



# OP-ED



THE FUTURE OF ACCREDITATION: THE CHEA OPINION SERIES

Issue 12 • October 2017

## Why Accreditors Can't and Shouldn't Adopt the Consumer Reports Methodology

Diane Auer Jones

**I**t is time to end the charade of crediting institutional quality for student outcomes that are driven mostly by admissions selectivity. If we want to provide more accurate and useful data to inform student choice, we must first develop more accurate methods to measure, compare and report institutional quality.

**H**ow many times have you heard the complaint that it is easier to find more information about the quality of a \$30 toaster than it is about a \$50,000 education? That question is typically followed by a call for accreditors to adopt a *Consumer Reports* rating methodology, as if students are machines and learning is manufactured by the institution a student attends.

**I** recently turned to *Consumer Reports* (CR) to inform my choice of a new toaster (my 30-year-old wedding gift finally met its fateful end) and in so doing, was reminded of why CR does not provide a relevant model for institutional quality rankings. To conduct their 2015 quality test of toasters, *Consumer Reports* subjected 600 slices of Freihofer's bread to toasting in a variety of two- and four-slice models. Apparently Freihofer's bread has just the right texture and consistency for precision tests of color, crunch and toaster speed.

**S**tudents, however, are not equally measured slices cut from identical loaves of humanity. *Consumer Reports* would never evaluate toasters by putting a fresh-from-the-oven slice of bread in a bargain brand, a week-old slice in a mid-level appliance and a frozen slice in a luxury model, because doing so would leave testers wondering if the bread or the toaster was responsible for the results. Most sixth graders would recognize the flaws in that experimental design (sixth grade is when the concepts of independent, dependent and controlled variables are typically introduced) and likely conclude it was invalid.

**S**imilarly, outcomes assessments that ignore the many confounding variables that influence student outcomes should not serve as the foundation for accreditation reviews – or performance-based funding schemes or practically any other public policy decision. Instead, these assessments should be the subject of intense scrutiny since in any other research context, such an invalid methodology would never pass the scientific method sniff test.

**O**ur current outcomes measures tell us much more about the institution's admissions criteria than they do about what happens once a student enrolls. In addition, these assessments may also, due to Simpson's Paradox, lead us to blatantly wrong conclusions. Simpson's paradox is a statistical phenomenon in which trends that appear in disaggregated groups disappear or reverse when these groups are combined. To illustrate Simpson's paradox, consider that Institution A posts a graduation rate of 80 percent and Institution B posts a rate of 33 percent. Today's quality assessment methodology would lead us to believe that Institution A is the superior institution.



**H**owever, when outcomes are disaggregated by student risk groups (i.e., traditional, minimally non-traditional, moderately non-traditional and highly non-traditional),<sup>1</sup> it turns out that Institution B achieves the same graduation rate as Institution A among low-risk students and Institution B achieves a slightly higher graduation rate among high-risk students, meaning that for each of those student groups, Institution B performs as well as or better than Institution A. Institution A's aggregate outcomes overstate its performance because the population it serves is comprised primarily of low-risk students whereas Institution B's population is comprised mostly of high-risk students. Therefore, Institution A isn't necessarily a higher quality institution, nor do its students – when considering matched comparison groups – graduate at a higher rate. Institution A is simply more selective.



**T**he negative consequences of invalid aggregate outcomes go beyond reputational damage. Institution B, for example, might be denied funding or renewal of accreditation (if accreditors are forced to adopt bright line standards), even though the disaggregate analysis shows that Institution B serves its students at least as well as Institution A. Even worse, these aggregate results could dissuade a student who would be well served by Institution B from enrolling there, and could dissuade an employer from hiring a graduate of Institution B, even if the graduate was a top student. When the “low quality” moniker is inappropriately applied to an institution that actually serves students well and maintains academic rigor, the achievements of hard working students are undervalued and their future opportunities are inappropriately limited.

**A**t Urban Institute, I am working with a group of pilot institutions and an outside data and analytics firm, MeasureOne, to test a new model we've developed, a more accurate assessment methodology that takes into account a number of important student and institutional characteristics, including per student spending. After all, *Consumer Reports* would not typically put a \$30 toaster and a \$500 model in the same comparison group. Our methodology enables data to be disaggregated so that students with similar risk profiles can be compared from one institution to the next as well as to national averages. It can also provide institutions with evidence-based outcomes targets that appropriately motivate better performance without incentivizing grade inflation or K-12 style social promotion. Our ongoing pilot study suggests that a number of laudable institutional results were simply hidden by the invalid methodology on which current accountability assessments are based. We welcome additional institutions to join the pilot project to test the validity and usefulness of risk-adjusted outcomes.

**W**hile those advantaged by the current methodology might suggest that risk adjusting simply excuses poor performance, I note that risk adjustment is already part of the Department of Health and Human Services' (HHS) methodology for evaluating the quality of managed care organizations. Risk adjusting patient outcomes allows HHS to more carefully examine the value added a healthcare organization provides rather than the zip code in which its health professionals practice. This suggests that HHS, rather than *Consumer Reports*, offers a better model for performing institutional quality assessments.

**R**isk-adjusted outcomes can tell us a great deal about the value added an institution provides, and it may be the most effective way to achieve our public policy goals of expanding access and reducing college costs. After all, the pressure on institutions to become more selective has widened the opportunity gap and fueled the amenities arms race that has contributed to soaring college costs.

<sup>1</sup> <https://nces.ed.gov/pubs/web/97578g.asp>

*Diane Auer Jones is Senior Fellow at Urban Institute and former Assistant Secretary for Postsecondary Education at the U.S. Department of Education.*



One Dupont Circle, Suite 510  
Washington, DC 20036  
Phone: 202-955-6126 Fax: 202-955-6129