

EDUCATION NEWS

Search

EDUCATION POLICY TECHNOLOGY HIGHER EDUCATION ONLINE SCHOOLS PARENTING INTERNATIONAL / UK

CAREER INDEX

CAREER INDEX

Phelps: OECD Encourages World to Adopt Failed US Ed Programs

Facebook Tweet Delicious Digg StumbleUpon Reddit



By Richard P. Phelps

In the aftermath of Hurricane Katrina, one of the US's largest insurers, refused to honor damage claims from customers living on the US Gulf Coast who submitted hurricane insurance claims, asserting that their

MONDAY
01 28, 2013

RICHARD P. PHELPS

FILED UNDER

[INTERNATIONAL EDUCATION](#)

[OECD RICHARD PHELPS](#)

PRINT

EMAIL

CAREER INDEX

Plan your career educator using online database information.

SUBJECT

STATE

STATE + SUBJECT

CITY + SUBJECT

ON TWITTER

@EdNews

Coding bootcamp offering more opportunities to and higher salaries
<https://t.co/RCS>
[#edtech](#) [#education](#) [#coding](#)

21
A new, unique partnership between Khan Academy and the San Francisco 49ers promotes [#education](#) [http://I1v8MQadcr](#)

property had not been damaged by hurricane, but by flooding. Only a high-stakes, high-profile, class-action lawsuit ultimately pried the insurance payments loose. Currently, this large US insurance company, with its own trust issues, is running a series of television commercials poking fun at an institution that it assumes is trusted by the public even less—the Internet. “They wouldn’t put it on the Internet if it wasn’t true” says the naïve foil who purchased allegedly inferior insurance after believing the promises in an Internet advertisement, presumably eliciting off-screen laughter in millions of living rooms.

Now suppose that you are responsible for learning the “state of the art” in the research literature on an important, politically-sensitive, and hotly-contested public policy topic. You can save money by hiring master’s level public policy students or recent graduates, though none with any particular knowledge or experience in the topic at hand—a highly specialized topic with its own doctoral-level training, occupational specializations, and vocabulary. You give your public policy masters a computer with an Internet browser and ask them to complete their reports within a few months. What would you expect them to produce?

You can see for yourself at the website of the Organisation for Economic Co-operation and Development (OECD). In 2009 the OECD launched the Review on Evaluation and Assessment Frameworks for Improving School Outcomes. Apparently the “Review” has not claimed an acronym. In my own interest, then, I give it one—REAFISO.

In its own words, REAFISO was created to:

“...provide analysis and policy advice to countries on the following overarching policy question: How can assessment and evaluation policies work together more effectively to improve student outcomes in primary and secondary schools?”



RT @VsauceTw
2 of Hyper Hist
never see this b
@googlearts [w
https://t.co/8U
https://t.co/...

2 r

ON FACEBOOK

 Like 6.4K

No related posts

RECENT NEWS

THURSDAY, SEP 24

Brazil Propose
More Flexible
Plans

WEDNESDAY, SEP

Muslim Studen
School Don't V
Single Sex Clas

TUESDAY, SEP 27,

Oxford Tops T
Global Higher
Rankings

POPULAR NEWS



NUMBER OF
HOMESCHOOLEI
GROWING NATI

92

[READ MORE](#)

To answer this question, the OECD intended to:

“...look at the various components of assessment and evaluation frameworks that countries use with the objective of improving the student outcomes produced by schools.... and

“...extend and add value to the existing body of international work on evaluation and assessment policies.”

REAFISO’s work interested me for two reasons. First, I once worked at the OECD, on fixed-length consulting contracts accumulating to sixteen months. I admired and respected their education research work and thoroughly enjoyed my time outside work hours. (The OECD is based in Paris.) I particularly appreciate the OECD’s education (statistical) indicators initiatives.

Second, I have worked myself and on my own time to address the overarching question they pose, ultimately publishing a meta-analysis and research summary of the effect of testing on student achievement. As I lacked the OECD’s considerable resources, it took me some time—a decade, as it turned out—to reach a satisfactory stage of completion. I hedge on the word “completion” because I do not believe it possible for one individual to collect all the studies in this enormous research literature.

To date, I have read over 3,000 studies, found about a third of them to be appropriate for inclusion in a summary of qualitative studies and meta-analyses of quantitative and survey studies. I looked at studies published in English between 1910 and 2010 that I could obtain and review before a self-imposed deadline in 2010. My coverage of the research literature is far from complete. It includes 244 qualitative studies (e.g., direct observations, site visits, interviews, case studies), 813 individual item-response group combinations from 247 survey studies (e.g., program evaluation surveys,



ALICE – IN WAL

[READ MORE](#)



**MATHEMATICS |
BEING OUTWITT
STUPIDITY** 80

[READ MORE](#)

opinion polls), and 640 separate measurements of effects from 177 quantitative research studies (e.g., regression analysis, structural equation modeling, pre-post comparison, experimental design, or interrupted time series design). In total, I analyzed 1,671 separate effects from 668 studies.

A summary is published in the International Journal of Testing (Phelps, 2012b). Source lists can be found at these three virtual locations:

<http://npe.educationnews.org/Review/Resources/QuantitativeList.htm>

<http://npe.educationnews.org/Review/Resources/SurveyList.htm>

<http://npe.educationnews.org/Review/Resources/QualitativeList.htm>

All but a couple of these several hundred sources were available to REAFISO as well. Yet, despite having many times the resources at their disposal, they managed to find just a few percent of what I found. Granted, our search parameters (as best I can discern theirs) were not exactly the same, but were far more alike than different. Not surprisingly, a review of a great expanse of the research literature, rather than just the selective, tiny bit covered by REAFISO, leads to quite different conclusions and policy recommendations.

Deficiencies of the OECD's REAFISO research reviews include:

- overwhelming dependence on US sources;
- overwhelming dependence on inexpensive, easily-found documents;
- overwhelming dependence on the work of economists and education professors;
- wholesale neglect of the relevant literature in psychology, the social science that invented cognitive assessment, and from practicing assessment and

measurement professionals; and wholesale neglect of the majority of pertinent research. Moreover, it seems that REAFISO has fully aligned itself with a single faction within the heterogeneous universe of education research—the radical constructivists. Has the OECD joined the US education establishment? One wouldn't think that it had the same (self-) interests. Yet, canon by canon by canon, REAFISO's work seems to subscribe to US education establishment dogma. For example, in her report "Assessment and Innovation in Education", Janet Looney writes:

"Innovation is a key driver of economic and social programs in OECD economies. If absent, innovation growth stalls; economies and communities stagnate...." (p. 6)

"...Teaching and learning approaches considered as innovation, on the other hand, are generally characterized as being 'student-centered', or 'constructivist'." (p. 8)

"This report has focused on [the] impact of high-stakes assessment and examinations on educational innovation. It has found significant evidence that such assessments and examinations undermine innovation." (p. 23)

First, Looney equates national economies and school classrooms. Then she adds the famous economist Joseph Schumpeter's definition of innovation as "creative destruction". For radical constructivists, and apparently for Looney, each teacher is a unique craftsperson in a unique classroom, and anything done to standardize their work inhibits their potential to guide each unique student in his or her own unique, natural discovery of knowledge. To radical constructivists, there are no economies of scale or scope in learning.

Whereas, innovation is a holy commandment for the

US education professoriate, critics charge that it leads to a continuous cycle of fad after fad after fad. After all, if innovation is always good, then any program that has been around for a while must be bad, no matter how successful it might be in improving student achievement. Moreover, if the pace of today's-innovation-replacing-yesterday's-innovation proceeds fast enough, evaluation reports are finished well after one program has been replaced by another, become irrelevant before they are published and end up unread. Ultimately, in a rapidly innovating environment, we learn nothing about what works. Some critics of the radical constructivists suspect that that chaotic, swirling maelstrom may be their desired equilibrium state.


A case in point is the sad and expensive 1990s saga of the New Standards Project in the United States and the most deliberate effort to implement its assessment formula in practice, the State of Maryland's MSPAP (for Maryland School Performance Assessment Program). REAFISO writer Allison Morris (p.16) cites Thomas Toch's (2006) erroneous assertion that cost considerations reversed the 1980s–1990s US trend toward more performance testing. Not so, the MSPAP and similar programs (e.g., CLAS [California Learning Assessment System], KIRIS [Kentucky Instructional Results Information System] failed because test reliability was so low, test scores were too volatile to be useful, feedback was too late and too vague to be useful, and parents thought it unfair when their children's grades depended on other students' efforts (in collaborative group activities).

Resounding public disgust killed those programs. But, ten years is a long time in the ever-innovating world of US education policy, long enough for the young REAFISO writers to be unaware of the fiascos. REAFISO now urges us to return to the glorious past of New Standards, MSPAP, CLAS, and KIRIS, dysfunctional programs that, when implemented, were overwhelmingly rejected by citizens, politicians, and

measurement professionals alike.

Richard P. Phelps is the author of [Standardized Testing Primer](#) (2007) and other books about testing and is the founder of the [Nonpartisan Education Review](#). He lives in Asheville, North Carolina.

[Bio](#) [Latest Posts](#)



Richard P. Phelps

0 Comments

Sort by [Oldest](#)



Add a comment...

 Facebook Comments Plugin

[Privacy Policy](#) [Education](#)