

# PISA Struggles to Measure All Participants

David Rutkowski & Leslie Rutkowski

The 2018 cycle of the Programme for International Student Achievement (PISA) featured nearly 80 system-level participants from all continents, excluding Antarctica. Participants include all OECD countries, accounting for most of the wealthiest countries in the world and a number of newcomers, such as Belarus whose, GDP per capita falls well below the OECD average. Such a heterogeneous collection of participating educational systems poses challenges in terms of deciding what should be measured and how to measure it in a comparable way.

## The PISA test is too difficult for many countries

In this post, we aim to give you a glimpse of the problem of having such a diverse set of economies served by one assessment. In other words, we will show how a test which was developed to measure wealthy and well-resourced educational systems may be too difficult for lower resourced systems. Based on an empirical analysis of science item difficulty in PISA 2015 mapped against examinee proficiency in several participating education systems, we found that large segments of the proficiency distributions in low performing countries are measured by few or no items<sup>1</sup>. In fact, we found that PISA is only currently designed with a small number of questions that can be used to meaningfully measure the lower performance levels. One consequence of an overly difficult test is that measurement precision is poorer for these participants than for educational systems that are well-matched to PISA.

---

David Rutkowski  
Indiana University, e-mail: [drutkows@iu.edu](mailto:drutkows@iu.edu)

Leslie Rutkowski  
Indiana University, e-mail: [lrutkows@iu.edu](mailto:lrutkows@iu.edu)

<sup>1</sup> <http://doi.org/10.1111/emip.12225>

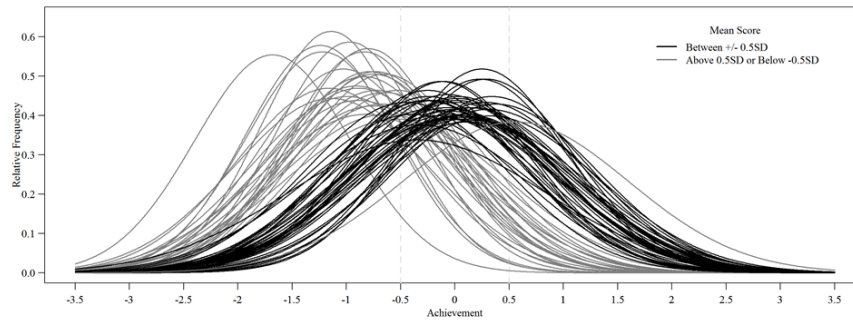
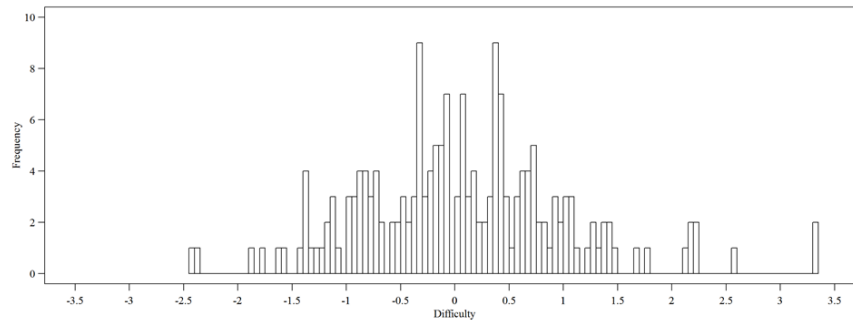
**Fig. 1** Empirical proficiency distributions by educational systems.**Fig. 2** Empirical item difficulty distribution for PISA 2015 science domain.

Figure 1 shows the standardized proficiency distributions across education systems participating in PISA 2015 (science domain). To highlight distribution differences across these educational systems, means that fall within half a standard deviation of the PISA science mean are marked in black. Distributions in gray exhibit means that are more than one half a standard deviation from the PISA average.

Figure 2 shows the distribution of item locations (or difficulty), which are on the same scale as achievement (represented in Figure 1). Figure 2 shows a high concentration of items, centered about zero, with far less representation away from this center point, suggesting that educational systems that differ meaningfully from zero will be measured by fewer items.

In other work we have shown that in some low performing countries more than half the population are not given items aimed at measuring their ability levels. In other words, PISA is not measuring low performing systems well<sup>2</sup>. Although, we are hopeful that as PISA moves to an adaptive design, that these issues can be ame-

<sup>2</sup> <http://doi.org/10.1080/0969594X.2019.1577219>

liorated in future cycles. We further argue that an expanded framework should also be adopted for low performing countries.

### **We do not even know if PISA measures the same construct in low performing countries**

Another issue that we have uncovered in our work is that the statistical tools used to determine whether items function equivalently in all countries are unable to detect poor fitting items in low performing countries<sup>3</sup>. All of these issues take on increased importance when we recognize that the number participating countries increases with each PISA cycle. Clearly, there is much work to be done to understand the degree to which international assessments are fit for purpose; however, our recent research calls into question the ability to apply a one size fits all assessment to a heterogenous world.

This text has been posted on the blog [international-education.blog](https://international-education.blog) and it is available in different languages on [international-education.blog](https://international-education.blog)

Rutkowski, D., & Rutkowski, L. (2020). PISA struggles to measure all participants. Retrieved from <https://international-education.blog/en/?p=1813>

The blog International-Education.Blog is an outcome of the project OCCAM. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 76500.



---

<sup>3</sup> <http://embracingheterogeneity.com/>