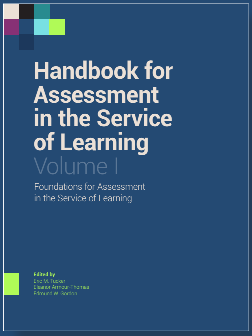
It's Time for a Paradigm Shift in Educational Measurement

Pamela Cantor, M.D. and Kate Felsen

UMassAmherst
University Libraries



Series Editors:
Edmund W. Gordon, Stephen G. Sireci, Eleanor Armour-Thomas, Eva L. Baker, Howard T. Everson, and Eric M. Tucker

ISBN: 978-1-945764-33-2

Suggested Citation:
Cantor, P., & Felsen, K. (2025). It's time for a paradigm shift in educational measurement. In E. M. Tucker, E. Armour-Thomas, & E. W. Gordon (Eds.), *Handbook for assessment in the service of learning, Volume I: Foundations for assessment in the service of learning.* University of Massachusetts Amherst Libraries.

It's Time for a Paradigm Shift in Educational Measurement

Pamela Cantor and Kate Felsen

Science and History

For nearly 1,500 years, from the time of Ptolemy until the 17th century, most people believed the Earth was the center of the universe and the planets and the sun revolved around it. Maps and almanacs depicted this geocentric system that seemed logical at the time. Logical, but also dead wrong. And that meant early navigation tools were wrong, so some ships never reached their destinations.

Then, in 1543, Nicolaus Copernicus proposed that the sun was the center of the universe and built mathematical models accordingly (Copernicus, 1543). Over the next century and a half, astronomers, including Galileo Galilei, refined telescopes and designed better navigation tools. Marine chronometers, sextants and octants improved and so did travel by ship. As the Copernican heliocentric view took hold, understanding of the nature of the universe grew.

In 1928, Alexander Fleming discovered that a mold in one of his petri dishes—something he called penicillin—could kill bacteria (Fleming, 1929). At that time many still believed that miasma or foul air caused disease. Few grasped the concept of germ theory posited by scientists such as Louis Pasteur and Robert Koch, that particles not visible to the human eye, whether bacteria or viruses, were wreaking havoc on human health (Koch, 1882; Pasteur, 1878).

But when penicillin was used to treat soldiers' infected wounds during World War II and saved millions of lives, germ theory took hold. This led to the development of new antibiotics, vaccines and antiseptics. Coupled with better optics in the form of microscopes that could observe pathogens, more accurate diagnoses and better interventions became possible. The field of medicine was changed forever.

The Structure of Scientific Revolutions

In 1962, Thomas Kuhn wrote a landmark essay called "The Structure of Scientific Revolutions" where he described the conditions that can bring about a seismic change, what he called a paradigm shift, where theories and applications are given up in favor of new and improved ones that lead to more and more coherent and innovative models (Kuhn, 1962).

Kuhn stated that three conditions precede scientific revolutions:

1. A flawed theory that leads to what he called mistakes, applications that are incorrect or insufficient and pile up because they are not based in solid science. These mistakes cause mounting frustration and failure, making anything other than incremental change impossible;
2. New instrumentation that enables new mathematical models and algorithms that better account for what is observed; and
3. An alternate theory that better explains what is observed in practice and in human beings.

Based on Kuhn's rubric, we are living in a time that is ripe for a scientific revolution in how we understand human growth and learning, human performance and potential, and how we measure it.

Flawed Theory

Flawed theories are not only alive and well today, they are also foundational principles for learning and measurement systems in the United States. One of the most damaging is the idea that intelligence is fixed, that our brains cease to develop at the time we start kindergarten. That is not true. The human brain is malleable, even highly plastic, and can grow, heal and change throughout our lifetimes (Cantor et al., 2019). All the more reason to measure an individual over time, rather than once or twice a year.

Consider the persistent myth about the power of genes. Many people still believe that genes are the drivers of who we become, including our intelligence. That is false. What matters most is not whether we have a particular gene, but what our genes are doing. Genes can be turned on and off like a dimmer switch. Context–the environments, relationships and experiences in our lives – controls this switch (Cantor & Osher, 2021).

Our genome is not just something we are born with. It develops, as we do, over our lifetimes because of the influence of context. The field of epigenetics describes this phenomenon; "epi" means what sits above—in this case, above the gene—and influences or drives its expression. The genome you have at birth is not the same as the one you will have when you die or the one you will pass on to your children (National Scientific Council on the Developing Child, 2010). David Moore's 2015 book, *The Developing Genome* describes this developmental process in fascinating detail (Moore, 2015).

A third flawed theory is that talent is scarce and distributed like a bell curve, with most of us falling in the middle and a smaller number on each of the tail ends. The notion that only a select few are capable of high levels of achievement has been hard to shake, leading to deeply ingrained inequities in education and employment that constrain people from achieving what they might under more equitable circumstances. This statistical model has been disproven for decades, most powerfully in Stephen Jay Gould's *The Mismeasure of Man* (Gould, 1981). The truth is that human variation is the norm in human development, not the exception (Rose, 2016). No group, race, or ethnicity has more talent or intelligence or more potential for developing it. Talent is everywhere. We don't always look for it, see it, or recognize it, but it is out there, and there are many pathways to develop it.

A fourth flawed theory concerns potential or developmental range. What's false is that human potential is knowable in advance, that we can test for it. What's true is that the fullest expression of a human being's potential depends on contextual factors to reveal and build it (Cantor et al., 2021). Every human skill has a developmental range of performance, and for the most part we all live without knowing what the upper end of that range might be.

By contrast, when a context is designed to reveal capabilities—in learning, in sports, in the workplace—we can discover potential and grow it. Kurt Fischer called this kind of context a "constructive web" in his 2007 paper, "The Dynamic Development of Action and Thought" (Fischer, 2007). When he wrote it, Fischer surely didn't have the 2024 Paris Olympic Games in mind, but every single story of athletic performance we witnessed there is a story of the bi-directional connection between developmental range and the power of context.

Take 9-time Olympic gold medalist Katie Ledecky. In a poolside interview with NBC Sports on July 31, 2024 she described what she was thinking as she pulled away from the field in the 1500 meter freestyle: "I let my mind wander during the race, thinking of all the people that have trained with me, just kind of saying their names in my head and thinking about them" (NBC Sports, 2024). Ledecky trained in contexts that included relationships with teammates, coaches, and physiotherapists who knew how to bring out her best and the best in each other. And she was reliant on data from underwater cameras about each stroke, kick, and training regimen to do it. The full web of experiences constructed around this one human has helped her become one of the greatest swimmers of all time.

New Developmental Insights Lead to a Dynamic Approach to Measurement-3D Measurement

Like Copernicus, who wanted to understand the whole universe, measurement in medicine seeks to understand the whole human being. To do so, medical professionals use three forms of measurement: population, differential, and person-specific (Parrish, 2010). Medicine also assesses contextual factors around each patient, including nutrition, sleep patterns, and living conditions (Cook et al., 2023).

When a doctor takes your blood and compares your white blood cell count to large populations to tell if yours is within the normal range, that is an example of population measurement. In sports, a Nordic skiing or biathlon coach might want to measure an athlete's resting heart rate and compare it to the general population. If the athlete is fit, it should be lower than the population average. If it's not, it might be a sign of inadequate conditioning or maybe illness.

Differential measurement is about comparisons, comparing something about you, such as high blood pressure, to people who are like you, say women in their 50s. When it comes to performance, we need to compare your score to the top scores to know how to shape the goal you are shooting for—such as VO2Max in sports with sprints or a five on an Advanced Placement exam.

The third form of measurement is person-specific, looking at an individual's progress over time. In medicine, this might mean looking at the range of motion in your wrist or ankle four weeks after surgery and comparing it two months later. In music, teachers of brass or wind instruments such as the trumpet or bassoon might measure lung capacity or diaphragm strength in their pupils, then tailor individual breathing exercises for them. Weeks later, they might repeat the student's measurements to see if the regimen has resulted in positive changes. In athletics, runners training for the 100-meter dash want to know if their training regimen is helping to lower their time from one meet to the next.

In most learning and workforce settings, however, we don't measure in three dimensions. We rely heavily on standardized assessments, many of which are interpreted using norm-referenced frameworks—comparing individuals to a generalized group rather than to relevant peers or performance benchmarks. This approach often ignores individual context, which is crucial for personalizing learning and growth (Cantor et al., 2021; Immordino-Yang et al., 2023).

No surprise, aspiring Olympians are highly attuned to how they feel physically and mentally. It's also true that they understand acutely that the interaction between their bodies and who and what's around them impacts whether they will stand on top of the podium or not. In other words, athletes perform differently at different times and in different contexts, and they need to know why to optimize performance. We do not do this for learning, at least not yet.

Educational measurement should be done in such a way that we learn more about the fit between an individual student and the learning context. This is doable today. It is fit that amplifies purpose and confidence. It is fit that primes performance. It is fit that produces cures. It is fit that unlocks human potential. The Rosetta Stone for measurement is about how close we can get to measuring the individual, understanding and measuring the context, and measuring the fit between the two (Cantor et al., 2021). If this existed today, all learners, and all of us, would understand what we need to reach the top of our developmental range, and what, in our community, school, work, team, troop, would enable us to perform at our best.

New Paradigms

Thomas Kuhn helped us to see that new models and innovations can flow out of changes in theories, mathematics, instrumentation, and methods. The mathematical psychologist Peter Molenaar has moved us to understand that averages do not tell us enough about individuals. In his 2004 paper, "A Manifesto on Psychology as Idiographic Science: Bringing the Person Back into Scientific Psychology, This Time Forever", he demonstrated that it is possible to measure human development in context and even gain insights about the specific impact of a specific aspect of context on a specific person at a specific moment in time (Molenaar, 2004).

Dr. Eli Van Allen is doing something similar at the level of the cell. He leads Clinical Computational Oncology at the Dana Farber Cancer Institute in Boston where his approach to data analytics has led to innovations. One of them is the ability to describe the relationship between a cancer cell and its context (Van Allen et al., 2016). Today, targeted immunotherapies can make changes to the context of individual cancer cells in a patient and destroy those cells without damaging surrounding tissue. This discovery is revolutionizing the treatment of highly lethal cancers.

Emerging technologies and dynamic measurement methods can help us reshape how we understand and nurture talent, learning, and performance. The integration of artificial intelligence (AI) into assessment methodologies is giving us new insights into how individuals learn, develop skills, and optimize their abilities over time. Along with machine learning and computational modeling, AI produces:

- **Personalized learning environments** that adapt to individual strengths, weaknesses, and learning styles
- **Real-time feedback mechanisms** that help learners and educators make data-driven decisions about instructional strategies.
- **Advanced predictive analytics** that identify potential barriers to success and recommend targeted interventions

By integrating AI-driven assessment tools, educators and employers can gain deeper insights into an individual's learning patterns, strengths, and areas for improvement. AI-powered adaptive learning platforms are already demonstrating success in tailoring instructional content to individual needs, optimizing the pace and complexity of material based on a learner's progress. In workforce development, AI-driven skills mapping and competency-based assessments can help match individuals to roles that align with their strengths and career aspirations. And by embracing dynamic measurement systems, we can move away from rigid evaluation models and toward a more personalized and adaptive approach to human development.

Can anyone become a gymnast like Simone Biles, a cellist like Yo-Yo Ma or a poet like Amanda Gorman? Likely no. But none of us can know what is encoded in our DNA without experiences that bring that code to life and without measurement capability that can see potential as it is unfolding. All of us could discover the top of our developmental range if only we had the tools to see it.

References

Cantor, P., Gomperts, N., Lerner, R. M., Pittman, K., & Chase, P. (2021). *Whole-child development, learning, and thriving: A dynamic systems approach.* Cambridge University Press.

Cantor, P., & Osher, D. (Eds.). (2021). *The science of learning and development: Enhancing the lives of all young people.* Routledge.

Cantor, P., Osher, D., Berg, J., Steyer, L., & Rose, T. (2019). Malleability, plasticity, and individuality: How children learn and develop in context. *Applied Developmental Science, 23*(4), 307–337. https://doi.org/10.1080/10888691.2017.1398649

Cook, C. E., Bailliard, A., Bent, J. A., Bialosky, J. E., Carlino, E., Colloca, L., Esteves, J. E., Newell, D., Palese, A., Reed, W. R., Vilardaga, J. P., & Rossettini, G. (2023). An international consensus definition for contextual factors: findings from a nominal group technique. *Frontiers in Psychology, 14,* Article 1537242.

Copernicus, N. (1543). *De revolutionibus orbium coelestium* [On the revolutions of the celestial spheres]. Johannes Petreius.

Fischer, K. W. (2007). Dynamic development of action and thought: The role of the constructive web. In W. Damon & R. M. Lerner (Eds.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (6th ed., pp. 313–399). Wiley.

Fleming, A. (1929). On the antibacterial action of cultures of Penicillium, with special reference to their use in the isolation of B. *influenzae. British Journal of Experimental Pathology, 10*(3), 226–236.

Gould, S. J. (1981). *The mismeasure of man.* W.W. Norton.

Immordino-Yang, M. H., Nasir, N. S., Cantor, P., & Yoshikawa, H. (2023). Weaving a colorful cloth: Centering education on humans' emergent developmental potentials. *Review of Research in Education, 47*(1), 1–45. https://doi.org/10.3102/0091732X231223516

Koch, R. (1882). Die Ätiologie der Tuberkulose [The etiology of tuberculosis]. *Berliner Klinische Wochenschrift, 19*(15), 221–230.

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press.

Molenaar, P. C. M. (2004). *A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. Measurement, 2*(4), 201–218.

Moore, D. S. (2015). *The developing genome: An introduction to behavioral epigenetics.* Oxford University Press.

National Scientific Council on the Developing Child. (2010). *Early experiences can alter gene expression and affect long-term development* (Working Paper No. 10). Center on the Developing Child, Harvard University.

NBC Sports (2024, July 31). *Katie Ledecky Olympic 1500m freestyle post-race interview* [Television broadcast]. NBC Sports.

Parrish, R. G. (2010). *Measuring population health outcomes. Preventing Chronic Disease, 7*(4), A71. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2901569/

Pasteur, L. (1878, April 29). *La théorie des germes et ses applications à la médecine et à la chirurgie* [Speech]. French Academy of Sciences, Paris, France.

Rose, T. (2016). *The end of average: How we succeed in a world that values sameness.* HarperOne.

Van Allen, E. M., et al. (2016). Genomic Approaches to Understanding Response and Resistance to Immunotherapy. *Clinical Cancer Research, 22*(23), 5642–5650.