



AUTHORS

Robin D. Anderson, Psy.D.
James Madison University

Nicolas A. Curtis, Ed.S.
James Madison University

Tracking the Evolution of Research & Practice in Assessment Through the Pages of RPA

Ten years ago, *Research & Practice in Assessment* (RPA) was born, providing an outlet for assessment-related research. Since that first winter issue, assessment research and practice has evolved. Like with many evolutions, the assessment practice evolution is best described as a change of emphasis as opposed to a radical revolution. Previous areas of exploration and challenge have given way to more complex areas of study as the technical skills of assessment coordinators have advanced and stakeholders have become more willing to incorporate assessment into their practices. However, the promise of learning improvement has been slow to follow. In addition, stakeholders are asking new questions regarding the impact of a college education across one's lifespan.

In the early years, researchers and practitioners at several institutions across the Commonwealth of Virginia were engaged in research to advance assessment practice. Some of this research involved investigating the impact of assessment strategies on the validity of assessment findings. For example, conducting research on issues such as how proctors, instructions, and testing conditions affected student motivation and efforts on institutions' student learning assessments became a regular part of the assessment practitioner's role on many college campuses. Those practitioners with measurement training were also conducting extensive studies on the psychometric properties of testing instruments and on the impact of rater characteristics on the reliability of ratings for performance assessments. Researchers and students were routinely conducting these scholarly studies, master's theses, and doctoral dissertations related to assessment practice. However, the authors were often challenged to find a journal that would publish such assessment studies, even if these studies were steeped in traditional research methodology. Research practitioners could present at the Virginia Assessment Group conference each fall and some also published articles in assessment-focused publications such as *Assessment Update*. However, the Virginia Assessment Group did not publish conference proceedings, and *Assessment Update*, while providing an open and professional venue for brief assessment pieces, did not feature full-length journal articles.

CORRESPONDENCE

Email
ander2rd@jmu.edu

In 2005, the forward thinking Virginia Assessment Group members addressed this scholarly gap. They converted the online newsletter to an online peer-reviewed journal that would disseminate evolving research coming out of institutions across not only Virginia but also the Mid-Atlantic and beyond. We knew that assessment practice was accelerating and here, ten years later, we see this is particularly true in the areas of measurement and “meta-assessment.”

Like with many evolutions, the assessment practice evolution is best described as a change of emphasis as opposed to a radical revolution.

In the winter of 2006, when RPA first hit the web, many stakeholders were focused on the quality of our institutions’ assessment practices. An unwillingness to make programmatic changes based on assessment findings was often couched in concerns regarding the quality of assessment instruments and the soundness of assessment procedures. Practitioners, under increasing pressure from unfunded mandates, were having to build or gear up their assessment initiatives at a mindblowing pace, requiring a heavy reliance on the adoption of pre-existing instruments or the hurried production of local assessments. In essence, we were too often building the plane while also trying to fly it. Practitioners understood that if stakeholders were to invest in the use of assessment findings to improve student learning, stakeholders would need to know that they could make valid inferences from the assessment results. For that to happen, practitioners would have to answer tough questions regarding the reliability of our measures and the validity of our inferences.

Accordingly, a significant portion of the assessment-related scholarship appearing in the early issues of RPA centered on issues of reliability and validity (Ridley & Smith, 2006; Pieper, Fulcher, Sundre, & Erwin, 2008). Assessment researchers needed to fill the gaps within the psychometric literature. These research practitioners wanted to ensure we could document appropriate forms of reliability including test-retest reliability, parallel forms reliability, internal consistency, and interrater reliability. We worked to understand differential item functioning (Thelk, 2008) and its impact on the interpretation of assessment findings. While at least limited reliability evidence was often available for measures of student learning, validity evidence was somewhat nonexistent. Over time, additional research was conducted to document validity evidence. Many of these studies took the form of expected group differences studies, concurrent validity studies, and studies examining the factor structure of new and existing assessment instruments (Smiley & Anderson, 2011). Others were working to encourage practitioners and stakeholders alike to think differently about how to interpret findings using effect sizes (Springer, 2006). Today, much of this work would be considered preliminary psychometric work. For example, we regularly engage in equivalence studies before we compare groups based on the findings from a given instrument. We do this to ensure that the scores do not simply demonstrate reliability and that our inferences are supported but also that these things hold true in a similar way for each of the groups involved.

As further evidence of the advancement in practice, one needs simply to look at how the training offerings for assessment practitioners have changed. When RPA first launched, trainings and workshops in structural equation modeling (SEM) and item response theory (IRT) were considered “advanced.” Today, there are graduate courses and professional workshops in *advanced* SEM, *advanced* IRT, hierarchical linear modeling (HLM), and Bayesian analyses. We have evolved past the days where basic psychometric work is what we do; instead it is a given, a necessity, a starting point from which we now build our research agendas. Finding the “right methodology” is still at the center of what we do as assessment practitioners; however, we now do so much more.

As we have moved beyond basic psychometric work, so too have we moved beyond the days of convincing programs that they must engage in student learning outcomes assessment. In 2006 when RPA was launched, many assessment practitioners were still working to convince academic programs on campus to even minimally engage in program assessment. As Marilee Brenciani (2011) stated “there are many others who do not believe the process [outcomes-based assessment] adds value to their day-to-day teaching or administrative duties” (p. 5). Thus it is not surprising that much of the scholarship conducted by assessment practitioners was aimed at convincing faculty that assessment was worth doing and could benefit faculty members’ efforts to improve student learning. Even

those in areas with a long history of program evaluation found themselves in the position of arguing for a shift toward a greater focus on student learning outcomes. In his piece focused on assessment in academic libraries, Eric Ackermann (2007), pointed out that despite academic libraries having a nearly 100 year history of conducting assessment, they were still using “metrics and reporting protocols that are meaningful primarily to other librarians” (p.23). Ackermann called for new processes that would examine the relationships between academic libraries and student learning, stating “such information could only help the efficacy of libraries across the country” (p. 23). Assessment experts from academic and students affairs alike were producing such argumentative manuscripts in an effort to make the case for assessment.

A decade later, at many institutions, it is a given that programs are engaged in the ongoing assessment of student learning outcomes. Today many of the efforts of assessment offices focus on supporting these program initiatives and assessing the assessment. An increasing number of institutions have systematized a process to gather, review, and evaluate the assessment processes being used in the programs. Many of these assessment-of-the-assessment processes, or meta-assessments, include a step by which reviewers or raters provide feedback to the programs for the purpose of improving programs’ assessment practices and hopefully also student learning outcomes.

While we have made huge strides in addressing many of the major challenges that face assessment researchers and practitioners at the time RPA was launched, we remain challenged to help assessment research and practice to evolve to the next level. While certainly no one can predict the future, we are seeing evidence that points to some of the major areas of focus in assessment practice for the coming years. These foci include learning improvement, lifelong impact of college, and the engagement of new groups in the assessment process. One of the criticisms we still hear regarding assessment is that institutions struggle to identify specific examples of where student learning has improved as a result of lessons learned through assessment. However, this is not overly surprising. As many freshmen psychology students would be able to tell you, basic needs need to be met before higher order needs can be addressed (thank you, Maslow). We argue that before widespread learning improvement might occur, we needed to address the basics. We needed to strengthen our assessment processes to ensure that those who are in the position to use our findings have confidence in the results. This means all those studies on the quality of our instruments (Barnes & Burchard, 2011) and on the impact of our systematic assessment procedures (Barry & Finney, 2009) on the validity of our inferences needed to become common place. Now, with a long record of disseminated research and research-based assessment practice becoming common, we are standing at yet another launching point. We believe that this next great advance will be the widespread institutionalization of learning improvement initiatives based on student learning outcomes assessment findings.

Keston Fulcher talks extensively regarding the learning improvement initiatives taking place at James Madison University (JMU) (Fulcher, Smith, Sanchez, Ames, & Meixner, 2017). These initiatives are grounded in a partnership with the University’s Center for Faculty Innovation, the teaching and learning center on campus. Such partnerships are proving to be the cornerstone of many learning improvement initiatives. In addition to partnerships like the one at JMU between the Center for Assessment and Research Studies and the Center for Faculty Innovation, we are observing an increase in partnerships between assessment practitioners and those faculty responsible for assessment within the disciplines. Much of the authors’ own research is done in partnership with longtime colleagues in our engineering department. It is not unusual for assessment practitioners, with their skills in measurement and evaluation, to serve as evaluators on grants. However, within our partnerships we often serve, not as evaluators, but as full co-principal investigators on major initiatives. We do not merely have a seat at the table or serve to represent the university as the “assessment person” on the committee, but we serve as full partners contributing at every stage of the project. The evolution of these partnerships between discipline faculty and assessment practitioners helps to merge the roles of content experts and measurement experts increasing the likelihood that findings will lead to improvements in student learning and an increase in interprofessional practice.

We believe that this next great advance will be the widespread institutionalization of learning improvement initiatives based on student learning outcomes assessment findings.

A review of more recent editions of RPA reveals an increasing number of improvement focused manuscripts. Scholarly pieces calling for a greater focus on learning improvements appeared in RPA as early as 2008. In the 2008 winter edition, Pieper et al. discuss four analytic strategies for addressing educational research questions that could help institutions “reap the maximum benefits from their assessment efforts” (p.4). However, starting in 2013, articles focusing on using data for improvement began to appear with increasing frequency in RPA (e.g., Gerstner & Finney, 2013). By winter 2015, at least three of the articles appearing in RPA focused on the use of assessment results for improvement purposes. As reflected in the pages of RPA, we see using assessment to inform changes that lead to learning improvement as an idea whose time has come.

**In our opinion,
research and assessment
practitioners should
be pleased with our
progress and excited
by our future.**

A second area emerging in our field is an interest in assessing the lifelong impact of the college experience. Efforts by organizations such as Gallup to gather data on lifelong satisfaction and the role college plays in how we view ourselves and the quality of our lives have expanded our view beyond the initial experience of college. Now the horizon is not the senior year, or even graduation, but life five, ten, or forty years beyond the graduation ceremony. With partners such as Gallup, institutions hope to be able to show that the truly meaningful outcomes of college expand well beyond the senior (or fifth) year. One such partnership between Gallup, the State Council of Higher Education for Virginia (SCHEV), and the Virginia Assessment Group resulted in a Summer 2015 Summit on Quality & Value in Virginia’s System of Higher Education. As a part of the summit, representatives from the University of Virginia, Virginia Tech, and George Mason University spoke about the importance of preparing students for both work and life. Ray and Kafka (2014) summarize the inaugural findings of the Gallup-Purdue Index stating “when it comes to being engaged at work and experiencing high well-being after graduation...the type of institution they [students] attended matters less than what they experienced there” (para. 1). Such findings open up an entirely new area of assessment practice and research, redefining alumni assessment. The Gallup-Purdue Index, and the importance the initial results place on college experiences, ties to our final thought on emerging assessment trends, student engagement.

As we discussed, in the “dark ages” (i.e. 10 years ago), many faculty were resistant to program assessment. While many things contributed to our current, more assessment-friendly, climate, one of the biggest contributors was the push to engage faculty at the beginning of the assessment process. Not only did this change the climate, but it also improved the assessment process itself. Now, we believe that a similar engagement evolution is on the horizon for a different audience, the students that we assess. If successful, this shift could again radically alter the value of assessment in higher education. If our goal is student learning improvement, why would we not partner with students at all levels of the assessment process? If our goal is student learning improvement, why would we try to interpret results and make decisions without students’ input. If our goal is student learning improvement, why should we not leverage students’ obvious stake in the outcomes of our assessments? This is not a radical idea. Charlie Blaich and Kathy Wise, of the Center for Inquiry at Wabash College, have been working with undergraduate students to train them to conduct peer-led groups focused on assessment results. These students have been able to help get at the “why and how” of the assessment data. The students are often able to help frame assessment results to make them more useful and compelling. We are optimistic and excited about the impact partnering with students may make on assessment practice. Research on such topics are just beginning to appear in publication. We believe that as RPA covers the next 10 years of assessment practice these will be among the topics that appear in its pages.

In our opinion, research and assessment practitioners should be pleased with our progress and excited by our future. For those of us who have been at this a while, it is important to occasionally remind ourselves just how far we have come. Given that the ultimate goal of assessment is to improve student learning, we believe we are just now positioned to see the fruits of our labors. Yet, those just coming into the field can rest assured that there are still discoveries to be made and challenges to be overcome.

References

- Ackermann, E. (2007). Program assessment in academic libraries: An introduction for assessment practitioners. *Research & Practice in Assessment, 2*, 18-23.
- Barnes, S.K. & Burchard, M.S. (2011). Quality and utility of the Multi-tiered Instruction Self-Efficacy Scale. *Research & Practice in Assessment, 6*(2), 22-42.
- Barry, C.L. & Finney, S.J. (2009). Does it matter how data is collected? A comparison of testing conditions and the implications of validity. *Research & Practice in Assessment, 4*, 17-26.
- Bresciani, M.J. (2011). Identifying barriers in implementing outcomes-based assessments program review: A grounded theory analysis. *Research & Practice in Assessment, 6*(1), 5-16.
- Fulcher, K. H., Smith, K. L., Sanchez, E. R. H., Ames, A. J., & Meixner, C. (2017). Return of the pig: Standards for learning improvement. *Research & Practice in Assessment, 11*(2), 10-27.
- Gerstner, J.J. & Finney, S.J. (2013). Measuring the implementation fidelity of student affairs programs: A critical component of the outcomes assessment cycle. *Research & Practice in Assessment, 8*(2), 15-28.
- Pieper, S.L., Fulcher, K.H., Sundre, D.L., & Erwin, T.D. (2008). "What do I do with the data now?": Analyzing a comparison of testing conditions and the implications for validity. *Research & Practice in Assessment, 3*, 4-10.
- Ray, J. & Kafka, S. (2014). Life in college matters for life after college. Retrieved from: <http://www.gallup.com/poll/168848/life-college-matters-life-college.aspx>
- Ridley, D.R. & Smith, E.D. (2006). Writing across the curriculum works: The impact of writing emphasis upon senior exit writing samples. *Research & Practice in Assessment, 1*, 14-17.
- Smiley, W. & Anderson, R. (2011). Measuring students' cognitive engagement on assessment tests: A confirmatory factor analysis of the short form of the cognitive engagement scale. *Research & Practice in Assessment, 6*(1), 17-28.
- Springer, R. (2006). Using effect size to assist in NSSE reporting. *Research & Practice in Assessment, 1*, 18-22.
- Thelk, A. (2008). Detecting items that function differently for two- and four- year college students. *Research & Practice in Assessment, 3*, 23-27.