UiO : CREATE – Centre for Research on Equality in Education
University of Oslo

Norwegian Centre of Excellence

Bringing it all together—

How International Large-Scale Assessments Can Inform Meta-Analyses in Education

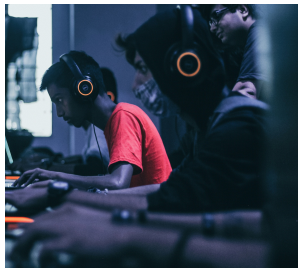
Ronny Scherer, CEMO & CREATE

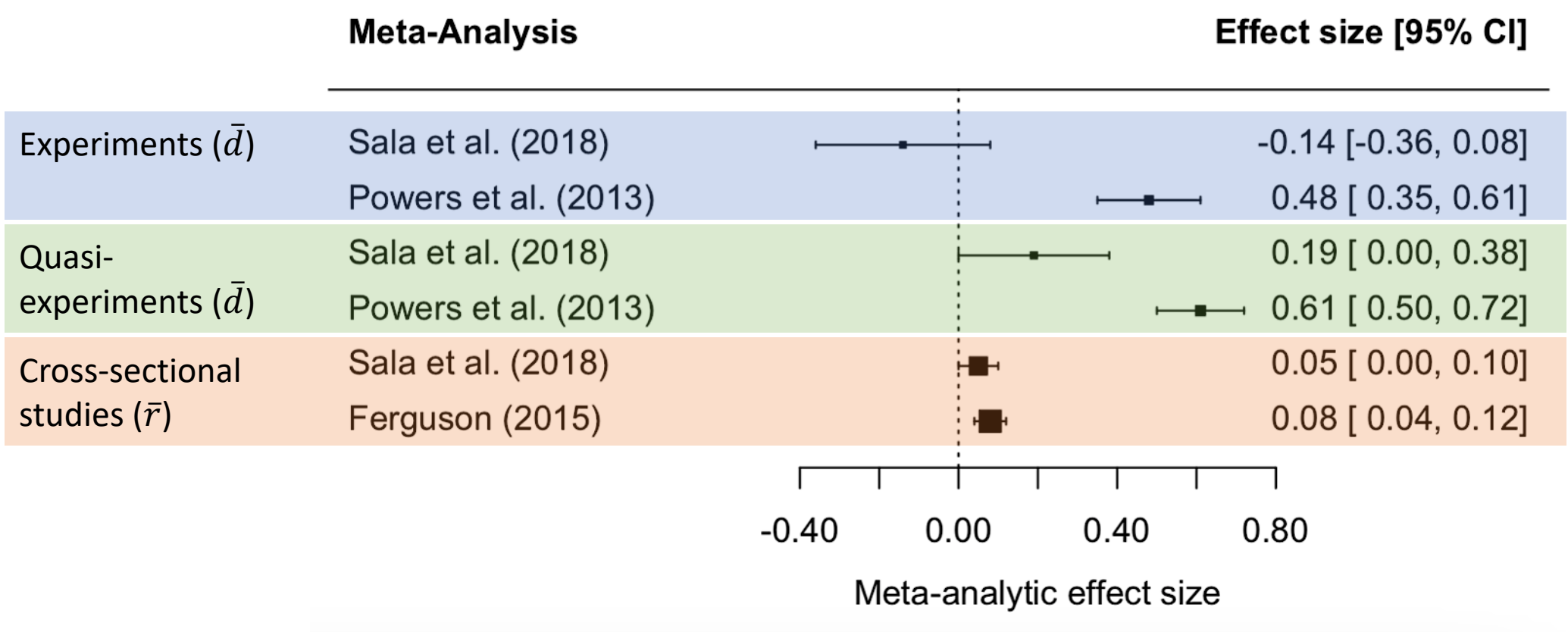
QRM Conference
Gothenburg, 11 June 2024

Video Gaming and Academic Performance

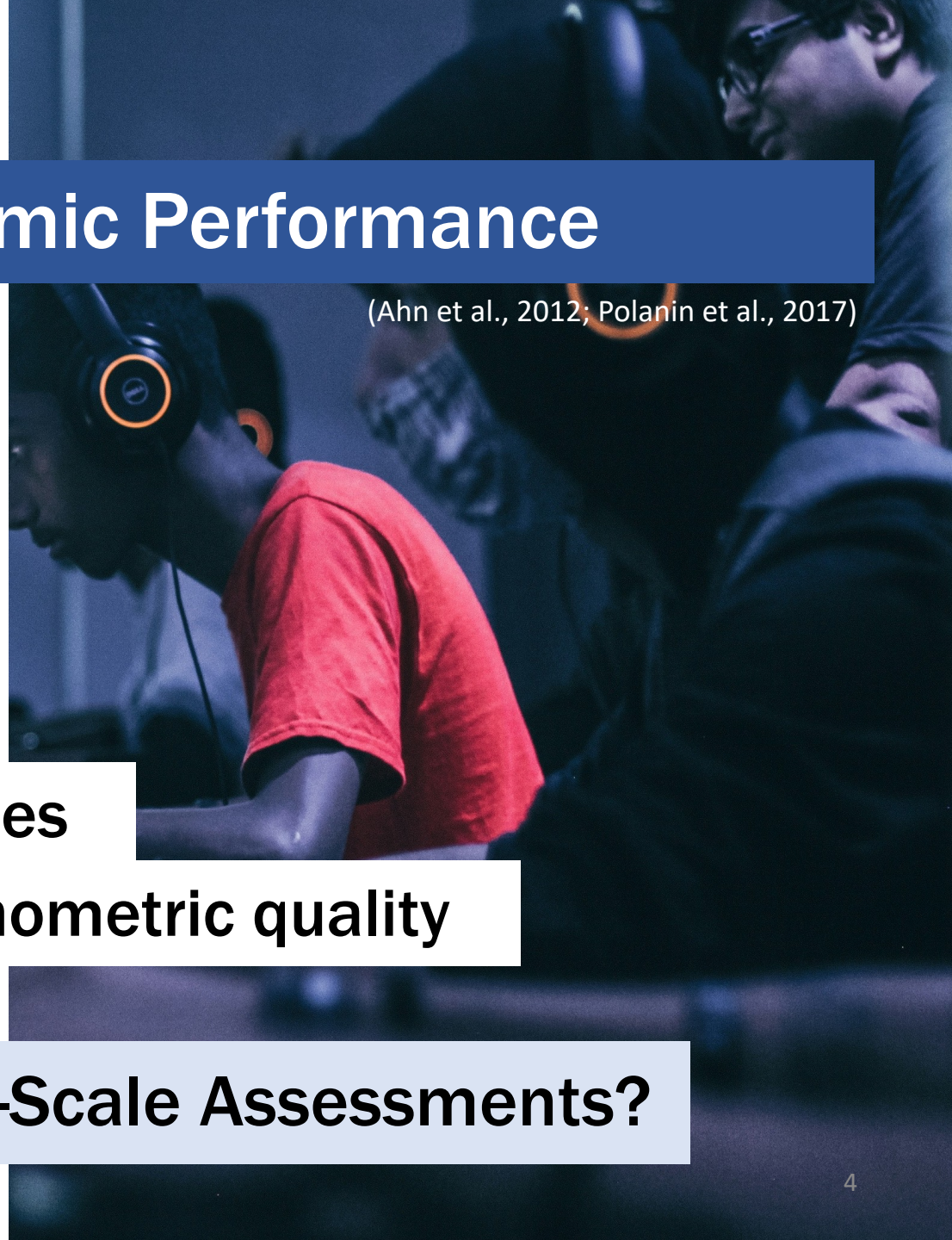
Video Gaming and Academic Performance

What Some Meta-Analyses Reveal





Video Gaming and Academic Performance



(Ahn et al., 2012; Polanin et al., 2017)

Some Meta-Analyses Differ

- Search, screening, and coding of the study data
- Meta-analytic modelling approach
- Features and quality of the included studies

Language/culture selection bias

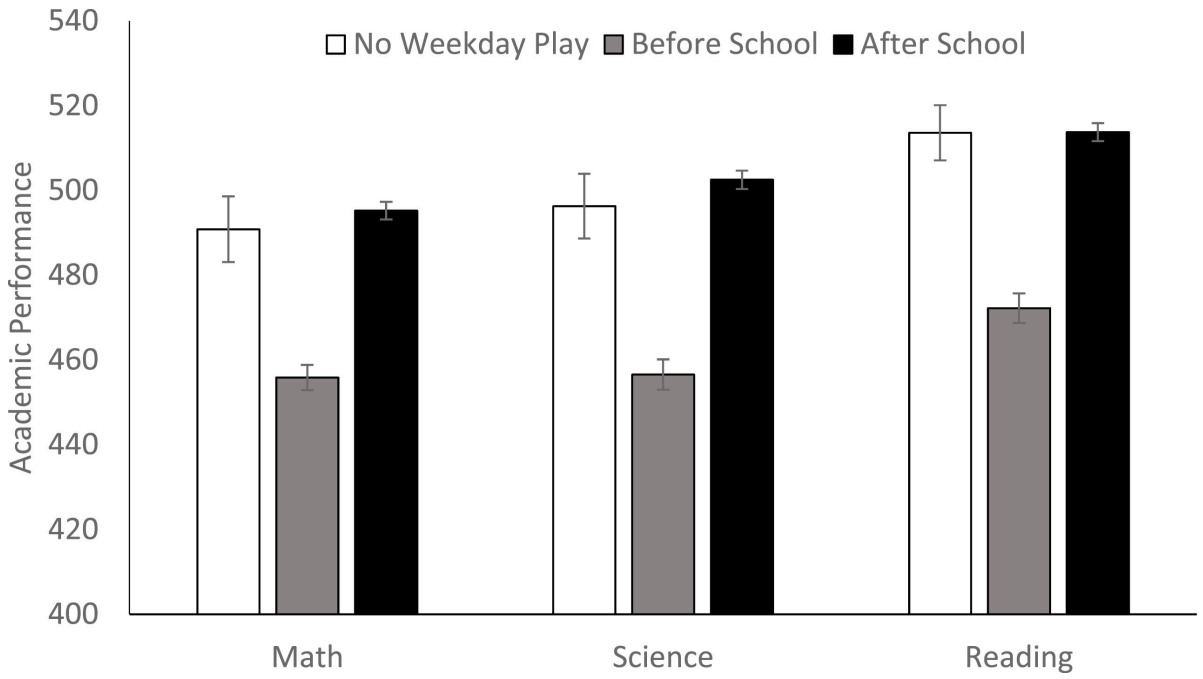
Small and non-representative samples

Measures with insufficient psychometric quality

Opportunity for International Large-Scale Assessments?

Video Gaming and Academic Performance

Drummond and Sauer (2020) used PISA 2015 data to clarify some differences.



(Source: Drummond & Sauer, 2020, p. 9)

Meta-Analyses and ILSAs as Evidence Sources

Benefits, Challenges, and Procedures of Meta-Analysing ILSA Data

1. More complex meta-analytic models
2. Improved generalizability and robustness
3. Better quality and reduced bias of meta-analytic evidence

International Large-Scale Assessments (ILSAs)

Key features

- Empirical & quantitative studies
- Large, international, representative samples
- Complex survey sampling design

OECD
BETTER POLICIES FOR BETTER LIVES

PISA
TALIS
PIAAC

IEA
Researching education, improving learning

TIMSS
PIRLS
ICILS

Meta-Analyses in Education

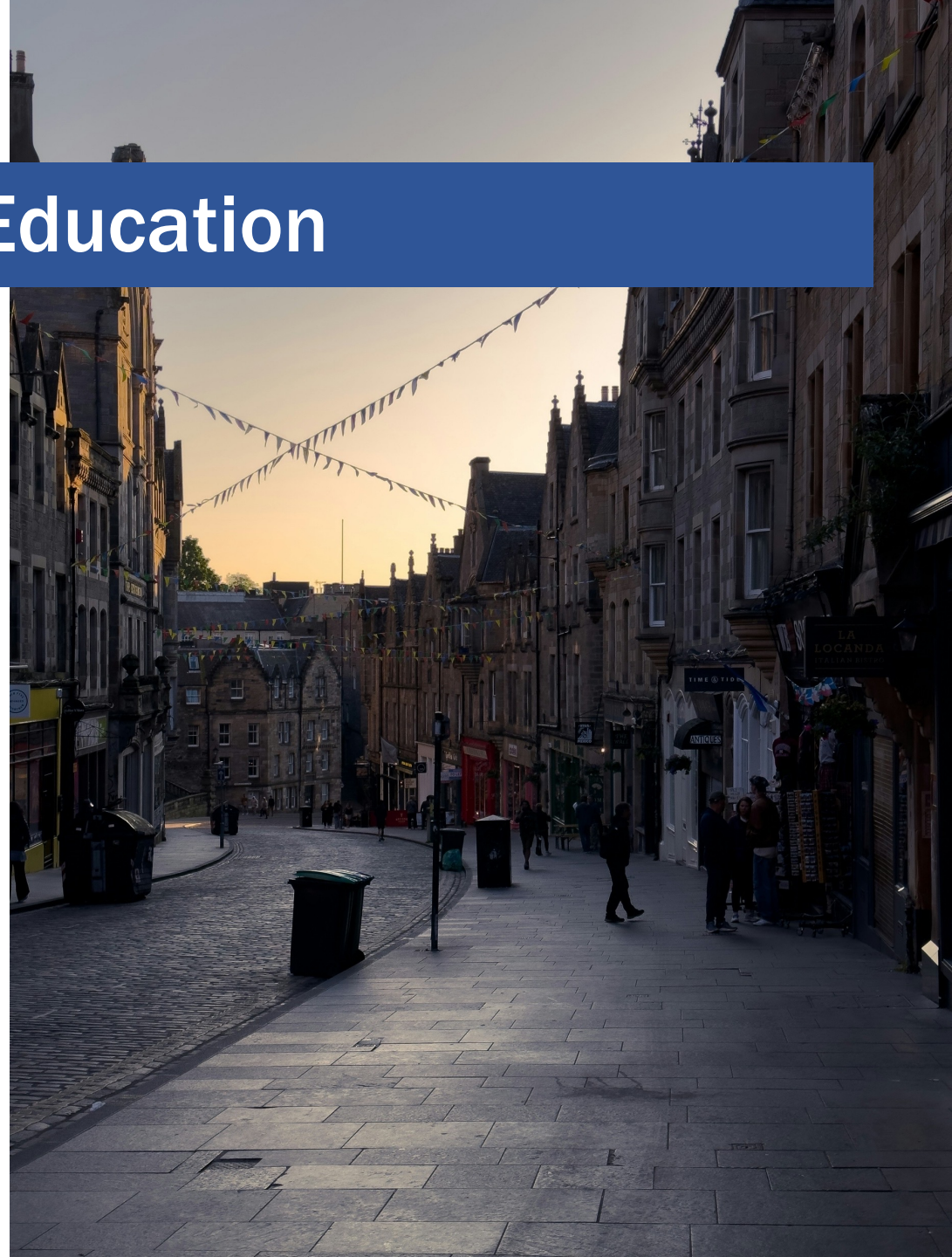
"Quantitative technique for synthesizing the results of multiple studies of a phenomenon into a single result."

(APA, 2018, p. 1)

Typical effect sizes

- Intervention effects
- Relations among constructs
- Group differences

https://dictionary.apa.org/meta-analysis



Meta-Analyses in Education

(Borenstein et al., 2009; Card, 2012)

Typical meta-analysis

m effect sizes
m studies

Study 1 θ_1 v_1
Study 2 θ_2 v_2
… … …
Study m θ_m v_m

Pooling & Weighting

$\bar{\theta}$

Weighted average effect size

τ^2

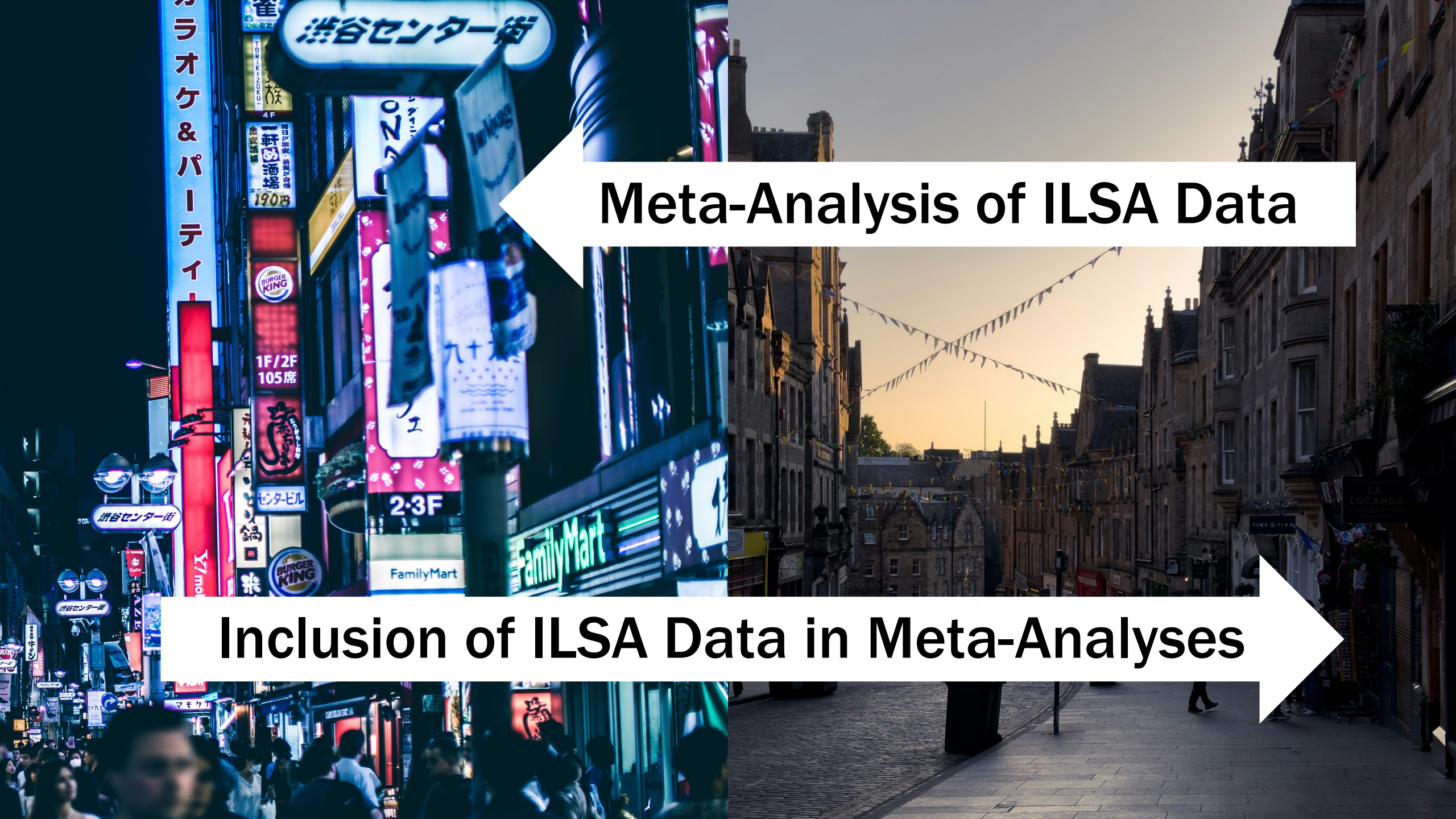
Between-study heterogeneity

Sampling variances v_i and possible covariances

x_i

Effect sizes θ_i
(e.g., correlations, standardized mean differences, regression coefficients)

Moderators
(e.g., study, sample, measurement, country characteristics)


Meta-Analysis of ILSA Data
Inclusion of ILSA Data in Meta-Analyses

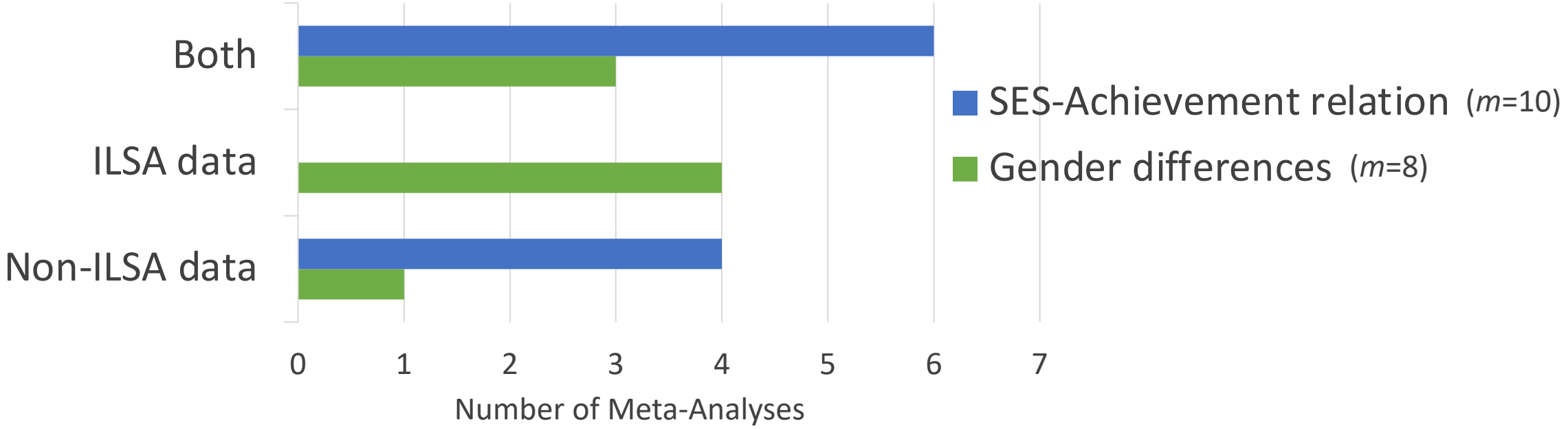
Status of Including ILSAs in Meta-Analyses

(Scherer, Siddiq, & Nilsen, 2024)

ILSAs cannot address all possible research questions in education.

Rapid systematic review

Meta-analyses on the relation between SES and student achievement and gender differences in achievement until 2022

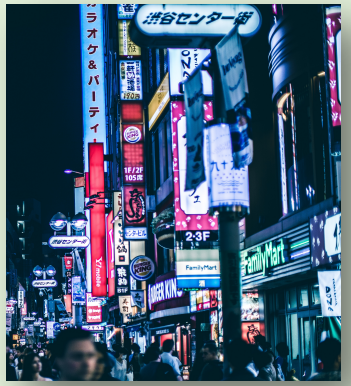


Utilizing ILSAs in Meta-Analyses

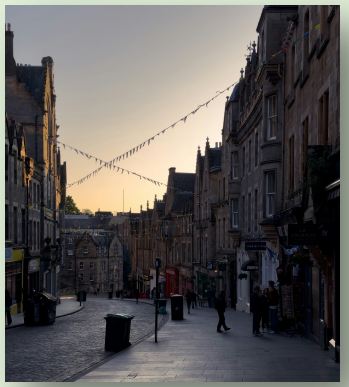
(Scherer, Siddiq, & Nilsen, 2024)

Several approaches

Meta-analysis of ILSA data

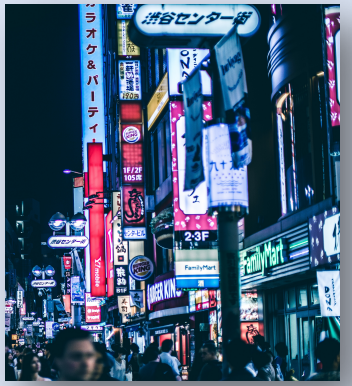


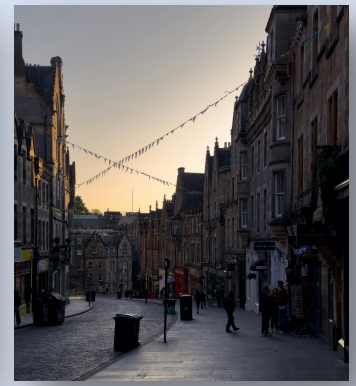
Meta-analysis of non-ILSA data



Separate Meta-Analyses

Meta-analysis of ILSA *and* non-ILSA data





Joint Meta-Analysis

Utilizing ILSAs in Meta-Analyses

(Campos et al., 2023; Brunner et al., 2023)

Two analytic stages

Weighting, stratification, multilevel structure, PVs, missing data

IPD

Extraction
Generation
Conversion

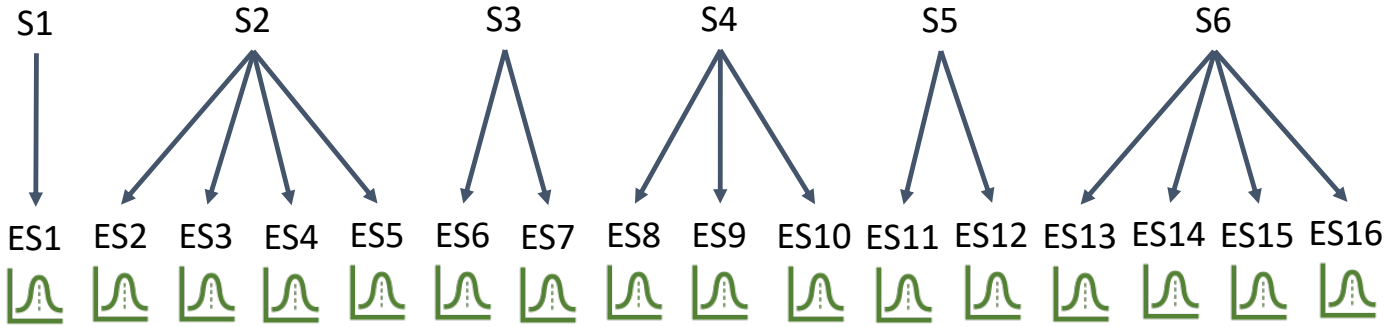
Effect sizes and sampling (co-)variances

Meta-analytic synthesis of effect sizes

Pooled effects
Heterogeneity
Moderators

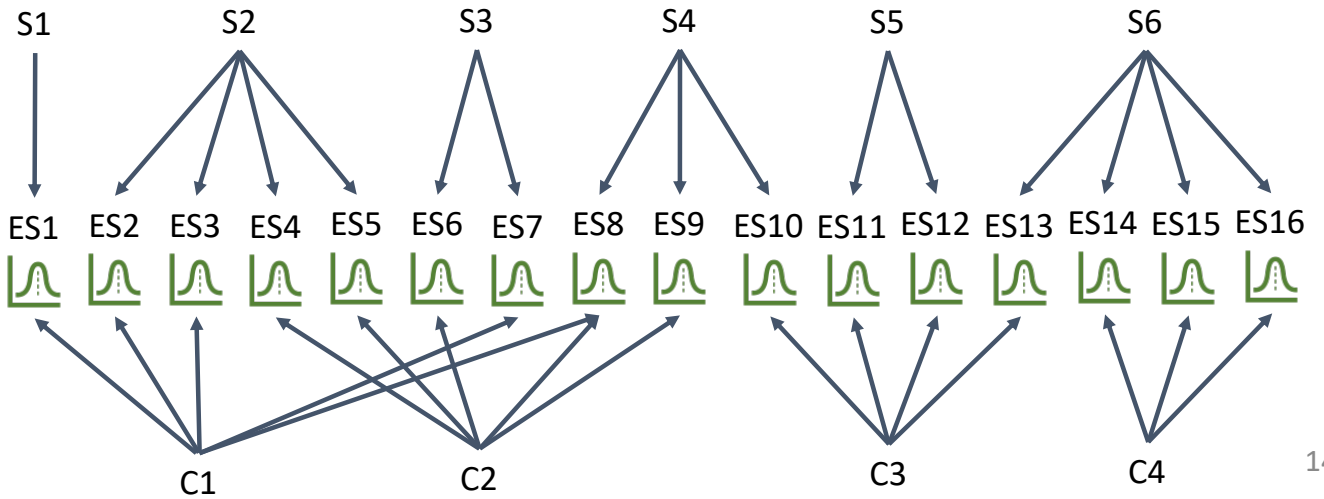
Effect Size Multiplicity

Multiple effect sizes per study (e.g., multiple samples or measures)

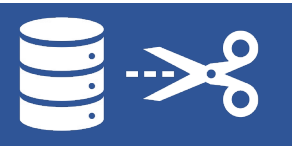


Multiple effect sizes nested in primary studies

Multiple effect sizes nested in primary studies and countries



Effect Size Multiplicity



(Cheung, 2015; Campos et al., 2023; Pustejovsky & Tipton, 2021)

- Random-Effects Meta-Analysis with Robust Variance Estimation

- Multilevel Random-Effects Meta-Analysis

i: Effect sizes, j: Studies

$$\theta_{ij} = \beta_R + u_{(2)ij} + u_{(3)j} + e_{ij}$$

$$e_{ij} \sim N(0, v_{ij}),\ u_{(2)ij} \sim N(0, \tau^2_{(2)}),\ u_{(3)j} \sim N(0, \tau^2_{(3)})$$

- Other multilevel and/or multivariate working models

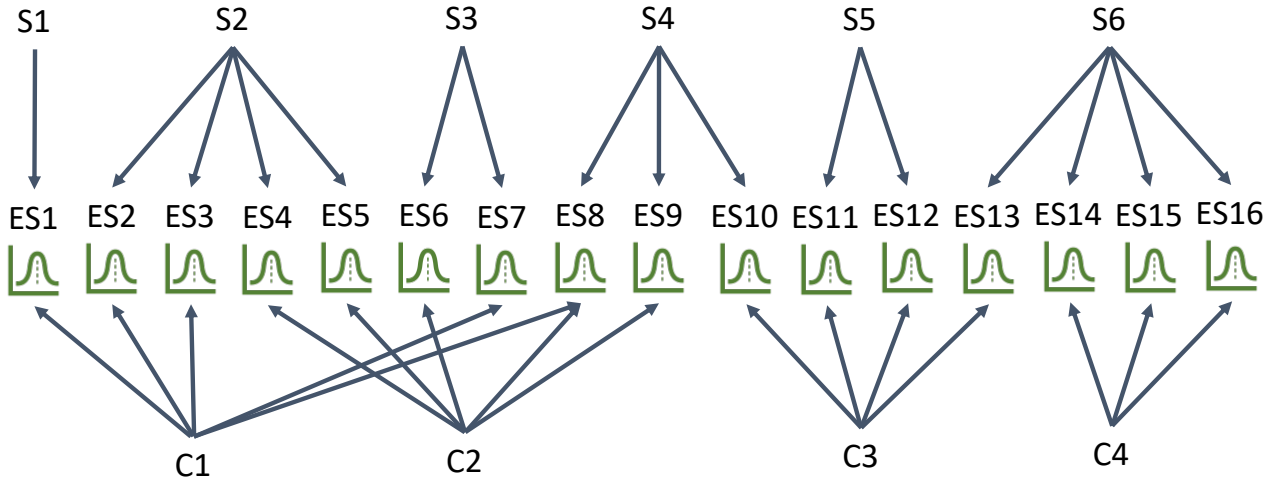
Effect Size Multiplicity

(Campos et al., 2023; Scherer et al., 2024)

Cross-Classified Random-Effects Meta-Analysis

$$\theta_{i(jk)} = \beta_R + u_{(2)ij} + u_{(3)j} + u_{(3)k} + e_{i(jk)}$$

$$e_{i(jk)} \sim N(0, v_{i(jk)}),\ u_{(2)ij} \sim N(0, \tau^2_{(2)}),\ u_{(3)j} \sim N(0, \tau^2_{(3a)}),\ u_{(3)k} \sim N(0, \tau^2_{(3b)})$$



Hierarchical *and* non-hierarchical data structure with two independent upper levels

i: Effect sizes, j: Studies, k: Countries

Meta-Analysis of ILSA Data

Psychological Bulletin
2010, Vol. 136, No. 1, 103–127

0033-2909/10/$12.00 DOI: 10.1037/a0018053

Gender differences in mathematics achievement

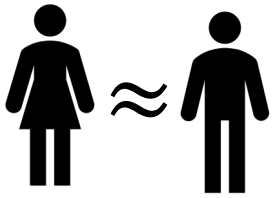
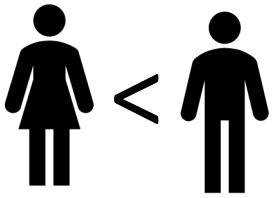
TIMSS & PISA 2003, 69 countries

Cross-National Patterns of Gender Differences in Mathematics: A Meta-Analysis

Nicole M. Else-Quest
Villanova University

Janet Shibley Hyde
University of Wisconsin—Madison

Marcia C. Linn
University of California, Berkeley

TIMSS 2003	PISA 2003
	
$\bar{d}$ = -0.01	$\bar{d}$ = 0.11
95% *CI* [-0.05, 0.03]	95% *CI* [0.09, 0.13]

Random-effects meta-analysis of standardized mean differences treating countries as “studies”.

Effect size: Cohen’s d

Meta-Analysis of ILSA Data

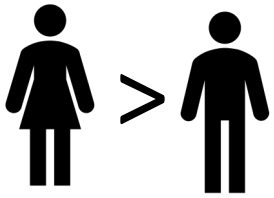
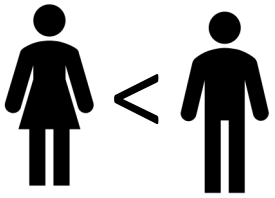
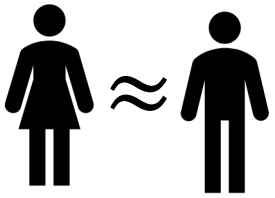
Gender differences in student achievement

PISA 2000-2015, 82 countries

AMERICAN PSYCHOLOGICAL ASSOCIATION
Journal of Educational Psychology
© 2021 American Psychological Association
ISSN: 0022-0663
2022, Vol. 114, No. 5, 966–991
https://doi.org/10.1037/edu0000685

Top-Performing Math Students in 82 Countries: An Integrative Data Analysis of Gender Differences in Achievement, Achievement Profiles, and Achievement Motivation

Lena Keller[1, 2], Franzis Preckel[3], Jacquelynne S. Eccles[4, 5], and Martin Brunner[1]

Reading	Mathematics	Science
		
$\bar{d}$ = -0.23	$\bar{d}$ = 0.05	$\bar{d}$ = 0.01
95% *CI* [-0.21, -0.25]	95% *CI* [0.03, 0.06]	95% *CI* [-0.01, 0.02]

Integrative data analysis with the same analytic protocol to generate model-based effect sizes.

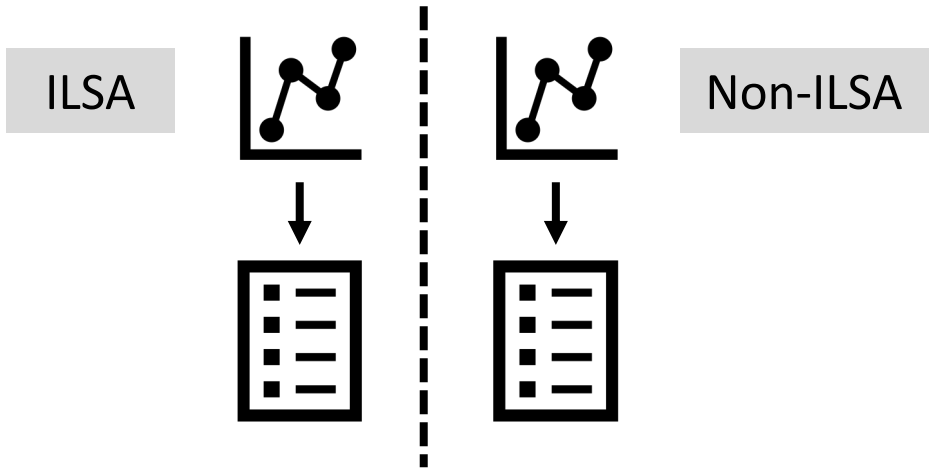
Effect size: Cohen's d

Separate Meta-Analyses of ILSA & Non-ILSA Data

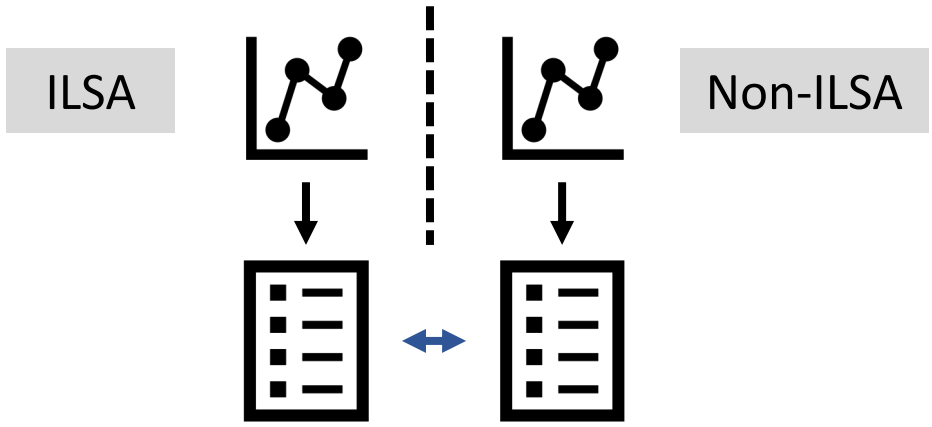
(Scherer, Siddiq, & Nilsen, 2024; Harrer et al., 2021)

Two approaches

Separate Independent Meta-Analyses



Separate Meta-Analyses Informing Each Other



Bayesian meta-analysis with the respective priors

Separate Meta-Analyses of ILSA & Non-ILSA Data

Relation between SES and student achievement

Non-ILSA data, 326 studies

All PISA, TIMSS, and PIRLS cycles up to 2019

Educational Psychology Review (2022) 34:2867–2896
https://doi.org/10.1007/s10648-022-09689-y

REVIEW ARTICLE

Check for updates

Socioeconomic Status and Academic Achievement in Primary and Secondary Education: a Meta-analytic Review

Juan Liu[1] · Peng Peng[2] · Baobao Zhao[1] · Liang Luo[3]

Non-ILSA data	ILSA data
$\bar{r}$ = 0.22	$\bar{r}$ = 0.28
95% *CI* [0.20, 0.23]	95% *CI* [0.28, 0.29]

Large-scale meta-analysis with small differences in the pooled effect sizes and precisions

Effect size: Pearson correlation r

Separate Meta-Analyses of ILSA & Non-ILSA Data

Relation between SES and student achievement

Educational Psychology Review (2022) 34:2867–2896
https://doi.org/10.1007/s10648-022-09689-y

REVIEW ARTICLE

Socioeconomic Status and Academic Achievement in Primary and Secondary Education: a Meta-analytic Review

Juan Liu[1] · Peng Peng[2] · Baobao Zhao[1] · Liang Luo[3]

Non-ILSA data

Table 3 Moderation effects on the correlations between SES and academic achievement for cross-sectional effect sizes

Measure	β	se	t	95% CI	p
Years					
Years2 (2000–2014) vs. Years1 (1990–1999)	**.05**	**.02**	**2.05**	**[.001, .10]**	**.04**
Years3 (2015–2021) vs. Years1 (1990–1999)	.06	.04	1.42	[−.02, .14]	.16
Years3 (2015–2021) vs. Years2 (2000–2014)	.01	.04	.21	[−.06, .08]	.84
GDP per capita	.02	.01	1.36	[−.01, .04]	.18
Net enrollment ratio	.01	.01	1.38	[−.01, .04]	.22
Duration of compulsory education	.01	.01	.51	[−.02, .03]	.61
GINI coefficient	.01	.01	.80	[−.02, .03]	.43
Grade level					
Secondary education vs. primary education	−.02	.02	−.67	[−.06, .03]	.51
Measurement of SES					
Education vs. occupation	−.03	.03	−.98	[−.09, .03]	.34
Family income vs. occupation	**−.09**	**.04**	**−2.19**	**[−.18, −.01]**	**.03**
Resources vs. occupation	−.01	.04	−.30	[−.10, .07]	.77
Composite SES index vs. occupation	−.02	.03	−.80	[−.09, .04]	.43
Education vs. family income	**.06**	**.03**	**2.01**	**[.00, .13]**	**.05**
Resources vs. family income	.08	.04	1.92	[−.01, .17]	.07
Composite SES index vs. family income	**.07**	**.03**	**2.14**	**[−.004, .13]**	**.04**
Education vs. Resources	−.02	.03	−.52	[−.09, .05]	.61
Composite SES index vs. Resources	−.01	.03	−.35	[−.08, .06]	.73
Composite SES index vs. education	.01	.02	.25	[−.04, .05]	.80
Subject of academic achievement					
STEM vs. language	.04	.02	1.96	[−.001, .09]	.05
General achievement vs. language	.03	.03	1.17	[−.02, .08]	.24
General achievement vs. STEM	−.01	.03	−.44	[−.07, .04]	.66

ILSA data

Table 4 Moderation effects on the correlations between SES and academic achievement for international large-scale assessments

Correlation	β	SE	t	95% CI	p
Years					
Years2 (2000–2014) vs. Years1 (1990–1999)	**.04**	**.01**	**5.48**	**[.03, .06]**	**<.001**
Years3 (2015–2021) vs. Years1 (1990–1999)	**.06**	**.01**	**7.72**	**[.05, .08]**	**<.001**
Years3 (2015–2021) vs. Years2 (2000–2014)	**.02**	**.01**	**3.65**	**[.01, .03]**	**<.001**
GDP per capita	−.003	.003	-.91	[−.01, .004]	.36
Net enrollment ratio	**.01**	**.004**	**2.78**	**[.003, .02]**	**.01**
Duration of compulsory education	**.02**	**.003**	**7.46**	**[.02, .03]**	**<.001**
GINI coefficient	−.001	.004	−.34	[−.01, .01]	.73
Grade level					
Secondary education vs. primary education	**.02**	**.01**	**2.32**	**[.002, .03]**	**.02**
Measurement of SES					
Education vs. occupation	**−.02**	**.003**	**−6.40**	**[−.03,−.01]**	**<.001**
Family income vs. occupation	.01	.01	1.06	[−.01, .03]	.29
Resources vs. occupation	**.01**	**.004**	**2.74**	**[.003, .02]**	**.01**
Composite SES index vs. occupation	**.06**	**.003**	**23.93**	**[.06, .07]**	**<.001**
Education vs. family income	**−.03**	**.01**	**−3.01**	**[−.05,−.01]**	**.003**
Resources vs. family income	−.001	.01	−.06	[−.02, .02]	.95
Composite SES index vs. family income	**.05**	**.01**	**5.27**	**[.03, .07]**	**<.001**
Education vs. resources	**−.03**	**.003**	**−9.17**	**[−.04,−.02]**	**<.001**
Composite SES index vs. resources	**.05**	**.004**	**12.48**	**[.05, .06]**	**<.001**
Composite SES index vs. education	**.08**	**.004**	**21.74**	**[.08, .09]**	**<.001**
Subject of academic achievement					
STEM vs. language	−.01	.005	−1.20	[−.02, .004]	.23

Differences in heterogeneity and moderator effects

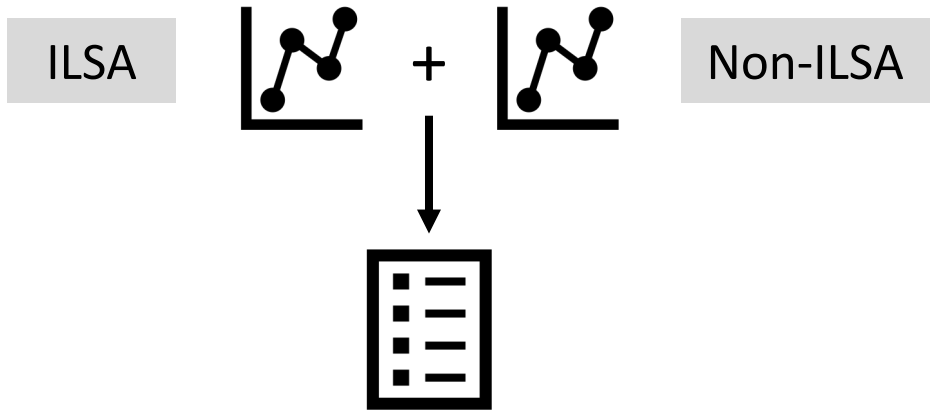
For ILSA data:
Moderation by year of assessment, duration of compulsory education, and the types of SES measures

Joint Meta-Analysis of ILSA and non-ILSA Data

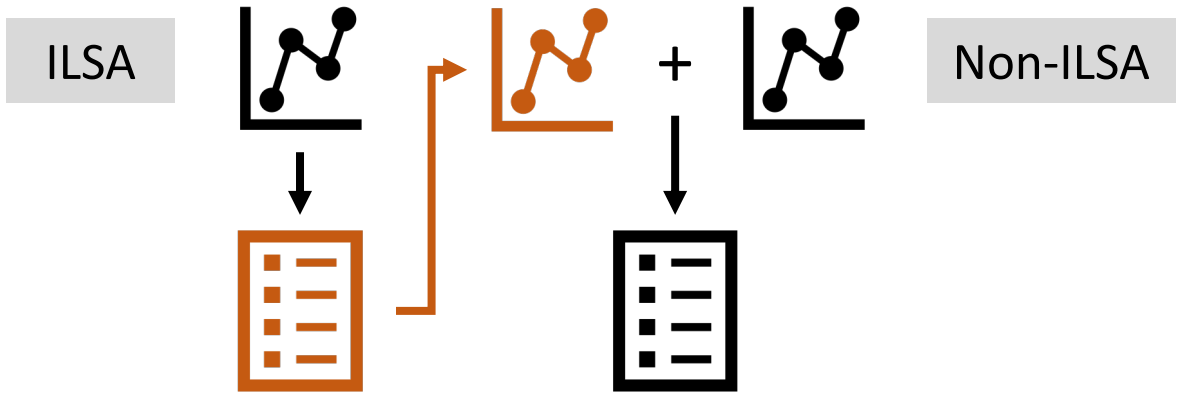
(Scherer, Siddiq, & Nilsen, 2024)

Two approaches

Direct inclusion of ILSA data in **one step**



Direct inclusion of ILSA data in **multiple steps**



Pooled effect size and sampling variance

Joint Meta-Analysis of ILSA & Non-ILSA Data

Gender differences in digital skills

Non-ILSA data, 22 studies
ICILS 2013 & 2018

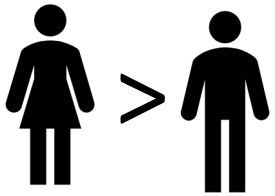
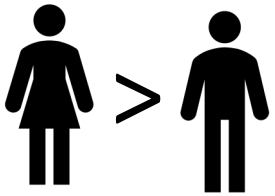
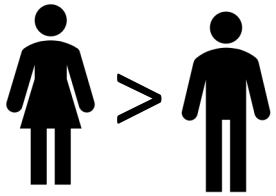
Scherer *et al.*
Large-scale Assessments in Education (2024) 12:4
https://doi.org/10.1186/s40536-024-00191-1

Large-scale Assessments in Education

REVIEW Open Access

The potential of international large-scale assessments for meta-analyses in education

Ronny Scherer[1,2*], Fazilat Siddiq[3] and Trude Nilsen[2,4]

Non-ILSA data	ICILS 2013	ICILS 2018
		
$\bar{g}$ = -0.12	$\bar{g}$ = -0.13	$\bar{g}$ = -0.21
95% *CI* [-0.04, -0.20]	95% *CI* [-0.10, -0.17]	95% *CI* [-0.15, -0.27]

Separate Meta-Analyses for Reference

Effect size: Hedges' g

Joint Meta-Analysis of ILSA & Non-ILSA Data

Gender differences in digital skills

Non-ILSA data, 22 studies
ICILS 2013 & 2018

Scherer *et al.*
Large-scale Assessments in Education (2024) 12:4
https://doi.org/10.1186/s40536-024-00191-1

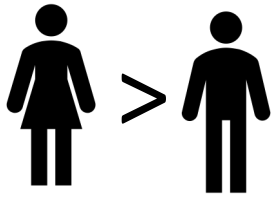
Large-scale Assessments in Education

REVIEW Open Access

The potential of international large-scale assessments for meta-analyses in education

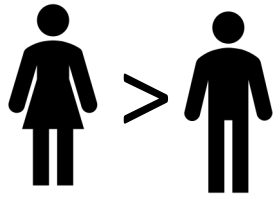
Ronny Scherer[1,2*], Fazilat Siddiq[3] and Trude Nilsen[2,4]

One-step inclusion



$\bar{g}$ = -0.13
95% *CI* [-0.05, -0.21]

Two-step inclusion



$\bar{g}$ = -0.12
95% *CI* [-0.05, -0.19]

Two approaches likely agree but have different levels of granularity.

Effect size: Hedges' g

Benefits and Challenges

(Scherer, Siddiq, & Nilsen, 2024; Campos et al., 2023)

Benefits

- Reduced language/culture bias
- Overcoming small-sample issues
- Representative samples
- Rigorous psychometric quality

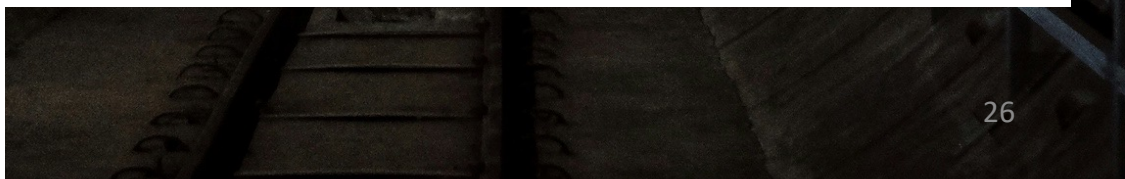
Challenges

- Complex extraction of effect sizes
- Different levels of inference
- Dominance of large-scale samples
- Complex meta-analytic models

Some Conclusions

1. Including ILSA data is a way of improving meta-analytic evidence.
2. Flexibility and range of meta-analytic approaches.
3. Meta-analyses including ILSA data can address new RQs.

→ Despite the analytic complexities, **necessity of considering ILSA data** in meta-analyses in education, *if* they can address the RQs.




Thank You.

Illustrative pictures were retrieved from unsplash.com

References

Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weaknesses. *Review of Educational Research, 82*(4), 436–476. https://doi.org/10.3102/0034654312458162

APA (2018). APA Dictionary of Psychology: Meta-analysis. https://dictionary.apa.org/meta-analysis

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to Meta-Analysis. John Wiley & Sons, Ltd.

Brunner M, Keller L, Stallasch SE, et al. 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. *Res Syn Meth*. 2023; 14(1): 5-35. https://doi.org/10.1002/jrsm.1584

Burgard, T., Bošnjak, M., & Studtrucker, R. (2021). Community-augmented meta-analyses (CAMAs) in psychology: Potentials and current systems.*Zeitschrift für Psychologie, 229*(1), 15–23. https://doi.org/10.1027/2151-2604/a000431

Campos, D. G., Cheung, M. W. -L., & Scherer, R. (2023). A primer on synthesizing individual participant data obtained from complex sampling surveys: A two-stage IPD meta-analysis approach. Psychological Methods. https://doi.org/10.1037/met0000539

Card, N. A. (2012). Applied meta-analysis for social science research. The Guilford Press.

Cheung, M. W.-L. (2015). Meta-Analysis: A Structural Equation Modeling Approach. John Wiley & Sons Ltd.

Drummond, A., & Sauer, j. (2020). Timesplitters: Playing video games before (but not after) school on weekdays is associated with poorer adolescent academic performance. A test of competing theoretical accounts. Computers & Education, 144, 103704. https://doi.org/10.1016/j.compedu.2019.103704

Ferguson, C. J. (2015). Do Angry Birds Make for Angry Children? A Meta-Analysis of Video Game Influences on Children's and Adolescents' Aggression, Mental Health, Prosocial Behavior, and Academic Performance. *Perspectives on Psychological Science*, *10*(5), 646-666. https://doi.org/10.1177/1745691615592234

Harrer, M., Cuijpers, P., Furukawa, T.A., & Ebert, D.D. (2021). *Doing Meta-Analysis with R: A Hands-On Guide*. Boca Raton, FL and London: Chapman & Hall/CRC Press. ISBN 978-0-367-61007-4. https://bookdown.org/MathiasHarrer/Doing_Meta_Analysis_in_R/#citing-this-guide

Polanin, J. R., Maynard, B. R., & Dell, N. A. (2017). Overviews in education research: A systematic review and analysis. *Review of Educational Research, 87*(1), 172–203. https://doi.org/10.3102/0034654316631117

Powers, K.L., Brooks, P.J., Aldrich, N.J. *et al.* Effects of video-game play on information processing: A meta-analytic investigation. *Psychon Bull Rev* **20**, 1055–1079 (2013). https://doi.org/10.3758/s13423-013-0418-z

Pustejovsky, J.E., Tipton, E. Meta-analysis with Robust Variance Estimation: Expanding the Range of Working Models. *Prev Sci* **23**, 425–438 (2022). https://doi.org/10.1007/s11121-021-01246-3

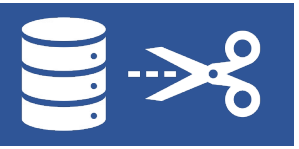
Sala, G., Tatlidil, K. S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin, 144*(2), 111–139. https://doi.org/10.1037/bul0000139

Scherer, R., Siddiq, F. & Nilsen, T. The potential of international large-scale assessments for meta-analyses in education. *Large-scale Assess Educ* **12**, 4 (2024). https://doi.org/10.1186/s40536-024-00191-1

Viechtbauer W (2010). "Conducting meta-analyses in R with the metafor package." Journal of Statistical Software, 36(3), 1–48. https://doi.org/10.18637/jss.v036.i03

Appendix

Effect Size Multiplicity



(Scherer et al., 2024; Viechtbauer, 2010)

Cross-Classified Random-Effects Meta-Analysis

Implementation in the R package `metafor`

```
## Model estimation
CCREM5 <- rma.mv(d,
                 vd,
                 random = list(~ 1 | StudyID/ESID,
                               ~ 1 | factor(Country)),
                 method = "REML",
                 data = dat)

## Model summary
summary(CCREM5)
```

Nesting of effects sizes in primary studies

Nesting of effects sizes in countries

Meta-Analysis of ILSA Data

Gray et al. Large-scale Assess Educ (2019) 7:2
https://doi.org/10.1186/s40536-019-0070-9

Large-scale Assessments in Education

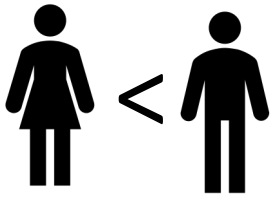
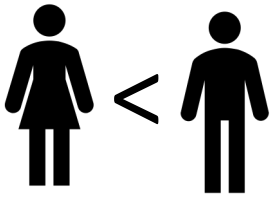
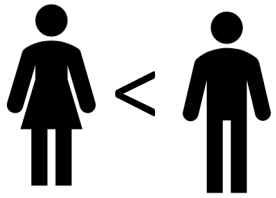
RESEARCH Open Access

Gender differences in variability in student achievement

All PISA, TIMSS, and PIRLS cycles up to 2015, about 2500 effect sizes

Sex differences in variability across nations in reading, mathematics and science: a meta-analytic extension of Baye and Monseur (2016)

Helen Gray[1], Andrew Lyth[1], Catherine McKenna[1], Susan Stothard[3], Peter Tymms[1] and Lee Copping[2*]

Reading	Mathematics	Science
		
$\overline{VR}$ = 1.15	$\overline{VR}$ = 1.12	$\overline{VR}$ = 1.13
95% *CI* [1.15, 1.16]	95% *CI* [1.11, 1.12]	95% *CI* [1.12, 1.13]

Large-scale meta-analysis of effect sizes beyond correlations and standard mean differences.

Effect size: Variance ratio VR