

Discussion: The use of test scores in secondary analysis

Edwin Leuven

University of Oslo

PIAAC Methodological Seminar, OECD Paris, 14th June
2019

Introduction

I will be making two broad points

1. we don't always know the model beforehand
2. we don't always know what we are measuring

and will argue that this should be reflected in data production

The starting point:

- ▶ how PIAAC provides test scores

The end point:

- ▶ an observation about secondary analysis

Item Response Theory

PIAAC scores are so-called plausible values which are constructs computed using IRT

But IRT is **a model (not a theory)**

$$\Pr(R = 1|\theta; a, b, c) = c + (1 - c) \frac{\exp(a(\theta - b))}{1 + \exp(a(\theta - b))}$$

What if

- ▶ the world is not like this model? or
- ▶ a researcher wants to estimate another model?

We don't always know the model beforehand

Plausible Values (PV) are:

Random draws from the posterior distribution of θ given item responses **and student background characteristics Z** :

$$PV \approx E[\theta | R_i, Z_i] + \text{noise}$$

We can regress PV on X as long as X is a part of Z

But what is in Z ?

We don't always know the model beforehand

Let's check Chapter 17, "Scaling PIAAC Cognitive Data", of the Technical report:

- ▶ p.4: citizenship, gender, immigrant status.
- ▶ p.5: gender, country of birth, education, occupation, employment status, reading practices, etc.
- ▶ p.9: gender, age, educational attainment and country of birth.
- ▶ p.20: demographic information, educational experiences, occupational experiences and skill use, among others.

What is in Z ?!

And what if I don't want to control for everything in Z ?!

What to do about scaling?

Ultimately scaling is a model dependent and **normative** choice

1. by changing the scaling we change the model (i.e. change the outcome or the treatment)
2. through scaling we put more or less weight on different parts of the distribution
 - ▶ what margins are important?

To make these choices we need to

1. know the model
hard to do beforehand & one-size-doesn't-fit-all
2. understand what we measure
documentation, test items

What are we measuring?

“Problem solving in technology rich environments”

Three broad dimensions

1. cognitive
2. task
3. technological
 - ▶ Hardware devices: Desktop or laptop computers, mobile phones, personal assistants, geographical information systems, integrated digital devices
 - ▶ Software applications: File management, Web browser, Email, Spreadsheet
 - ▶ Commands, functions: Buttons, Links, Textboxes, Copy/Cut-Paste, Sort, Find
 - ▶ Representations: Texts, Sound, Numbers, Graphics (fixed or animated), Video

What are we measuring?

“Problem solving in technology rich environments”

Unit 10 - Part 1

You are looking for a job and have located these five websites.

You want to use a site that does not require you to register or pay a fee.

Bookmark all the sites that meet your requirements.

Once you have bookmarked the sites, click Next to go on.

The screenshot shows a web browser window titled "Web". The address bar contains the URL "www.websearch.com/jobsearch". The browser has a menu bar with "File", "Edit", "Bookmark", and "Help". Below the menu bar are navigation icons: back, forward, home, search, and refresh. The main content area is titled "Web Search" and features a magnifying glass icon and a search box labeled "Job search". Below the search box, there are five job search websites listed with their descriptions and URLs:

- [Find Your Job - JobSearch.com](#)
The best job search site on the web. Check with us first!
[www.jobsearch.com](#)
- [Work Links](#)
We connect you with the best jobs on the web.
[www.worklinks.com](#)
- [Looking for a job?](#)
Start your job search here.
[www.careerstarters.com](#)
- [Connections.com](#)
We provide access to the best jobs
[www.connections.com](#)
- [The best jobs online](#)
If you are looking for the perfect job, start right here.
[www.greatjobs.com](#)

At the bottom of the browser window, there is a status bar with the word "Web" and some icons.

Why care about your PIAAC score?



Is it only another

product?

(like IRT?)

Or is it something more?

Why care about your PIAAC score?

Can we care if we do not know how PIAAC scores fit in the **causal chain**:

- ▶ how does parenting, education, work, ... affect PIAAC scores?
- ▶ how important are innate factors?

What matters?

- ▶ **Observed skills:**
I know how to use a web browser
- ▶ **Potential skills:**
I can learn how to use a web browser

Where is the secondary analysis?

Returns to skills around the world: Evidence from **PIAAC**

[EA Hanushek](#), [G Schwerdt](#), [S Wiederhold](#)... - [European Economic ...](#), 2015 - Elsevier

☆ [🔗](#) Cited by 309 [Related articles](#) [All 9 versions](#)

PIAAC numeracy: A conceptual framework

[PIAAC Numeracy Expert Group](#) - 2009 - [repositorio.minedu.gob.pe](#)

☆ [🔗](#) Cited by 98 [Related articles](#) [All 10 versions](#) [🔗](#)

[HTML] **PIAAC**: A new strategy for assessing adult competencies

[A Schleicher](#) - [International Review of Education](#), 2008 - Springer

☆ [🔗](#) Cited by 104 [Related articles](#) [All 14 versions](#)

Technical report of the survey of adult skills (**PIAAC**)

[Organisation for Economic Co-operation and ...](#) - 2013 - [voced.edu.au](#)

☆ [🔗](#) Cited by 55 [Related articles](#) [🔗](#)

PIAAC problem solving in technology-rich environments: A conceptual framework

[PIAAC Expert Group in Problem Solving in Technology ...](#) - 2009 - [oecd-ilibrary.org](#)

☆ [🔗](#) Cited by 40 [Related articles](#) [All 5 versions](#)

[HTML] Measuring occupational mismatch: overeducation and overskill in Europe —evidence from **PIAAC**

☆ [🔗](#) Cited by 45 [Related articles](#) [All 15 versions](#)

Summary

PIAAC leans heavily on IRT/PV

- ▶ downsides: model dependent, obscures issues, complicated, limits researcher degrees of freedom, expensive
- ▶ upsides: ?

Economists are of-the-shelf users of test scores

- ▶ make test material more readily available
- ▶ educate us on the substantive nature of the test

Economists are producers of models

- ▶ provide plug-in estimates of individual competencies