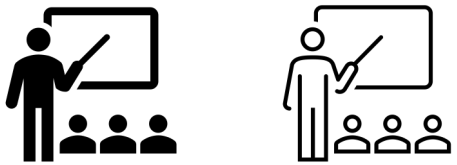
**Survey Weights**



**Rotated Questionnaire Design**

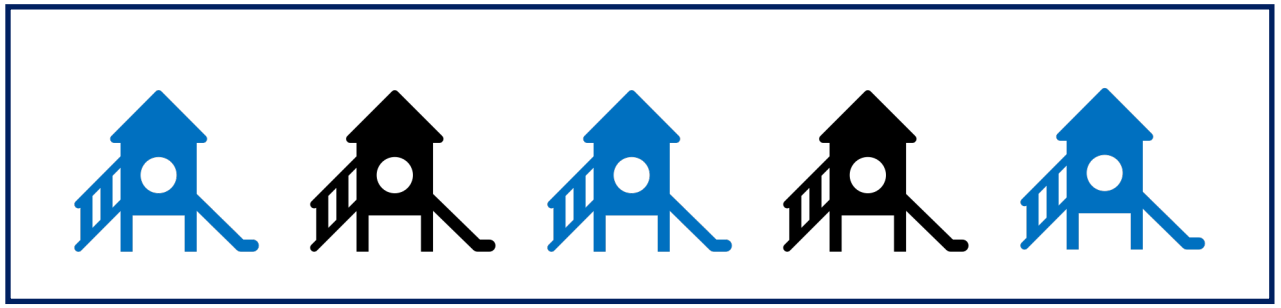
Rutkowski, Gonzalez, Joncas & von Davier (2010)

# What Are Complex Surveys?



**Multistage Sampling**

**1st stage**



**2nd stage**

Stapleton & Thomas (2008)

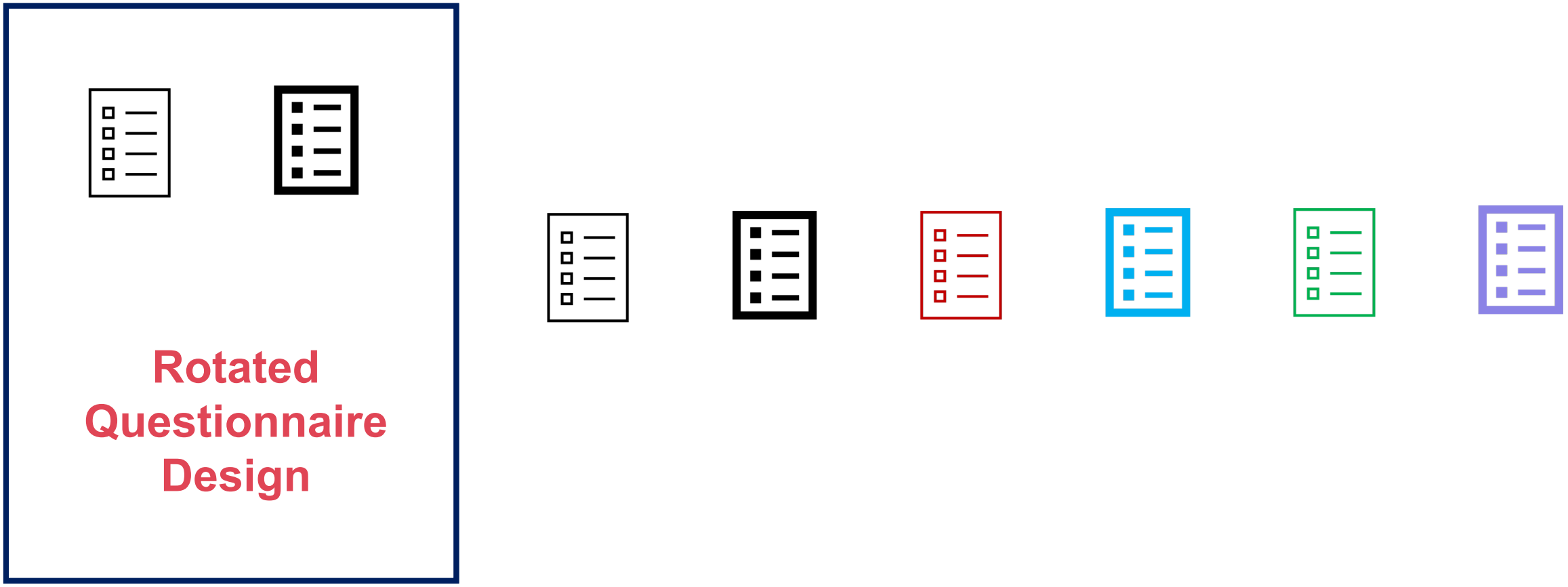
# What Are Complex Surveys?

**Survey Weights**

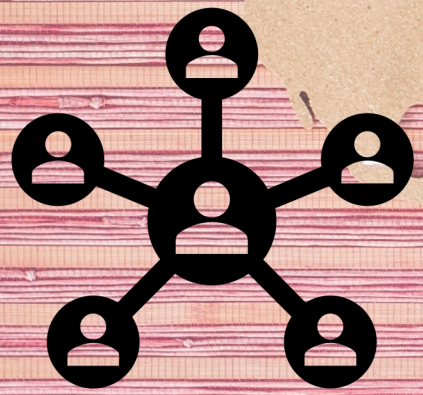
Total weights refer to the weight components that reflect the inclusion probability of a school and a student of being selected.

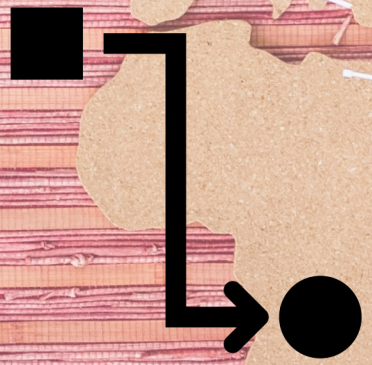
$$W_{ij} = \frac{1}{p_{ij}}$$

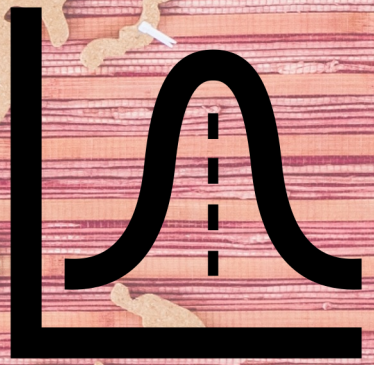
# What Are Complex Surveys?



# Large and Representative Samples





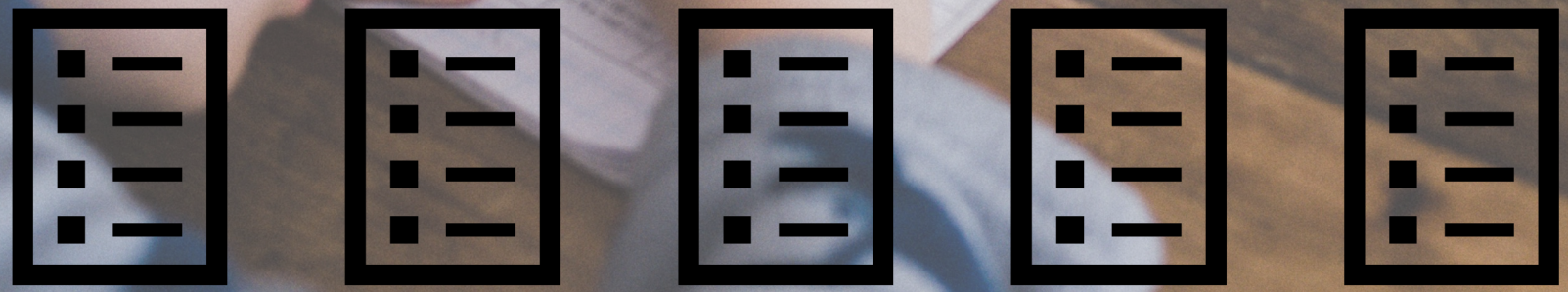


UNIVERSITY
OF OSLO

Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

Photo by Louis Hansel on Unsplash

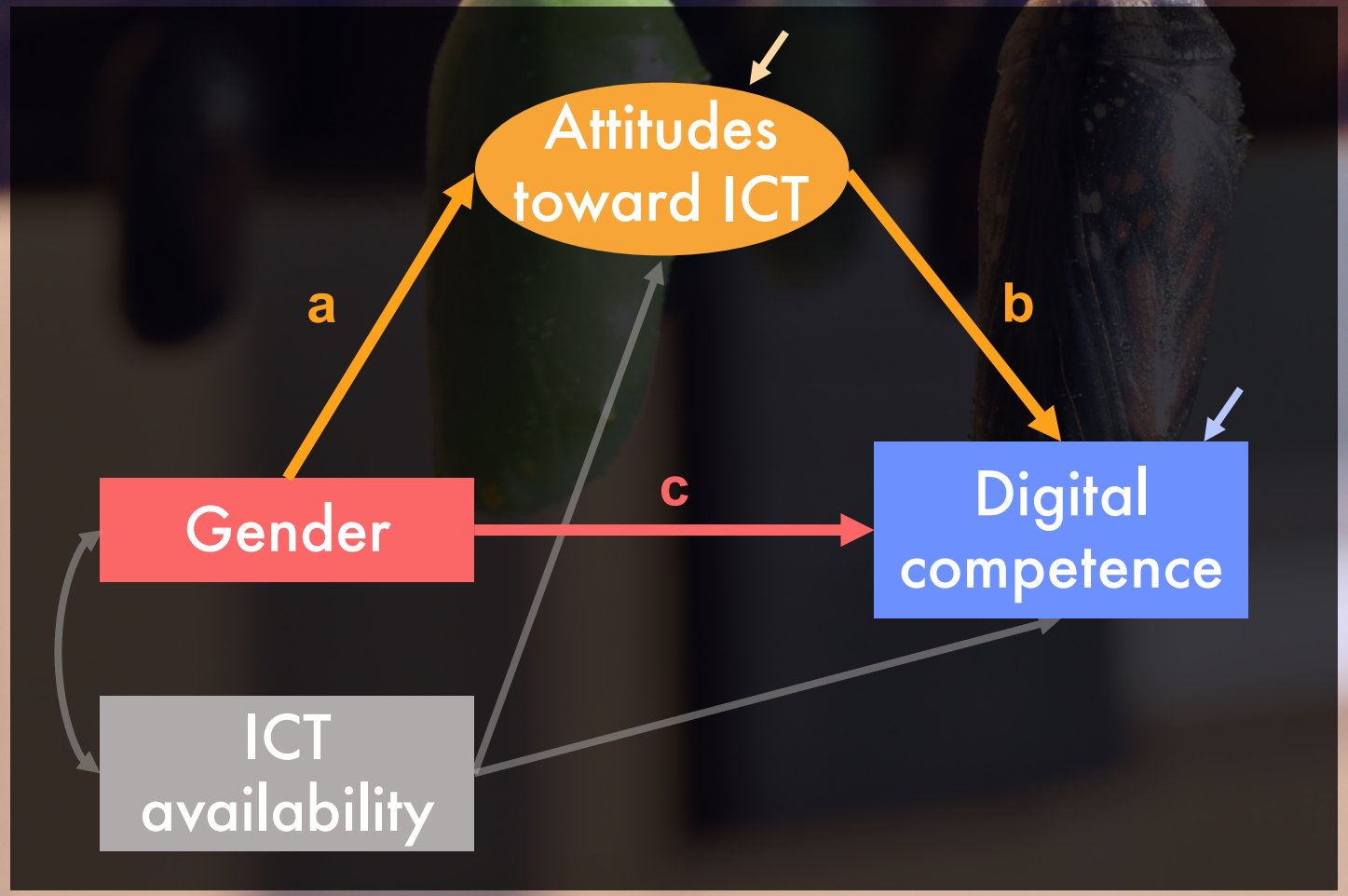
# High Quality Measures



UNIVERSITY
OF OSLO

Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

# Raw Data




UNIVERSITY OF OSLO

Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

# Challenges of Meta-Analysis in Educational Research

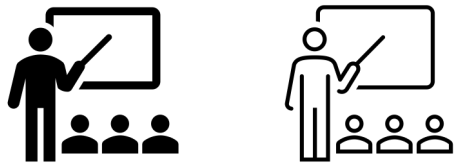
Small samples in primary studies

Study characteristics that may affect the quality and magnitude of effects

Insufficient psychometric quality of outcome measures

Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

# Challenges in the Synthesis of Complex Survey Data

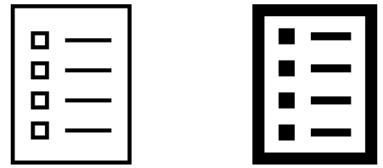


**Multistage Sampling**

Stapleton et al., (2016)

**Survey Weights**

Rutkowski, Gonzalez, Joncas & von Davier (2010)

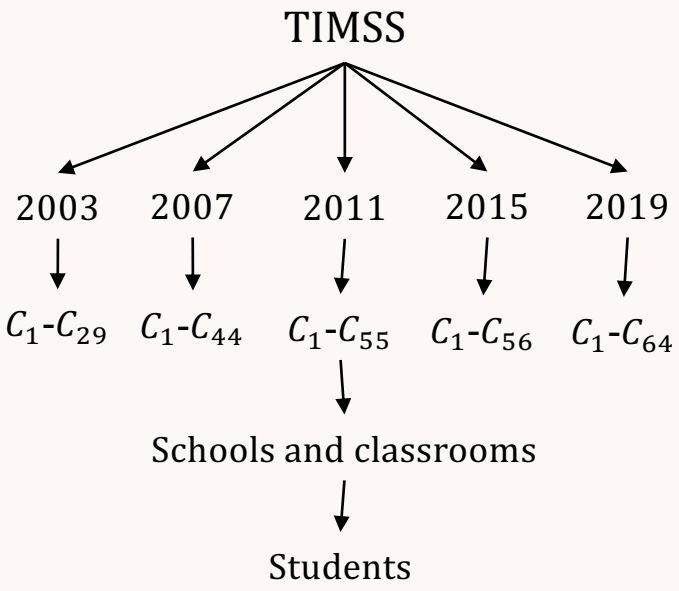


**Rotated Questionnaire Designs**

von Davier et al., (2009)

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 1: Raw Data Analysis




- Hierarchical structure
- Sampling weights
- Plausible values
- Measurement Invariance

*i*: Students, *j*: Classrooms, *k*: Countries

## Stage 2: Meta-Analysis

- Multivariate meta-analysis
- Multilevel meta-analysis
- Mixed-effects meta-regression

# Two-Stage Individual Participant Data Meta-Analysis

## Advantages

Level 1 (individual participants):

$$Y_{ijk} = \beta_{0jk} + \beta_{1jk}X_{ijk} + e_{ijk}$$

Level 2 (clusters):

$$\beta_{0jk} = \beta_{00k} + \beta_{01k}Z_{jk} + u_{0jk}$$

$$\beta_{1jk} = \beta_{10k} + u_{1jk}$$

Level 3 (primary studies):

$$\beta_{00k} = \gamma_{000} + v_{00k}$$

$$\beta_{01k} = \gamma_{010} + v_{01k}$$

$$\beta_{10k} = \gamma_{100} + v_{10k}$$

$\beta_{01k} = \theta_k + r_k$ with $r_k \sim N(0, \sigma^2_{r_k})$.

$\theta_k = \mu + u_k$ with $u_k \sim N(0, \sigma^2_u)$.

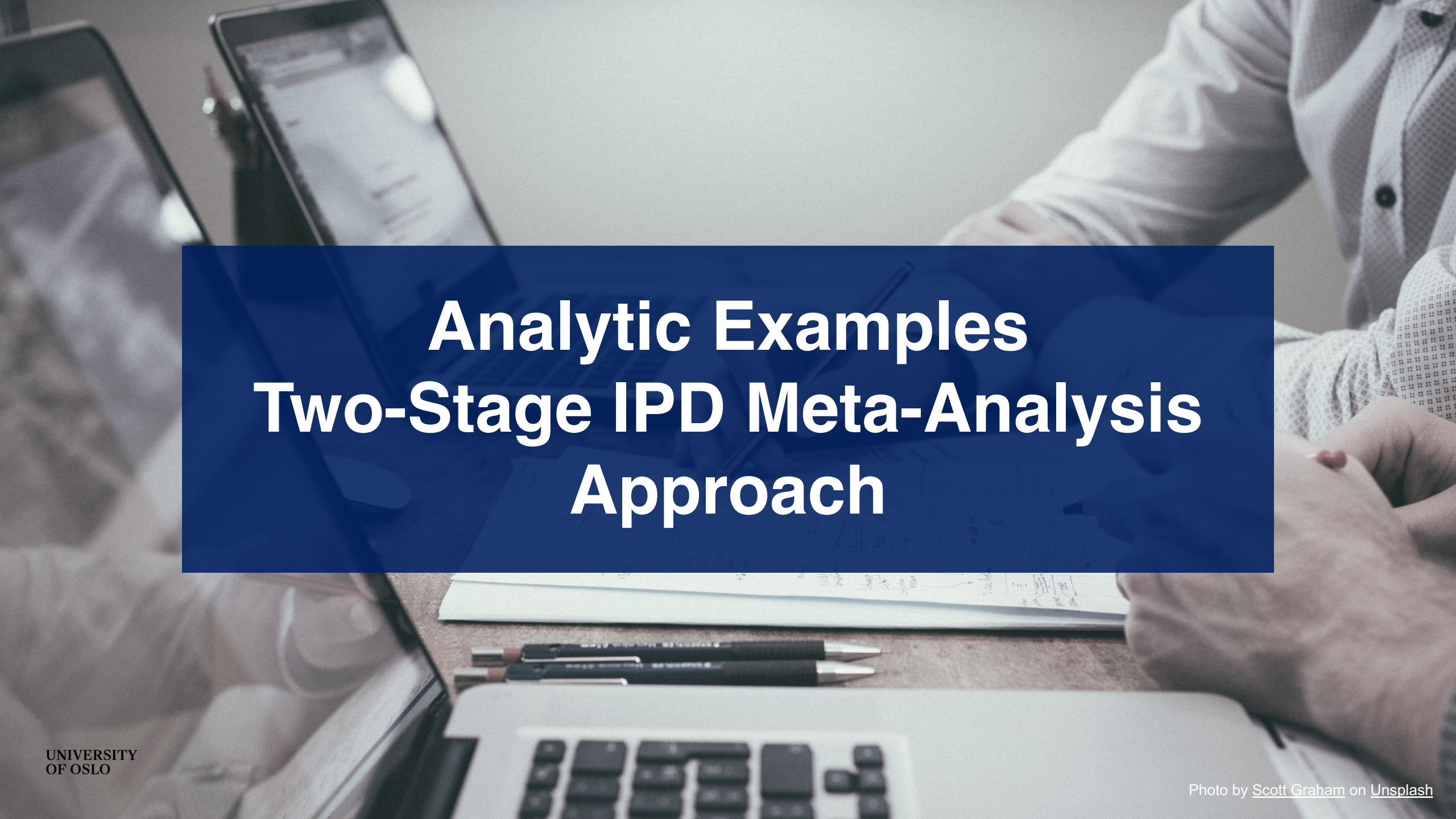
Complex Survey Designs

Measurement Heterogeneity

Multilevel Structures



Campos, Cheung & Scherer (2022)
Brunner, et al. (2022)



# Analytic Examples Two-Stage IPD Meta-Analysis Approach

UNIVERSITY
OF OSLO

Photo by Scott Graham on Unsplash

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 1

*To what extent do girls and boys in secondary education differ in their digital skills?*

ICILS 2013

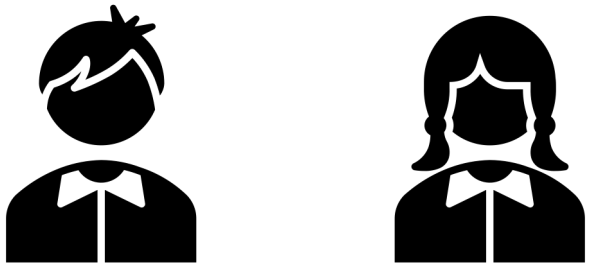
ICILS 2018

| | |
|---|---|
| Constructs of interest | • Gender<br>• CIL |
| Complex data structure | • Study ID<br>• Country ID<br>• School ID<br>• Student ID<br>• Student weights<br>• Jacknife codes |

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 1

*To what extent do girls and boys in secondary education differ in their digital skills?*



$$d = \frac{\bar{X}_G - \bar{X}_B}{\sqrt{\frac{(n_G-1)SD_G^2+(n_B-1)SD_B^2}{n_G+n_B-2}}} \text{ with } v_d = \frac{n_G+n_B}{n_G n_B} + \frac{d^2}{2(n_G+n_B)}$$

$$g = \left(1 - \frac{3}{4(n_G+n_B-2)-1}\right) \cdot d \text{ with } v_g = \left(1 - \frac{3}{4(n_G+n_B-2)-1}\right)^2 \cdot v_d$$

# Two-Stage Individual Participant Data Meta-Analysis
## Stage 2

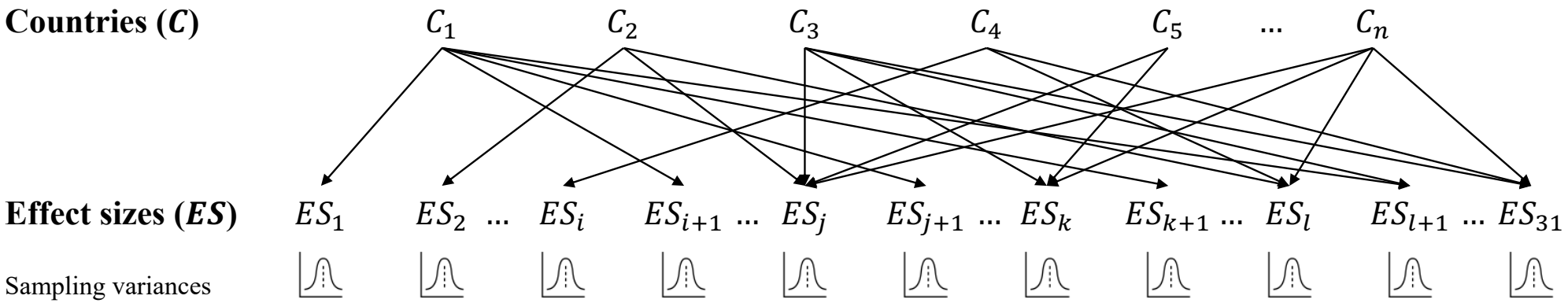
**Three-level random-effects model with effect sizes nested in countries**

Level 1 (sampling variance): $\beta_{jk} = \theta_{jk} + r_{jk}$

Level 2 (within countries): $\theta_{jk} = \lambda_k + q_{jk}$

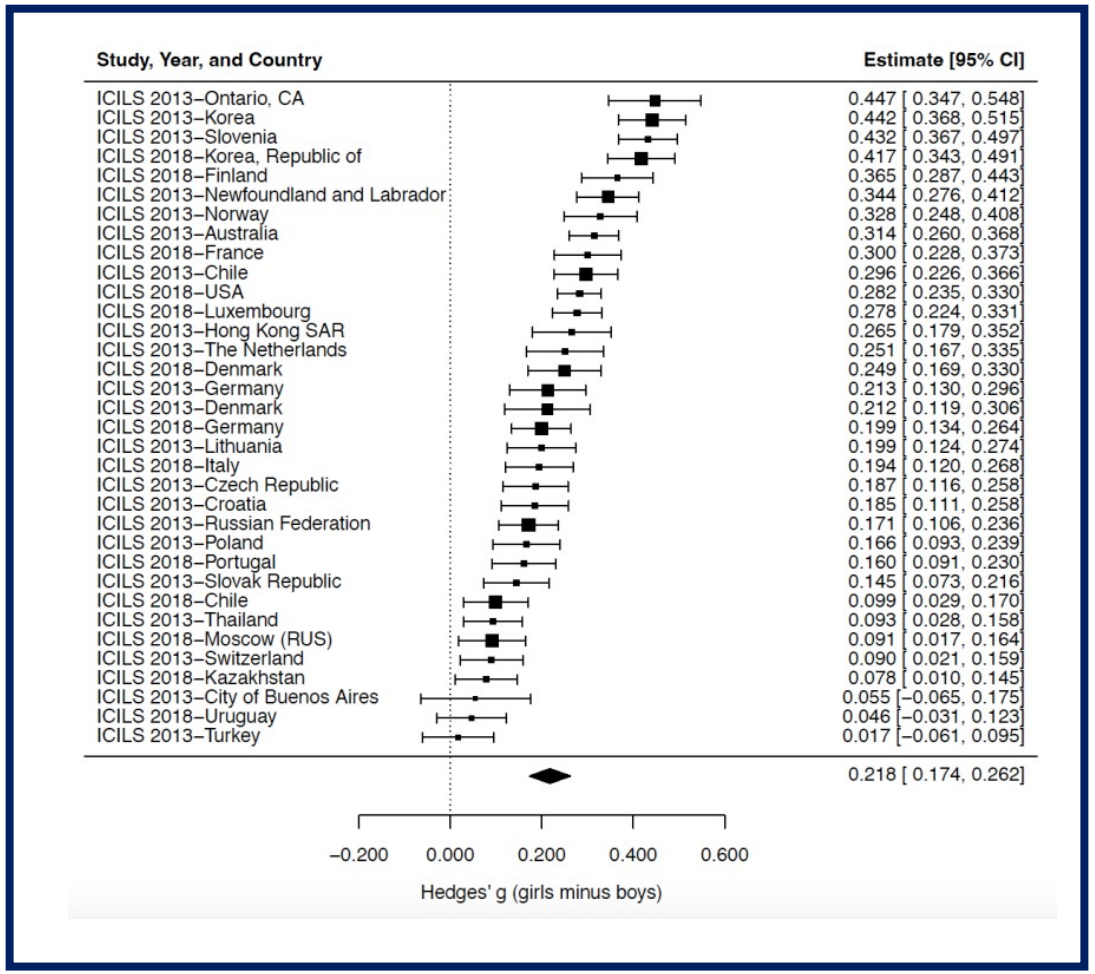
Level 3 (between countries): $\lambda_k = \mu + u_k$

$r_{jk} \sim N(0, \sigma_r^2)$, $q_{jk} \sim N(0, \sigma_q^2)$, $u_k \sim N(0, \sigma_u^2)$



# Two-Stage Individual Participant Data Meta-Analysis

## Individual Participant Data

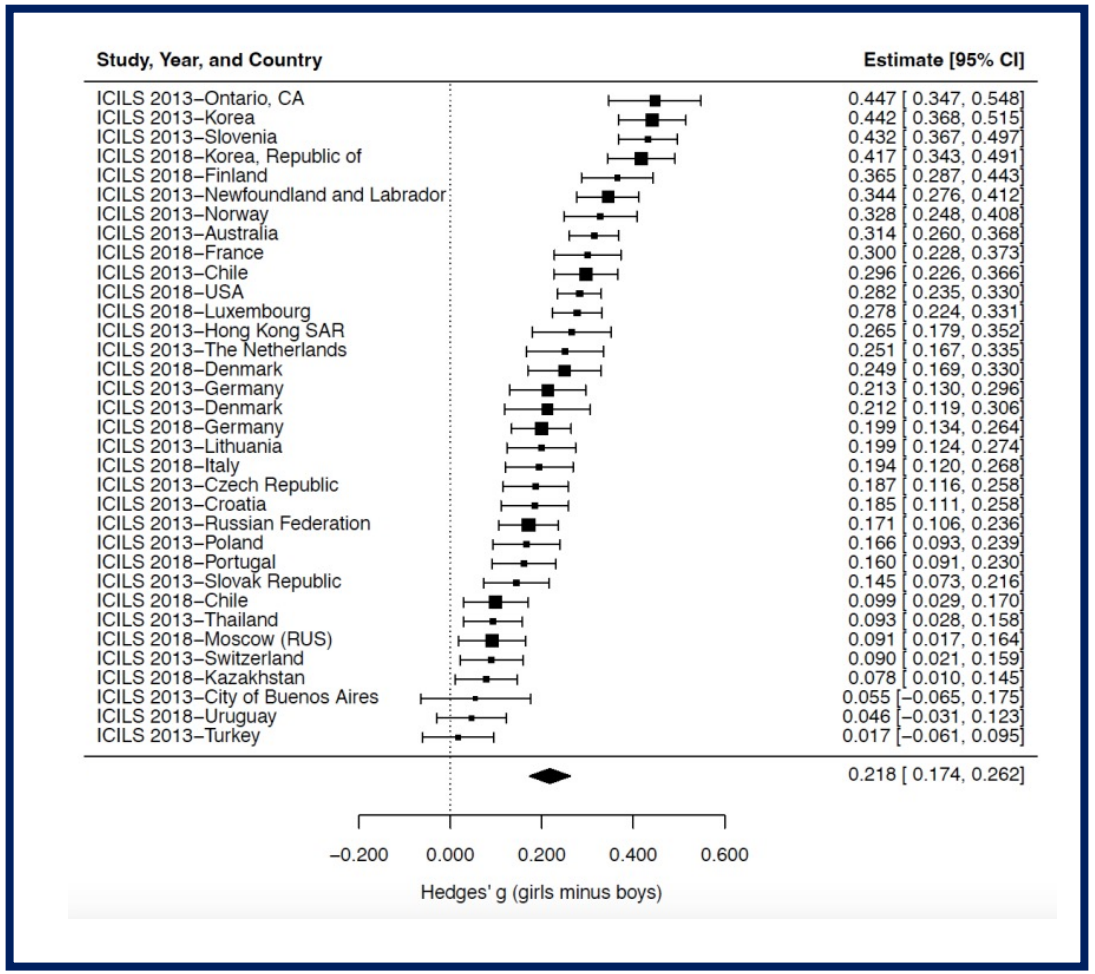



$\bar{g}_{\text{IPD}}$ = 0.218 (95% CI [0.174, 0.262])

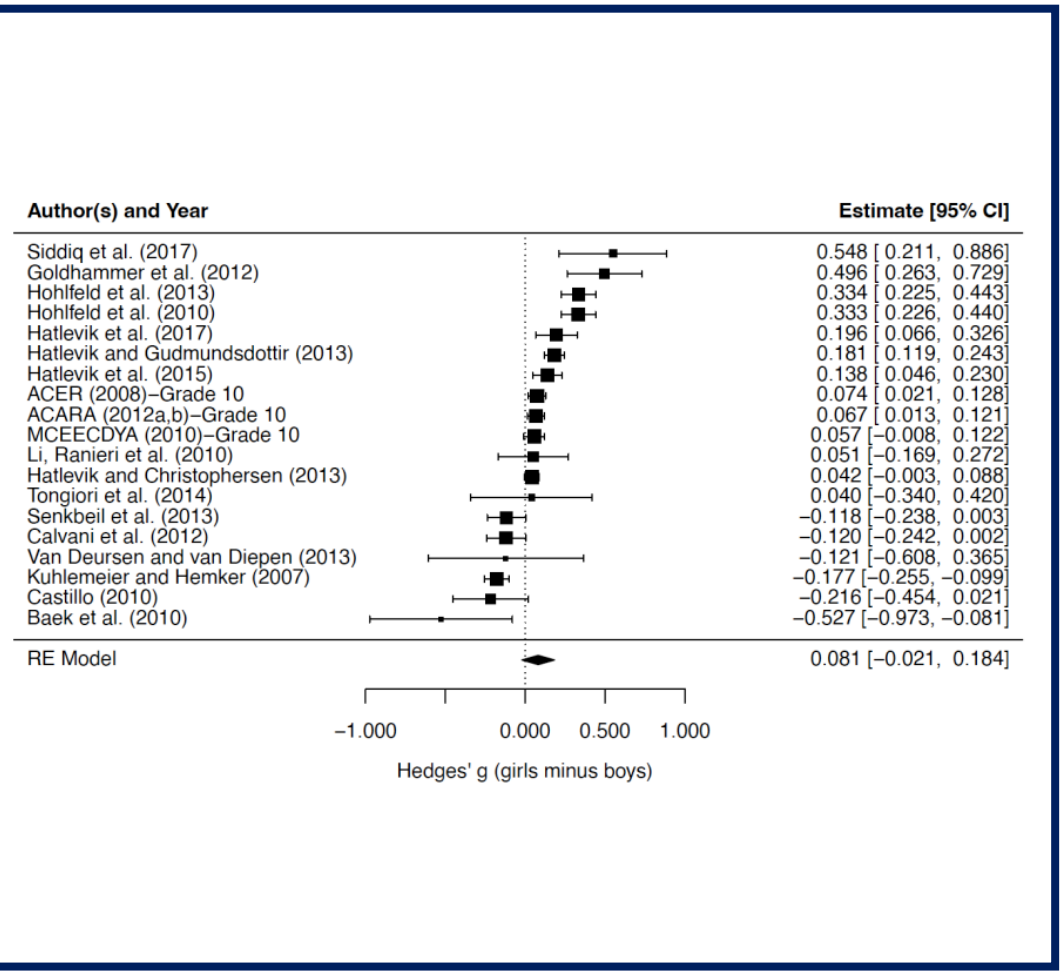
Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

# Two-Stage Individual Participant Data Meta-Analysis

## Individual Participant Data




## Aggregate Data

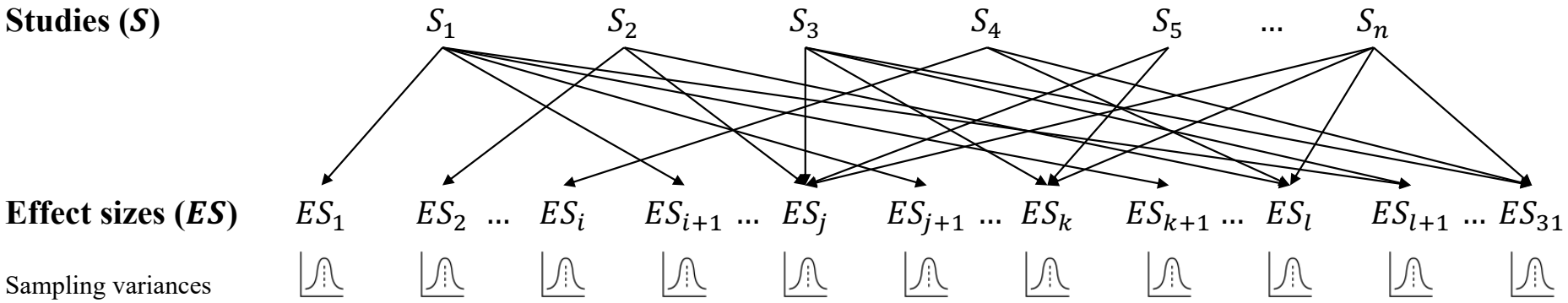



Scherer, Siddiq & Nilsen (2021)
Campos, Cheung & Scherer (2022)

# Two-Stage Individual Participant Data Meta-Analysis

## Combining AD and IPD

**Three-level random-effects model with effect sizes nested in countries/studies**




IPD ? AD

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2 - Moderation Analysis AD vs. IPD

| Baseline model | $\hat{\mu}$ [95 % $CI$] | $B$ [95 % $CI$] | $\tau^2_{IPD}$ [95 % $CI$] | $\tau^2_{AD}$ [95 % $CI$] | $\gamma^2_{IPD}$ | $\gamma^2_{AD}$ |
|---|---|---|---|---|---|---|
| **Model 1:** Data-specific effect sizes, data-specific between-country, and data-specific between-sample residual heterogeneity | .074 [-.003, .177] | .144 [.032, .256] | .003 [.001,.013] | .035 [.011,.093] | .009 [.001, .020] | .002 [.000, .073] |

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2 - Moderation Analysis AD vs. IPD

| Baseline model | $\hat{\mu}$ [95 % $CI$] | $B$ [95 % $CI$] | $\tau^2_{IPD}$ [95 % $CI$] | $\tau^2_{AD}$ [95 % $CI$] | $\gamma^2_{IPD}$ | $\gamma^2_{AD}$ |
|---|---|---|---|---|---|---|
| **Model 1:** Data-specific effect sizes, data-specific between-country, and data-specific between-sample residual heterogeneity | .074 [-.003, .177] | .144 [.032, .256] | .003 [.001,.013] | .035 [.011,.093] | .009 [.001, .020] | .002 [.000, .073] |
| **Model 2:** Data-specific effect sizes, overall between-country heterogeneity, and data-specific between-sample residual heterogeneity | .031 [-.079 .140] | .187 [.080, .295] | .003 [.000 .016] | .035 [.013, .095] | .008 [.000, .020] | |

| | df | AIC | BIC | AICc | logLik | LRT | pval | QE | tau^2 |
|---|---|---|---|---|---|---|---|---|---|
| Full | 6 | -40.5818 | -28.9909 | -38.6728 | 26.2909 | | | 479.8761 | NA |
| Reduced | 5 | -41.3649 | -31.7057 | -40.0315 | 25.6824 | 1.2170 | 0.2700 | 479.8761 | NA |

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2 - Moderation Analysis AD vs. IPD

| Baseline model | $\hat{\mu}$ [95 % $CI$] | $B$ [95 % $CI$] | $\tau^2_{IPD}$ [95 % $CI$] | $\tau^2_{AD}$ [95 % $CI$] | $\gamma^2_{IPD}$ | $\gamma^2_{AD}$ |
|---|---|---|---|---|---|---|
| **Model 1:** Data-specific effect sizes, data-specific between-country, and data-specific between-sample residual heterogeneity | .074 [-.003, .177] | .144 [.032, .256] | .003 [.001,.013] | .035 [.011,.093] | .009 [.001, .020] | .002 [.000, .073] |
| **Model 2:** Data-specific effect sizes, overall between-country heterogeneity, and data-specific between-sample residual heterogeneity | .031 [-.079 .140] | .187 [.080, .295] | .003 [.000 .016] | .035 [.013, .095] | .008 [.000, .020] | |
| **Model 3:** Data-specific effect sizes, overall between-country heterogeneity, and overall between-sample residual heterogeneity | .065 [-.012, .142] | .158 [.070, .245] | .002 [.000, .013] | | .002 [.000, .013] | |

# Two-Stage Individual Participant Data Meta-Analysis

## Example 2

*To what extent is class-average student achievement in mathematics related to individual students' mathematics self-concept after controlling for students' individual mathematics performance in primary school (BFLPE)?*

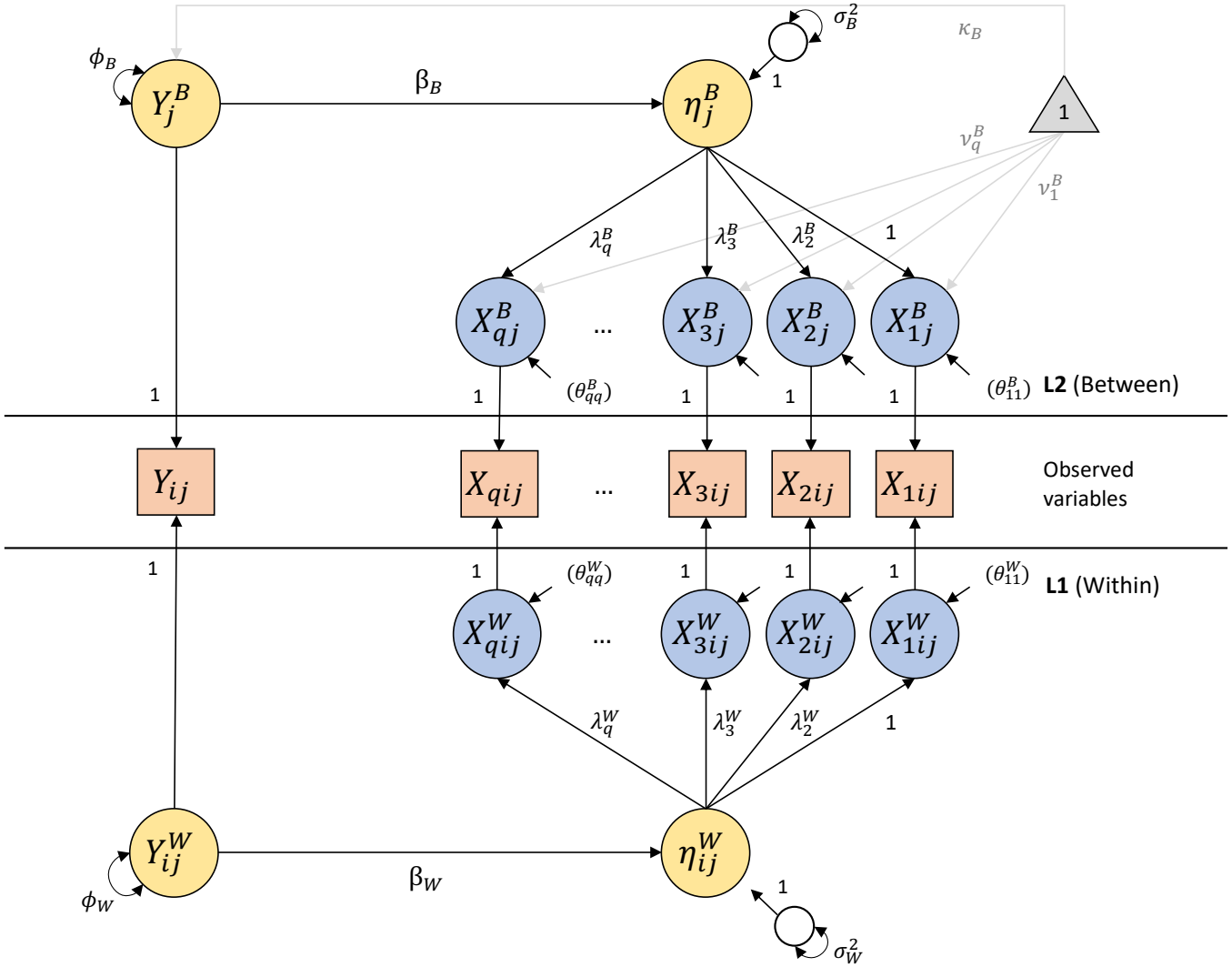

| | |
|---|---|
| Constructs of interest | • Mathematics achievement<br>• Self-concept |
| Complex data structure | • Study ID<br>• Country ID<br>• Classroom ID<br>• Student ID<br>• Student weights<br>• Classroom weights |

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 1

Multilevel SEM to estimate contextual effects

$$ES_{BFLPE} = (\beta_B - \beta_W)$$



# Two-Stage Individual Participant Data Meta-Analysis

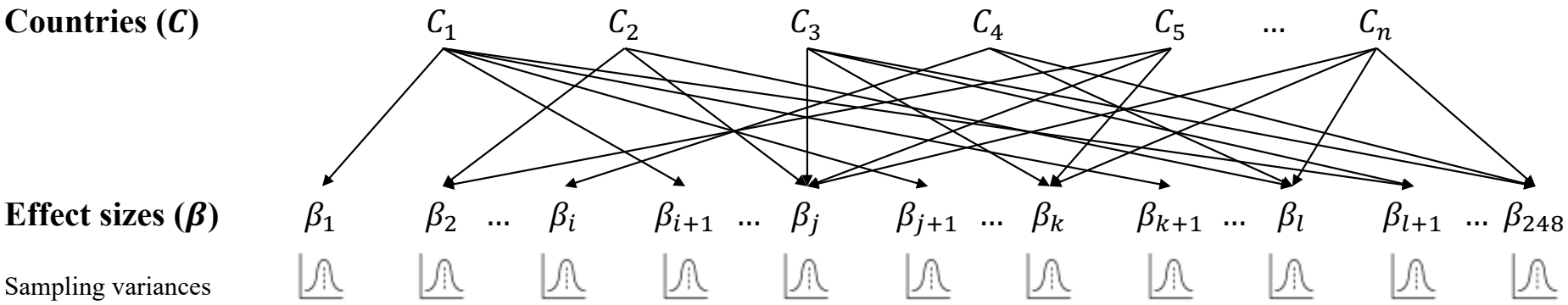
## Stage 2

### Three-level random-effects model with effect sizes nested in countries

Level 1 (sampling variance): $\beta_{jk} = \theta_{jk} + r_{jk}$ $r_{jk} \sim N(0, \sigma_r^2)$, $q_{jk} \sim N(0, \sigma_q^2)$, $u_k \sim N(0, \sigma_u^2)$

Level 2 (within countries): $\theta_{jk} = \lambda_k + q_{jk}$

Level 3 (between countries): $\lambda_k = \mu + u_k$



# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2

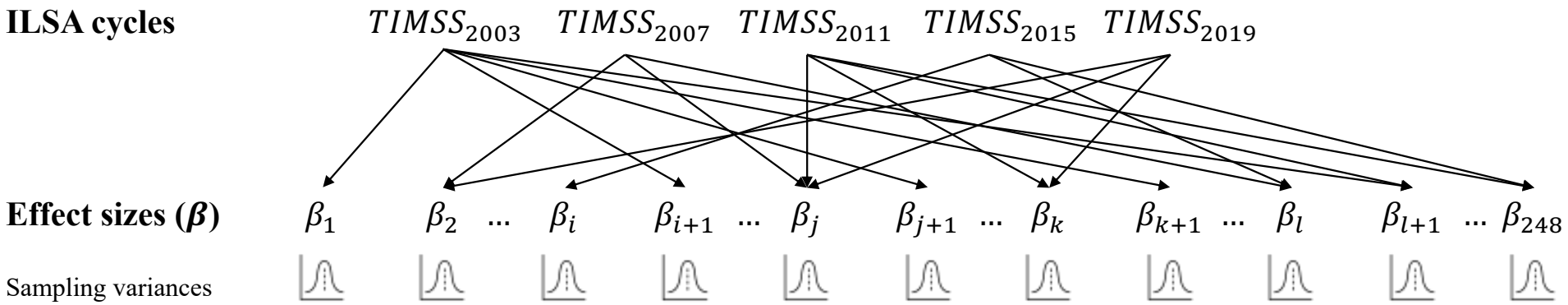
### Three-level random-effects model with effect sizes nested in cycles

Level 1 (sampling variance): $\beta_{jk} = \theta_{jk} + r_{jk}$

Level 2 (within cycles): $\theta_{jk} = \lambda_k + q_{jk}$

Level 3 (between cycles): $\lambda_k = \mu + u_k$

$r_{jk} \sim N(0, \sigma_r^2)$, $q_{jk} \sim N(0, \sigma_q^2)$, $u_k \sim N(0, \sigma_u^2)$



# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2

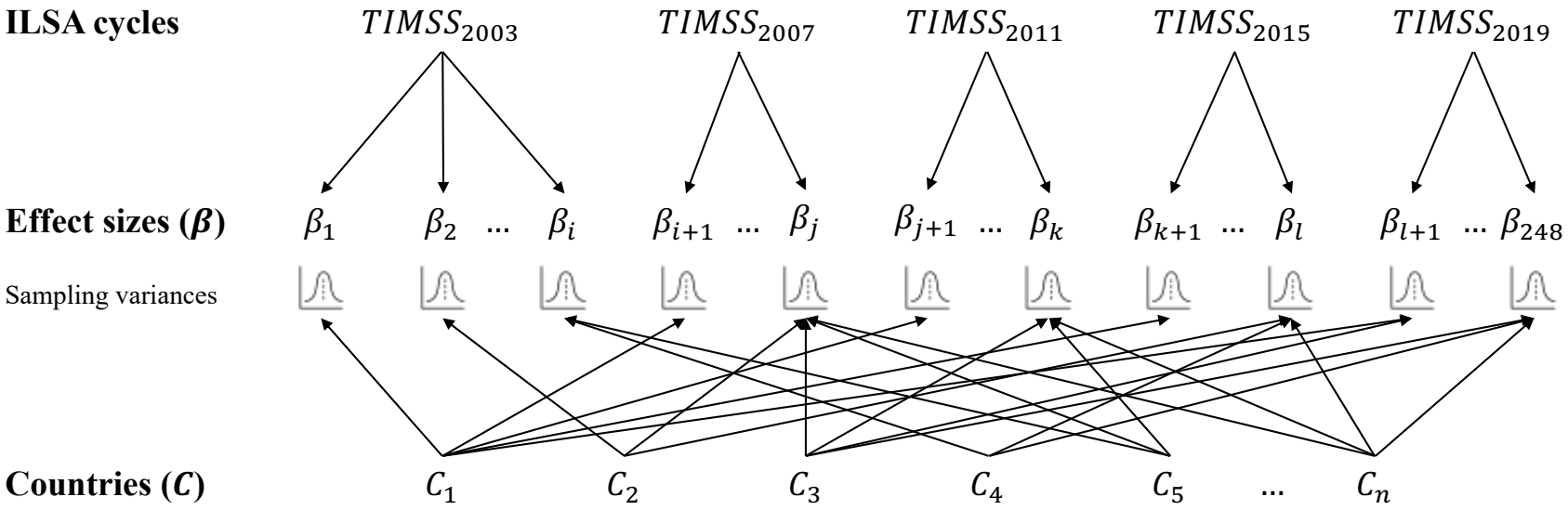
### Four-level cross-classified random-effects model with countries and cycles

Level 1: $\beta_{j(kl)} = \theta_{j(kl)} + r_{j(kl)}$

Level 2: $\theta_{j(kl)} = \lambda_{(kl)} + q_{j(kl)}$

Levels 3 and 4: $\lambda_{(kl)} = \mu + u_k + p_l$

$r_{j(kl)} \sim N(0, \sigma_r^2)$, $q_{j(kl)} \sim N(0, \sigma_q^2)$,
$u_k \sim N(0, \sigma_u^2)$, $p_l \sim N(0, \sigma_p^2)$



# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2

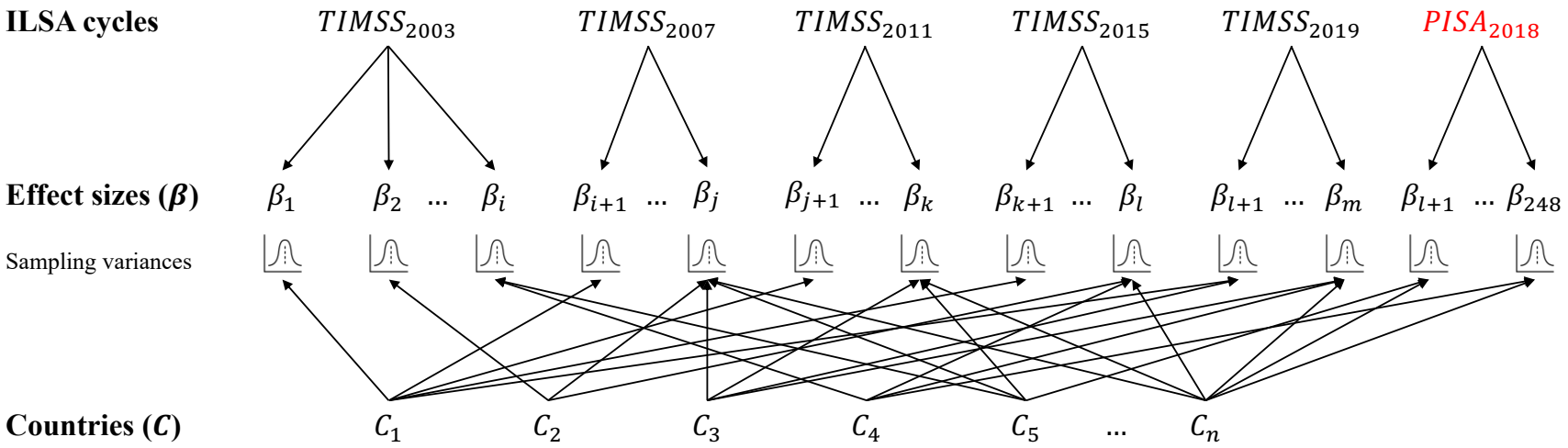
### Four-level cross-classified random-effects model with countries and cycles

Level 1: $\beta_{j(kl)} = \theta_{j(kl)} + r_{j(kl)}$

Level 2: $\theta_{j(kl)} = \lambda_{(kl)} + q_{j(kl)}$

Levels 3 and 4: $\lambda_{(kl)} = \mu + u_k + p_l$

$r_{j(kl)} \sim N(0, \sigma_r^2)$, $q_{j(kl)} \sim N(0, \sigma_q^2)$,
$u_k \sim N(0, \sigma_u^2)$, $p_l \sim N(0, \sigma_p^2)$



# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2

| Baseline model | $\hat{\mu}$ [95 % $CI$] | $\tau^2_{ES}$ [95 % $CI$] | $\sigma^2_C$ [95 % $CI$] | $\sigma^2_S$ [95 % $CI$] | $I^2_{ES}$ | $I^2_C$ | $I^2_s$ |
|---|---|---|---|---|---|---|---|
| **Model 1:** Standard random-effects model | -.458 [-.481, -.436] | .025 [.020, .030] | - | - | 82.4 % | - | - |
| **Model 2:** Three-level random-effects model with effect sizes nested in countries | -.451 [-.486, -.416] | .004 [.002, .006] | .021 [.014, .031] | - | 12.1 % | 70.0 % | - |
| **Model 3:** Three-level random-effects model with effect sizes nested in TIMSS cycles | -.458 [-.484, -.432] | .025 [.020, .031] | - | .000 [.000, .002] | 82.4 % | - | 0.0 % |
| **Model 5:** Four-level cross-classified random-effects model | -.452 [-.489, -.415] | .003 [.001, .006] | .021 [.014, .031] | .000 [.000, .002] | 10.9 % | 70.6 % | 0.7 % |

# Two-Stage Individual Participant Data Meta-Analysis

## Stage 2

| Moderator | Three-level mixed-effects meta-regression | | | |
|---|---|---|---|---|
| | Model 2a | | Model 2b | |
| | *B* | *SE* | *B* | *SE* |
| Intercept | -0.137 | 0.174 | -0.812 | 0.225 |
| *Cultural dimensions* | | | | |
| PDI | -0.002 | 0.002 | - | - |
| IDV | -0.002 | 0.001 | - | - |
| MAS | -0.001 | 0.001 | - | - |
| UAI | -0.001 | 0.001 | - | - |
| LTO | 0.000 | 0.001 | - | - |
| IVR | 0.000 | 0.002 | - | - |
| *Economic development* | | | - | - |
| HDI | - | - | 0.426* | 0.258 |
| *Moderator test* | | | | |
| $Q_M$(df) | 8.6 (6), $p$ = .20 | | 5.5 (1), $p$ = .02 | |
| *Variance explanation* | | | | |
| $R^2_{ES}$ | 0.0% | | 2.5% | |
| $R^2_C$ | 5.5% | | 12.3% | |

# Potential of Two-Stage IPD Meta-Analysis

1. Enlarge the generalizability of meta-analytic conclusions
2. Two-stage IPD meta-analysis enables researchers to synthesize information from complex surveys studies
3. Standardized analyses across studies
4. Direct and model-based generation of the effect sizes of interest
5. Appropriate handling of statistical dependencies in meta-analytic data sets from complex sampling surveys

Campos, D, Cheung, W.-L.M., & Scherer, R. (2022). A Primer on Synthesizing Individual Participant Data Obtained From Complex Sampling Surveys: A Two-Stage IPD Meta-Analysis Approach. (Accepted for Publication: *Psychological Methods*)


Mastodon: diegc_@nerdculture.de

LinkedIn: Diego Campos

Email: diegogc@uio.no

Webpage: www.diegocampos.co


UNIVERSITY OF OSLO

# References


- Brunner, M., Keller, L., Stallasch, S. E., Kretschmann, J., Hasl, A., Preckel, F., Lüdtke, O., & Hedges, L. V. (2022). Meta-Analyzing Individual Participant Data From Studies With Complex Survey Designs: A Tutorial on Using the Two-Stage Approach for Data From Educational Large-Scale Assessments. *Research Synthesis Methods*, *n/a*(n/a). https://doi.org/10.1002/jrsm.1584
- Cooper, H., & Patall, E. A. (2009). The relative benefits of meta-analysis conducted with individual participant data versus aggregated data. *Psychological Methods*, *14*(2), 165. https://doi.org/10.1037/a0015565
- Curran, P. J., & Hussong, A. M. (2009). Integrative Data Analysis: The Simultaneous Analysis of Multiple Data Sets. *Psychological Methods*, *14*(2), 81–100. https://doi.org/10.1037/a0015914
- Cheung, M. W. L., & Jak, S. (2016). Analyzing big data in psychology: A split/analyze/metaanalyze approach. *Frontiers in Psychology, 7(MAY)*, 1–13. https://doi.org/10.3389/fpsyg.2016.00738
- Fernández-Castilla, B., Maes, M., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2019). A demonstration and evaluation of the use of cross classified random-effects models for meta-analysis. *Behavior Research Methods, 51*(3), 1286–1304. https://doi.org/10.3758/s13428-018-1063-2
- Marsh, H. W., Lüdtke, O., Robitzsch, A., Trautwein, U., Asparouhov, T., Muthén, B., & Nagengast, B. (2009). Doubly-Latent Models of School Contextual Effects: Integrating Multilevel and Structural Equation Approaches to Control Measurement and Sampling Error. *Multivariate Behavioral Research*, *44*(6), 764–802.
- Pustejovsky, J. E., & Tipton, E. (2022). Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prevention Science*, *23*, 425–438. https://doi.org/10.1007/s11121-021-01246-3
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International Large-Scale Assessment Data: Issues in Secondary Analysis and Reporting. *Educational Researcher, 39*(2), 142–151. https://doi.org/10.3102/0013189X10363170
- Scherer, R., Siddiq, F., & Nilsen, T. (2021). *The Potential of International Large-Scale Assessments for Meta-Analyses in Education*. PsyArXiv. https://doi.org/10.31234/osf.io/bucf9
- Stapleton, L. M. (2013). Incorporating sampling weights into single-and multi-level models. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 363–388). Chapman and Hall/CRC. https://doi.org/10.1201/b16061