

The Problem with Proxies

Steven Khan



A proxy is something (for example, a document) or someone (for example, an attorney) that is used as a stand-in or substitute for another. In scientific research or professional practice, when a phenomenon or object can't be measured directly one often resorts to an indirect measure, indicator or proxy until more direct measures can be devised. A classic example is estimating the height of a tall object by accurately measuring the length of the shadow cast by the object, measuring one's own height and shadow length at the same time, assuming that the angle between you and the sun and the object and the sun remains the same in the time it takes to do the measurements, and using similar triangles to calculate an approximation of the height. In this case the system is simple, the mathematical

relationships well established and valid with an acceptably small error (remember the surface of the earth is curved though experienced by us as flat) and the use of shadow length as a proxy for an unavailable direct measure is legitimate. But this is not always the case.

Take the digestive system for example. The human gut is not simply a biochemical or electromechanical system. Our guts contain a variety of living microorganisms whose activities contribute to our well-being and health. These constitute the gut microbiome. Traditionally, fecal samples, or scat, have been used as a proxy for the gut microbiome and to infer health and functioning. This makes sense. Like the shadow cast by the tree, we can get a rough indication of what is happening in the gut by identifying and measuring

The diameter of a cylindrical waterslide is 1.5 m. The water running along the bottom of the slide measures 1 m across, as shown in the diagram below.

5. To the nearest hundredth of a metre, what is the depth of the water, x , at its deepest point?

A. 0.19 m
 B. 0.25 m
 C. 0.40 m
 D. 0.50 m

Alberta Education, Provincial Assessment Sector 21 Mathematics 9

FIGURE 1. Item from Alberta Grade 9 Assessment Highlights Mathematics Provincial Achievement Test 2017–2018 (Alberta Education 2018)

which organisms come out of the gut. The math makes sense, right? In a beautifully designed, executed and reported study (Zmora et al 2018), immunologist Eran Elinav, of the Weizmann Institute of Science in Israel, and colleagues, put this belief and common research practice to the test in studies examining probiotic (products marketed as containing healthy microorganisms claimed to restore gut health in everyone) colonization patterns using a more direct method involving colonoscopic and endoscopic sampling. In brief, they found that the microbiome represented in the fecal samples was not representative of what was in the gut and that the pattern of colonization was “highly individualized” and not predictable based on scat analysis; that is, you couldn’t tell whose gut had responded in what way or even whether they had received placebos based on the microbes in their poo! Their findings show quite strongly that, “solely relying on stool sampling as a proxy of mucosal [gastro-intestinal] composition and function may yield limited conclusions” (p 1400). Or as bluntly reported in an interview with *New Scientist*, “Relying on faecal

samples as an indicator of what goes on inside the gut is inaccurate and wrong” (George 2018).

We cannot tell by looking at a single score how much “value” an individual teacher added for an individual learner or class, nor can we tell by looking at that score what particular pedagogical approaches were used (or not), with what fidelity and to what effect.

The complexity of what goes on in classrooms at any level is probably, at least, of the same order as that occurring in our guts (and likely higher). A standardized common achievement test score is the traditional single proxy often used by education reporters and activists and upon which multiple claims regarding what happened or didn’t happen in classrooms is based. They are probably as valid proxies as the fecal samples that have been traditionally used as indicators of what happens in the gut—behaving less like shadows in a well-defined physical system underpinned by

well-understood and valid mathematical models and more like . . . a gross index of overall diet. As with Elinav's study, we cannot tell by looking at a single score how much "value" an individual teacher added for an individual learner or class, nor can we tell by looking at that score what particular pedagogical approaches were used (or not), with what fidelity and to what effect. Standardized common achievement test scores are poor proxies for these activities. Such popular and populist claims that rely primarily on such scores are likely to be of limited practical value at best, are most likely inaccurate and wrong, and dangerous contributors to truth decay (Kavanagh and Rich 2018) in public discourse at worst. Like probiotic colonization, the effects of teaching

What goes on in classrooms and into learning across the lifespan is complex.

and what "sticks" to be "measured" on the traditional paper-based tests are likely to be highly individualized and dependent on a vast array of factors.

Some of those factors are identified in a pair of recent companion articles (Cantor et al 2019 and Osher et al 2020) by researchers from the Harvard Graduate School of Education and the American Institutes for Research Pamela Cantor, David Osher and colleagues. They identify and synthesize the foundational knowledge across multiple disciplines of relevance to education of how human beings develop in contexts (learn). They zero in on features of relationships and contexts that are drivers and delimiters of human development and articulate the strong convergences among the sciences studying learning. These ideas are reiterated and elaborated on in the most recent (2018) National Academies of Sciences, Engineering and Medicine consensus report, *How People Learn II: Learners, Contexts and Cultures* (HPL II) a follow-up to the hugely influential report *How People Learn: Brain, Mind, Experience, and School: Expanded Edition* published in 2000. Below are the key findings of HPL II:

1. Human development depends on the ongoing, reciprocal relations between individuals' genetics, biology, relationships, and cultural and contextual influences.
2. Each individual's development is a dynamic progression over time.
3. The human relationship is a primary process through which biological and contextual factors mutually reinforce each other.
4. All children are vulnerable. In addition to risks and adversities, micro- and macroecologies provide

assets that foster resilience and accelerate healthy development and learning.

5. Students are active agents in their own learning, with multiple neural, relational, experiential and contextual processes converging to produce their unique developmental range and performance. This holistic, dynamic understanding of learning has important implications for the design of personalized teaching and learning environments that can support the development of the whole child.

These findings are drawn from syntheses across multiple domains of research involving contexts ranging from homes to schooling from early years onward and involve such constructs and phenomena as epigenetics, neural malleability and plasticity; integrated complex skill development, human variability, relationships and attachment; self-regulation; science of learning; dynamics of stress, adversity and resilience; relational patterns, attunement, cognitive flexibility, executive function, working memory, sociocultural context, constructive development, self-organization, dynamic skill development, neural integration, relational pattern making, adverse childhood experiences, poverty, racism, families, communities, schools and peers in supporting and/or undermining the healthy development of children and youth; relationships with parents, siblings, peers, caregivers and teachers; chronic stress, institutionalized racism, stereotype threat and racial identity.

In short, what goes on in classrooms and into learning across the lifespan is complex. Standardized common assessment tests are likely not a good measure of and a poor proxy for this complexity and contribute little to shifting public discourse toward more reasoned, informed and evidence-based discussions about what is happening, not happening, should be happening and might be possible in classrooms.

Closer to home here in Canada, Eizadirad (2019), in *Decolonizing Educational Assessment*, has argued that the EQAO standardized common achievement tests administered in Ontario are "culturally and racially biased as it promotes a Eurocentric curriculum and way of life privileging white students and those from higher socio-economic status while simultaneously lacking relevant connections to the identity and lived experiences of racialized students and families from lower socio-economic backgrounds" (p 205). In Khan (2020) I have argued that standardized common achievement tests are over-represented as a genre among forms of assessment that limit human and multispecies' flourishing as ends for education. They are not just poor proxies; they are demonstrably harmful to an increasingly large number of students.

Eizadirad (2019, 205) goes on to argue for equity in addition to equality framing his recommendations through decolonization. He argues that such an approach must begin by

assessing the needs of students and the local communities; how students socio-emotionally and culturally enter the learning environment and the power dynamics embedded in the community and the learning space. This involves practices such as validating the histories and lived experiences of students as a form of valuable knowledge; recognizing their interests, passions...; and recognizing what resources are available and what new resources can be secured through cooperation and synergic collaborations with external organizations. Above all, a decolonized education focuses on providing support to teachers in numerous ways in order for them to assess in multiple ways and in different contexts the potential and competencies of students in relation to their unique needs ... and personalities.

This work has taken on new significance and urgency during the COVID-19 pandemic with many school jurisdictions locally and internationally temporarily suspending the use of grades and standardized common achievement tests (AAC 2020; Vernell 2020) upon recognizing and prioritizing students' and teachers' well-being and health. The inability to provide a secure and "standardized" testing environment, the irrelevance of such tests during a period where survival, dignity, belonging to community and managing everyday challenges during a period of mass vulnerability provides an opportunity to carefully reconsider and design alternatives for a future that may be radically different.

To translate the recommendations through the digestive system analogy used earlier: We need to ensure that all learners are provided appropriate, nourishing meals; we need to support those who grow, procure, transport and sell the necessary ingredients needed to prepare those meals; we need to support and value those essential workers whose labour brings them to the table; we need to observe the short- and long-term effects of these meals, not on stool composition and probiotic colonization alone but on individual/collective physical and psychical well-being and see if our practices also are increasing the overall well-being of our communities and all the members of the multispecies world (that is, all inspired beings) therein.

Note: A shorter version of this article originally appeared on the Medium platform on November 6, 2018. https://medium.com/@skkhan_87562/the-problem-with-proxies-1bd6f9c569f1 (accessed December 17, 2020).

References

- Alberta Assessment Consortium (AAC). 2020. "The 'New Normal' Is Not Normal So Let's Stop Pretending That Grades Matter." April 21. AAC website. <https://aac.ab.ca/the-new-normal-is-not-normal-so-lets-stop-pretending-that-grades-matter> (accessed December 17, 2020).
- Alberta Education. 2018. *Grade 9 Assessment Highlights Mathematics Alberta Provincial Achievement Testing 2017–2018*. https://education.alberta.ca/media/3795489/04-math9-assess-highlights-2018_20181025.pdf (accessed December 16, 2020).
- Cantor, P, D Osher, J Berg, L Steyer and T Rose. 2019. "Malleability, Plasticity, and Individuality: How Children Learn and Develop in Context." *Applied Developmental Science* 23, no 4: 307–37, DOI: 10.1080/10888691.2017.1398649.
- Eizadirad, A. 2019. *Decolonizing Educational Assessment: Ontario Elementary Students and the EQAO*. New York: Palgrave Macmillan, 205.
- George, A. 2018. "Probiotics Are Mostly Useless and Can Actually Hurt You." *New Scientist* 3195, September 6. www.newscientist.com/article/2178860-probiotics-are-mostly-useless-and-can-actually-hurt-you/ (accessed December 17, 2020).
- Kavanagh, J, and M D Rich. 2018. *Truth Decay: An Initial Exploration of the Diminishing Role of Facts and Analysis in American Public Life*. Santa Monica, Calif: RAND Corporation. <https://doi.org/10.7249/RR2314> (accessed December 17, 2020).
- Khan, S K. 2020. "After the M in STEM: Towards Multispecies' Flourishing." *Canadian Journal of Science, Mathematics and Technology Education* 20: 230–45. <https://doi.org/10.1007/s42330-020-00089-4> (accessed December 17, 2020).
- National Academies of Sciences, Engineering, and Medicine. 2018. *How People Learn II: Learners, Contexts, and Cultures*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/24783> (accessed December 17, 2020).
- Osher, D, P Cantor, J Berg, L Steyer and T Rose. 2020. "Drivers of Human Development: How Relationships and Context Shape Learning and Development." *Applied Developmental Science* 24, no 1: 6–36, DOI: 10.1080/10888691.2017.1398650.
- Vernell, S. 2020. "Assessment: Time to Close Down the Exam Factory." April 3. www.tes.com/news/assessment-time-close-down-exam-factory (accessed December 17, 2020).
- Zmora, N, G Zilberman-Schapira, J Suez, U Mor, M Dori-Bachash, S Bashiardes, E Kotler, M Zur, D Regev-Lehavi, R Ben-Zeev Brik, S Federici, Y Cohen, R Linevsky, D Rothschild, A E Moor, S Ben-Moshe, A Harmelin, S Itzkovitz, N Maharshak, O Shibolet, H Shapiro, M Pevsner-Fischer, I Sharon, Z Halpern, E Segal and E Elinav. 2018. "Personalized Gut Mucosal Colonization Resistance to Empiric Probiotics Is Associated with Unique Host and Microbiome Features." *Cell* 174, no 6: 1388–1405. <https://doi.org/10.1016/j.cell.2018.08.041> (accessed December 16, 2020).

Steven Khan is an assistant professor of mathematics education and computational thinking in the Department of Elementary Education at the University of Alberta.