Impossible generalisations: metaanalyses of education interventions

Edoardo Masset 19th June 2019 Washington DC RISE Conference

Meta-analysis of education studies

- They tell 'what works' in education
- But find different results
 - because looking at different studies
 - because the evidence base in heterogeneous







Misunderstandings of meta-analyses

- 1st misconception: meta-analyses identify a universal effect applicable to all contexts
- 2nd misconception: the mean of a metaanalysis predicts the effect to be observed by a future study

Fixed-effect meta-analysis

- Assume that all interventions have the same effect
- The mean is converging to the 'true' universal effect
- The confidence interval is the result of sampling variation



Sources of heterogeneity

Sources of heterogeneity in education

 Differences in populations
 Differences between interventions
 Differences in outcomes, biases and measurement error

Random-effects meta-analysis



- Assume that interventions have different effects
- The mean is simply our estimate of the mean effect

Confidence intervals and prediction intervals

- Confidence intervals describe the accuracy of the mean
- The impact of a new intervention similar those included studies is given by a prediction interval
- Prediction intervals rarely used, overstating impacts in metaanalyses

$$CI = \mu \pm Z^{\alpha} \sqrt{\sigma_{\mu}^2}$$

 $PI = \mu \pm t^{\alpha}_{df} \sqrt{\sigma^2_{\mu} + \sigma^2_b}$

Education systematic reviews

- 3ie systematic review: 238 studies, 216 programmes, 20 intervention categories
- CCT have largest impact on participation outcomes, while SF is 'promising'
- Structured pedagogy have largest impact on learning, while meritbased scholarships, SF, extra time and remedial education are 'promising'
- Some interventions have zero impact





School attendance

School attendance confidence intervals



School attendance Predicted intervals



School completion

School completion confidence intervals

Intervention

Merit-based scholarships		
PPP	+	
Multi-component	-	
Pedagogy		
Cash transfers		
CAL		
Community-based monitoring	+	
School-based management	+	
Teacher hiring	+	
Teacher incentives	•	
User fee reduction	+	
School feeding	•	
	52 0 .2 .5 1	

School completion prediction intervals



Maths test scores

Maths test scores confidence intervals

Intervention

Construction of new schools Remedial education Multi-component SBH - Malaria control Pedagogy Merit-based scholarships School feeding Teacher hiring Community-based monitoring Extra time Teacher incentives CAL SBH - Micronutrient PPP SBH- Deworming Tracking Diagnostic Feedback School-based management Cash transfers Providing materials				
	5	2 0 .2	.5	1

Maths test scores prediction intervals

Intervention



Reading test scores

Reading test scores confidence intervals

Intervention

Pedagogy Extra time Remedial education Community-based monitoring Tracking School feeding Teacher hiring Merit-based scholarships Multi-component PPP SBH - Malaria control Construction of new schools CAL Diagnostic Feedback School-based management SBH- Deworming		← ← ← ←	
	52 0	.2 .5	1

Reading test scores prediction intervals

Intervention



What works in education?

- No intervention is predictably more effective or more promising
- Heterogeneity is very high for all outcomes
- Heterogeneity is underestimated by prediction intervals:
 - New studies likely to be different from those included
 - Publication bias

Conclusions for education metaanalyses

- More studies will not reduce heterogeneity
 Heterogeneity likely to increase with number of
 - Heterogeneity likely to increase with number of studies
- Redefining intervention categories:
 - More precise categories will make more homogeneous groups

Addressing heterogeneity

- The grand mean is not very useful, can be misleading
- Analyse heterogeneity

 extent by category
 sources of heterogeneity
- Lessons for single studies: explore mechanisms do not just estimate effect sizes

Thank you