Inclusion at Risk?

Striving for More Inclusiveness: Understanding and Addressing Student Exclusion in TIMSS

Umut Atasever, Sabine Meinck,
Rolf Strietholt (IEA)
Matthias von Davier, Bethany Fishbein (BC)

06 October 2026, Warsaw

Content

- Introduction & context
- Trends in exclusion rates in ILSA
- The R&D project
- Ways forward
- Discussion

Introduction: UNCRPD

2008: UN Convention on the Right of Persons with Disabilities came into force, resulting in a shift from segregation to inclusion in education

- Right to inclusive education
- Provision for accommodation
- Accessibility improvements
- …

Inclusiveness in IEA Studies?

- Despite global developments, inclusiveness in IEA studies has *decreased*
- In recent IEA studies:
  - One third to half the countries missed the exclusion rate threshold (5%)
  - Within-school exclusions rising while school-level exclusions remain stable

(Atasever et al., 2023)

Within-school exclusions: Recap

- Students with functional disabilities
- Students with intellectual disabilities
- Students facing a severe language barrier

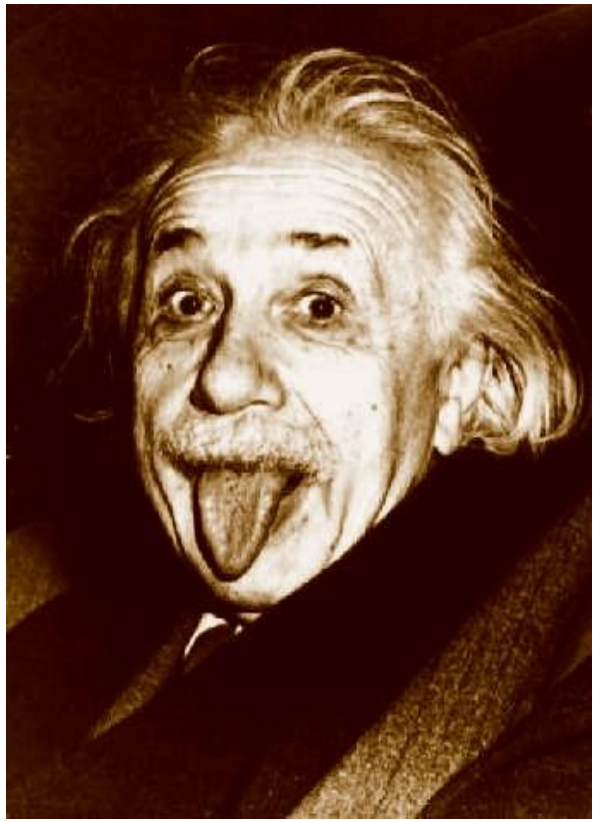
Only students who really cannot participate at all!

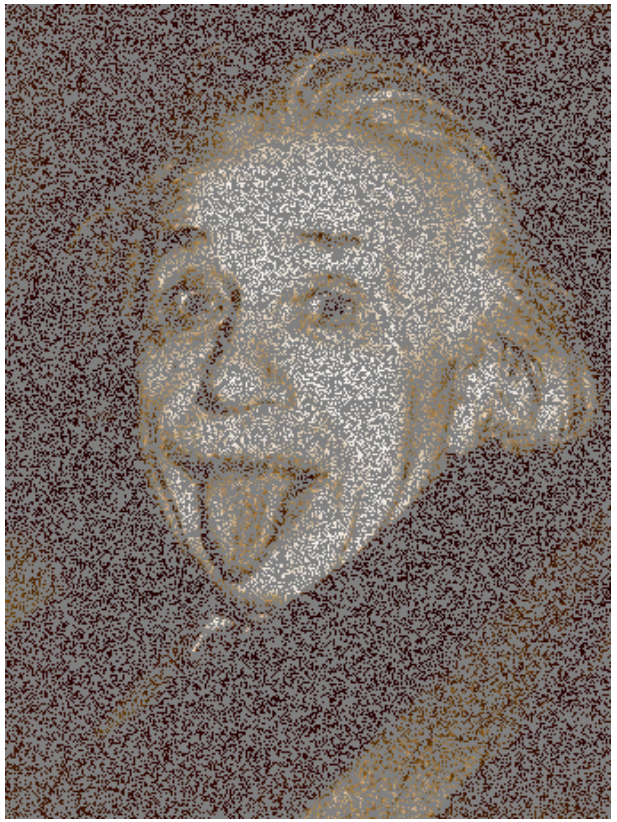
Potential Reasons for Rising Exclusion Rates in ILSAs

- UNCRPD
  - More students with special education needs (SEN) now attend regular schools
    - BUT there is no reason why the overall number of SEN would have changed!
  - New regulations and resources for SEN students may result in more being flagged for exclusion(?) (Weber & Meinck, 2019)
- New arrivals of migrants are increasing the number of students excluded for language barriers
  - Military conflicts, economic crisis, climate change

    (Broken Chalk, 2023; UNICEF, 2023)

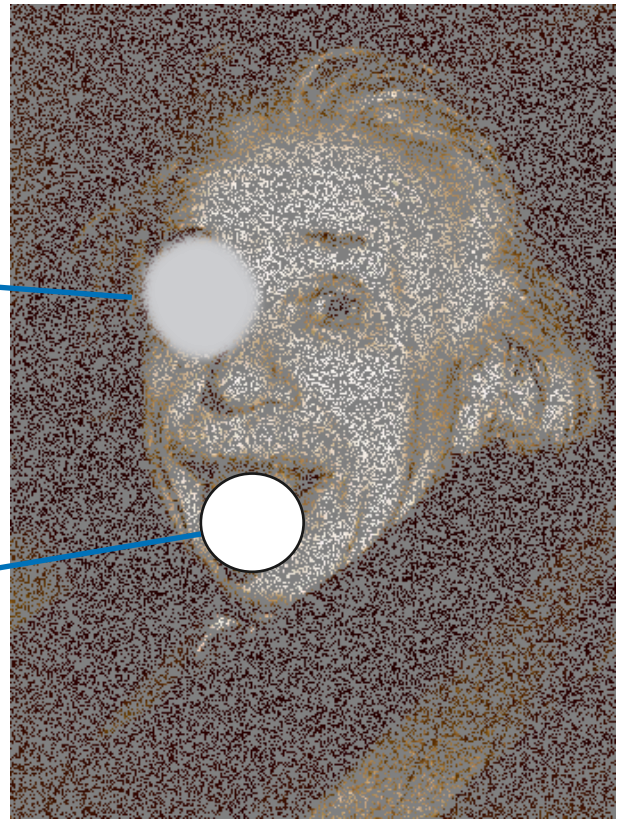
Are Rising Exclusion Rates a Problem?





Readers of our reports think we cover **all** students in **all** countries.

Are Rising Exclusion Rates a Problem?



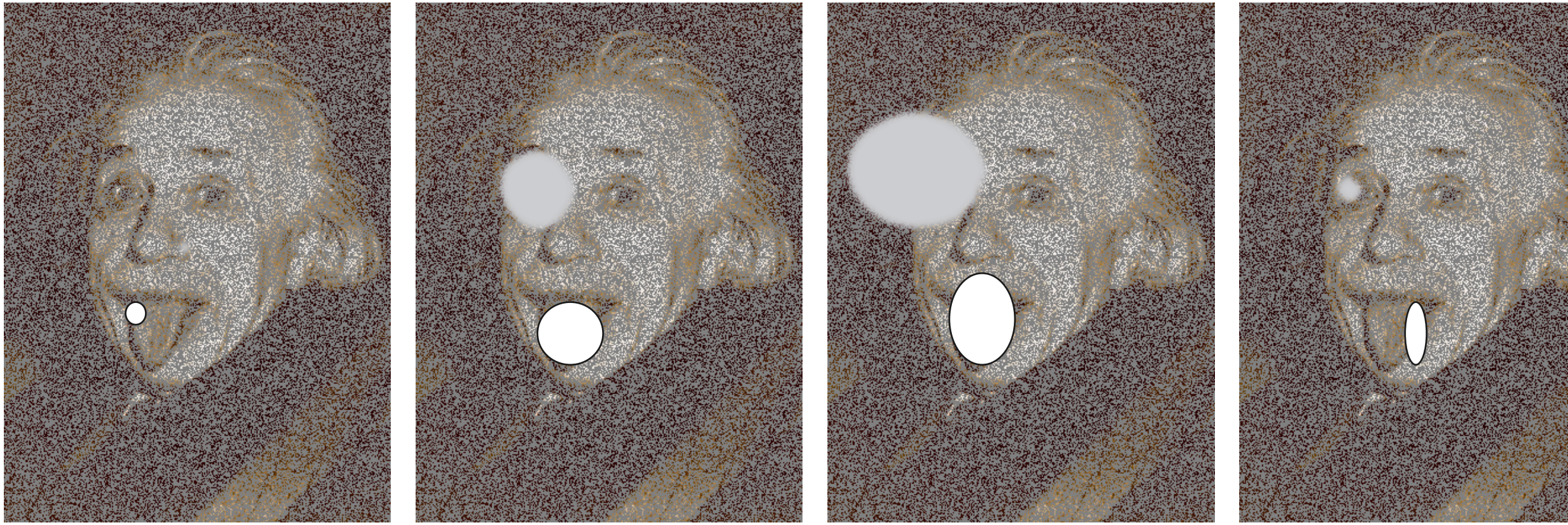
Within-school exclusions

School exclusions

Fuzziness of picture: sample data is already subject to uncertainty

Readers of our reports think we cover **all** students in **all** countries.

Are Rising Exclusion Rates a Problem?



Increasing exclusion rates are a potential threat on comparability.

Previous Studies: Threats Towards Representativity

We know:

- Selective participation inflates achievement
  (Micklewright et al., 2012; Durrant & Schnepf, 2018)
- Register data comparisons show overestimation
  (Anders et al., 2021; Borger et al., 2024)
- Bias risks increase with non-response (Meinck et al., 2023)

But:

- Most evidence concerns non-response
- Mostly country-specific
- Rarely uses observed achievement data
- **No multi-country empirical evaluation of excluded students**

Anders, J., Has, S., Jerrim, J., et al. (2021). Is Canada really an education superpower? The impac
results from PISA 2015. Educational Assessment, Evaluation and Accountability, 33(2), 22
https://doi.org/10.1007/s11092-020-09329

Atasever, U., Jerrim, J., & Tieck, S. (2023). Exclusion rates from international large-scale assessm
years of IEA data. Educational Assessment, Evaluation and Accountability. https://doi.org
09416-3

Atasever, U., Meinck, S. (2025). Striving for more inclusiveness: Understanding and addressing s
studies. Symposium paper presented at IRC 2025, Symposium: Exclusions and Participati
Striving for More Inclusiveness. Rome, Italy.

Borger, L., Johansson, S., & Strietholt, R. (2024). How representative is the Swedish PISA sample
register data. Educational Assessment, Evaluation and Accountability. https://doi.org/10.

Durrant, G., & Schnepf, S. 2018. Which schools and pupils respond to educational achievement
English Programme for international student assessment sample. Journal of the Royal Sta
181(4): 1,057–1,075.

Education Datalab. 2017. Why does Vietnam do so well in PISA? An example of why naive inter
rankings is such a bad idea Accessed 05/05/2022 from https://ffteducationdatalab.org.u
vietnam-do-so-well-in-pisa-an-example-of-why-naive-interpretation-of-international-ran

Meinck, S., Heine, H.-J., Mang, J., Nagy, G. (2023). Bias risks in ILSA related to non-participation
from a longitudinal large-scale survey in Germany (PISA Plus). Educational Assessment, Ev
https://doi.org/10.1007/s11092-023-09422-5

Micklewright, J.; Schnepf, S. V. and Skinner, C. J. 2012. Non-response biases in surveys of schoo
English PISA samples. Journal of the Royal Statistical Society. Series A 175: 915–938.

Kennedy, A. I., & Strietholt, R. (2025). Has Student Absenteeism Risen Since the COVID-19 Pan
TIMSS and PIRLS. Symposium paper presented at IRC 2025, Symposium: Using Internatio
Assessments to Understand Changes to Education During the COVID-19 Pandemic. Rom

Spaull, N. (2018). Who makes it into PISA? Understanding the impact of PISA sample eligibility u
(PISA 2003–PISA 2012). Assessment in Education: Principles, Policy & Practice, 26(3), 39
https://doi.org/10.1080/0969594X.2018.1504742

Pereira, M. 2011. An analysis of Portuguese students' performance in the OECD programme for
assessment (PISA). Accessed 05/05/2022
from https://www.bportugal.pt/sites/default/files/anexos/papers/ab201111_e.pdf.

Weber, S., Meinck, S. (2019). Students with special education needs: suggestions to improve IEA
accommodate related research. Paper presentation at the 63rd Annual Conference of the
International Education Society (CIES), April 14-19, USA, San Francisco.

The R&D Project INCLUDE



- Exclusion rates in ILSAs have increased in recent cycles
  (Atasever, Jerrim & Tieck, 2023; Atasever & Meinck, 2025)
- Yet we lack empirical evidence on what exclusion means for achievement

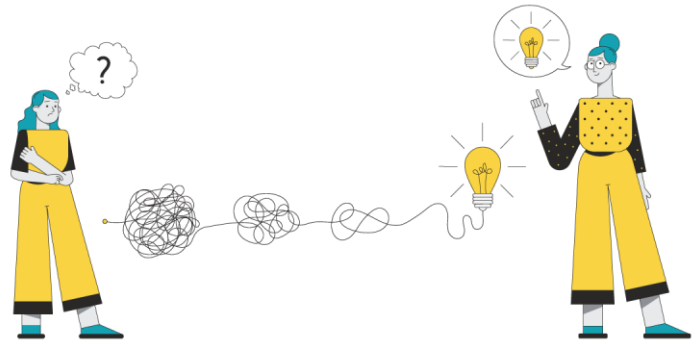
Specifically:

- What happens when excluded students complete the assessment?
  - Are they truly unable to participate?
  - Or are we excluding students who could participate?
- Has their in-/exclusion an impact on achievement results?

TIMSS 2023 Procedures: An Opportunity for Research

In TIMSS 2023:

- Some students officially marked as excluded completed the assessment for classroom inclusiveness
- Generated achievement and background data
- But excluded from international weighting & scaling



This enabled us to:

- Include “excluded” students
  - Re-weight them into to the main sample
  - Estimate plausible values (achievement scores)
- Investigate whether they “meaningfully” participated in the assessment
- Quantify the bias introduced by exclusion

Research Questions

RQ1 – Measurement Feasibility

- Do excluded students provide sufficient information for stable proficiency estimation within the TIMSS scaling model?

RQ2 – Achievement Differences – National Impact

- How do included and excluded students differ in achievement?
- How does the inclusion of previously excluded students affect national estimates?

Data: TIMSS 2023



Sample

- Grade 4: 17 countries + Grade 8: 8 countries
- ≥ 50 excluded students with achievement data
- Limitation: Assessment inclusion of excluded students was an arbitrary decision of countries or even schools

Subject: Mathematics + Science

Student groups in this study

- **Included:** Incorporated into official international results 
- **Excluded:** Marked as excluded but completed the assessment 
- **Overall (Included + Excluded):**
  Reweighted and rescaled using TIMSS procedures 

Exclusion Rates Before and After Including "Excluded"

Country	Before (%)	After (%)
Armenia	**6.1**	5
Australia	3.8	2.7
Belgium (French)	**8.9**	**5.7**
Brazil	**6.5**	4.4
Canada	**14.3**	5.2
Canada (Ontario)	**13.2**	3
Canada (Quebec)	**8.8**	3
Chinese Taipei	2.6	0.9
Czech Republic	**7.3**	**6**
Denmark	**9.3**	5.3
Hungary	5.1	3.6
Italy	**7**	3.4
Oman	2.4	1.6
Poland	**6.5**	3.3
Türkiye	**21.9**	**19**
United Arab Emirates	1.7	1.5
United Arab Emirates (Dubai)	3.6	2.8

11 countries out of 17 missed the 5% threshold.

Only 3 countries would still have missed it if they'd included excluded students with data.

Note: Denmark finally decided to include excluded students with data in the TIMSS IDB.

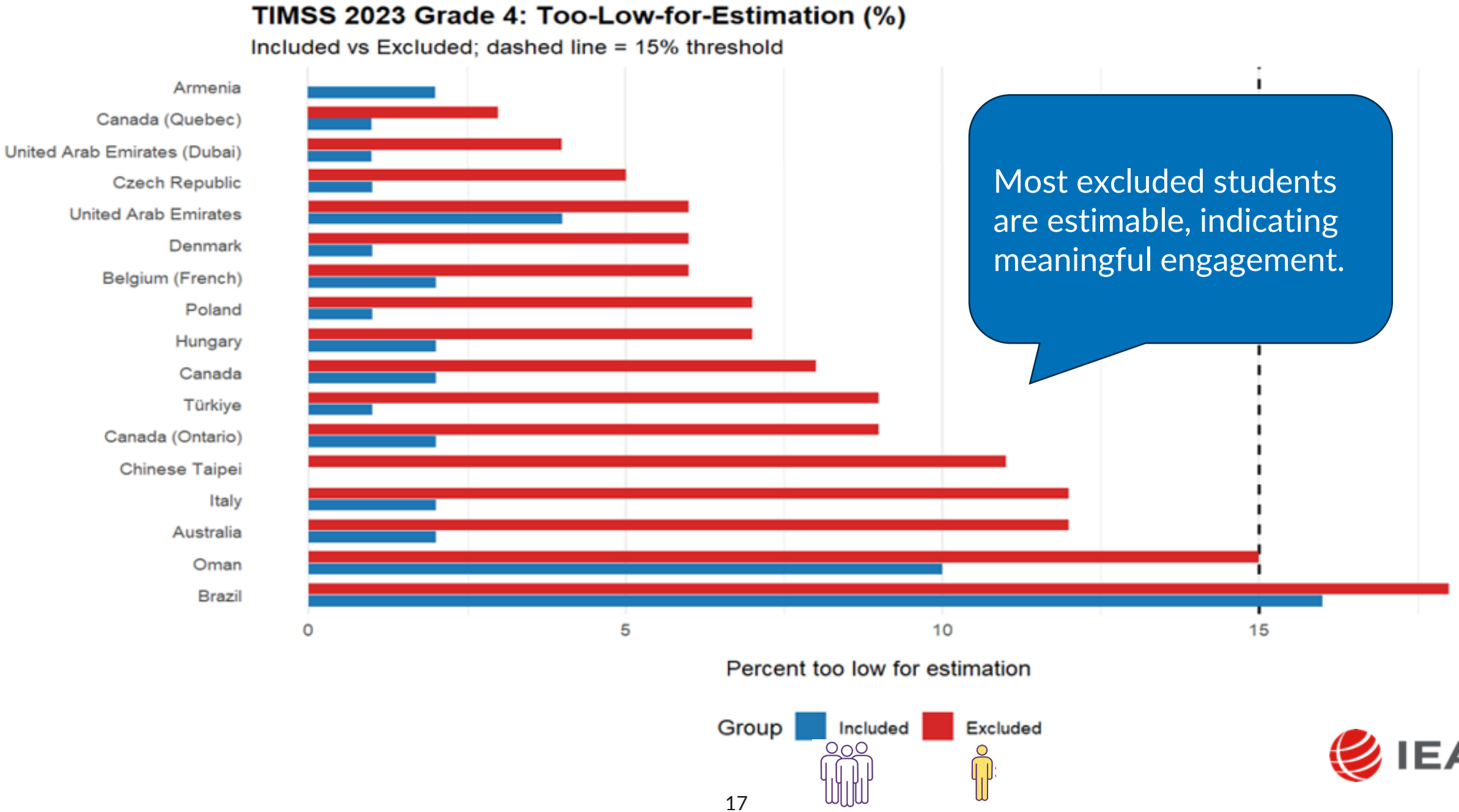
Analysis/Results

RQ1 – Measurement Feasibility

- Do excluded students provide sufficient information for stable proficiency estimation within the TIMSS scaling model?

TIMSS classifies students as *too low for estimation (TLE)* if their performance is no better than random guessing on multiple-choice items.

- We examined the percentage of TLE students among included and excluded student groups.
- Note: TIMSS flags systems if TLE exceeds **15%**.

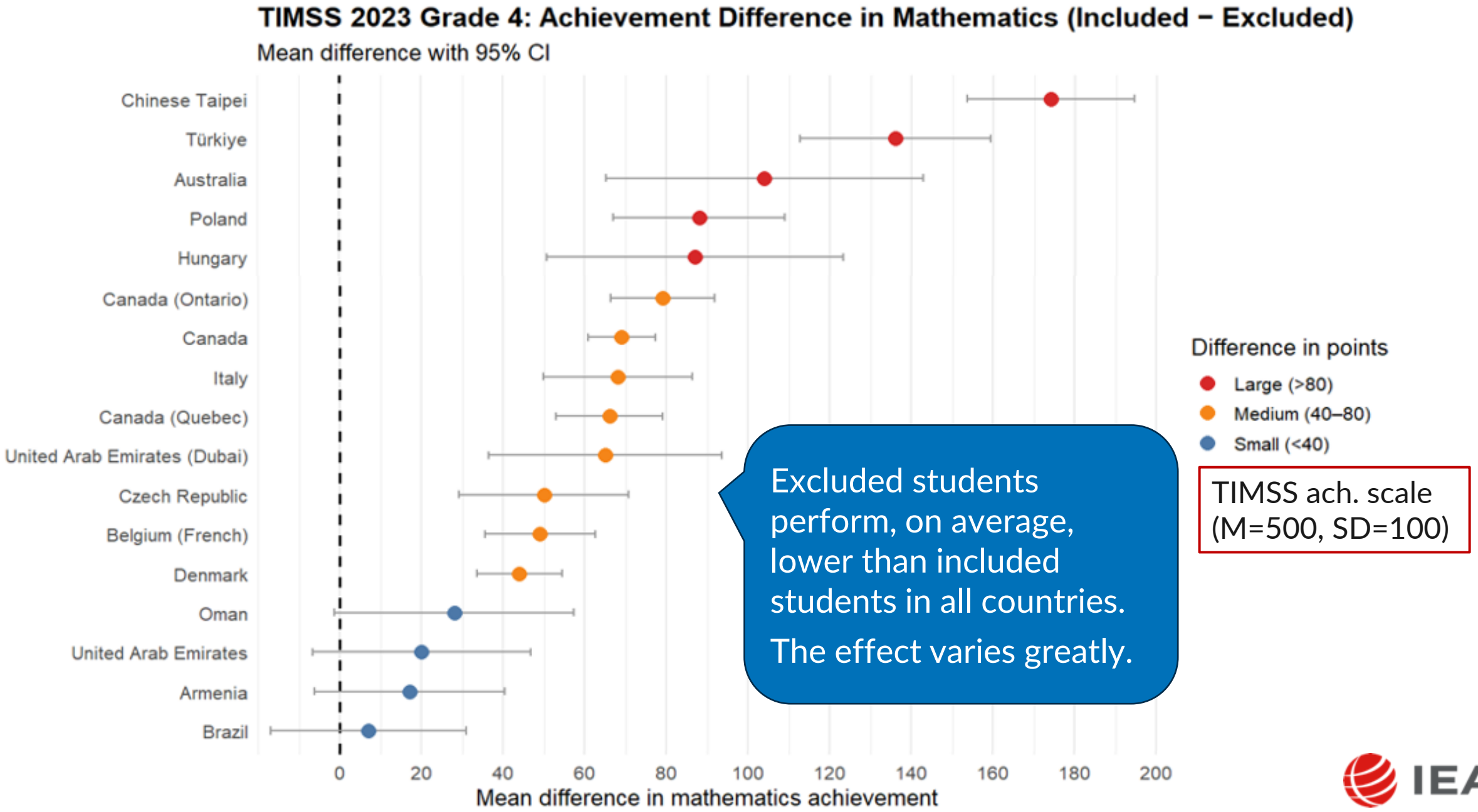

TIMSS 2023 Grade 4: Too-Low-for-Estimation (%)
Included vs Excluded; dashed line = 15% threshold
Armenia
Canada (Quebec)
United Arab Emirates (Dubai)
Czech Republic
United Arab Emirates
Denmark
Belgium (French)
Poland
Hungary
Canada
Türkiye
Canada (Ontario)
Chinese Taipei
Italy
Australia
Oman
Brazil
0
5
10
15
Percent too low for estimation
Group
Included
Excluded
Most excluded students are estimable, indicating meaningful engagement.

Analysis/Results

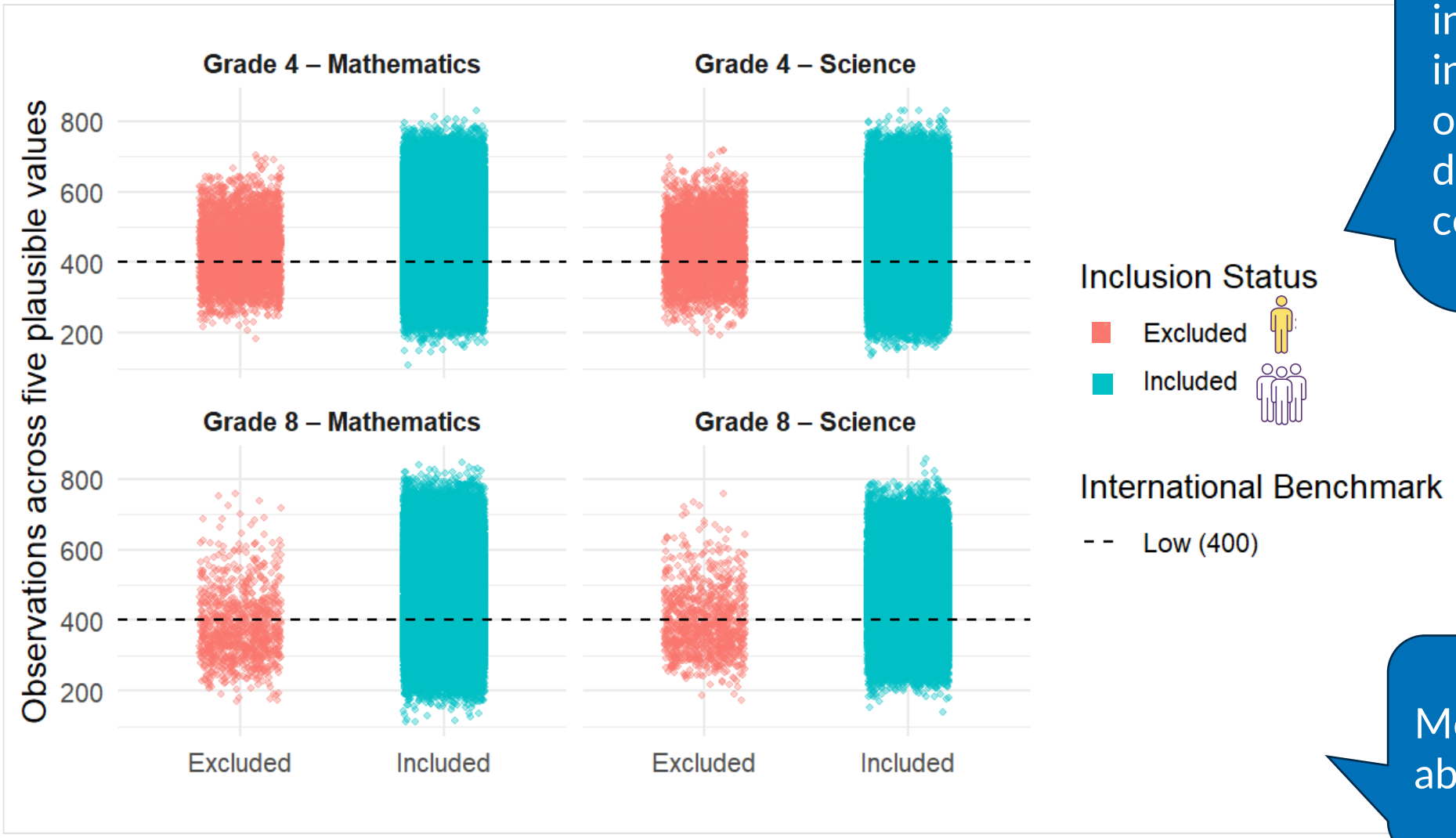
RQ2 – Achievement Differences – National Impact

- How do included and excluded students differ in achievement?
- How does the inclusion of previously excluded students affect national estimates?

- We computed the average scores by group after re-weighting and re-scaling.


TIMSS 2023 Grade 4: Achievement Difference in Mathematics (Included − Excluded)
Mean difference with 95% CI
Chinese Taipei
Türkiye
Australia
Poland
Hungary
Canada (Ontario)
Canada
Italy
Canada (Quebec)
United Arab Emirates (Dubai)
Czech Republic
Belgium (French)
Denmark
Oman
United Arab Emirates
Armenia
Brazil
0
20
40
60
80
100
120
140
160
180
200
Mean difference in mathematics achievement
Difference in points
Large (>80)
Medium (40–80)
Small (<40)
Excluded students perform, on average, lower than included students in all countries.
The effect varies greatly.
TIMSS ach. scale (M=500, SD=100)

Mathematics and science achievement distributions by inclusion status, grade, and subject



Evidence for inconsistent and incorrect application of exclusion definitions across countries

Most students perform above low benchmark.

Analysis/Results

RQ2 – Achievement Differences – National Impact

- How do included and excluded students differ in achievement?
- How does the inclusion of previously excluded students affect national estimates?

- We compared the average scores before and after including the excluded students.

Example with Three Countries

	Difference in Achievement (Included – Excluded)	Difference in Achievement (Included-Overall)	Rate of Excluded Students (those with data, %)
Poland	**88 (10.5)**	3 (2.0)	3.3
Canada (Ontario)	**79 (6.5)**	**8 (3.5)**	**10.1**
Chinese Taipei	**174 (10.4)**	3 (1.6)	1.7

Significant results are shown in **bold**.
Exclusion rates exceeding the TIMSS 5% threshold are highlighted in **bold**.

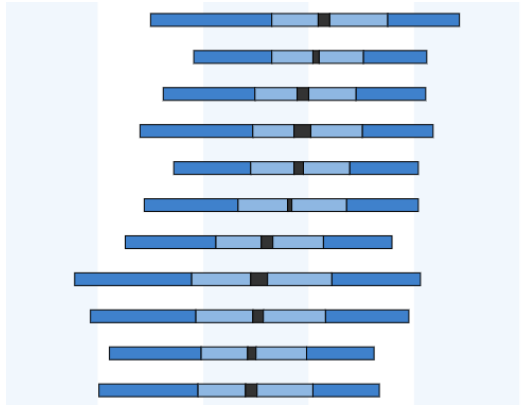
Effect depends on interplay of achievement difference and exclusion rate.

Across all countries: Change in average score always negative, but mostly small, and significant only in Canada.

5% threshold seems quite reasonable!

Discussion: What Do These Findings Mean?

1. **Exclusion assumptions are not uniformly supported**
   - Many excluded students are psychometrically estimable
   - A non-trivial share demonstrates meaningful achievement
   - Exclusion criteria seem to be applied too broadly and/or inconsistently
2. **Exclusion affects national estimates**
   - Leaving out low performing students increases artificially national means
   - Bias depends on both exclusion rate and achievement gap
3. **Implications for ILSA validity**
   - Jeopardizes cross-national comparability



Conclusion: We Need to be More Inclusive!

1. **Refine exclusion definitions**
   - Most excluded students perform well above benchmarks
   - Check why interpretation of definitions is too broad or inconsistent (wording, translation, national context)
2. **Strengthen accommodation provision and guidance**
   - Accommodation may substitute for exclusion (e.g., extended time, screen magnification, accessible interface options)
   - Both international study centers and national centers must make efforts
3. **Leverage test design**
   - Assign less difficult items as accommodation
   - Expand adaptive and performance-based routing

Aside: R&D Project ADAPT



Purpose:

- ADAPT investigated the provision and use of accommodations for students with disabilities across countries and studies

Results:

- Inclusion requires all students to have the necessary conditions to participate meaningfully
- Substantial variation across countries in how they approach accommodation and inclusion
- This may reflect differences in national contexts, perspectives, and goals, but also different interpretations of rather vague guidelines
- Some countries appear to manage inclusion successfully

(Mejia-Rodriguez et al., 2026)

Ways Forward…

Efforts for strengthening inclusivity need to be made at both ISC's and NC's.



- Increase awareness & accountability
- Revisit exclusion process
- Accommodate rather than exclude
- Install early warning & monitoring systems
- Consider including all students participating in the test

References

Anders, J., Has, S., Jerrim, J., et al. (2021). Is Canada really an education superpower? The impact of non-participation on results from PISA 2015. Educational Assessment, Evaluation and Accountability, 33(2), 229–249. https://doi.org/10.1007/s11092-020-09329

Atasever, U., Jerrim, J., & Tieck, S. (2023). Exclusion rates from international large-scale assessments: An analysis of 20 years of IEA data. Educational Assessment, Evaluation and Accountability. https://doi.org/10.1007/s11092-023-09416-3

Atasever, U., Meinck, S. (2025). Striving for more inclusiveness: Understanding and addressing student exclusion in IEA studies. Symposium paper presented at IRC 2025, Symposium: Exclusions and Participation Rates in IEA Studies: Striving for More Inclusiveness. Rome, Italy.

Borger, L., Johansson, S., & Strietholt, R. (2024). How representative is the Swedish PISA sample? A comparison of PISA and register data. Educational Assessment, Evaluation and Accountability. https://doi.org/10.1007/s11092-024-09438-5

Durrant, G., & Schnepf, S. 2018. Which schools and pupils respond to educational achievement surveys? A focus on the English Programme for international student assessment sample. Journal of the Royal Statistical Society, Series A, 181(4): 1,057–1,075.

Education Datalab. 2017. Why does Vietnam do so well in PISA? An example of why naive interpretation of international rankings is such a bad idea Accessed 05/05/2022 from https://ffteducationdatalab.org.uk/2017/07/why-does-vietnam-do-so-well-in-pisa-an-example-of-why-naive-interpretation-of-international-rankings-is-such-a-bad-idea/.

Meinck, S., Heine, H.-J., Mang, J., Nagy, G. (2023). Bias risks in ILSA related to non-participation: evidence from a longitudinal large-scale survey in Germany (PISA Plus). Educational Assessment, Evaluation and Accountability https://doi.org/10.1007/s11092-023-09422-5

Mejía-Rodríguez, A. M., Eryilmaz, N., Kunina-Habenicht, O., & Strietholt, R. (2026). *Participation and accommodation Policies in IEA-Assessments: Patterns of Inclusion and Exclusion.* https://doi.org/10.58150/rd_4_6

Micklewright, J.; Schnepf, S. V. and Skinner, C. J. 2012. Non-response biases in surveys of school children: the case of the English PISA samples. Journal of the Royal Statistical Society. Series A 175: 915–938.

Spaull, N. (2018). Who makes it into PISA? Understanding the impact of PISA sample eligibility using Turkey as a case study (PISA 2003–PISA 2012). Assessment in Education: Principles, Policy & Practice, 26(3), 397–421. https://doi.org/10.1080/0969594X.2018.1504742

Pereira, M. 2011. An analysis of Portuguese students' performance in the OECD programme for international student assessment (PISA). Accessed 05/05/2022 from https://www.bportugal.pt/sites/default/files/anexos/papers/ab201111_e.pdf.

Thank you!
Questions, comments?

umut.atasever@iea-hamburg.de

sabine.meinck@iea-hamburg.de