

## Metrics Mania

An essay submitted to

Scholar-Practitioner  
NEXUS

<http://scholarpractitionerexus.com/>

by

Helen M. Hazi, Ph.D.  
Professor Emerita  
West Virginia University

Metrics can be a method and product of measurement. The use of numbers to quantify human reality emerged in post-Renaissance Europe (Neylan, 2005). Numbers began to replace words when merchants needed reliable measures of distance, time and trade (Labaree, 2011b). A government also needed knowledge to govern and one of its bureaucratic tools was large banks of official statistics about the social needs of its citizens used to identify priorities of the state. Statistics, “knowledge of the state,” (Neylan, 2005, p. 24) allows the state to surveil patterns and irregularities of death, disease, and scarcity to plan for and manage the needs of its citizens, and at the same time to monitor and control them. Quantification also emerged as a reaction to a medieval society dominated by religious rule.

Statistics supports the bureaucratic practice of evidence-based decision to support social policy. This view “takes a positivist approach to what constitutes evidence, preferring the numerical outputs of research such as statistical data, over non-numerical knowledge such as values, intuition and practical know-how” (Neylan, 2005, p. 25). A key reason for the rise of social quantification is the bureaucrat’s greatest fear—that the public believes they have acted

unfairly and politically. Numbers give the appearance of neutrality, objectivity, and precision, and communicate that the government is honest and fair in its allocation of resources.

Numbers are ideally constructed from systematic measurement processes that capture predictable properties. Social entities generally are not of this character....social entities are often unpredictable and possess little of the repeatable, systematic nature of a physical entity such as distance or weight. Even so, we find that the most common way of signifying social entities is through numbers. (Neylan, p. 24)

“In quests for objectivity quantification becomes most important where elites are weak, private negotiation is suspect and trust is lacking” (Neylan paraphrasing Porter, 1995, p. 37). The single figure, especially, gives the appearance of stability to highly subjective entities and is difficult to dispute. Since *A Nation at Risk* (National Commission on Excellence in Education, 1983), trust in public schooling has most recently been lacking, and since *The Widget Effect* (Weisberg, 2009), trust in teacher evaluation and teacher quality has also been lacking.

While testing was ushered into education first with IQ testing in the 1900s, our metrics mania has been further fueled by many policy aims to include: state and international comparisons, the search for effective (and improved) schools, and the identification of weak teachers. Leaders equate test scores with individual and national success (Zhao, 2016). Our current preoccupation with testing was fueled by the *No Child Left Behind* era. With the addition of multiple measures in the post-NCLB era, the metrics now require complicated weighting in what some call “multimetric accountability” (ASCD, 2016).

In the current teacher evaluation arena, a continuum of metrics is in use. The single figure is used for students (the test score) and teachers (the overall rating). A complex and opaque algorithm is used to calculate student growth gains. The controversial metric is the use of

student test scores in teacher evaluation. It is most contentious when its weighting is at least 50% or “significant” (Hazi, 2016).

Educators have turned to metrics because psychometricians provide the tools to measure. These tools, however, allow us “to focus on what they can measure statistically rather than on what is important” and “to be methodologically sophisticated at exploring educational issues that do not matter” (Labaree, 2011b, p. 625). Educators are experiencing metrics mania, the overvalued use of numbers to gauge and measure complex social phenomena that results in oversimplified understandings and unintended consequences. “While numbers can help to simplify complex phenomena, move policy makers to more precise thinking, and remove the appearance of bias in decision making, numbers do have the potential to convey unintended meanings and mislead a fuller understanding of complex events, situations and practices” (Hazi & Garman, 2007, p.1).

It appears that educators will remain a prisoner of metrics mania. Reports continue that despite changes made to teacher evaluation in the states, very few teachers receive low ratings (Brown, 2016; Sawchuk, 2016). Unfortunately, policy makers believe that when students score poorly on tests, then teacher ratings should mirror student achievement. They believe that more teachers should be rated as unsatisfactory and more should be dismissed in these low-performing schools (Butrymowicz, 2014). Some blame principals for not being honest (e.g., Brown, 2016) instead of believing that the overall metrics schemes fail to capture teacher quality.

## References

- ASCD. (2016). *Multimetric accountability systems.: A next-generation vision of student success.* Alexandria, VA: Author. Retrieved from <http://www.ascd.org/public-policy/multimetric-accountability.aspx>
- Brown, E. (2016, February 29). Very few teachers receive poor job ratings and new evaluations haven't changed that. *The Washington Post*. Retrieved from <https://www.washingtonpost.com/news/education/wp/2016/02/29/very-few-teachers-receive-poor-job-ratings-and-new-evaluations-havent-changed-that/>
- Butrymowicz, S. (2014, June 16). How many bad teachers are there? Not many according to new—and expensive—evaluations. Retrieved from [http://hechingerreport.org/content/many-bad-teachers\\_16397/](http://hechingerreport.org/content/many-bad-teachers_16397/)
- Hazi, H.M. (2016). Educators challenging teacher evaluation in a high stakes climate in the states. A paper presented to the Division L Educational Policy & Politics, Section 6: Human Capital, AERA, Washington DC, April 8-12, 2015.
- Hazi, H. M., & Garman, N. B. (2007, April). “Teach by Numbers”: Is technological momentum breeding a culture of compliance, surveillance and seduction in American schools? A paper presented to the SIG: Supervision and Instructional Leadership at the annual meeting of the American Educational Research Association, Chicago.
- Labaree, D. (2011). The lure of statistics for educational researchers. *Educational Theory*, 61(6), 621-632.
- The National Commission on Excellence in Education. (1983). *A nation at risk: The imperative of educational reform.* Washington, DC: Author. Retrieved from <https://www2.ed.gov/pubs/NatAtRisk/risk.html>
- Neylan, J. (2005). Quantifying social entities: An historical-sociological critique. *Journal of*

*Sociology and Social Welfare*, XXXII(4), 23-40.

Sawchuk, S. (2016, March 1). Despite teacher-evaluation changes, the 'widget effect' is alive and well. Retrieved from

[http://blogs.edweek.org/edweek/teacherbeat/2016/03/teacher\\_evaluation\\_results\\_the.html?cmp=enl-enl-eu-news2](http://blogs.edweek.org/edweek/teacherbeat/2016/03/teacher_evaluation_results_the.html?cmp=enl-enl-eu-news2)

Wachter, R. (2016, January 16). How measurement fails doctors and teachers. *The New York Times*. Retrieved from [http://www.nytimes.com/2016/01/17/opinion/sunday/how-](http://www.nytimes.com/2016/01/17/opinion/sunday/how-measurement-fails-doctors-and-)

[measurement-fails-doctors-and-](http://www.nytimes.com/2016/01/17/opinion/sunday/how-measurement-fails-doctors-and-teachers.html?emc=edit_tnt_20160117&nliid=28255360&tntemail0=y&r=0)

[teachers.html?emc=edit\\_tnt\\_20160117&nliid=28255360&tntemail0=y&r=0](http://www.nytimes.com/2016/01/17/opinion/sunday/how-measurement-fails-doctors-and-teachers.html?emc=edit_tnt_20160117&nliid=28255360&tntemail0=y&r=0)

Weisberg, D., Sexton, S., Mulhern, J., & Keeling, D. (2009). The widget effect: Our national failure to acknowledge and act on differences in teacher effectiveness. Brooklyn, NY: The New Teacher Project. Retrieved from [www.widgeteffect.org](http://www.widgeteffect.org)

Zhao, Y. (2016). Numbers can lie: The meaning and limitations of test scores. In Y. Zhao (Ed.) *Counting what counts: Reframing education outcomes* (pp. 13-29). Bloomington, IN: Solution Tree Press.