

***The Disconnect Between Recent Research and Education Policy:
Basing High-Stakes Decisions on Tests That Have Not Proven to Be Valid***

Introduction

High stakes tests have been a fixture in American education for many years. The College Entrance Examination Board administered the first standardized exam in 1901 and the first SAT multiple choice test was administered in 1926. (History of the Tests).

Such testing always has had its detractors. In recent decades, the higher education community expressed concern that high stakes testing may not be as valid as was formerly believed. The main concern had focused on the cultural biases in the placement of minority students, though some also have voiced concern that the tests may not measure higher-level thinking skills (e.g., problem-solving). Nevertheless, the sales of high-stakes college placement tests increased through the decades with the expansion of the college-bound student pool.

The K-12 world also has embraced high stakes testing through the years. With the initiation of the No Child Left Behind (NCLB) legislation, such testing at the K-12 level increased substantially. While many states developed their own tests, some turned to standardized testing companies like ACT to provide them with components for a state test. Since NCLB testing focused on reading and mathematics, one ACT advantage over its main competitor, the College Board's SAT, was that ACT had offered a reading test since 1989 (it previously had been called the social science test). SAT did not begin to offer a test specifically named "reading" until 2005 (it previously had been termed "verbal").

There are 22 states with state-funded ACT partnerships, ten states that have almost universal ACT testing (one with 97%, the rest with 100%), and five states offering the ACT and/or the WorkKeys as components of their official NCLB test. (ACT and Statewide Testing). "Almost 20% of ACT volume" is from state testing. (Discussion Topic: State Adoptions)

Due in part to such partnerships, ACT's expansion during the past ten years has been substantial. The use of the ACT has expanded greatly, with 1,666,017 Class of 2012 graduating seniors taking the test, an increase of 49.2% from the Class of 2002. From 2004 to 2009 (the latest year for which data are available), ACT's revenue from its testing programs and related services increased by over 63%, from approximately \$155,000,000 to \$253,000,000. (ACT: The First Fifty Years, 1959–2009)

In contrast, College Board's SAT's expansion has been much more modest, though still substantial, with 24.4% more Class of 2012 graduating seniors taking the test (1,664,479) than for the Class of 2002 (1,337,999). Only one state, Delaware, had universal SAT testing, with no other state having more than 93% of its students taking the test.

The community colleges provide another standardized test stronghold. Comparable use trends over time are not published for the ACT-produced COMPASS and College Board's ACCUPLACER. However, their use is very widespread. According to a 2008 survey of assessment practices, 62% of community colleges use ACCUPLACER, and 46% use COMPASS (some institutions use both, depending on the subject). (Primary Research Group, Inc.)

Similarly, no expansion or sales volume data are available for the WorkKeys post-secondary workforce readiness test. ACT did report that they have issued 1.5 million National Career

Readiness Certificates and have "statewide or regional credentialing initiatives in more than 40 states." The certificates are issued for passing all three WorkKeys tests at a certain level. In fact, the ACT website recently announced that "The American Council on Education (ACE) has recommended that postsecondary educational institutions award college credit in applied critical thinking to individuals who earn the ACT National Career Readiness Certificate (NCRC™). ACE recommends that three credit hours be awarded to recipients who earn specific levels of the credential." (i.e., silver, gold, or platinum score levels). (ACT National Career Readiness Certificate Recommended for Three Semester Hours of College Credit). Inasmuch as SAT and ACT performance alone does not alone earn college credit, such a recommendation is quite unprecedented.

So the college placement companies, particularly ACT, are experiencing substantial usage and growth from secondary and post-secondary institutions, and have sufficient credibility to warrant the expenditure of hundreds of millions of dollars.

Yet several national institutions with strong independent credentials have conducted and published studies in the past two years which would call into question the current usage of such tests. This report will review the major studies focused on the ACT and COMPASS/ACCUPLACER tests in the past two years. It also will examine the research behind WorkKeys. Since SAT has not had very recent independent studies of its validity for placement use, it will not be a focus of this report. While both ACT and SAT conduct a fair amount of research on their own tests, we will not examine such studies, since they cannot be considered to be at all impartial or unbiased. The one exception to omitting test vendor studies from this report is the referencing of an ACT study that inadvertently communicated information supporting a major independent study's reservations on two ACT subject tests.

ACT

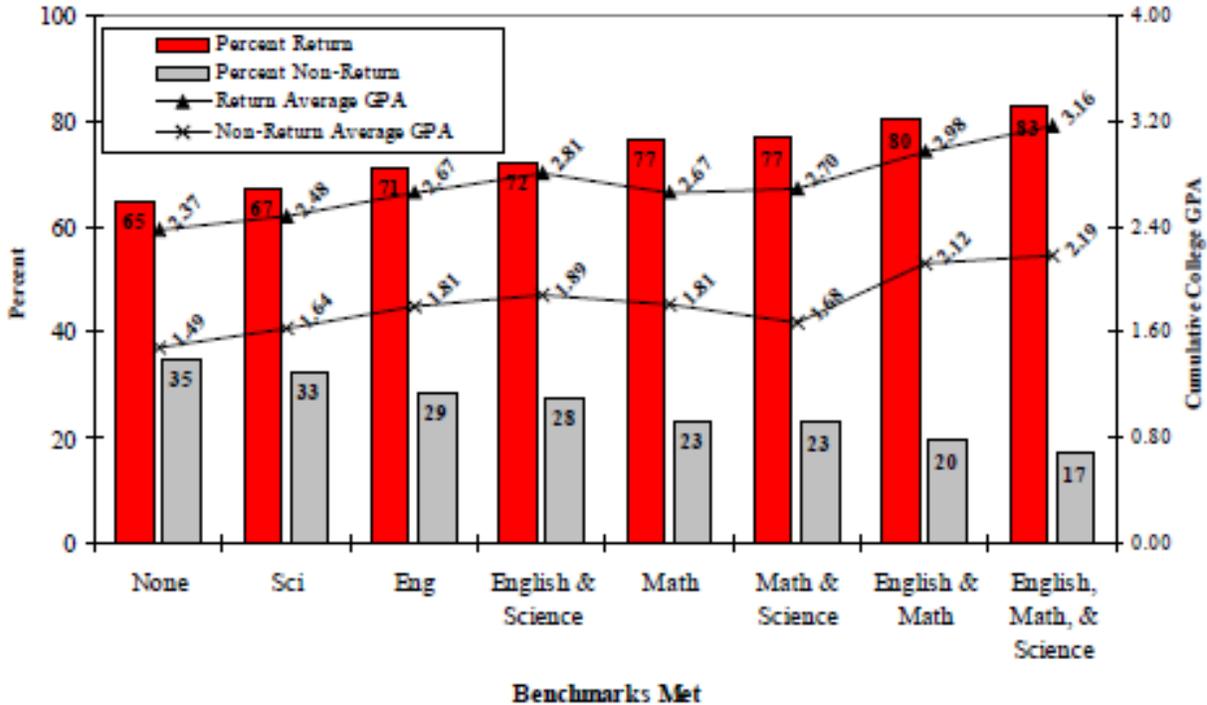
The ACT test was the subject of a large research study produced by the National Bureau of Economic Research (NBER). "Improving College Performance and Retention the Easy Way: Unpacking the ACT Exam" was released in June 2011, conducted by Eric Bettinger and Brent Evans of Stanford and Devin Pope of the University of Chicago. Their abstract stated that: "Colleges rely on the ACT exam in their admission decisions to increase their ability to differentiate between students likely to succeed and those that have a high risk of under-performing and dropping out. We show that two of the four subtests of the ACT, English and Mathematics, are highly predictive of positive college outcomes while the other two subtests, Science and Reading, provide little or no additional predictive power. This result is robust across various samples, specifications, and outcome measures. We demonstrate that focusing solely on the English and Mathematics test scores greatly enhances the predictive validity of the ACT exam." Also, "By introducing noise that obscures the predictive validity of the ACT exam, the reading and science tests cause students to be inefficiently matched to schools, admitted to schools that may be too demanding -- or too easy -- for their levels of ability."

The national media responded to the study, and ACT responded to the media. In *EdWeek*, Jon Erickson, the current (then interim) president of ACT's Education Division stated that: "ACT is an achievement-based test that is used for multiple goals and purposes beyond just admissions or predicting overall student success, such as college GPA or retention." (Adams) *USA Today* reprinted an article from *Inside Higher Education*, which related comparable information. (Jaschik) Erickson's dismissing of the role of the ACT in "admissions or predicting overall college success" is surprising,

since ACT's own website describes the ACT test as "the college admissions and placement test." (What We Do).

The most important validation for the NBER study came, ironically, from ACT's own research. "ACT College Readiness Benchmarks, Retention, and First-Year College GPA: What's the Connection?" was a 2005 study conducted by ACT in support of their benchmarks. It unintentionally cast doubt on the value of the benchmarks by showing how most students who did not meet them persisted successfully in their first year of college. But more importantly, it showed that meeting the science benchmark was a weak predictor of college persistence and GPA, and it left out references to the reading benchmark entirely. The graph depicting the findings shows that the persistence level is only very minimally increased by the addition of the meeting of the science benchmark, and the findings from reading are omitted entirely (the graph is reproduced below. When I met with the ACT research team in Iowa City in 2008 to question the benchmarks, the frequency printouts provided from the original study also omitted the reading benchmark data. Such omissions should be interpreted as a demonstration that the reading test does not have predictive value for college placement.

Figure 1: Public Four-Year Retention and College GPAs by ACT College Benchmarks



So why the disconnect between such findings and education policy? The NBER authors offered the following: "Personal conversations suggest that most admission officers are simply unaware of the difference in predictive validity across the tests and have limited time and resources to analyze the predictive power of its various components at their institution. An alternative explanation is that schools have a strong incentive -- perhaps due to highly publicized external rankings such as those compiled by U.S. News & World Report, which incorporate students' entrance exam scores -- to admit students with a high ACT composite score, even if this score turns out to be unhelpful."

Reading is, of course, critical to academic success. But this does not mean that a test labeled "reading" is necessarily a good measure of reading. In fact, reading already is measured by all ACT subject tests. And science is much too broad a field to be comprehensively tested in 35 minutes.

One study alone, no matter how large, cannot refute the value of a long-established test. But when ACT's own research validates that study's findings, and when ACT's own president of Educational Services acknowledges the limitation, then the test should not be considered suitable for either college admission screening or statewide student assessment. Additionally, my own studies of the ACT products (EXPLORE and PLAN) at the high school level, and other past studies, also have validated the limitations of both the science and reading tests from ACT, and no independent studies in recent years contradict such findings.

The implications for such findings are the following:

- The much-touted ACT College Readiness Benchmarks are not valid indicators of college readiness. Most students not meeting all four benchmarks did so due to the inability to meet the inconsistently high science benchmark. If the science benchmark (recently lowered from a score of 24 to a 23) is based upon a test that has little predictive value for college persistence or classroom performance, and the reading benchmark (recently raised from 21 to 22) appears to have almost no predictive value, then such benchmarks cannot measure college readiness. There are many other issues with the benchmarks, but these considerations alone invalidate them.
- The state standards that depend to a significant degree on the ACT reading test need to find another instrument. There is no logic to basing a state standard for student performance on a test that has been proven to have little explanatory value for post-secondary student classroom performance.
- Secondary and post-secondary institutions need to research placement at their own schools and reconsider their use of the ACT (and PLAN and EXPLORE) science and reading tests. Unless local research clearly demonstrates otherwise, placement should focus on the English and math test findings, and pay minimal attention to the science and reading.

COMPASS/ACCUPLACER

The community college placement tests from ACT and College Board have been the subject of two major studies released in February, 2012 from the Community College Research Center at Columbia University's Teachers College. A third and not-yet-completed study has been conducted through a collaboration between Harvard and Rand Corporation staff members (Daugherty and Long); its preliminary findings align with the Columbia studies. The Columbia studies, one conducted in a statewide community college system, the other in a large (42,000 student) urban one, have expressed concern over the value of the COMPASS and ACCUPLACER in placing students. Each study found that more than one-fourth of the students placed in remedial classes due to the use of the two tests should have been able to pass college-credit bearing courses with a grade of B or better. They also found that the tests resulted in overplacement; i.e., students being placed in courses that were too advanced for them, but these were much less of a concern.

From the conclusion of Judith Scott-Clayton's "Do High-Stakes Placement Exams Predict College Success?" (based upon the COMPASS test): "The predictive power of placement exams is in a sense quite impressive given how short they are (often taking about 20–30 minutes per subject/module). But overall the correlation between scores and later course outcomes is relatively weak, especially in light of the high stakes to which they are attached. Given that students ultimately succeed or fail in college-level courses for many reasons beyond just their performance on placement exams, it is

questionable whether their use as the sole determinant of college access can be justified on the basis of anything other than consistency and efficiency." Additionally, she commented that: "...the incremental validity of placement tests relative to high school background predictors of success is weak, even in math."

From the abstract of Clive Belfield and Peter Crosta's "Predicting Success in College: The Importance of Placement Tests and High School Transcripts" (based upon both the COMPASS and ACCUPLACER): "We find that placement tests do not yield strong predictions of how students will perform in college. Placement test scores are positively—but weakly—associated with college grade point average (GPA). When we control for high school GPA, the correlation disappears...In contrast, high school GPAs are useful for predicting many aspects of students' college performance."

The national media again responded to the studies. The New York Times lamented that "less than a quarter of those who start in remedial classes go on to earn two-year degrees or transfer to four-year colleges." (Lewin) EdWeek stated that "The way colleges are using standardized placement tests such as the College Board's ACCUPLACER, ACT's Compass, and others can misidentify students, and secondary schools and universities should work to develop a more comprehensive profile of students' strengths and weaknesses in performing college-level work." (Sparks) Representatives from both ACT and College Board acknowledged that misplacement is likely if the schools use only the test score without additional information from other academic indicators, particularly high school grades, and that preparation for the tests will reduce misplacement.

The study authors stressed the need for additional information. Belfield stated that "Where you have both a test and a high school transcript, the best thing is to use both together."

As mentioned above, both ACCUPLACER and COMPASS are computer-based and computer-adaptive (i.e., matching subsequent item difficulty to success on prior items). So they may be unlike any other test that the student has taken. Difficulty with the format may affect the score and reduce the validity of the test as a measure of content knowledge. So preparation for the test may help increase the relationship between test scores and subsequent college classroom performance.

Unlike the ACT and SAT, the COMPASS and ACCUPLACER tests are often used with minimal consideration for a student's high school grades. Additionally, they often are used in conjunction with the ACT and SAT, with the latter tests used as a screener for requiring the taking of the COMPASS or ACCUPLACER. For example, in Illinois, the COMPASS is given to most community college students despite the fact that there is virtually universal testing of Illinois high school 11th graders with the ACT. Given the strong relationships between performance on the general college admissions test and the community college test (e.g., students who do well on the ACT do well on the COMPASS), it is unlikely that the latter will provide a significant level of additional information over the former.

Additionally, with so much emphasis being placed on only one marginally valid (at best) indicator, schools can get desperate to improve their student's performance. A two-high school Chicago suburban school district has gone so far as to add a new course next year, with one semester solely dedicated to preparing for the COMPASS exam. Using valuable classroom time for test preparation at the expense of content knowledge is very questionable, and a sad comment on the effect of putting so much emphasis on any single test, let alone a very weak one.

So again, why the disconnect? Community colleges like to use the COMPASS and ACCUPLACER for placement use. They are relatively simple to administer (requiring computer station access), inexpensive (\$2.00 to \$8.00 per test module, usually paid for by the student), and can yield a score instantly (and in as short a time as completing only eight items). Unfortunately, the brevity of a test usually is inversely related to the reliability (and therefore validity) of a test. Additionally, computerized testing, particularly computer-adaptive testing, often is unfamiliar to many students.

An additional incentive for colleges is that labeling students as needing remediation will generate more revenue and more jobs. Adding an additional course or two to a student's degree will require more classes, teachers, and tuition. However, that advantage may be offset by the aforementioned high attrition rate among those taking remedial courses. With schools being increasingly accountable for graduation rates, these old incentives may be vanishing.

Hence, while there generally is a positive relationship between community college classroom performance and COMPASS or ACCUPLACER scores, the relationship is weak, and may cease to exist if prior grades already have been considered. Many factors may positively relate to college classroom performance, but that does not mean that they are sufficiently strong enough to justify their use for college placement. For example, if you substituted some frivolous measure, like daily minutes of watching TV or crossword puzzle time completion, you probably would find a relationship to college performance. But it does not mean that students should be identified as college ready because they watch less TV or are good with crossword puzzles. Placement cannot be based upon weak indicators.

The implications for such findings are the following:

- The test should not be used as the sole or even primary or indicator for placement. The tests predict too little future student classroom performance and are much less predictive than prior student classroom performance.
- There is no research that the tests discriminate among lower performing students better than the ACT or SAT, which the students often already have taken. With Common Core assessments looming on the horizon, there may be uniform regional testing that also would render the COMPASS redundant and/or less needed.

WorkKeys

As discussed above, the ACT and the COMPASS/ACCUPLACER tests have been the subject of independent validation studies by major institutions of research that have examined whether they predict college success. In contrast, WorkKeys research has been substantially more limited, with no major independent research entities conducting such prediction research. Studies that have been found have been limited to content analysis, ACT-sponsored/partnered case studies, and one study by an analyst in a state regional office.

All major tests are subject to content analysis in the course of their development. In 2010, ACT commissioned a pair of experts to conduct a content validation of its WorkKeys tests, and, not surprisingly, they found that "The validity of the use of WorkKeys tests as selection criteria may properly be shown by using a content validation approach, which is the approach that is typically used by employers utilizing the WorkKeys job profiling process." (External Review Addresses Validity) In its endorsement process, the American Council on Education (ACE) also conducted such an analysis, explaining to me that they "contract with college and university professors to evaluate the test content" when they make evaluations of tests.

ACT dismissed the need for users to further validate the test. In the WorkKeys FAQ section, the ACT maintained the following position: "Question: Does an employer's use of the National Career Readiness Certificate or any three or more of the WorkKeys foundational skills assessments require a local validation study? Answer: No. The generally accepted principles of validity generalization support use of well-studied cognitive ability measures such as WorkKeys without having to conduct a local validation study."

While analyzing content validity is useful and necessary for test development, it is no substitute for seeing if the test actually works in practice. The WorkKeys content is supposed to measure skills that are more focused upon the workplace. The items are workplace-oriented, e.g., test questions related to ringing up a sale in a store. But, as explained above, the "common sense" attitude that such test question will predict workplace performance better than, for example, correctly identifying a number sequence, may be totally flawed.

The relationship of the ACT science and math tests, discussed earlier, is an example of such a common sense-defying reality. ACT has conducted years of psychometric validation in the refinement of its ACT science test. The content is meaningful and aligns to skills needed in a science class. And if we had no other measure with which to predict college science performance, the science test might be a useful measure. However, we know that the science test is a very weak predictor of future science class performance, and the ACT math test is a much better predictor of science class performance. Using the math test renders the science test largely useless, since the latter adds no additional information. This occurs despite the fact that, unlike the science test, the math test does not have items that specifically align to science classroom skills. This is a major difficulty of test psychometrics - measures that seem to specifically relate to a behavior may not predict it at all. That is why predictive research is necessary.

ACT also has showcased an array of its own case studies from its product users. While these technically are prediction studies, they also are marketing materials. As mentioned in the introduction, vendor studies conducted to promote vendor products are extremely susceptible to bias. These case studies are not listed in ACT's extensive website research study files.

I personally am very familiar with case studies by ACT, since our district was the subject of a glowing 2004 ACT case study (still available on the ACT website) presumably due to our heavy usage of their products. Ironically, the study covered a period of mediocre and stagnant overall performance, with substantial improvement occurring after that time frame.

The sole potentially independent study is supportive of WorkKeys. It was produced by a research analyst at the Business Solutions Center in the Workforce Investment Board of Southwest Missouri. The institutional credibility of such a study is not remotely of the level of the National Bureau of Economic Research study of ACT or the Columbia University COMPASS/ACCUPLACER studies. This alone does not mean that the findings are invalid, but when you examine the study, its credibility diminishes.

The brief study contains is not a research study in the scientific sense, with no examination of whether findings are statistically significant, no references, no explanation of conspicuous data gaps, and conclusions that are supported only by selecting specific data and ignoring other equally compelling data. Given such selective data use, it cannot at all be considered impartial research. Additionally, some of its findings conflict markedly with one of ACT's own case studies. So there

appears to be no independent scientific predictive research study available to validate (or invalidate) WorkKeys.

Human behavior is hard to predict, particularly by using a relatively brief test. The best ACT subject tests (English and Math) predict only a small percentage of the variance in future classroom performance. Most future performance is not able to be predicted. Additionally, classroom track record has almost always been found to be a better predictor of future academic performance than a test. This is why most four-year colleges use a combination of high school grades and test scores to place students.

Workplace behavior is even harder to predict than future classroom performance, since classroom performance already is in part measured by test performance, and workplace behavior generally is not. So even less of workplace performance (i.e., having the skills and attitudes to perform job tasks over time, measured by likelihood of being hired, persistence, earning power) will be measured by a test. Asking questions on a test related to work may only predict a tiny percentage of the variance in such important workplace behaviors.

There probably is a positive relationship between workplace performance and WorkKeys performance. The ACT testimonials to WorkKeys value are not without some basis. Any systematic use of a measuring tool such as a test should improve job placement. If a workplace had used no other measures or systematic screening system, using WorkKeys probably would be useful. But many factors can predict future workplace performance. The important questions to ask are "how much can they predict" and "can another predictor do better?"

Given the many millions of dollars spent on WorkKeys, it seems surprising that the test would be accepted on such little supporting evidence. In Illinois alone, one of the five states that administers WorkKeys universally, students have been mandated to take two of the test sections since 2002, and all three will be mandated in 2014. With approximately 150,000 students and \$5.50 for each test section (the figure includes a \$.50 discount from ACT's \$6 rate), this means that the administration of the test cost Illinois \$1,650,000 in 2013 and will cost \$2,475,000 in 2014. All of these students also take the ACT.

Anecdotes, test content validations, and case studies from vendors will not suffice. We still do not know with any certainty how good WorkKeys is, how much variance in workplace behavior it predicts, and whether it is any better than already available information; e.g., ACT tests and high school grades in Illinois. We unbiased scientific research studies to provide such answers.

Incidentally, in reference to the ACE recommendation of college credit being awarded for certain performance level achievement on the WorkKeys, none of the staff members contacted were aware of, or could find information on, the endorsement, and I was unable to find any mention of the ACT National Career Readiness Certificate on the ACE website. Given the lack of definitive evidence on the predictive value of the test, and the lack of precedent for giving college credit on more established tests like the ACT and SAT, the ACE recommendation is utterly without justification.

Conclusion

Standardized tests can be very useful in identifying the overall performance of students at a school. We can be rather certain that students at a school with a composite average ACT score of 25 will be higher performing overall than students at a school with 20. We can even assume that it is likely

(though not as certain) that an individual student with an ACT composite score of 25 will do better in college coursework than one with a score of 20. However, recent research, which clearly is supported by ACT's own research, shows that some ACT tests (math and English), are useful predictors of college performance, and some (reading and science) are not. Given the expanding use of the ACT reading test in state assessments under NCLB, there is a serious disconnect between the research and state policy. Use of the reading test (rather than, for example, the English test) for such purposes is not supported. This does not mean that reading is not integral to student success, but that it already is measured by all ACT subject tests.

Recent studies (as well as past ones) also are clearly demonstrating that the COMPASS and ACCUPLACER are only minimally useful, and useful only as an adjunct to prior classroom performance. Their current overly-frequent use as a single indicator of college readiness is highly inappropriate and unsupported by independent research. Institutions also need to examine whether other already-administered tests (whether currently administered ACT's or SAT's, or a future Common Core assessment) would be equal or superior to the task.

Finally, the validity of WorkKeys in terms of screening students for the workplace is unsupported except by superficial and/or potentially biased research. Unbiased scientific research is needed on whether the test has sufficient value in predicting workplace performance. As with COMPASS and ACCUPLACER, research also is needed on whether the test provides any additional prediction information over already-mandated tests or high school transcripts. Until that time, there is insufficient scientific evidence justifying the use of WorkKeys.

Public education students already are seriously over-tested and money is very limited. Spending many millions of dollars and hours of student and staff time on yet more tests cannot be justified without first making sure that the tests really provide additional useful information.

Students and schools will continue to have to pay attention to such tests as the ACT reading and science, the Compass, the ACCUPLACER, and the WorkKeys when high stakes decisions are being made upon them. But given the availability of good tests like the ACT English and math, as well as the as-yet-not-refuted SAT tests, we have much better tests available to answer those high-stakes questions.

This report is not meant to be an exercise in psychometric nit-picking. The high stakes tests discussed in this report provide measurements that can be seriously life-altering for the students who take them. When there is compelling recent research from nationally established institutions that challenge the validity of such tests, and when the findings from those studies validate each other (in the case of the Columbia studies) or are validated by the vendor themselves (the NBER study), we need to step back and reconsider the use of such tests. Similarly, we must reconsider the use of measures like WorkKeys that have not been independently validated. The educational and research community need to connect placement policy with contemporary research.

References

- ACT and Statewide Testing. *State Services*. (2013). Retrieved from http://www.act.org/stateservices/index_ie.html.
- ACT College Readiness Benchmarks, Retention, and First-Year College GPA: What's the Connection? (2005). *ACT Research and Policy Issues*. Retrieved from <http://www.act.org/research/policymakers/pdf/2005-2.pdf>.
- ACT National Career Readiness Certificate Recommended for Three Semester Hours of College Credit. (2012). Retrieved from <http://www.act.org/newsroom/releases/view.php?lang=english&p=2660>.
- ACT: The First Fifty Years, 1959–2009. (2009). Retrieved from http://media.act.org/documents/ACT_History.pdf.
- Adams, C. (2011). New Study Questions Validity of Two Parts of ACT. *EdWeek*. Retrieved from http://blogs.edweek.org/edweek/college_bound/2011/06/new_study_questions_validity_of_two_parts_of_act.html?utm_source=twitterfeed&utm_medium=twitter
- Allen, J. & Sconing, J. (2005). Using ACT Assessment Scores to Set Benchmarks for College Readiness. *Research and Policy Issues: ACT Research and Policy Reports*. Retrieved from http://www.act.org/research/researchers/reports/pdf/ACT_RR2005-3.pdf.
- Belfield, C. & Crosta, P. (2012) Predicting Success in College: The Importance of Placement Tests and High School Transcripts. *Community College Research Center, Teachers College, Columbia University*. Retrieved from <http://ccrc.tc.columbia.edu/Publication.asp?UID=1026>.
- Bettinger, E., Evans, B., Pope, D. (2011, June). Improving College Performance and Retention the Easy Way: Unpacking the ACT Exam. *National Bureau of Economic Research*. Retrieved from <http://www.nber.org/papers/w17119>.
- Daugherty, L. & Long, B.T. *The Effectiveness of an Intervention to Address Misplacement in Developmental Education*. Paper presented at the 2013 Association for Educational Finance and Policy (AEFP) Annual Meeting, March 15th, 2013.
- Discussion Topic: State Adoptions/ Partnerships. *ACT State Organizations*, 2012. Retrieved from http://www.act.org/stateorgs/docs/Discussion_State_Adoptions.doc.
- External Review Addresses Validity of ACT's WorkKeys Program for Use in Selection. (2010). Retrieved from <http://www.act.org/workkeys/validity.html>.
- History of the Tests (2009). <http://sat.collegeboard.org/about-tests/history-of-the-tests>.
- Jaschik, S. College Entrance Exam ACT's Validity Questioned. (2011). *Inside Higher Ed*. Retrieved from http://usatoday30.usatoday.com/news/education/2011-06-21-act-test-validity-questioned_n.htm#.TgOCEgw5A7c.email.

Lewin, T. (2012) Colleges Misassign Many to Remedial Classes, Studies Find. *New York Times*. Retrieved from http://www.nytimes.com/2012/02/29/education/colleges-misassign-many-to-remedial-classes-studies-find.html?_r=0.

Lichtenberger, Eric J. (2011). Reverse Transfer Students and Postsecondary Outcomes: A Potential Opportunity. *Illinois Education Research Council Policy Research 2011-5*. Retrieved from http://ierc.siue.edu/documents/2011-5_Reverse_Transfer.pdf.

Olsen, J. & House, J. (1997) *Student Achievement in Science and Mathematics General Education Courses: An Assessment of the Effects of Academic Background and Noncognitive Variables*. Paper presented at the Illinois Association for Institutional Research.

Primary Research Group, Inc. (2008). *Survey of assessment practices in higher education*. New York, NY.

Scott-Clayton, J. (2012) Do High-Stakes Placement Exams Predict College Success? *Community College Research Center, Teachers College, Columbia University*. Retrieved from <http://ccrc.tc.columbia.edu/Publication.asp?UID=1026>

Sparks, S. (2013). Many Students Don't Need Remediation, Studies Say. *EdWeek*. Retrieved from <http://blogs.ajc.com/get-schooled-blog/2013/02/27/19506/?cp=2>.

What We Do (2013). Retrieved from <http://www.act.org/about-us/what-we-do/>

Wiley, A., Wyatt, J. & Camara, W. (2010). The Development of a Multidimensional College Readiness Index. *College Board Research Report #2010 3*. Retrieved from http://professionals.collegeboard.com/profdownload/pdf/10b_2084_DevMultiDimenRR_WEB_100618.pdf