Middlesex University

Reading the PIAAC Results: what to look out for

Jeff Evans j.evans@mdx.ac.uk

ACAL Annual Conference, UTS Sydney, October 2013

Plan

0. Brief history of PIAAC (Project for International Assessment of Adult Competencies, aka Survey of Adult Skills)

1. Global context of developments in educational policy

2. Concept of adult numeracy and other methodological issues

3. Examples of results (from Australia 2013)

4. Presentation of results and possible effects

Introduction

PIAAC builds on earlier IALS (1990s) and ALLS (2002-06)

- *Literacy, Numeracy*, and *Problem solving in TRE*
- Unlike earlier surveys, uses computer administration, allows **'adaptive routing**', to find appropriate "level" of respondent
- methodological & fieldwork improvements, e.g. regulation of sampling and fieldwork standards.

First round in 2011-12, full results available in Oct. 2013

- larger sample of 24 "industrial" countries (more languages)
- Unlike PISA (& TIMSS): combines household survey methodology with educational testing
- adults usually 16-65: 5000+ sampled per country

Educational policy context

Educational policy currently developing on a world-wide scale

- Lifelong Learning (LLL) contested (Evans, Wedege, Yasukawa, 2013): OECD /PIAAC vs. UNESCO /LAMP approaches

OECD aims to promote broad view of learning:

- development of knowledge, enabling citizens' active participation in all spheres of social / economic life
- stress on citizens' need to acquire *and update* abilities, attitudes, knowledge & qualifications over life-course
- shift in focus of learning from "what people know" to "what they can do": *competence* → *performance* (Bernstein, 2000)
- weakening distinction between formal and informal learning

Role of international organisations (1)

EU working closely with OECD on PIAAC, disseminating ideas & practices that strongly influence national policy making around the world (Dale & Robertson, 2009)

- in context of 'Lisbon agenda', EU aims for both social cohesion and economic competitiveness: termed 'inclusive liberalism'
- Nations seek competitive advantage, defined in terms of national education and training systems judged by international standards, measured by international surveys

(Tsatsaroni & Evans, 2013)

Role of international organisations (2)

OECD and EU strategies:

- constructing 'the skills and competencies agenda' … (Ozga et al., 2011)
- promotion of "expertise in creating comparable datasets … countries can measure *relative* success of education systems & shift policy orientations accordingly" (Grek, 2010).
- new forms of 'soft governance' of national educational systems, via dissemination of knowledge, publication of comparative social indicators, country & thematic reviews: 'governing by data' (Mahon & McBride, 2008; Ozga, 2009)

PIAAC aims

Education Directorate at OECD: helping countries to:

- Identify and measure differences between individuals and across countries in key "competencies"
- Relate measures of skills based on these competencies to: *individual outcomes* , e.g. labour market participation / further learning / earnings; *aggregate outcomes*, e.g. econ. growth, or social equity in the labour market
- Assess performance of education / training systems, to enhance competencies through formal educational system – or in the work-place, through incentives (Schleicher, 2008)

Thus, a "human capital" approach + social concerns … presupposes competitive international economy

(Evans, Wedege & Yasukawa, 2013)

PIAAC concepts and measures (1)

OECD: *competencies:* internal mental structures, i.e., abilities, capacities or dispositions embedded in the individual […] cognitive skills & knowledge base are critical elements,

[but] important […] to include other aspects such as motivation and value orientation.

Numeracy: the ability to access, use, interpret, communicate mathematical information & ideas, to engage in / manage mathematical demands of a range of situations in adult life.

(PIAAC Numeracy Expert Group, 2009)

PIAAC concepts and measures (2)

To produce measures, must characterise *Numerate behaviour:* dimensions used in construction / validation of set of items:

- context (4 types): everyday life, work, societal, further learning
- response (to mathematical task – 3 main types): identify / locate / access (information); act on / use; interpret / evaluate.
- mathematical content (4 main types): quantity & number, dimension & shape, pattern & relationships, data & chance.
- representations (of mathematical / statistical information): e.g. text, tables, graphs.

Also Background Questionnaire: demographic & attitudinal information

And Job-Related Assessment: use of / need for skills at work

Methodology (1)

- the *content validity* of the definitions of numeracy and numerate behaviour ['types' of items]
- the *measurement validity* of the items presented, including the administration and scoring procedures ['qualities' of items]
- the *reliability* of the measurement procedures
- the *external validity*, or representativeness, for the national population of interest, of the results produced from the sample.

… Similar dilemmas for most educational assessment.

- *and for both Qual. and Quant. educational research*

Methodology (2)

Content validity: the extent to which a measure represents all facets of a given concept: … Here definition of numeracy based on 4 dimensions of numerate behaviour stipulated: *context, content, response, representation*.

Each item can be categorised on these four dimensions, and the proportion of items falling into each category can be controlled over the scale, so as to enhance the transparency of the operational definition.

However, a transnational definition …(*generalising*) ... How well does it "fit" adults' lives in any particular country?

Further, the four types of *context* (everyday, work, society and community, further learning) are *under-specified*: rather general to refer to any actual specific social practice or social context, in *any particular respondent's* everyday life.

(Evans Wedege & Yasukawa, 2013)

Methodology (3)

Measurement validity: extent to which person's responses to set of items actually capture what the conceptualisation of numeracy specifies

- depends on the *actual range of items used*: see below 3 illustrative items presented by OECD (2013) / on websites (e.g. CSO Ireland)
- Also requires procedures designed for administration of the survey to be standardised in advance across all countries, e.g. design specs. of the laptops & software to be used, and rules for access to calculators and other aids.
- Full appreciation of the validity of procedures requires assurance of how these procedures are followed in the field … even more crucial when results are compared across countries using different fieldwork teams.

Methodology (4)

External validity: includes representativeness of sample for the "population "… check a country's sample design + other fieldwork aspects, e.g. *incentives* for completing interview … & judgments depend on knowing about actual field practices.

SO *any summaries*, e.g. mean scores, or gender differences, are **sample-based estimates for the population value (**of the mean or size of gender difference or ...) for the country

These ***interval estimates not exact***, but show ***a margin of error*** [say, ***2* standard errors***, on either side -* depends on the ***level of confidence*** desired in the estimate] →surprises

e.g. PIAAC numeracy: overall country results 2013

Japan = 288 Finland = 282 NL / BELG = 280

286 to 290 280 to 284 278 to 282 (overlap !)

Methodology (5)

Reliability of test administration across countries and across interviewers, especially assuring same standards / practices in marking (problem with past international surveys) …

Computer presentation and marking will help greatly.

But it may tend to undermine *construct validity*, if it reduces that range of types of question that can be asked (example)…

And, increasing the reliability may lead to concerns about *ecological validity*, whether the setting of the research is representative of those to which one wishes to generalise the results. For example, *on-screen presentation* may limit this?

Presentation of Results (1)

Adult's performance not just expressed as the 'proportion correct', since adaptive routing → some got 'harder' items

So Item Response Theory (IRT) used to ('psychometrically') estimate a standardised score (e.g. mean 250, std dev 50)

Then, numerical score commonly related to one of 5 general 'levels' of literacy or numeracy to make it meaningful …

BUT simple and one-dimensional sense … e.g. "levels embody predetermined assumptions about progression and relative difficulty" (Gillespie (2004) referring to UK Skills for Life)

•Partly because many adults have different "spiky profiles", due to distinctive life experiences: some find type A items ("data & chance") more difficult; others find type B items ("dimension & shape").

Presentation of Results (2)

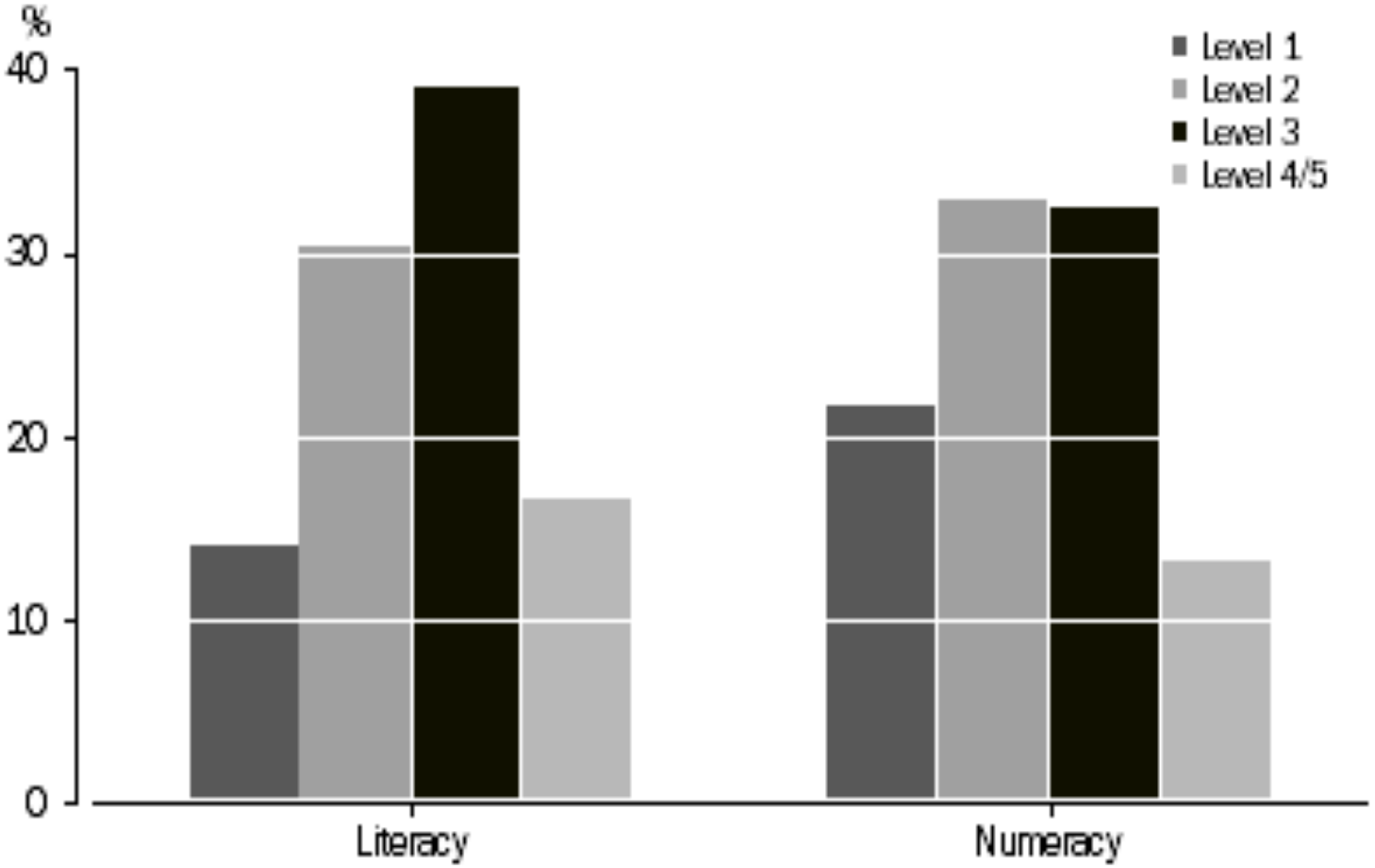
OR some policy makers attempt to stipulate “minimum level of numeracy needed to cope with the demands of adult life” in particular country (BUT see Black & Yasukawa, 2013 on Australia)…

HOWEVER, this

conflates adults with different work, family and social situations

+ tends to assume demands are the same across the countries studied.

PIAAC preliminary results - Australia



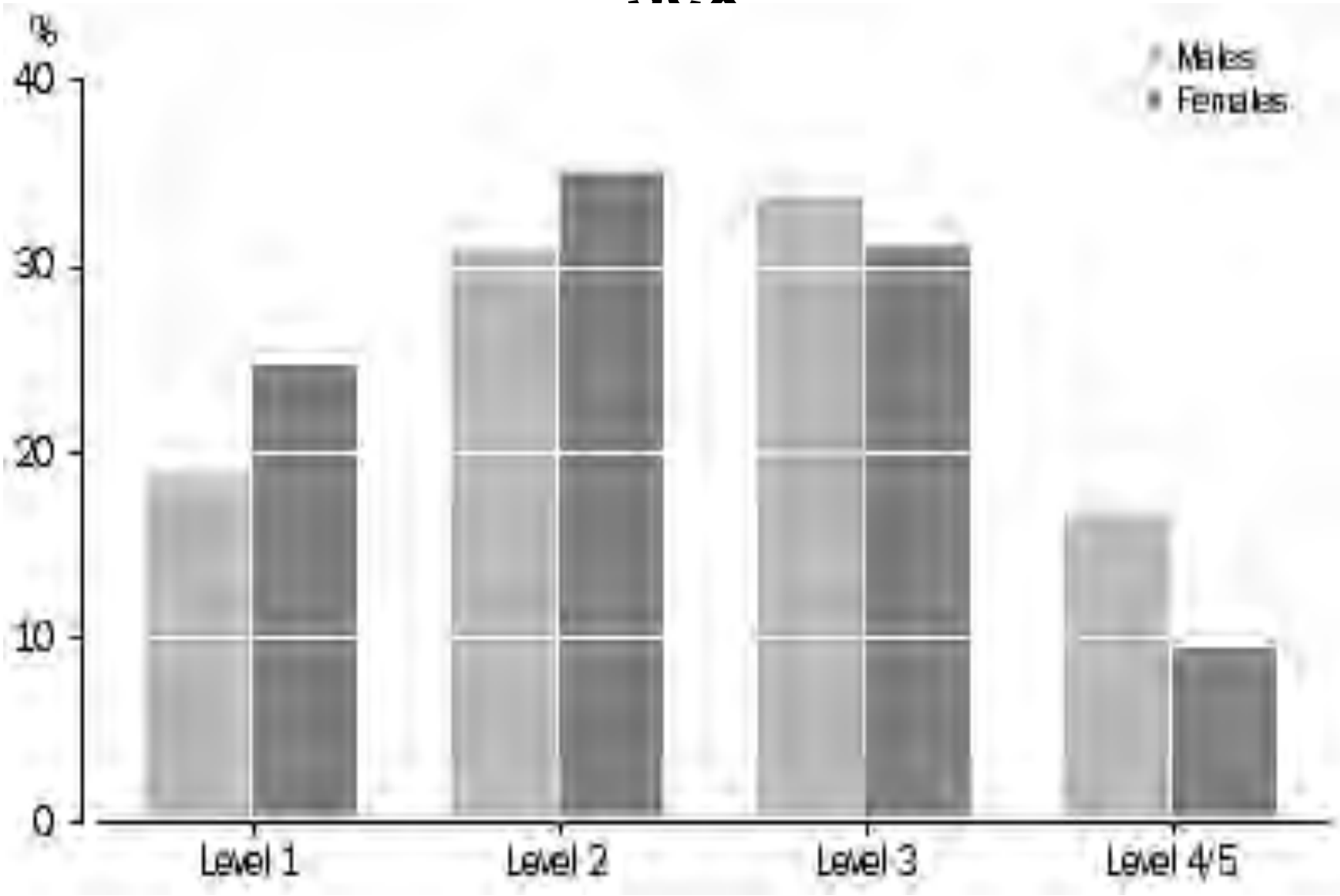
Proportions at different skills levels

Approximately 7.3 million (44%) Australians aged 15 to 74 years had literacy skills at Levels 1 or 2, a further 6.4 million (39%) at Level 3 and 2.7 million (17%) at Level 4/5. For the numeracy scale, approximately 8.9 million (55%) Australians were assessed at Level 1 or 2, 5.3 million (32%) at Level 3 and 2.1 million (13%) at Level 4/5.

One might ask about regional variation …

For numeracy, the Australian Capital Territory recorded the highest proportion at Level 4/5 (23%). ... (Why?)

Proportion at each numeracy level, by sex



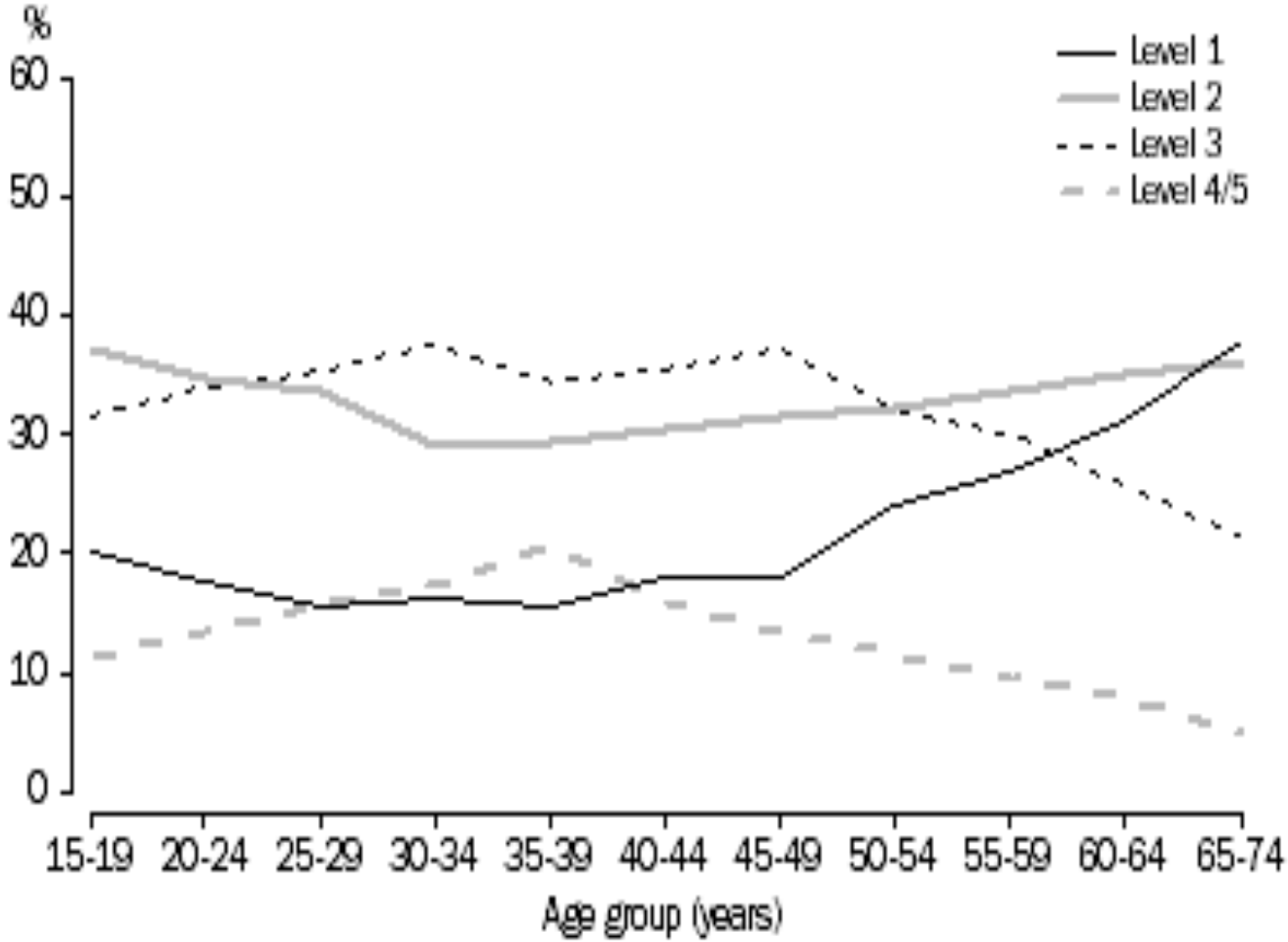
Interpretation

... little difference in the proportion of males and females at each level of the literacy scale.

However a higher proportion of males (17%) attained scores at Level 4/5 on numeracy scale compared with females (9%).

... Why?

Proportion at each numeracy level, by age



Interpretation

The proportions of people at Level 1 are higher among the oldest age groups (people aged 60 years and older) than in the younger age groups (people aged 15 to 29 years) for numeracy skills [and also literacy, not shown here].

... Other questions?

Does this indicate, as sometimes claimed, "a person's skills deteriorate over the life-course"?

Other results: early impressions

1. Within-country results complex → field day for media, politicians, spin-doctors, but

1a. ... Some laudable and some regrettable finding for almost everyone

2. Much discussion of age /intergenerational differences – patterns vary widely

3. Beware: many of the interesting findings are **correlations**, but not necess'ly causal

4. Between country results 'striking' but even harder to control for … cultural differences

Discussion

1. Bernstein (2000) et al.: critique a type of global pedagogic discourse, that asserts adults' need for certain *generic skills*, and countries' need to assess these same.

Earlier, we thought there was "a strong possibility that PIAAC could reinforce this type of pedagogic discourse, and the international surveys could work as a *curriculum type* which would indirectly prescribe *what knowledge the adult populations in all societies should value, strive to acquire*, and demonstrate" (Tsatsaroni & Evans, 2013).

But later Christine Pinsent-Johnson's paper (*Literacy as Numbers*, June 2013) argues this has already materialised in *Essential Skills* (Canadian provinces)….

Danger is that such a curriculum may 'crowd' out desirable *particularities* of a more 'local' one.

Discussion

2. Possible effects of such high stakes performance surveys on future research

- pathologisation of 'low-performing' countries
- unequal distribution of disciplinary (rather than 'functional') forms of knowledge may help to reassert the division and social distinction between those who are knowledgeable, creative in mathematics and those who are destined to fail, or need constant retraining

2a. may encourage *more narrow notions of numeracy*

... and may distract people from seeing the importance of '*powerful (mathematical) knowledge'* (Young, 2010)

e.g. the idea of *conditional probability* (Gigerenzer, 2003)

Discussion

3. Opportunities for further research linked to int'l surveys:

- OECD policy is helpful … to make available national datasets from the surveys on PIAAC website at the same time as overall results.
- Though results anonymous at individual level, much potential for relating of performances to *categories of respondents* – via demographic and attitudinal data, and data on "use of skills" at work (Background Q're & Job Requirements Assessment, respectively)

Discussion

3. (cont'd) Examples of future research topics:

3a. Why are numeracy levels higher in Canberra than in other states & territories in Australia ? E.G. Higher educational qualifications – or higher levels of numerate experience at work?

3b. Why do a higher proportion of males (17%) attain scores at Level 4/5 in Australia on numeracy scale compared with females (9%)? ???

3c. Why are the proportions of people at Level 1 highest in the oldest age groups (people aged 60 years and older)? *Does this indicate, as sometimes claimed, that "a person's skills deteriorate over the life-course"?*

Discussion

3d (cont'd). Probably need independent national studies or linking studies between PIAAC and national surveys

3e. Much scope for local qualitative studies, to supplement/ probe Background Q're results [e.g.]

e.g. I like learning new things: Not at all / Very little / To some extent / To a high extent / To a very high extent / Does not know / Refusal

4. Some evidence of use of results from earlier international studies e.g. PISA, TIMSS (Tsatsaroni & Evans, 2013, *ESM*)

Appendix

The current 24 participating countries in PIAAC include: 17 EU members, plus USA, Can., Aus, Japan, Korea, possibly Russian Federation. Developing countries are not involved in Round 1, including BRIC (except Russia),

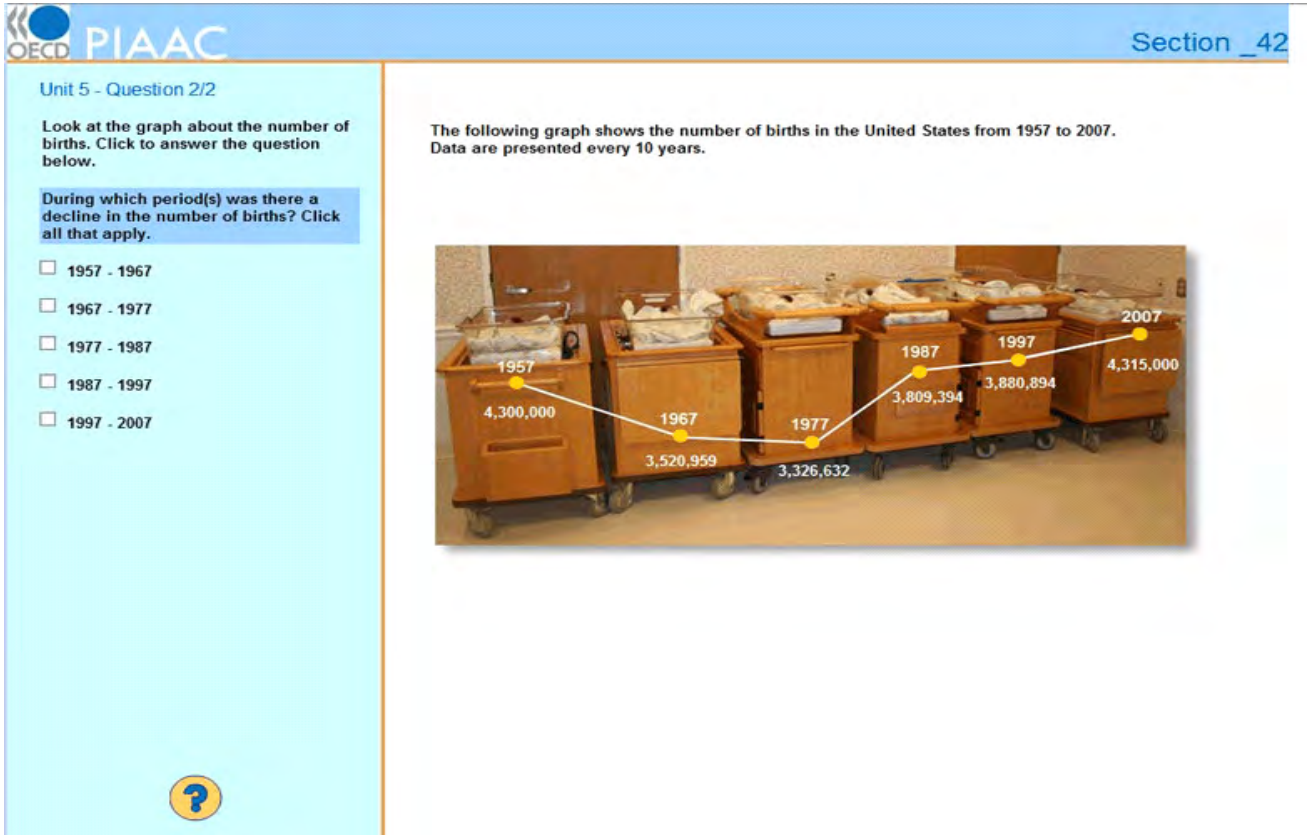
And Round 2 includes: New Zealand, Slovenia, Greece, Portugal, Chile, Mexico, Indonesia, Turkey, Singapore. Results expected in 2016.

Illustrative items (next 3 slides): taken from OECD (2013), *The Survey of Adult Skills: Reader's Companion*, pp28-30. Available online.

Numeracy – Sample Item 1

This sample item (of difficulty level 3)focuses on the following aspects of the numeracy construct:

Content	*Data and chance*
Process	*Interpret, evaluate*
Context	*Community and society*

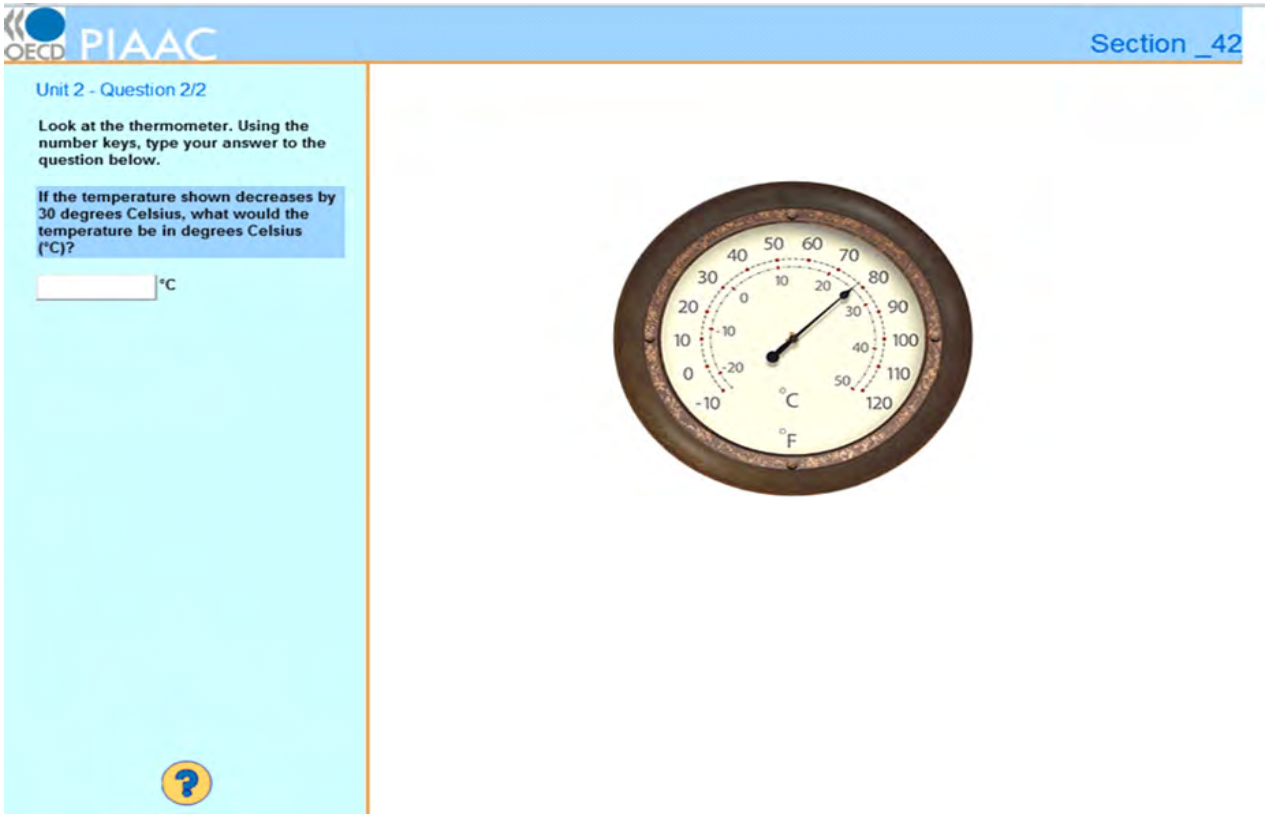


Correct Response: 1957 - 1967 and 1967 – 1977

Numeracy – Sample Item 2

This sample item (of difficulty level 1)focuses on the following aspects of the numeracy construct:

Content	*Dimension and shape*
Process	*Act upon, use (measure)*
Context	*Every day or work*

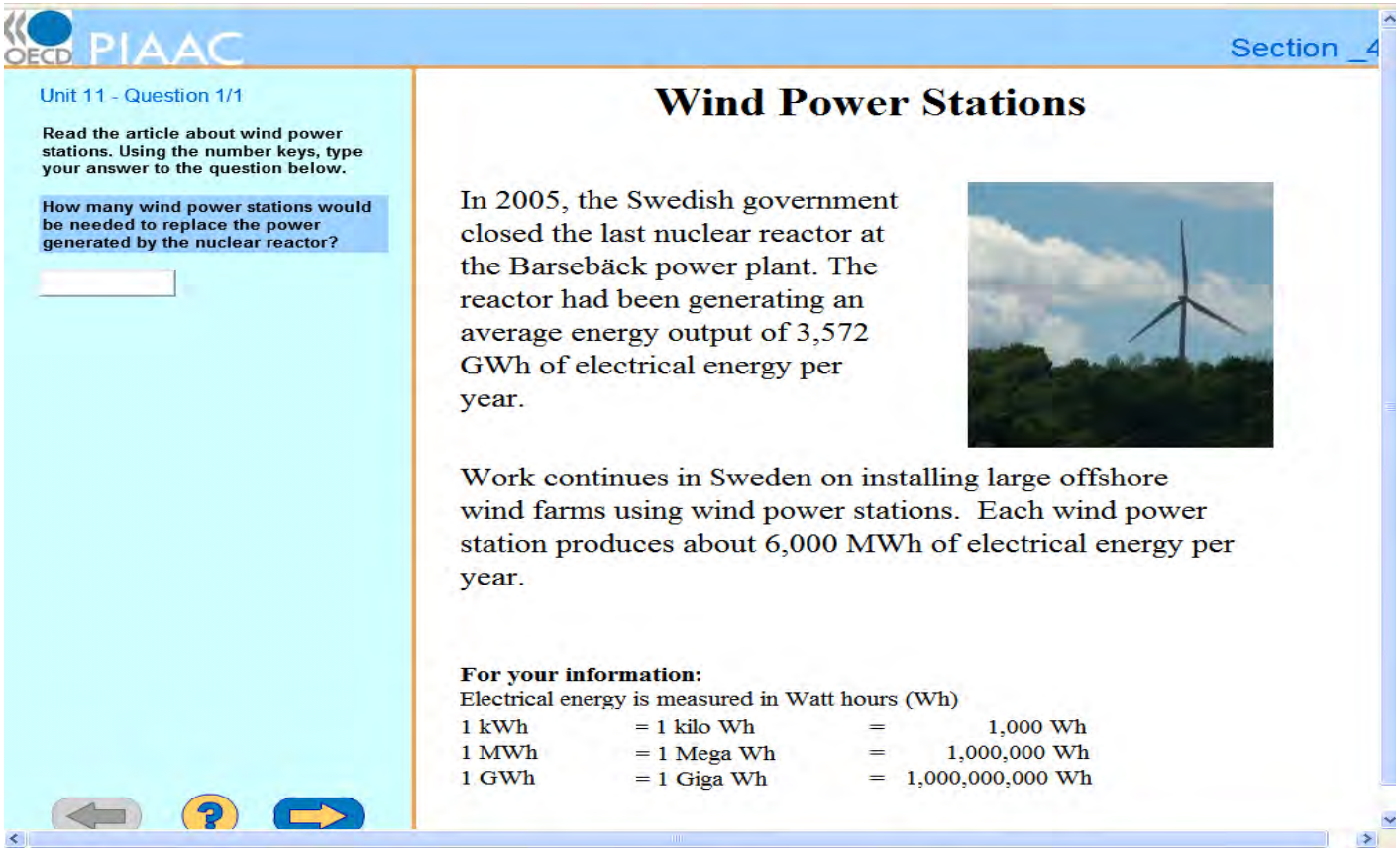


Correct Response: Any value between -4 and -5

Numeracy – Sample Item3

This sample item (of difficulty level4) focuses on the following aspects of the numeracy construct:

Content	*Quantity and number*
Process	*Act upon, use (compute)*
Context	*Community and society*



Correct Response: One of the three values (no values between): 565, 596 or 600.

Acknowledgements

Tsatsaroni, A. & Evans, J. (2013), Adult Numeracy and the Totally Pedagogised Society: PIAAC and other international surveys in the context of global educational policy, *Educational Studies in Mathematics, Special Issue on Social Theory in Mathematics Education* [DOI: 10.1007/ s10649-013-9470-x].

Evans, J., Wedege, T. & Yasukawa, K. (2013). Critical Perspectives on Adults' Mathematics Education; in M. A. Clements, A. Bishop, C. Keitel, J. Kilpatrick and F. Leung (Eds.), *Third International Handbook of Mathematics Education,* New York: Springer.

Additional References

Hamilton, M. (2011). *Literacy and the Politics of Representation*. London: Routledge.

OECD (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills.* Paris: OECD. Online: ... .

OECD (2013). *The Survey of Adult Skills: Reader's Companion*. Paris: OECD. Online: ... .

Pinsent-Johnson, C. (2013). How learning inequality is made to happen: Curricularizing an international literacy test and classifying adult learners. Paper given at *Literacy as Numbers* conference, Univ. London Institute of Education. Online: