

EDUCATION NEWS

Search

EDUCATION POLICY TECHNOLOGY HIGHER EDUCATION ONLINE SCHOOLS PARENTING INTERNATIONAL / UK

CAREER INDEX

CAREER INDEX

'One Size Fits All' National Tests Not Deeper or More Rigorous

CAREER INDEX

Plan your career educator using online database information.

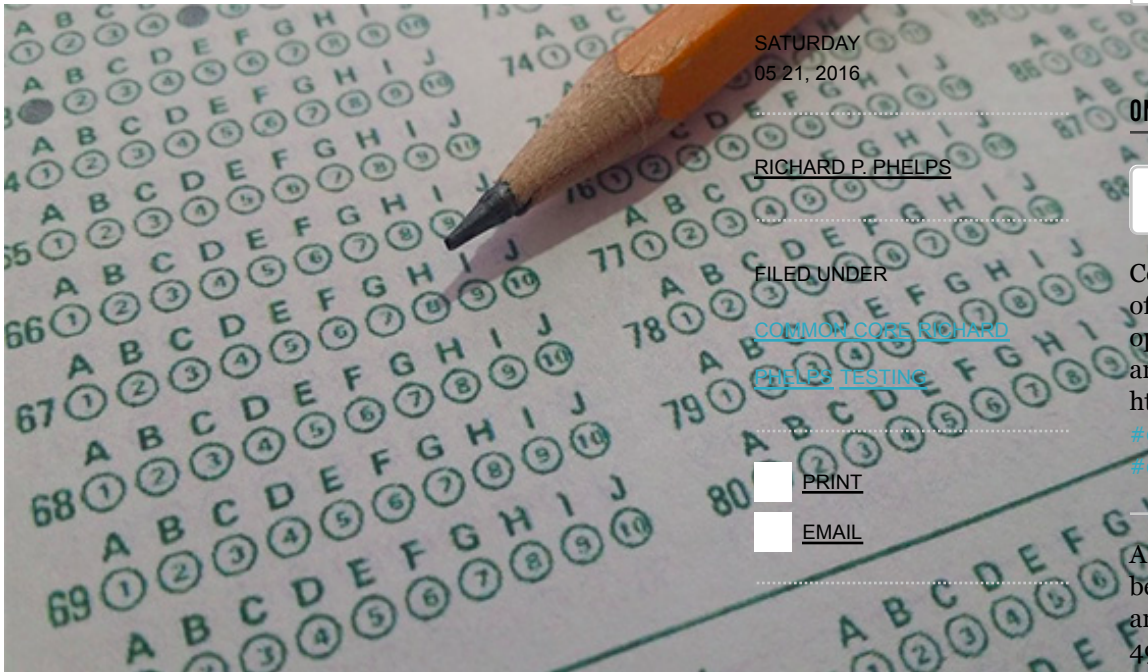
SUBJECT

STATE

STATE + SUBJECT

CITY + SUBJECT

Facebook Tweet Delicious Digg StumbleUpon Reddit



SATURDAY
05 21, 2016

RICHARD P. PHELPS

FILED UNDER

[COMMON CORE](#) [RICHARD PHELPS TESTING](#)

PRINT

EMAIL

ON TWITTER

@EdNews

Coding bootcamp offering more opportunities to and higher salaries https://t.co/RCC #edtech #education #coding

A new, unique partnership between Khan Academy and the San Francisco 49ers promotes #education http://11v8MQadcr...

(Photo: Pixabay, Creative Commons)

By Richard Phelps

Some say that now is a wonderful time to be a psychometrician — a testing and measurement professional. There are jobs aplenty, with high pay and great benefits. Work is available in the private sector at test development firms; in recruiting, hiring, and placement for corporations; in public education agencies at all levels of government; in research and teaching at universities; in consulting; and many other spots.

Moreover, there exist abundant opportunities to work with new, innovative, “cutting edge”, methods, techniques, and technologies. The old, fuddy-duddy, paper-and-pencil tests with their familiar multiple-choice, short-answer, and essay questions are being replaced by new-fangled computer-based, internet-connected tests with graphical interfaces and interactive test item formats.

In educational testing, the Common Core Standards Initiative (CCSI), and its associated tests, developed by the Smarter-Balanced Assessment Consortium (SBAC) and the Partnership for Assessment of Readiness for College and Careers (PARCC), has encouraged the movement toward “21st century assessments”. Much of the torrential rain of funding, burst forth from federal and state governments and clouds of wealthy foundations, has pooled in the pockets of psychometricians.

At the same time, however, the country’s most authoritative psychometricians—the very people who would otherwise have been available to guide, or caution against, the transition to the newer standards and tests—have been co-opted. In some fashion or another, they now work for the CCSI. Some work for the SBAC or PARCC consortia directly, some work for one or more of the many test development firms hired by the consortia, some help the CCSI in other capacities. Likely, they have all signed confidentiality



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

ON FACEBOOK

Like 6.4K

No related posts

RECENT NEWS

FRIDAY, SEP 30, 2016

Judge Orders 1
Access for Ohio
Transgender S

THURSDAY, SEP 29, 2016

Holcomb Prop
Build on Pence
Reforms in Inc

THURSDAY, SEP 29, 2016

Texas Board of
Continues Batt
Evolution in B

POPULAR NEWS



NUMBER OF
HOMESCHOOLERS
GROWING NATIONALLY

92

READ MORE



agreements (i.e., “gag orders”).

Psychometricians who once had been very active in online chat rooms or other types of open discussion forums on assessment policy no longer are, unless to proffer canned promotions for the CCSI entities they now work for. They are being paid well. They may be doing work they find new, interesting, and exciting. But, with their loss of independence, society has lost perspective.

Perhaps the easiest vantage point from which to see this loss of perspective is in the decline of adherence to test development quality standards, those that prescribe the behavior of testing and measurement professionals themselves. Over the past decade, for example, the International Test Commission (ITC) alone has developed several sets of standards.

Perhaps the oldest set of test quality standards was established originally by the American Psychological Association (APA) and was updated most recently in 2014—the Standards for Educational and Psychological Testing (AERA, NCME, APA). It contains hundreds of individual standards. The CCSI as a whole, and the SBAC and PARCC tests in particular, fail to meet many of them.

The problem starts with what many professionals consider the testing field’s “prime directive”—Standard 1.0 (AERA, NCME, APA, p.23). It reads as follows:

“Clear articulation of each intended test score interpretation for a specified use should be set forth, and appropriate validity evidence in support of each intended interpretation should be provided.”

That is, a test should be validated for each purpose for which it is intended to be used before it is used for that purpose. Before it is used to make important decisions. And, before it is advertised as serving that purpose.



ALICE – IN WAL

[READ MORE](#)



**MATHEMATICS |
BEING OUTWITT
STUPIDITY 80**

[READ MORE](#)

Just as states were required by the Race to the Top competition for federal funds to accept Common Core standards before they had even been written, CCSI proponents have boasted about their new consortium tests' wonderful benefits since before test development even began. They claimed unproven qualities about then non-existent tests because most CCSI proponents do not understand testing, or they are paid not to understand.

In two fundamental respects, the PARCC and SBAC tests will never match their boosters' claims nor meet basic accepted test development standards. First, single tests are promised to measure readiness for too many and too disparate outcomes—college and careers—that is, all possible futures. It is implied that PARCC and SBAC will predict success in art, science, plumbing, nursing, carpentry, politics, law enforcement ...any future one might wish for.

This is not how it is done in educational systems that manage multiple career pathways well. There, in Germany, Switzerland, Japan, Korea, and, unfortunately, few jurisdictions in the U.S., a range of different types of tests are administered, each appropriately designed for their target professions. Aspiring plumbers take plumbing tests. Aspiring medical workers take medical tests. And, those who wish to prepare for more advanced degrees might take more general tests that predict their aptitude to succeed in higher education institutions.

But that isn't all. SBAC and PARCC are said to be aligned to the K-12 Common Core standards, too. That is, they both summarize mastery of past learning and predict future success. One test purports to measure how well students have done in high school, and how well they will do in either the workplace or in college, three distinctly different environments, and two distinctly different time periods.

PARCC and SBAC are being sold as replacements for

state high school exit exams, for 4-year college admission tests (e.g., the SAT and ACT), for community college admission tests (e.g., COMPASS and ACCUPLACER), and for vocational aptitude tests (e.g., ASVAB). Problem is, these are very different types of tests. High school exit exams are generally not designed to measure readiness for future activity but, rather, to measure how well students have learned what they were taught in elementary and secondary schools. We have high school exit exams because citizens believe it important for their children to have learned what is taught there. Learning Civics well in high school, for example, may not correlate highly with how well a student does in college or career but many nonetheless consider it important for our republic that its citizens learn the topic

High school exit exams are validated by their alignment with the high school curriculum, or content standards. By contrast, admission or aptitude tests are validated by their correlation with desired future outcomes—grades, persistence, productivity, and the like in college—their predictive validity. In their pure, optimal forms, a high school exit exam, a college admission test, and vocational aptitude tests bear only a slight resemblance to each other. They are different tests because they have different purposes and, consequently, require different validations.

Let's assume for the moment that the Common Core consortia tests, PARCC and SBAC, can validly measure all that is claimed for them—mastery of the high school curriculum and success in further education and in the workplace. The fact is no evidence has yet been produced that verifies any of these things. And, remember, the proof of, and the claims about, a new test's virtues are supposed to be provided before the test is used purposefully.

Sure, Common Core proponents claim to have just

recently validated their consortia tests for correlation with college outcomes , for alignment with elementary and secondary school content standards, and for technical quality . The clumsy studies they cite do not match the claims made for them, however.

SBAC and PARCC cannot be validated for their purpose of predicting college and career readiness until data are collected in the years to come on the college and career outcomes of those who have taken the tests in high school. The study cited by Common Core proponents uses the words “predictive validity” in its title. Only in the fine print does one discover that, at best, the study measured “concurrent” validity—high school tests were administered to current rising college sophomores and compared to their freshman-year college grades. Calling that “predictive validity” is, frankly, dishonest.

It might seem less of a stretch to validate SBAC and PARCC as high school exit exam replacements. After all, supposedly they are aligned to the Common Core Standards so in any jurisdiction where the Common Core Standards prevail, they would be retrospectively aligned to the high school curriculum. Two issues tarnish this rosy picture. First, the Common Core Standards are subjectively narrow, just mathematics and English Language Arts, with no attention paid to the majority of the high school curriculum.

Second, common adherence to the Common Core Standards across the States has deteriorated to the point of dissolution. As the Common Core consortia’s grip on compliance (i.e., alignment) continues to loosen, states, districts within states, and schools within districts are teaching how they want and what they want. The less aligned Common Core Standards become, the less valid the consortium tests become as measures of past learning.

As for technical quality, the Fordham Institute, which is paid handsomely by the Bill & Melinda Gates Foundation to promote Common Core and its consortia

tests, published a report which purports to be an “independent” comparative standards alignment study. Among its several fatal flaws: instead of evaluating tests according to the industry standard Standards for Educational and Psychological Testing, or any of dozens of other freely-available and well-vetted test evaluation standards, guidelines, or protocols used around the world by testing experts, they employed “a brand new methodology” specifically developed for Common Core and its copyright owners, and paid for by Common Core’s funders.

Though Common Core consortia test sales pitches may be the most disingenuous, SAT and ACT spokespersons haven’t been completely forthright either. To those concerned about the inevitable degradation of predictive validity if their tests are truly aligned to the K-12 Common Core standards, public relations staffs assure us that predictive validity is a foremost consideration. To those concerned about the inevitable loss of alignment to the Common Core standards if predictive validity is optimized, they assure complete alignment.

So, all four of the test organizations have been muddling the issue. It is difficult to know what we are going to get with any of the four tests. They are all straddling or avoiding questions about the trade-offs. Indeed, we may end up with four, roughly equivalent, muddling tests, none of which serve any of their intended purposes well.

This is not progress. We should want separate tests, each optimized for a different purpose, be it measuring high school subject mastery, or predicting success in 4-year college, in 2-year college, or in a skilled trade. Instead, we may be getting several one-size-fits-all, watered-down tests that claim to do all but, as a consequence, do nothing well. Instead of a skilled tradesperson’s complete tool set, we may be getting four Swiss army knives with roughly the same features. Instead of exploiting psychometricians’ advanced

knowledge and skills to optimize three or more very different types of measurements, we seem to be reducing all of our nationally normed end-of-high-school tests to a common, generic muddle.

References

McQuillan, M., Phelps, R.P., & Stotsky, S. (2015, October). How PARCC's false rigor stunts the academic growth of all students. Boston: Pioneer Institute.

<http://pioneerinstitute.org/news/testing-the-tests-why-mcas-is-better-than-parcc/>

Nichols-Barrer, I., Place, K., Dillon, E., & Gill, B. (2015, October 5). Final Report: Predictive Validity of MCAS and PARCC: Comparing 10th Grade MCAS Tests to PARCC Integrated Math II, Algebra II, and 10th Grade English Language Arts Tests. Cambridge, MA: Mathematica Policy Research.

<http://econpapers.repec.org/paper/mprmprrs/a2d9543914654aa5b012e4a6d2dae060.htm>


Phelps, R.P. (2016, February). Fordham Institute's pretend research. Policy Brief. Boston: Pioneer Institute. <http://pioneerinstitute.org/featured/fordhams-parcc-mcas-report-falls-short/>
Reference

American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME). (2014). Standards for Educational and Psychological Testing, Washington, DC: AERA.

 **Richard P. Phelps**

0 Comments

Sort by **Oldest**

 Add a comment...

 Facebook Comments Plugin

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