



PISA: Myths and Accuracy



‘Let me tell you a story’

*Using PISA, but retaining complexity and sensitivity in
domestic policy formation*

Tim Oates CBE

Cambridge Assessment

Key points

Three points of argument

- 1 for policy formation, we need to look backwards in time as much as looking forward
- 2 we must avoid observer bias
- 3 assertions on 'future skills' need to be better theorized and underpinned
- 4 the Dipf model may be one worth emulating

Methodological concerns

Item quality and translation, equating method (Hodgen J; Benton T; Solan-Flores G; OECD)

Sampling: stratification (Benton T; OECD); exclusion (Wuttke J; OECD)

Measurement model (Jerrim J; Goldstein H; Kreiner; OECD)

Mode effects: on screen testing (Heller Sahlgren G; Jerrim J; OECD)

TIMSS-PISA contrasts (Gronmo L; Oates T; Kleime E)

Extrapolation (Benton T; Oates T)

Limitations

A single measurement point at 15 years' of age

Multiple layers of extrapolation

Extrapolation – theory laden – research referenced

Partial observation – information loss – observer bias

Complex systems and relationships

Critical realist perspectives

Bhaskar; Sayers; Lawson; Hodgson

Limitations



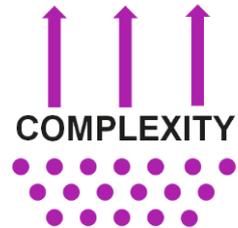
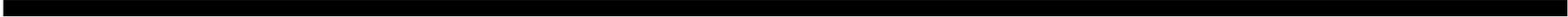


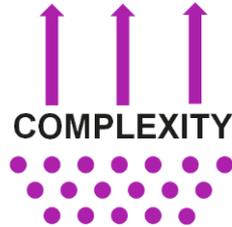






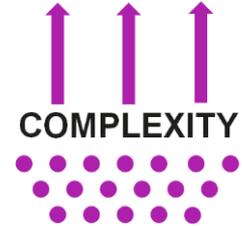
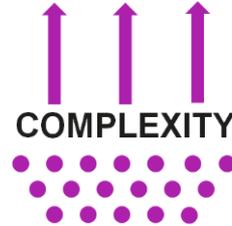
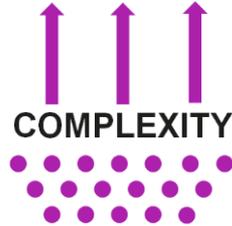
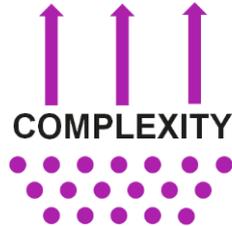
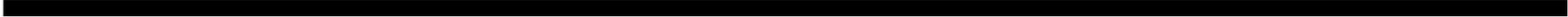








AUTONOMY



Autonomy

A topline message from OECD; widely discussed

Three versions present in the discourse

1 Autonomy is associated with high performance

2 Autonomy with accountability

3 Autonomy with accountability and a means of disseminating good practice

Effects in UK associated with autonomy are not present with re-analysis (Benton T)

Accountability can exist in very diverse and subtle ways (Oates T)

Association of high autonomy and high performance not consistent with historical evidence (Oates T, Heller Sahlgren G)

Looking backwards as well as forwards

Did significant improvement in education systems happen prior to the advent of PISA and TIMSS?

Germany, USA, France, England – (Green A)

Japan – (Stigler J; Dore R; Saito H)

Shanghai – (Jinjie X; Shen X)

Modern cases

Singapore

Massachusetts

Finland

Massachusetts

The first editions of the Massachusetts Curriculum Frameworks were adopted by the Board in:

1996 for Mathematics (Math), Science & Technology Engineering, Arts, Comprehensive Health, and World Language

1997 for English Language Arts (ELA) and History/Social Science

1999 for Foreign Languages

2003 for English Language Proficiency Benchmarks and Outcomes

2006 for Vocational Technical

2008 for Kindergarten Learning Experiences

The standards for mathematics were significantly updated in 2000 and 2004 and ELA in 2001 and 2004 based upon lessons learned since implementation of the original frameworks and updated research. While a combination of federal grants and state appropriation provided some support, upgrades to local curricula and associated professional development were led and supported by local school districts.

Singapore

Trace policy model to 1947 Ten Year Programme and 1950-60s policy on linking educational development to economic development

16th place of 26 nations in SISS 1984 to 1st place in TIMSS 1995

Strong emphasis on ‘curriculum coherence’ in Bill Schmidt’s strict definition of alignment of instruction, instructional materials and curriculum content/aims

Global observation, constant innovation, close observation

Singapore Maths is a very interesting case study of curriculum development and improvement

High levels of interaction and ‘interplay’ between central bureaucracy and schools – it is not consistent with a simplistic ‘high autonomy’ model

Massachusetts

US PISA 2012

31st in Maths

24th in Science

21st in Reading

Mass

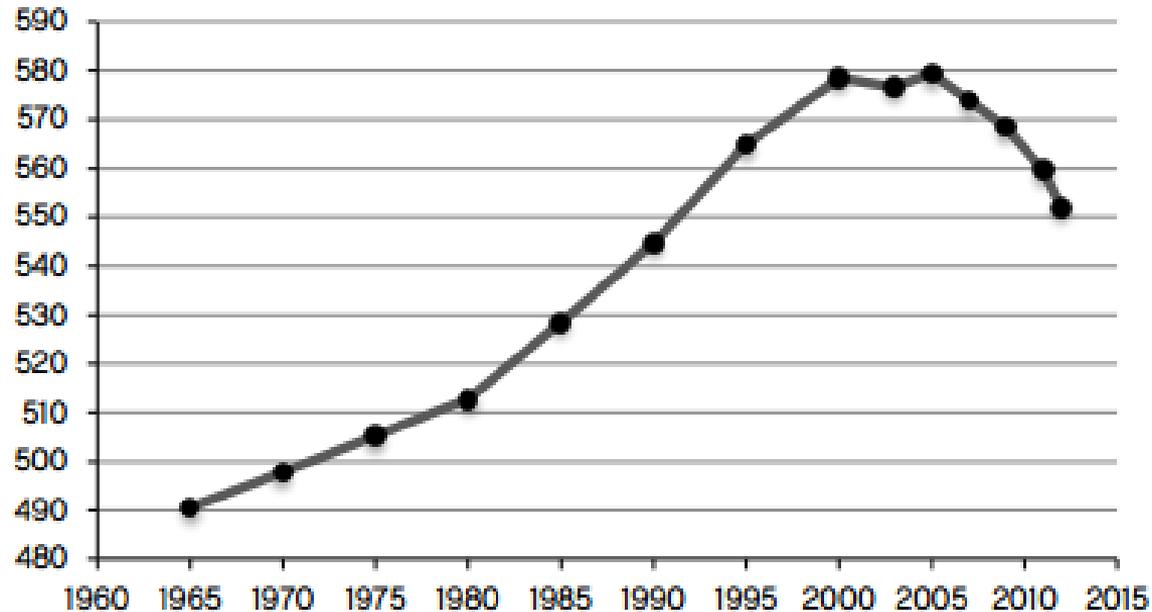
9th in Maths – tied with Japan and one place after Switzerland

4th in Reading – tied with Hong Kong and one place after Finland

In TIMMS Science Mass second to top-ranked Singapore

Finland

Chart 1: Finnish lower-secondary pupil performance in international assessments over time



Finland – a story

Full system reform – pedagogic and curriculum content

First phase From 1968, fundamental reform based on fully comprehensive model, highly centralised, heavy State involvement. Revision of teacher training, grade tests, State-approved textbooks, heavy school inspection.

Second phase Strategic move to higher institutional autonomy, office for textbook approval closed in early 90s, inspection eased, data submission on school performance continued – phase culminated in superlative performance in PISA 2000

Third phase Decay in attainment, large programme of school closure, urban choice issues, introduction of project-based cross-curriculum learning (20pc)

Throughout, Abitur fundamentally unchanged.

Control Factors

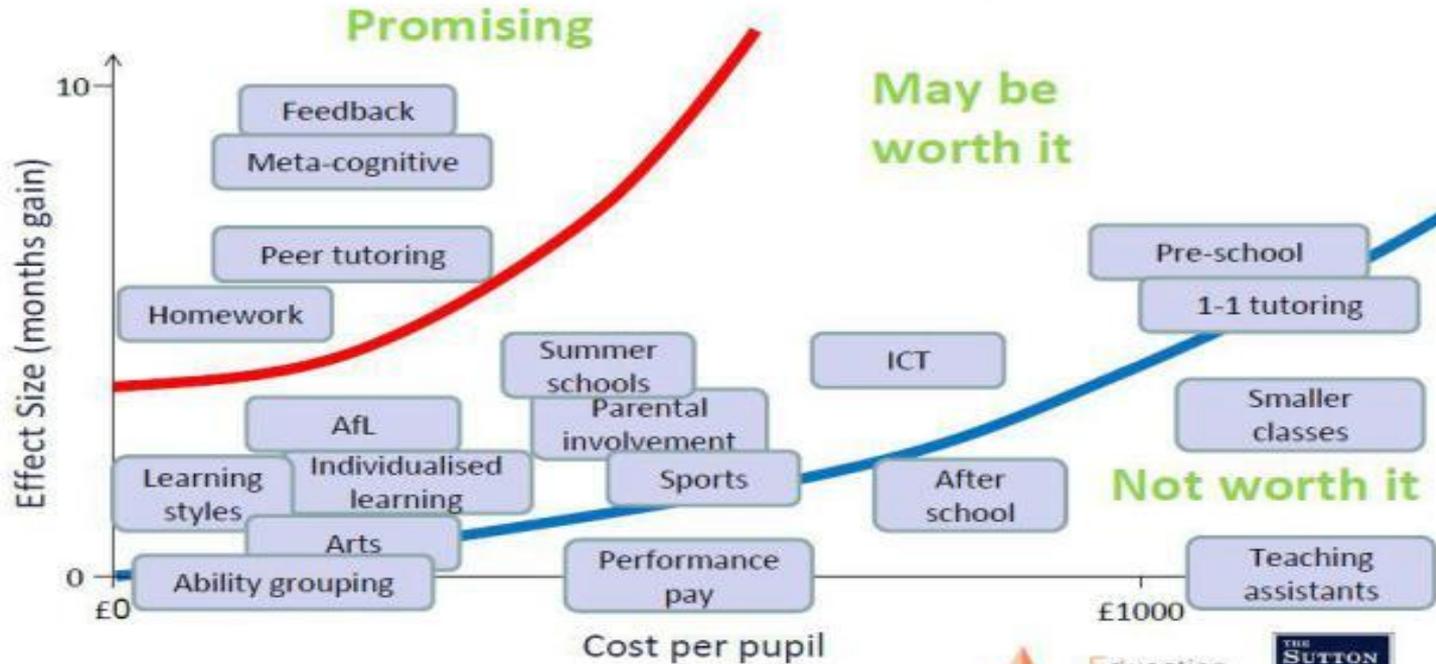
1. Curriculum content
2. Pedagogy
3. Assessment and qualifications
4. Institutional development
5. Institutional forms and structures
6. Governance
7. Professional development
8. Accountability
9. Inspection
10. Funding
11. National framework
12. Election and gatekeeping
13. Information and guidance about routes and choices
14. Allied social measures

Explanatory factors

- 1 global economy
- 2 domestic economy
- 3 culture
- 4 political structures and commitments
- 5 historical contingencies
- 6 natural environment

EEF (UK) analysis of policy focii and their impact

Toolkit findings



Looking forward

'Future skills'

'As well as; not instead of...' (Singapore; Hong Kong; Shanghai)

21C Skills are neither (Suto I)

Confusion between concept and context

Specific and general goods of education – curriculum distribution

Contamination by emphasis on academic education – vocational routes

Absence of strong evidence – existence of counterfactuals eg patent registration

Analysis of the areas of deficit in first year undergraduates – course modification

Labour market returns – maths, science and foreign languages

Longitudinal studies (Bynner J on Personal Capitals)

Looking forward

Dangers

A set of indicators or a set of curriculum imperatives – clarity regarding purposes and intention

Observation bias turning into curriculum aims and curriculum content

Washback into the curriculum (Boyle W; Gipps C; Stobart G)

The dangers of 'The Texas Test Effect' – (William D)

Remedy

Sophisticated, multifaceted and well-theorised transnational comparison

Research synthesis but attending to national context and aspirations

Sophisticated policy formation – the German Case

Dipf

Deutsches Institut für Internationale Pädagogische Forschung

German Institute of International Educational Research

Frankfurt

Founded 1951 as Hipf – became Dipf in the mid-1960s

Commissions to Dipf as a result of ‘PISA shock’ in 2000

Retention of precise knowledge of arrangements in Germany and local complexity

Detailed insights into PISA method and wider measurement by PISA surveys

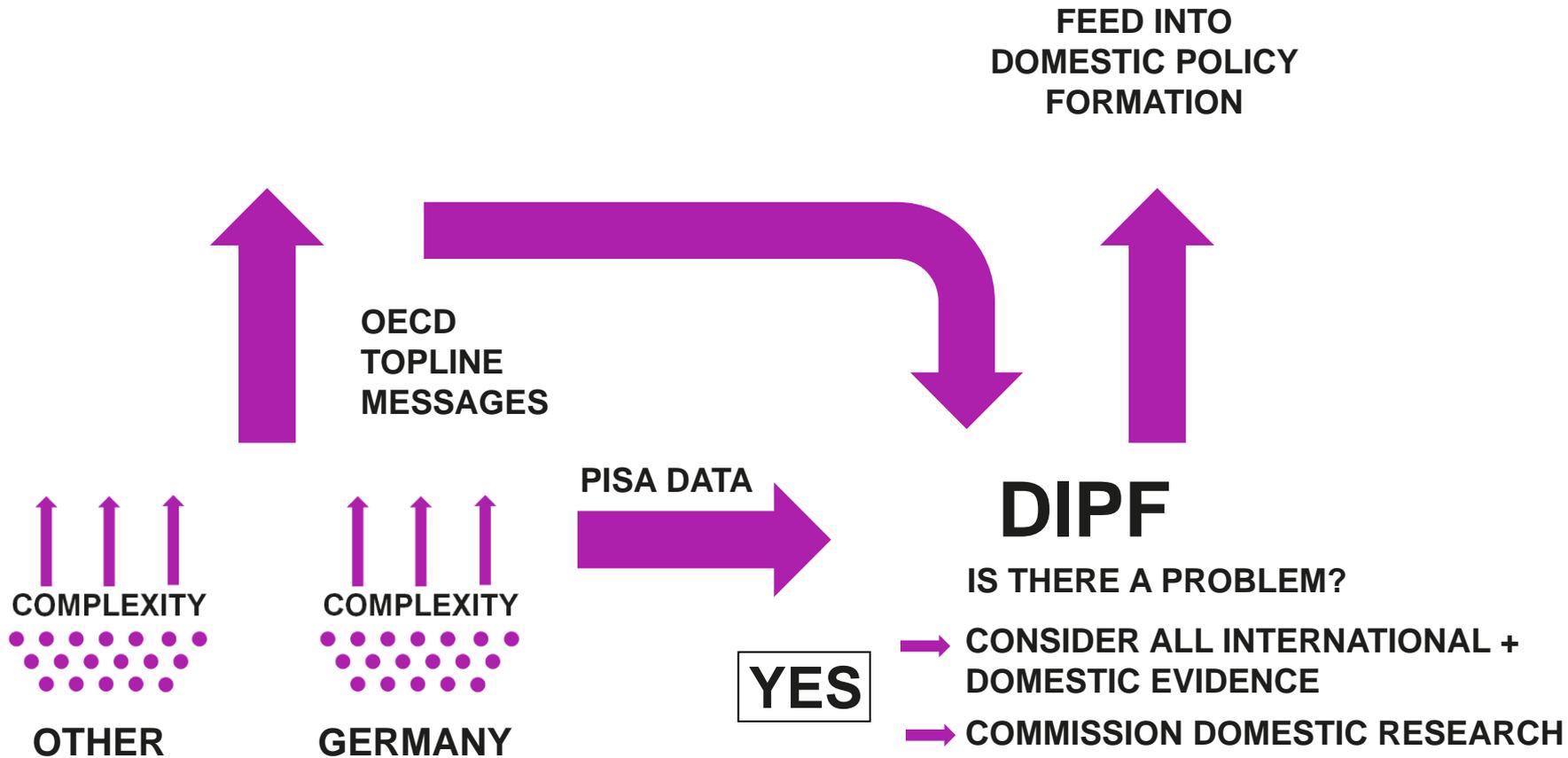
Commissioned to examine issues highlighted in the PISA survey but also by other measures

Undertakes further research to contribute to domestic policy formation

Sophisticated combination of domestic and international research

Carefully managed relationship with policy formation

The DIPF Model



Conclusions – beyond PISA

PISA provides valuable measurement - PISA provides one set of measurement amongst many possible measures of system performance

The topline messaging from PISA can include and concentrate various biases and limitations, and need to be treated with care

Sensitive, sophisticated analysis of previous periods of improvement are essential to effective interpretation of PISA and other measurements – the past matters, narrative matters

Dipf provides a very important model for country action on policy formation

And finally: **change costs** – it disrupts, it swiftly can lose subtle assets which have built up over long periods of time