GENDER GAPS IN EDUCATION THROUGH THE LENS OF INTERNATIONAL LARGE-SCALE ASSESSMENTS

Francesca Borgonovi – University College London & OECD

Francesca.Borgonovi@oecd.org
f.borgonovi@ucl.ac.uk

Outline of my talk

- What are International Large-Scale Assessments?
- Details of the OECD-led PISA study
- How can ILSAs be used to examine gender gaps in education? Examples from my own research

International-Large Scale Assessments in Education

- ***System level instruments*** designed to ***measure*** and ***allow comparisons*** in how well different systems do in equipping young people with ***knowledge*** and ***ability.***

- Low-stakes

- Include self-reports as well as actual tests

- They are called large scale because they involve many people in many countries

- Repeated cross-sections

- (Mostly) focus on school-age population

Curricular focus vs. problem-solving

- Frameworks:
  - Emphasis on curricular content and coverage
  - Emphasis on problem-solving and transversal skills
- Target populations:
  - Grade-based
  - Age-based

OECD-led and IEA-led large-scale assessments

- **OECD**
  - [PISA](#): reading, math, science, + innovative domains (15-year-old students).
  - [PIAAC](#): literacy, numeracy + problem solving (16 to 65 year olds).
  - [IELCWS](#) social and emotional skills (5-year-olds).
  - [SSES](#) social and emotional skills (15-year-old students + option 10-year-old students).

- **IEA**
  - [TIMSS](#) mathematics and science (4th and 8th grade students + last year secondary).
  - [PIRLS](#) reading literacy (4th grade students).
  - [ICCS](#) civic and citizenship competences (8th grade students).
  - [ICILS](#) digital competences (8th grade students).

Function of ILSAs in benchmarking and global monitoring

4 QUALITY EDUCATION

Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.

TARGET 4.1 — FREE PRIMARY AND SECONDARY EDUCATION

By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes.

TARGET 4.4 — INCREASE THE NUMBER OF PEOPLE WITH RELEVANT SKILLS FOR FINANCIAL SUCCESS

By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment,

TARGET 4.5 — ELIMINATE ALL DISCRIMINATION IN EDUCATION

By 2030, eliminate gender disparities in education and ensure equal access to all levels of education and vocational training for the vulnerable, including persons with disabilities, indigenous peoples and children in vulnerable situations.

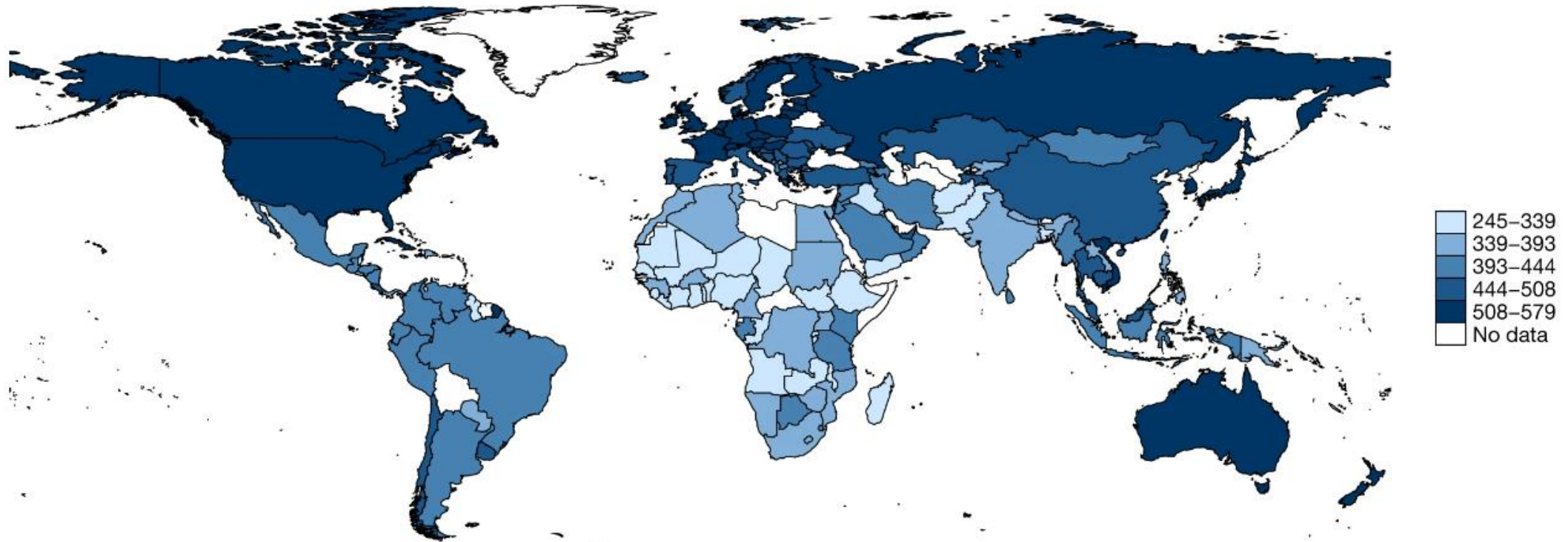### TARGET 4.6 — UNIVERSAL LITERACY AND NUMERACY

By 2030, ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy.

Function of ILSAs in benchmarking and global monitoring

Fig. 1: Average learning (2000–2017).

From: Measuring human capital using global learning data

Legend:
- 245–339
- 339–393
- 393–444
- 444–508
- 508–579
- No data

Source: Angrist, N., Djankov, S., Goldberg, P.K. *et al.* Measuring human capital using global learning data. *Nature* **592**, 403–408 (2021). https://doi.org/10.1038/s41586-021-03323-7

Programme for International Student Assessment (PISA)

- Low stakes assessment
- Length: 2 hours test + questionnaire
- System level assessment
- Target population: 15-year-old students
- Problem-based

PISA

- **Triennial** study since 2000 (no 2021 but 2022 because of COVID).
- Target **15-year-old students** in over 80 countries and economies
- Assessed in science, mathematics, reading, + innovative domains (problem solving, collaborative problem-solving, financial literacy, global competence, creativity)
- **Sample size** (minimum ~4500 students per country) from at least 150 schools but flexibility (allows for oversampling/regional sampling/additions of grade-based samples)
- Study conducted in the **language of instruction**

PISA

- ✓ Paper-based until 2012, computer-based since 2015 (influence on skill domains, test items, test design – limited adaptive, infrastructure needs).
- ✓ Rich background questionnaire administered to students
  - ✓ Socio-economic indicators
  - ✓ Learning experiences in school
  - ✓ Attitudes and behaviours (subject specific + domain general)
  - ✓ Well-being
- ✓ Questionnaires of school principals and parents
- ✓ Ad hoc student questionnaires: ICT, educational careers

PISA Databases and Analysis

https://www.oecd.org/pisa/data/2022database/

PISA Database

The PISA database contains the full set of responses from individual students, school principals and parents. These files will be of use to statisticians and professional researchers who would like to undertake their own analysis of the PISA data. The files available on this page include background questionnaires, data files in ASCII format (from 2000 to 2012), codebooks, compendia and SAS™ and SPSS™ data files in order to process the data.

2022   2018   2015   2012   2009   2006   2003   2000

PISA Data Analysis Manual: SPSS and SAS, Second Edition

These two publications are essential tools for researchers, as they provide all the information required to understand the PISA databases and perform analyses in accordance with the complex methodologies used to collect and process the data.

How to prepare and analyse the PISA database

This note summarises the main steps of using the PISA database. It describes the PISA data files and explains the specific features of the PISA survey together with its analytical implications. This document also offers links to existing documentations and resources (including software packages and pre-defined macros) for accurately using the PISA data files.

IDB analyzer (quick reproduction of the PISA results)

PISA Data Explorer

The data Explorer allows you to create your own analyses and build reports from the PISA data sets. You can access it, clicking at this link.
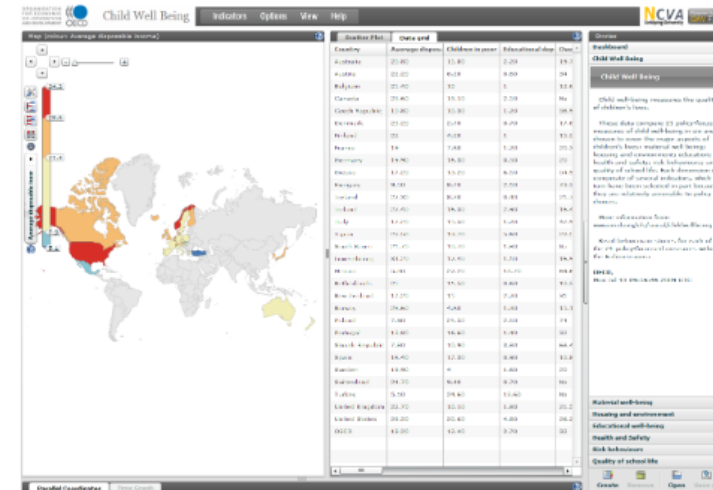
PISA Databases and Analysis

Questionnaires

Below are the master questionnaires administered to students who sat the computer-based assessment; each country translated these questionnaires into their own language(s).

- Student questionnaire administered to students - computer based (and paper based version)
- Student questionnaire (*une heure* version) administered to low-performing students - computer based (and paper based version)
- Well-being questionnaire administered to students
- Financial literacy questionnaire administered to students
- ICT familiarity questionnaire administered to students
- School questionnaire administered to school principals - computer based (and paper based version)
- Teacher questionnaire administered to teachers
- Parent questionnaire administered to parents

Download all the national versions of the questionnaires

Codebook and Compendia

Code book

Compendia

- Questionnaire items
- Cognitive items

Tables and Figures in Volumes I and II

StatLinks of Tables and Figures

PISA Databases and Analysis

SAS (TM) Data Files (compressed)

Creative Thinking data and Financial Literacy data will be uploaded in June 2024.

- School questionnaire data file
- Student questionnaire data file
- Teacher questionnaire data file
- Cognitive item data file
- Questionnaire timing data file

SPSS (TM) Data Files (compressed)

Creative Thinking data and Financial Literacy data will be uploaded in June 2024.

- School questionnaire data file
- Student questionnaire data file
- Teacher questionnaire data file
- Cognitive item data file
- Questionnaire timing data file

Codes to compute PISA Volumes I and II

You can access to the codes to compute PISA Volumes I and II here.

Rescaled Indices for Trend Analyses

Rescaled indices of economic, social and cultural status (ESCS) for use with the PISA 2012, 2015 and 2018 datasets.

- Rescaled indices in CSV format (22MB)

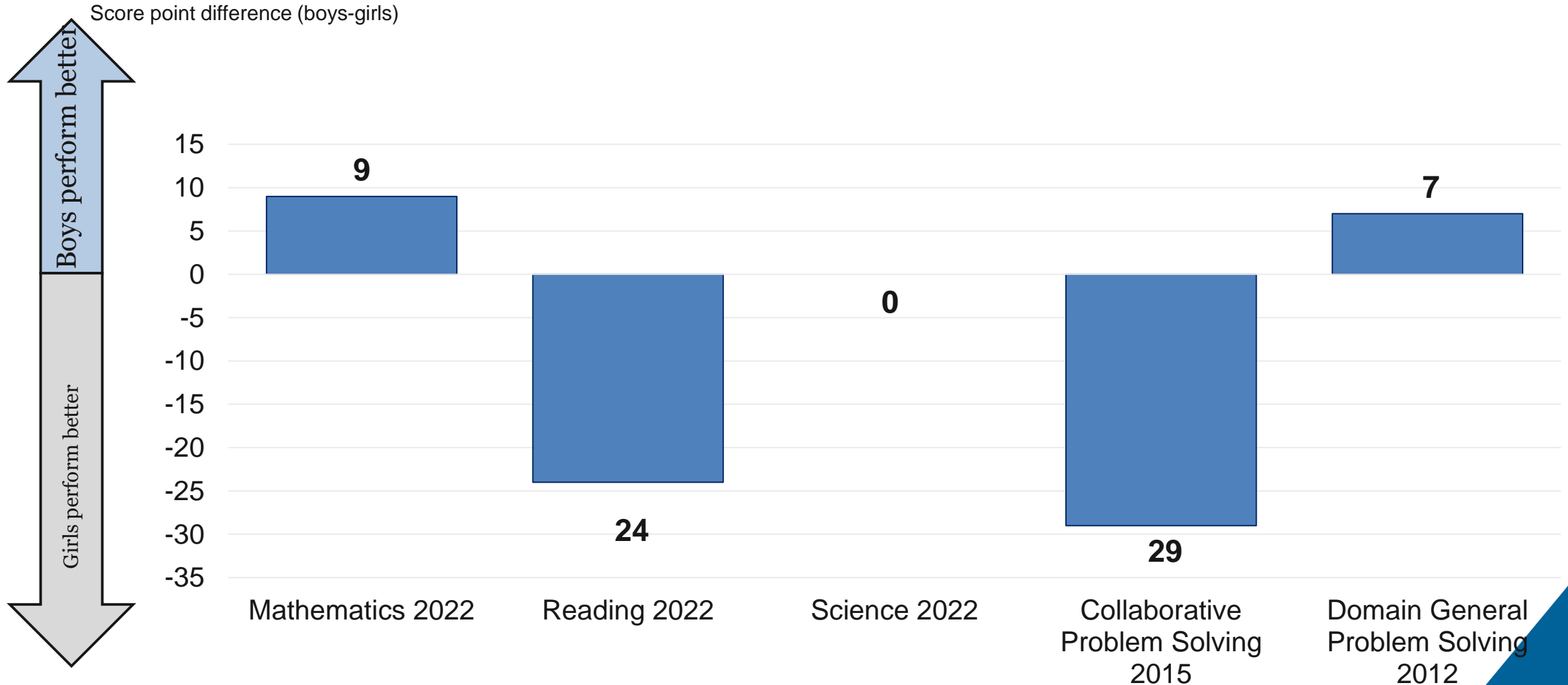FROM GENDER **GAP** TO GENDER **GAPS** IN
WHAT TO GENDER **GAPS** IN **WHAT, WHEN,**
AND **WHERE**

WHAT

Gender **gaps** in achievement at age 15

Score point difference (boys-girls)

Boys perform better

Girls perform better

	Mathematics 2022	Reading 2022	Science 2022	Collaborative Problem Solving 2015	Domain General Problem Solving 2012
Value	9	24	0	29	7

Source: OECD, PISA Databases 2022, 2015 and 2012.

Gender **gaps** in achievement at age 15

Source: OECD, PISA Databases 2022, 2015 and 2012.
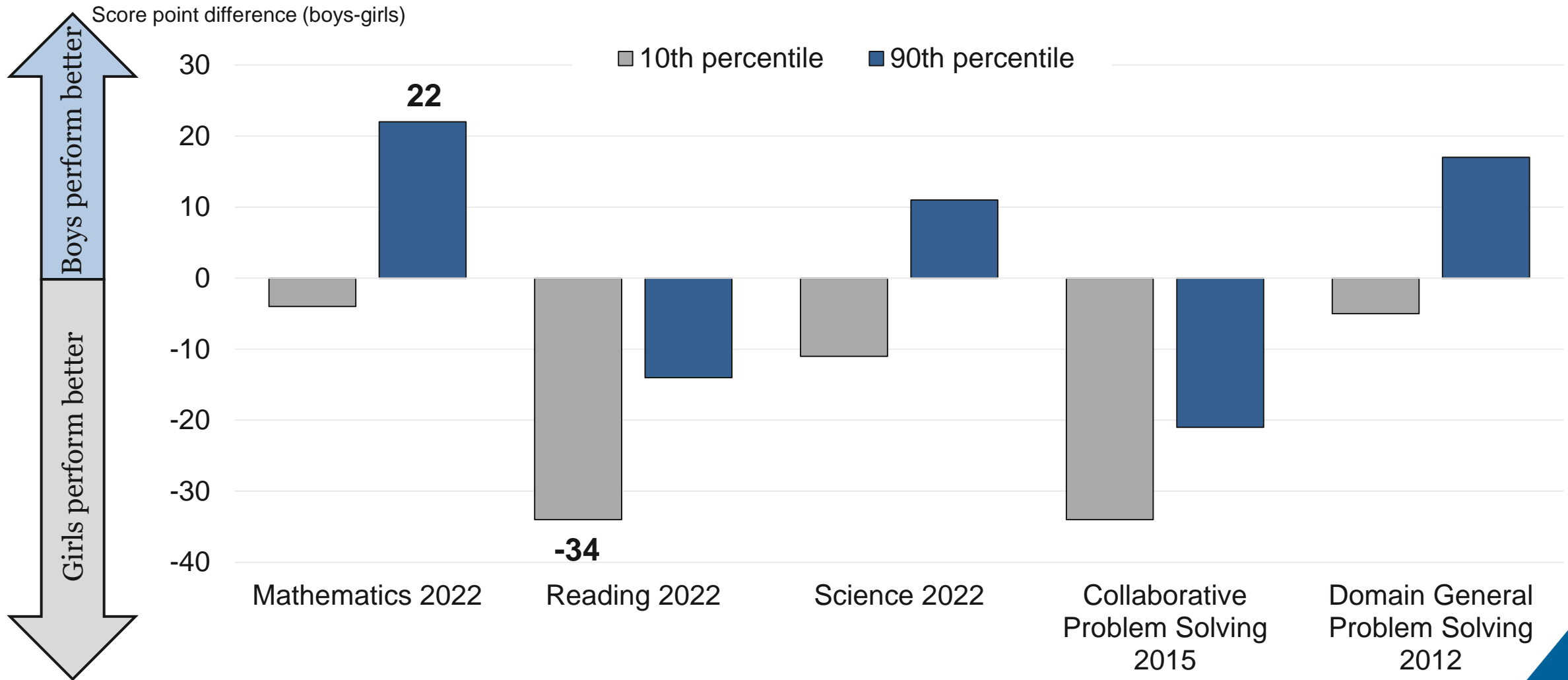
WHEN

Evolution of gender gaps between primary and young adulthood

BORGONOVI, F., CHOI, A. & PACCAGNELLA, M. (2021) The evolution of gender gaps in numeracy and literacy between childhood and young adulthood. *Economics of Education Review*
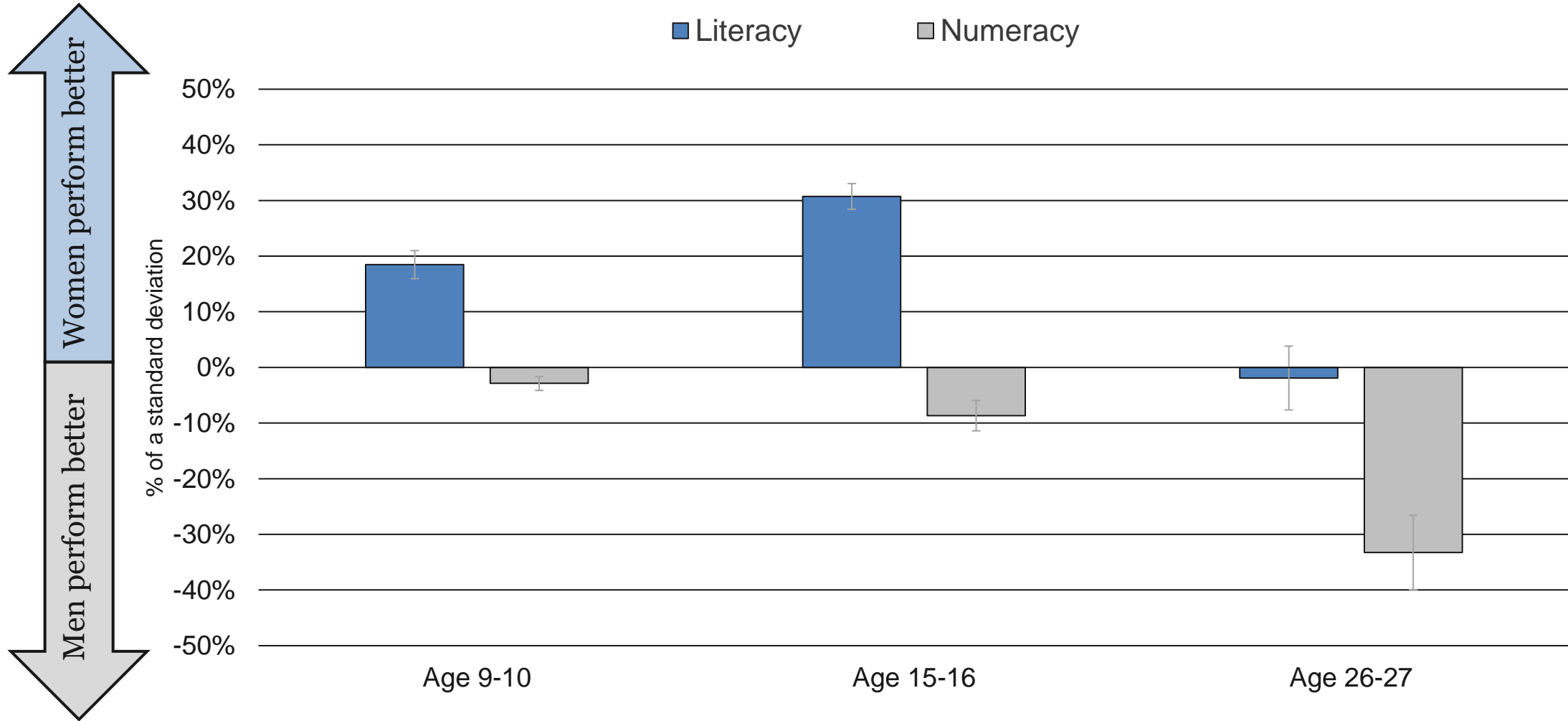
Evolution of numeracy gender gaps across the distribution

BORGONOVI, F., CHOI, A. & PACCAGNELLA, M. (2021) The evolution of gender gaps in numeracy and literacy between childhood and young adulthood. *Economics of Education Review*
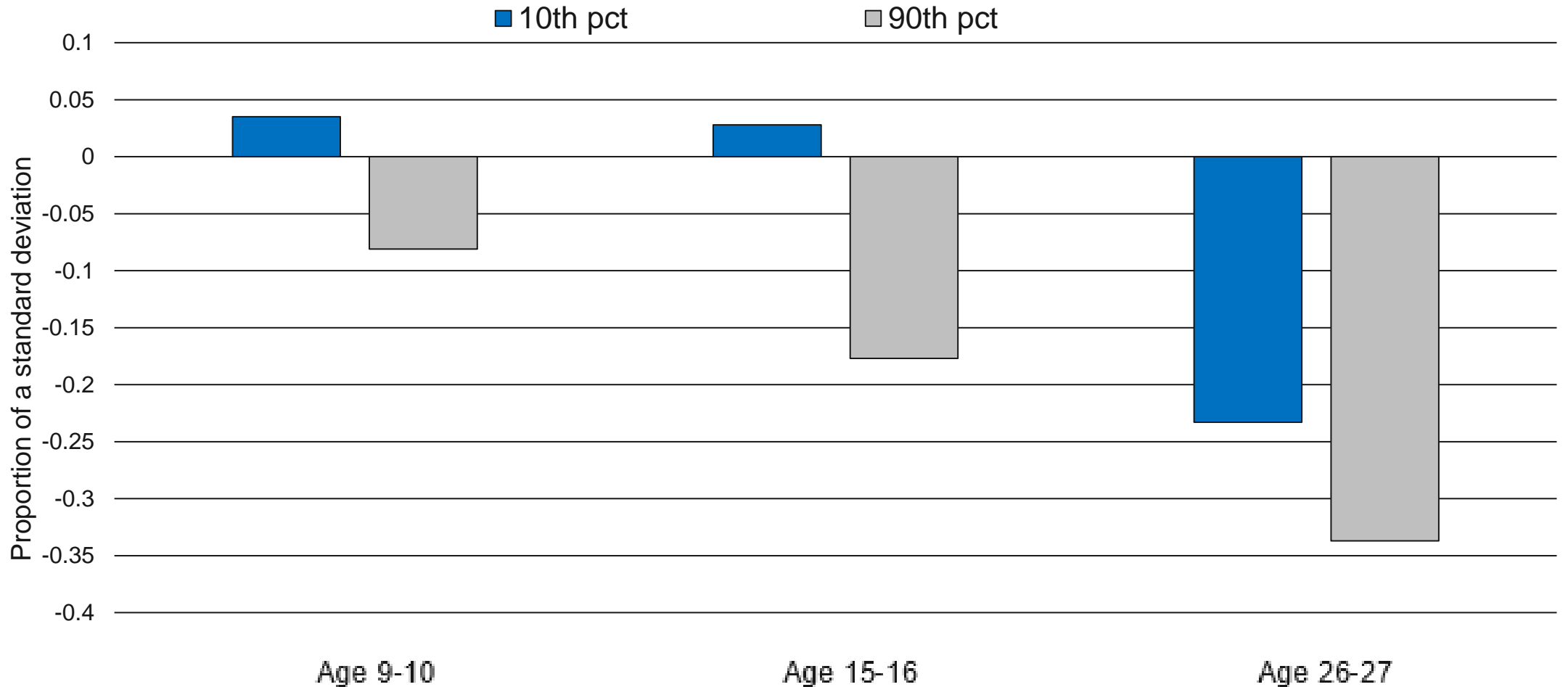
Gender differences in performance decline in reading, PISA 2009
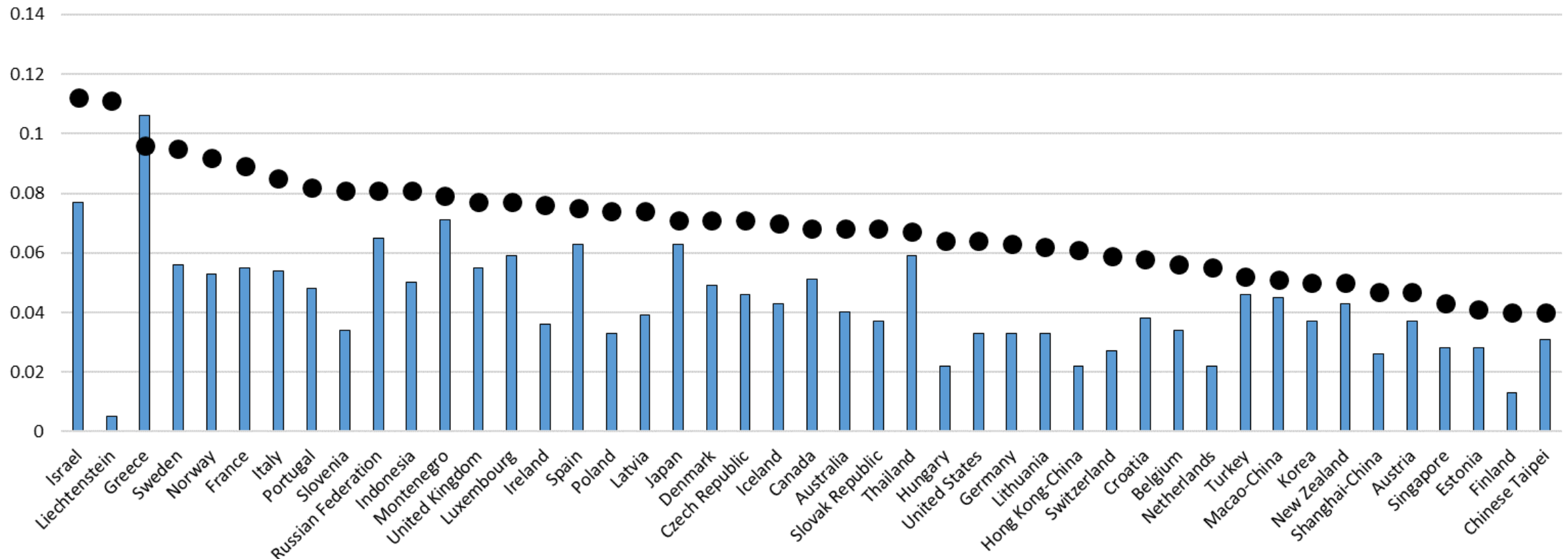
BORGONOVI, F., & BIECEK, P. (2016) An international comparison of students' ability to endure fatigue and maintain motivation during a low-stakes test. *Learning and Individual Differences* 49:128-137.

Performance decline predicts educational qualifications among both genders, PISA 2000 and longitudinal follow-ups

Legend:
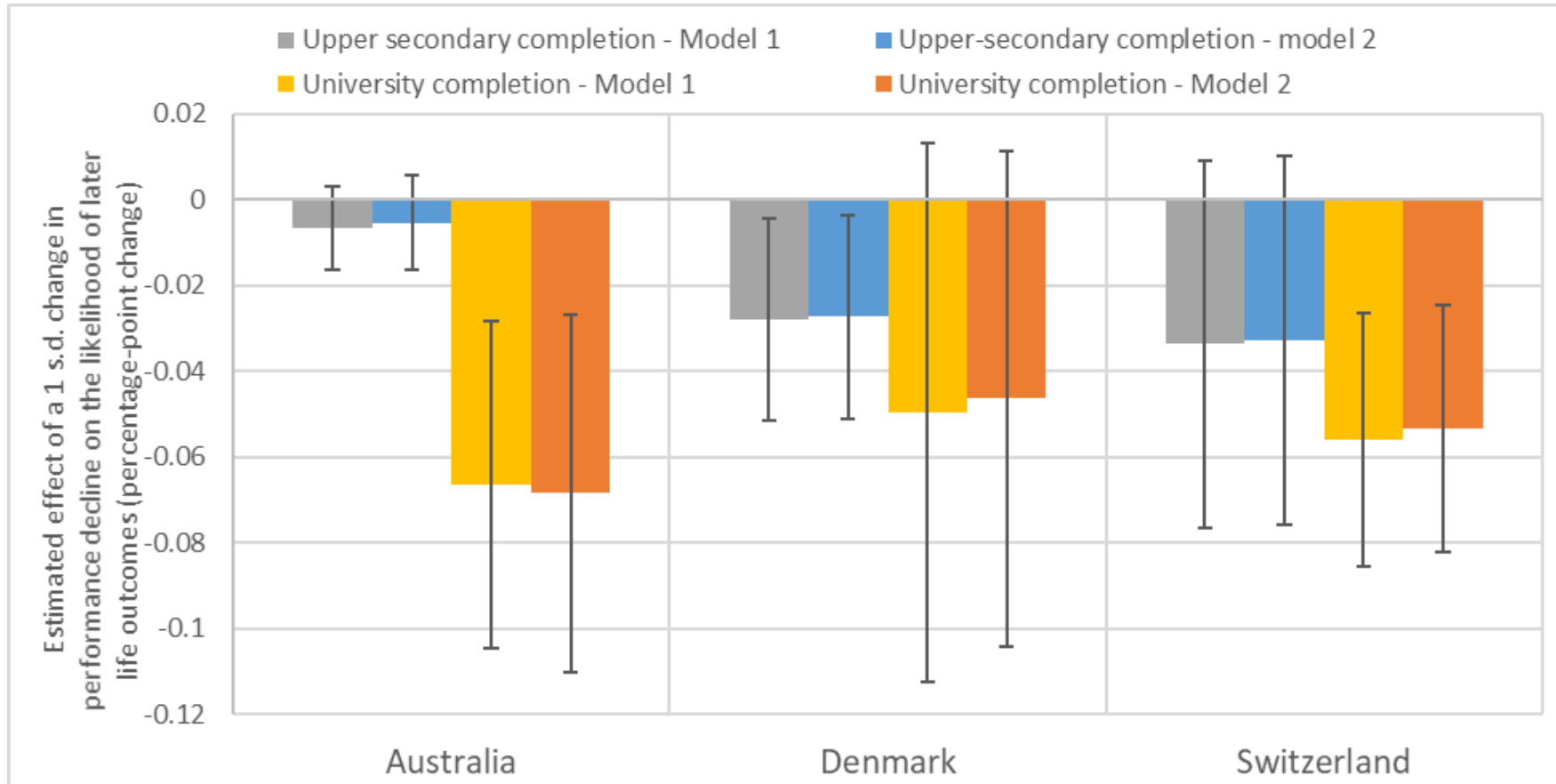- ■ Upper secondary completion - Model 1 (grey)
- ■ Upper-secondary completion - model 2 (blue)
- ■ University completion - Model 1 (yellow)
- ■ University completion - Model 2 (orange)

Y-axis: Estimated effect of a 1 s.d. change in performance decline on the likelihood of later life outcomes (percentage-point change)

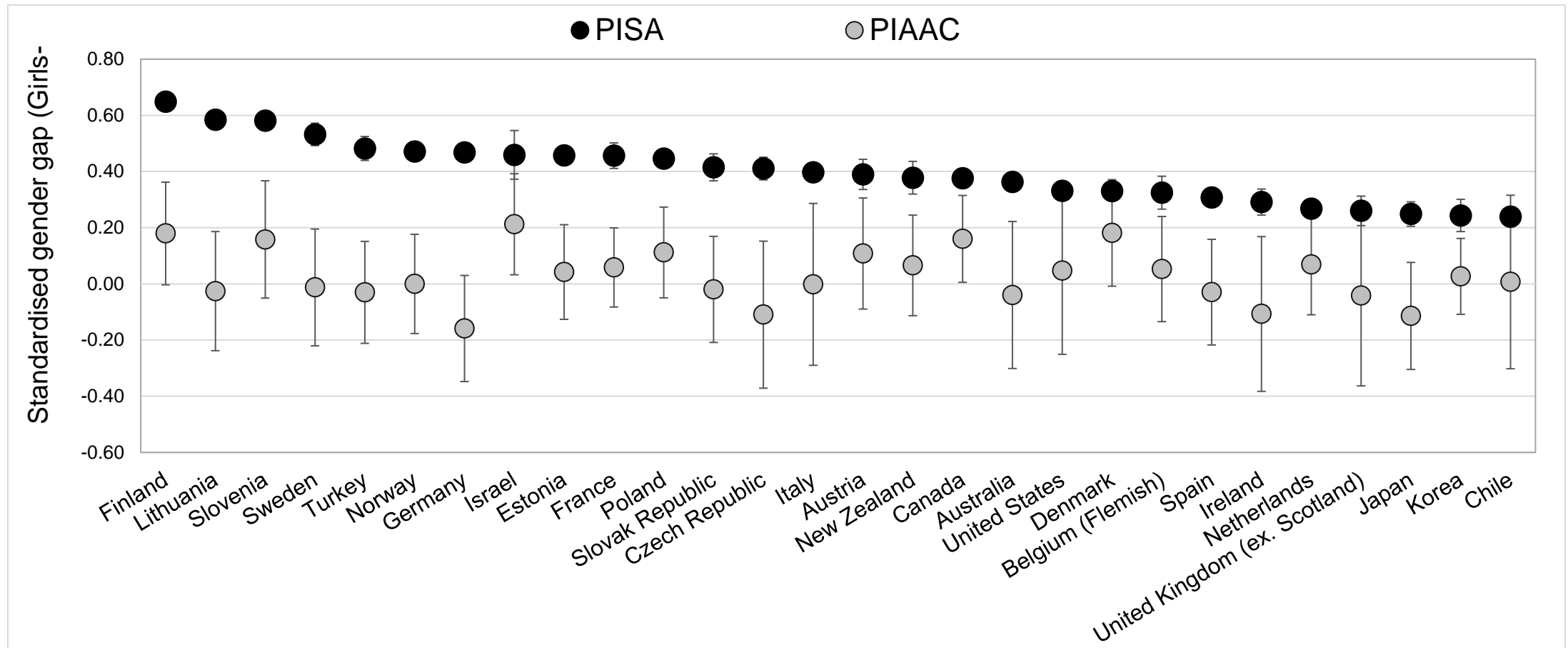X-axis categories: Australia, Denmark, Switzerland

Notes: The figure shows estimates of the effect of a 1 standard deviation change in the performance decline index. Estimates are presented with 95 percent confidence intervals. Model 1 controls for students' gender, immigrant background, age at test, school grade at test, whether they spoke the language of instruction at home, parental occupational class (isei score) and whether they had a parent with tertiary education. Model 2 also controls for self-reported perseverance for Denmark and Switzerland and for self-reported effort in the test for Australia.

BORGONOVI, F., FERRARA, A., & PIACENTINI, M. (2021) Socio-Emotional Skills and Educational Attainment: Cross-Country Longitudinal Evidence. *Journal of Adolescence*

WHERE

The gender gap in literacy, PISA and PIAAC 2012

Note: The PISA main sample refers to 15 and 16-year-olds. The PIAAC main sample refers to 16-20 year-olds (to increase the country specific sample size for PIAAC). The dark dot represents the estimated gender gap (F-M) expressed as a Cohen's d in PISA. The light dot represents the estimated gender gap (F-M) expressed as a Cohen's d in PIAAC. Confidence Intervals at the 95% level for each estimate are presented.

Borgonovi, F. (2022). Is the literacy achievement of teenage boys poorer than that of teenage girls, or do estimates of gender gaps depend on the test? A comparison of PISA and PIAAC. *Journal of Educational Psychology, 114*(2), 239–256. https://doi.org/10.1037/edu0000659

NEW MEASUREMENT POSSIBILITIES

Log-files: A by-product of computer-based delivery

- Log files **store traces** of the **communications** between a user interface and a server
- Can **potentially record all interactions** between the respondent and the testing platform
- **XML language** – strings of text
  - Messy, unformatted, difficult to analyse
  - Mostly designed to serve the purposes of software developers
- Increasing **interest** in the use of log files from **large-scale assessments**
  - Methodological interest
  - Substantive interest

To keep in mind…

- The analysis of log files in social research in education is **opportunistic**

- So far, log files were not collected and designed to be analysed in this way and they do not respond to any **theoretical model of human cognition** (with some exceptions – see VOTAT items in PISA 2012)

Why are log-files valuable?

- **Enrich** the set of information we can collect and to do so **cheaply**

- Enable to "observe" **test-takers' behaviour**

    – Problem-solving strategies and 'non-cognitive skills'

- Can **help item developers** to understand why some items do not work

- Useful to **monitor fielding** of the survey in **real time**

Weaknesses

- **Not everything** we would like to observe is **recorded** in log files

- Content and interpretation of log files inherently **item-specific**

- **Adaptive testing**: respondents assigned to any single items are not representative of any population

- Use and interpretation depend on **access to confidential items**

Complementing log files with physical measurements and observations

MADDOX, B., BAYLIS, A.P., FLEMING, P., ENGELHARDT, P.E., EDWARDS, S.G., & **BORGONOVI, F. (2018)** Observing response processes with eye tracking in international large-scale assessments: evidence from the OECD PIAAC assessment. *European Journal of Psychology of Education*, 33(3): 543-558.

THANK YOU