

College Admission Tests and Social Responsibility

Mladen Koljatic, *Pontificia Universidad Católica de Chile*, Mónica Silva, *Escuela de Administración, Pontificia Universidad Católica de Chile*, and Stephen G. Sireci, *Center for Educational Assessment, University of Massachusetts Amherst*

Abstract: *In this article we address the mounting criticism and rejection of standardized tests used in the selection of students for college or university education. Admission tests are being increasingly demonized in many parts of the world and many colleges and universities are dropping tests for selection purposes, claiming the tests are detrimental to fair selection. The testing industry is at the center of this criticism and is accused of maintaining, and even facilitating, the social ills associated with admissions testing, much like iconic business corporations were accused of supporting unfair labor practices in the 1990s. The response of some business corporations to those criticisms was to embrace corporate social responsibility and increase transparency and accountability in their operations. Unfortunately, such acceptance of responsibility and increased transparency have not emerged in the testing industry. We believe the legitimacy of admission tests will continue to be challenged until the testing industry adopts a new way of conducting their business to regain the goodwill of relevant stakeholders in society that so far have been largely ignored.*

Keywords: college admissions, high-stakes tests, social responsibility, standards, testing industry

Admission testing practice is being challenged in the educational arena and in court. Litigation regarding test use is escalating worldwide. In the United States over 1,600 colleges and universities have dropped the requirement of test scores for their admission processes (Fairtest, 2020), and in California, the most populated state in the United States, public universities have been prohibited from using the ACT or SAT admissions tests to admit students or to award scholarships, due to concerns regarding adverse impact against black, Latinx, and other minority students. This decision by the University of California system to eliminate ACT and SAT score requirements for admission or for scholarship consideration was made even though a blue-ribbon University Committee advised against abandoning use of these test scores in the admissions process (Systemwide Academic Senate, University of California, 2020). In Chile, riots in 2020 against the national university admission test (PSU) were successful in disrupting the annual selection process (Nugent, 2020; Sireci, 2020). The angry mob that stormed the building in which test booklets were stored, stole them, and burned them in the streets, did not come out of the blue. The authorities and experts in charge of the test knew about the test's serious equity shortcomings through two international reports—by Educational Testing Service (2005) and Pearson Education (2013)—and failed to act promptly to correct it (Koljatic & Silva, 2020).

To some, well-constructed tests provide valuable information about individuals and groups and merit to be considered “among the most important contributions of cognitive and behavioral sciences to our society” (American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in

Education [NCME], 2014, p. 1). To others, standardized tests should be banned for contributing to the perpetuation of society's enduring inequalities (Marmol, 2016). Why are college admissions tests and other standardized tests becoming increasingly seen as a major cause of these inequalities? As we describe in this article, it is because the testing industry has turned a blind eye to the negative consequences of educational testing for far too long.

The Standardized Testing Controversy

The controversy over the value of standardized tests, particularly in high-stakes contexts, has been present for decades (Cai, 2020), with two distinct camps: staunch rejecters (e.g., Au, 2010, 2018) and equally determined supporters (e.g., Phelps, 2005). One of the most salient arguments against the use of tests refers to its use in groups with restricted access to educational opportunities. Those opposed to testing claim tests favor society's well-positioned elites under the guise of “merit” (e.g., Marmol, 2016; Sacks, 1997), and are detrimental to historically marginalized groups of students who remain underrepresented in higher education and in white-collar jobs. On the other hand, defenders of tests claim achievement gaps across groups result from educational and societal inequities, not from bias in the test itself (The Economist, 2020), and that any well-constructed test will inevitably reflect educational inequality associated with background and prior experience (Cai, 2020).

Recent reactions from college admissions testing companies in the United States to the University of California's decision to eliminate ACT and SAT score requirements illustrate industry-level arguments in support of college

admissions tests in the face of adverse impact against underrepresented minorities. For example, a spokesperson from the College Board stated, “the notion that the SAT is discriminatory is false... any objective measure of student achievement will shine a light on inequalities in our education system” (Hoover, 2019), and ACT’s CEO tweeted “It’s inappropriate to blame admissions testing for inequities in society. We don’t fire the doctor or throw away the thermometer when an illness has been diagnosed. Test scores... expose issues that need to be fixed [in our educational system]” (Roorda, 2019). The Statement issued by NCME, *Misconceptions about Group Differences in Average Test Scores* (NCME, 2019) similarly rejected the criticism that tests are the cause of group differences in achievement.

We understand the argument that college admissions tests are not the *cause* of adverse impact on the tests; however, we believe the lack of attention to issues of educational inequity by the testing industry has perpetuated this inequitable system, and so makes the testing industry culpable for this problem. Essentially, the ultimate message from the testing industry and test advocates has been tests are fine and it is the responsibility of politicians and policy makers to level the playing field in the K–12 education system. It appears the testing industry has concluded that remotely equitable academic outcomes among racially and socioeconomically diverse students are not possible (Finn, 2020), and thus evade responsibility for the consequences of the effects of college admissions tests on individuals, universities, and society.

From Corporate Deniability to Corporate Responsibility

The denial of responsibility of negative testing consequences by the testing industry is similar to the denial of responsibility of adverse social consequences accrued from the operation of numerous business corporations in the 1990s. At the time, business boards and managers of large corporations were accused of turning a blind eye to negative societal consequences of their business such as environmental damage, and violations of human rights. To continue operating, many well-established companies had to change their way of conducting business. The footwear and sports clothes industry, particularly Nike, Inc., has been extensively studied by business analysts, and some lessons from this research can be applied to the testing industry. These lessons include the need to work on transparency and accountability (i.e., acknowledge and take responsibility for negative externalities of doing business), listen to stakeholders, and take a proactive stance on social responsibility. In this context, social responsibility is understood as the notion that corporations should behave ethically and aim to contribute deliberately to the welfare of society. In this view, society comprises all the relevant communities and stakeholders with whom they operate and interact.

Social responsibility applies to all organizational entities—whether a government, an industry, a corporation, an institution, or an organization dealing with society at large when conducting core commercial activities (Planken, 2013). The reaction of Nike, Inc. to criticisms regarding the negative externalities of their business operation provides helpful context for understanding how the testing industry needs to change to better serve societal interests.

Nike Inc.: An Example of a Corporation’s Adaptation to Social Responsibility

In the early 1990s Nike was a corporate success story. Selling a combination of basic footwear and sports clothes, it rapidly gained consumer support and increased its revenues. Like other companies in the industry, Nike saved costs by following an active strategy of outsourcing the manufacturing of athletic footwear and much of its apparel to independent contractors located in low-cost Asian countries. However, in the late 1990s, the company’s image was tarnished over allegations regarding the treatment of workers in the contractors’ factories. Labor activists and NGOs accused Nike of abuse of workers—particularly women—poor working conditions, low wages, and use of child labor through the media (Spar & La Mure, 2003).

Despite mounting criticism, Nike insisted labor conditions in its contractors’ factories were not—could not—be Nike’s concern or its responsibility. As a spokeswoman put it, “we’re about sports, not manufacturing 101,” and Nike’s general manager in Jakarta stated, “I don’t know that I need to know” (Spar & La Mure, 2003, p. 90). Nike’s company line on the issue was clear and stubborn: the company simply could not be held responsible for the actions of independent contractors. Legally speaking they were right, but the public considered there was a moral responsibility in the negative externalities of their business operations. Public reaction against Nike ignited in the form of protests and boycotts. Activists and students at universities that had apparel and footwear contracts with Nike demanded all contracts cease until labor practices were rectified.

University administrators heeded the student protests and ceased contracts. “Before long, student protests spread to campuses where Nike had no merchandising contracts... Previously apathetic college students stormed university buildings to protest sweatshop labor and the exploitation of foreign workers... Activists took over buildings at Duke, Georgetown, the University of Michigan and the University of Wisconsin, and staged sit-ins at countless other colleges and universities ” (Spar & Burns, 2002, p. 11).

At the time, Nike had begun expanding its chain of giant retail stores, only to find that each newly opened site came with an instant protest rally, complete with shouting protesters, sign waving picketers, police barricades, and attention from the press. Under public pressure, the company decided to hire Andrew Young, the respected civil rights leader and former mayor of Atlanta, to conduct an independent evaluation. Nike’s CEO granted Young “blanket authority... to go anywhere, see anything, and talk with anybody in the Nike family” (Spar & La Mure, 2003, p. 91).

When the report was released, rather than calming Nike’s critics, Young’s report had the opposite effect. Critics were outraged by the report’s research methodology and conclusions. In addition, they argued Young had failed to address the issue of factory wages, which was for many observers the crux of the issue, and had spent only 10 days interviewing workers. During these interviews, moreover, Young relied on translators provided by Nike, a major lapse in accepted human rights research techniques (Spar & Burns, 2002).

By 1998 Nike faced weak demand for its products and had to lay off workers. Its CEO began to realize the need to shift gears in the way of doing business (Nisen, 2013). That same year Nike announced a series of changes affecting its contract workforce in Asia, including an increase in the minimum

age, a tightening of air quality standards, a pledge to allow independent inspections of factories and the allocation of 25 employees to focus solely on corporate social responsibility. Nike also agreed to participate in an initiative—a multistakeholder task force convened by the President of the United States at the time—to address the apparel industry problems (Spar & Burns, 2002).

Nike learned the lesson, adopted an active stance in corporate social responsibility and worked hard on restoring its public image. For the sake of transparency and accountability, at present all Nike suppliers are identified and listed in its website, so independent observers can assess its labor and human rights record.

What the Testing Industry Should Learn from Nike

The parallels between the public outcry against Nike in the 1990s and against the testing industry today are striking. The response of the testing industry to calls for more attention to the social consequences of testing has been largely academic, with little evidence of research or changes in testing practices (Poggio, Ramler, & Lyons, 2018). Now that public criticism is having a financial impact on the testing industry, it appears the industry must quickly pivot to address social concerns, or it will continue to falter in both public support and financial resources. Thus, it is important to ask the question, “Are there lessons the testing industry can learn from Nike?”

Using Nike as an example, Zadek (2004) described the adoption of corporate social responsibility in terms of five stages of growing involvement in responsible business practices. His model is germane to organizations, corporations, or whole industries and can be applied to the testing industry as well. The five stages (and corresponding mottos) are: (1) defensive (“not my problem”), (2) compliance (“we’ll do just as much as we have to”), (3) managerial (“it’s the business, stupid”), (4) strategic (“it gives us a competitive edge”), and (5) civil (“we need to be sure everybody does it”) (Zadek, p. 126). He states the learning pathways of organizations are not always linear as they move along the learning curve with progress and setbacks, yet they invariably go through the five stages.

Judging from some official statements by members of prestigious testing organizations, the testing industry seems to be at stage 1, where Nike was in the 1990s, dwelling in the defensive stage and claiming “it’s not our job to fix that.” Thus, when confronted with the disparate impact of standardized tests across demographic subgroups, testing industry officials argue the tests are fair and the problem lies in the unequal educational system, just as Nike originally rationalized unfair labor practices as the problem of the countries in which its manufacturing contractors were based. This “hand washing” strategy may work for COVID-19, but it is a poor strategy for gaining public support and may explain why so many colleges and universities have become “test optional,” and the University of California system has banned the use of the SAT and ACT.

At stage 2, “compliance,” a corporate policy must be established and observed, usually in ways that can be visible to critics to protect the company’s reputation and reduce the risk of litigation. Codes of conduct are typical of stage 2. The AERA et al. (2014) *Standards for Educational and Psychological Testing* can be thought of as a professional code of conduct, although they are not legally enforced, and most testing programs are not even audited with respect to their adherence to

them. Thus, adherence to the *Standards* is recommended or prescriptive, not required. In a sense, they can be perceived as a venue to protect the testing industry’s reputation and reduce the risk of litigation. Still, a conflict is likely to arise when any company or organization merely offers compliance to minimum standards, while the public clearly demands a far greater commitment.

At stage 3, “managerial,” the organization realizes it is facing a long-term problem, one unlikely to be swatted away with attempts at compliance or a public relations strategy. The company will have to assume responsibility for the problem and invest heavily, if necessary, in its solution. For example, rather than simply claim they follow established statistical procedures to screen items for differential item functioning (DIF), testing companies should commit to invest in exploring new ways of testing to effectively reduce performance gaps. Testing companies should know their future depends upon their ability to develop tests that are seen as more fair and result in little to no adverse impact.

At stage 4, “strategic,” the corporation begins to realize it needs to realign its way of conducting business, to develop a consciousness and level of responsibility about how their products affect their customer base. Moving into this stage requires integrating the stakeholder perspective in the business model. To enhance economic and social value in the long term, testing companies need to go one step beyond what they are doing today and integrate educational improvement and gap reduction as part of their business model. Doing so will require rethinking their strategy and innovating their processes to focus on integrating the various stakeholders who have been largely ignored, such as teachers, parents, and students. Successful realignment of their strategy to address socially responsible practices will eventually become a necessity for any testing company to operate.

In the final stage, “civil,” the whole industry needs to promote collective action to address global concerns, adopting the motto “we need to make sure everybody does it” (Zadek, p. 127). In a global world, what happens in other nations affect others as well. Dissatisfaction with admission tests is rampant in different parts of the world and it needs to be addressed. Student riots against test use, (e.g., massive protests, strikes and demands for a moratorium on test use, the “opt-out” movement, etc.) may replicate in different parts of the world, whether promoted by activists or through the copycat effect. The testing industry could work together with civil society organizations and professional associations—such as AERA, APA, NCME, and the International Test Commission—to establish codified business practices that promote social responsibility in testing worldwide.

Discussion

The idea of a greater commitment of the testing industry to the improvement of social ills is not new. Gordon (2020), among others, has argued “educational assessments can and should be used to inform and improve teaching and learning processes and outcomes” (p. 72). However, the testing industry has become complacent and entrenched in its own culture without attending to legitimate criticism from relevant stakeholders. By doing so, it has failed to address pressing and pervasive problems in assessments (Sireci, 2020). This omission is particularly serious in the case of equity issues of high-stakes tests.

Even accepting that any educational test will expose differences in the quality of education received by test takers, the question remains as to whether testing agencies and psychometricians are doing their best to level the field in selection and promotion of fair testing practices. Statistical analyses such as DIF and differential predictive validity are clearly not enough to address social justice concerns—testing agencies need to do more (Helms, 2006). Specifically, experts at testing agencies should be concerned with assessing if the test represents students' performance and potential accurately, and in this context opportunity to learn is crucial (Reshetar & Pitts, 2020).

Taking Steps toward Social Responsibility in Testing

With respect to college admissions testing, there are immediate steps testing agencies could take to promote social justice in testing. First, testing experts should conduct studies to determine the knowledge and skills students need to master to succeed in college and ensure these knowledge and skill areas are built into the test. Loading the test with excessive and unnecessary content—as is done, for example, when admission tests are also expected to serve the purpose of assessing curricular coverage—can result in construct irrelevance with respect to the primary admissions purpose (Lorié, 2020). In trying to serve multiple goals, validity may be compromised, and tests may end up serving none of the multiple goals satisfactorily. Particularly in nations where the quality of the education received depends heavily on income, control of curricular coverage should be kept separate from admission testing to avoid favoring wealthy elites (Heyneman, 1987; Lewis & Dunder, 2002).

The psychometric community and testing agencies should take direct responsibility in this regard by collecting evidence that test-takers are provided with the curriculum and instruction that incorporate the constructs addressed by their tests, even if the costs entailed in test construction increase and their revenues decrease. Although opportunity to learn has been legally mandated for high-stakes tests in the United States (Sireci & Parker, 2006), too often opportunity to learn is assumed for all students, rather than studied as a potential source of test bias for some students. One suggestion we offer is to measure students' exposure to the curriculum (an opportunity to learn index, in a sense, for each student) that can be taken into account when interpreting a student's test performance.

Greater Transparency and Accountability

The public has lost confidence in educational assessments, and quite rightly so (Sireci, 2021). To regain it the testing industry urgently needs to work on transparency and accountability. For example, in the United States, the most recent version of the SAT—focused on the assessment of the common core curriculum—was hastily developed and released according to Manuel Alfaro, a high official of the College Board. He charged the College Board had made false claims about its tests when bidding for public contracts and misled the states about the process used to create questions for the new SAT, “resulting in an inferior exam” (Dudley & Schiffman, 2016). Whistleblowing is not inevitable if organizations provide adequate channels for detecting and correcting problems instead of attempting to ignore them, or worse, suppress the information (Davis, 1989). The testing industry needs to learn that transparency and accountability are the best pre-

vention for the long-term damage associated to whistleblowing.

For too long the psychometric community has been committing what Popham (2003) labeled “psychometric sins,” one of which is ignoring problems and looking the other way. “Most of the measurement-caused problems now seen in American schools arose because members of the measurement community were mute during the years the evaluative misuses of traditional tests became increasingly common” (p. 46). Pretending to be mute becomes a “mortal sin” (p. 45) when dealing with high-stakes tests that have important individual consequences for test-takers, such as admission tests. Turning a blind eye, remaining silent or refusing to assume responsibility for test problems and the social consequences of testing may have been tolerable in the past, but besides being unethical, it is also a risky stance for the future of testing.

Conclusions

The testing industry is currently being challenged, much the same way as business companies became questioned decades ago for the way they conducted their business. Large corporations were accused of turning a blind eye to the societal consequences of their business (and they were!). In the corporate world, managers had to change their way of conducting business to earn a social “license to operate,” particularly in certain industries where the local community distrusted their practices and criticized the externalities they had to pay for their operation (Demuijnck & FASTERLING, 2016; MORRISON, 2014). Businesses today are held accountable to a variety of stakeholders. Even law-abiding businesses are being prodded by the public and by civil society to take a step further in promoting common good. Avoiding negative externalities is a bare minimum. Corporate citizenship entails not only becoming evil-avoiders, but good-seekers.

Today, it is the testing industry's turn to be criticized for their negative externalities; in particular, restricting underrepresented minorities from attaining their educational and occupational goals. It is also the testing industry's turn to become more socially responsible. The AERA et al. (2014) *Standards* represents a starting point for the development of guidelines for socially responsible educational and psychological assessment. Extending these *Standards* to promote social responsibility in testing will involve multiple stakeholders working together to ensure the negative externalities of testing are minimized, and that an enforcement mechanism is in place to hold all testing agencies accountable to that principle.

Since the 1950s, the *Standards for Educational and Psychological Testing* have addressed minimum requirements for proper development and use of tests through guidance for practitioners. The higher the stakes of tests—and admission tests rank among the highest-stakes tests—the more attention should be devoted to guaranteeing they serve their purpose and are as free of bias as possible. The question is “can psychometricians in testing agencies do more?” We believe the answer is “yes.”

The need to promote transparency and accountability in the testing industry is crucially important, particularly since in many parts of the world, the testing industry is unregulated. Countries differ in the degree, if any, of statutory control they can exercise over the use of testing and its consequences for those tested (International Test Commission, 2013). In unregulated industries ethics may

in fact be the sole option to correct undesirable practices that occur in the absence of regulation (Norman, 2012). The AERA et al. (2014) *Standards for Educational and Psychological Testing* operate as a self-regulatory mechanism for testing professionals and are often invoked in courts to substantiate cases related to testing practices (Buckendahl & Hunt, 2005; Sireci & Parker, 2006). However, the *Standards* represent minimum requirements to be met, have no formal enforcement mechanism, and often psychometricians simply ignore them (Sireci, 2020).

To bring back legitimacy into test use, testing experts need to change their mindset about test development and adopt a proactive stance to social responsibility that entails going beyond meeting the AERA et al. (2014) *Standards*. In the same way businesses had to adapt to societal demands and enlarge their focus from profit maximization to value creation—economic and social—testing agencies need to adapt to a changing and demanding environment. Going about “business as usual” will likely reduce test use, or worse, will result in a moratorium of test use following test boycotts. Just as the corporate world was compelled to address social ills as part of their business model to be granted a “license to operate” in society, testing agencies will have to be more responsive to social inequalities in testing by embracing a proactive stance in social responsibility, if they wish to continue to operate. The change toward greater social responsibility in testing ultimately requires the adoption of a new perspective as to the role of testing agencies in society, and time is running short. Fortunately, it is not too late, and we look forward to witnessing the testing industry embrace the ideology of socially responsible assessment.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Lanham, MD: American Educational Research Association.

Au, W. (2010). The idiocy of policy: The anti-democratic curriculum of high-stakes testing. *Critical Education*, 1(1), 1–16. <http://m1.cust.educ.ubc.ca/journal/v1n1>

Au, W. (2018). The socialist case against the SAT. *Jacobin* [magazine]. Retrieved from <https://www.jacobinmag.com/2018/04/against-the-sat-testing-meritocracy-race-class>

Buckendahl, C., & Hunt, R. (2005). Whose rules? The relation between the “rules” and the “laws” of testing. In R. Phelps (Ed.), *Defending standardized testing* (pp. 147–158). New York: Lawrence Erlbaum. <https://doi.org/10.4324/9781410612595>

Cai, L. (2020). Standardized testing in college admissions: Observations and reflections. *Educational Measurement: Issues and Practice*, 39(3), 34–36. <https://doi.org/10.1111/emip.12389>

Davis, M. (1989). Avoiding the tragedy of whistleblowing. *Business and Professional Ethics Journal*, 8(4), 3–19. <https://doi.org/10.5840/bpej19898419>

Demuijnck, G., & Fasterling, B. (2016). The social license to operate. *Journal of Business Ethics*, 136(4), 675–685. <https://doi.org/10.1007/s10551-015-2976-7>

Dudley, R., & Schiffman, J. (2016). *Exclusive: FBI raids home of ex-College Board official in probe of SAT leak*. Reuters. Retrieved from <https://www.reuters.com/article/us-college-sat-fbi-idUSKCN112009>

The Economist. (2020). Are test scores the backbone of meritocracy or the nexus of privilege? Retrieved from <https://www.economist.com/united-states/2020/02/08/are-test-scores-the-backbone-of-meritocracy-or-the-nexus-of-privilege>

Educational Testing Service. (2005). Evaluación externa de las pruebas de selección universitaria (PSU). Retrieved from <https://ciperchile.cl/wp-content/uploads/informe-ets.pdf>

Fairtest. (2020). College Board sued by students, fair-testing advocates for access and technology failures on computerized AP exams. Retrieved from <https://www.fairtest.org/college-board-sued-students-fairtesting-advocates>

Finn, C. (2020). Eliminating the ACT and SAT won't fix what's wrong with education. *The Dispatch*. Retrieved from <https://thedispatch.com/p/eliminating-the-act-and-sat-wont>

Gordon, E. W. (2020). Toward assessment in the service of learning. *Educational Measurement: Issues and Practice*, 39(3), 72–78. <https://doi.org/10.1111/emip.12370>

Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, 61(8), 845–859. <https://doi.org/10.1037/0003-066X.61.8.845>

Heyneman, S. P. (1987). Uses of examinations in developing countries: Selection, research, and education sector management. *International Journal of Educational Development*, 7(4), 251–263. [https://doi.org/10.1016/0738-0593\(87\)90023-X](https://doi.org/10.1016/0738-0593(87)90023-X)

Hoover, E. (2019). University of California faces bias lawsuit over act/sat requirement. *The Chronicle of Higher Education*. Retrieved from <https://www.chronicle.com/article/u-of-california-faces-bias-lawsuit-over-act-sat-requirement/>

International Test Commission. (2013). ITC guidelines on test use. Retrieved from https://www.intestcom.org/files/guideline_test_use.pdf

Koljatic, M., & Silva, M. (2020). Chile's admission tests: Pending changes and revisions. In M.E. Oliveri & C. Wendler (Eds.), *Higher education admission practice: An international perspective*. Cambridge: Cambridge University Press (pp. 145–161). <https://doi.org/10.1017/9781108559607>

Lewis, D.R., & Dundar, H. (2002). Equity effects of higher education in developing countries: Access, choice, and persistence. In C. W. Chapman & A. E. Austin (Eds.), *Higher education in the developing world: Changing context and institutional responses* (pp. 169–196). Westport, CT: Greenwood Press.

Lorié, W. (2020). *Addressing the alignment challenge associated with the use of college admissions tests under ESSA* [Policy Brief]. Doer, NH: National Center for the Improvement of Educational Assessment.

Marmol, E. (2016). The undemocratic effects and underlying racism of standardized testing in the United States. *Critical Intersection in Education*, 4(Winter), 1–9. <https://jps.library.utoronto.ca/index.php/cie/article/view/26430/20714>

Morrison, J. (2014). *The social license to operate: How to keep your organization legitimate*. London: Palgrave MacMillan. <https://link.springer.com/book/10.1057%2F9781137370723>

National Council on Measurement in Education. (2019). Misconceptions about group differences in average test scores. Retrieved from <https://www.ncme.org/publications/statements/new-item2>

Nisen, M. (2013). *How Nike solved its sweatshop problem*. Business Insider. Retrieved from <https://www.businessinsider.com/how-nike-solved-its-sweatshop-problem-2013-5>

Norman, W. (2012). Business ethics as self-regulation: Why principles that ground regulations should be used to ground beyond-compliance norms as well. *Journal of Business Ethics*, 102(1), 43–57. <https://doi.org/10.1007/s10551-011-1193-2>

Nugent, C. (2020). Why Chile's SATs have become the new frontline of inequality protests. Retrieved from <https://time.com/5770308/chile-student-protests/>

Pearson Education. (2013). Evaluation of the Chile PSU—final report. Retrieved from https://www.educacion2020.cl/sites/default/files/2013013110575540.chile_psu-finalreport.pdf

Phelps, R. P. (2005). *Defending Standardized Testing*. Mahwah, NJ: Erlbaum.

Planken, B. (2013). Definitions of corporate social responsibility. In S. Idowu, N. Capaldi, L. Zu, & A. Das Gupta (Eds.), *Encyclopedia of social responsibility* (pp. 768–772). Springer, Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-28036-8_476

Poggio, J., Ramler, P., & Lyons, S. (2018). *Consequential validation: Where are we after 25 years of effort?* [Presentation]. 2018 Annual Meeting, National Council on Measurement in Education, New York.

- Popham, W. J. (2003). Seeking redemption for our psychometric sins. *Educational Measurement: Issues and Practice*, 22(1), 45–48. <https://doi.org/10.1111/j.1745-3992.2003.tb00117.x>
- Reshetar, R., & Pitts, M. (2020). General Academic and subject-based examinations used in undergraduate higher education admissions. In M.E. Oliveri & C. Wendler (Eds.), *Higher education admission practice: An international perspective* (pp. 237–255). Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108559607>
- Roorda, M. [@martenroorda]. (2019). It's inappropriate to blame admissions testing for inequities in society. We don't fire the doctor or throw away the thermometer Tweeter. Retrieved from <https://twitter.com/MartenRoorda/status/1204465574111105024>
- Sacks, P. (1997). Standardized testing: Meritocracy's crooked yardstick. *Change*, 29(2), 25–31. <https://doi.org/10.1080/00091389709603101>
- Sireci, S. G. (2020). Psychometricians in the hands of an angry mob. NCME 2020 Presidential Address [Video]. <https://www.youtube.com/watch?v=5X2uYYax9yw>
- Sireci, S. G. (2021). Valuing educational measurement. *Educational Measurement: Issues and Practice*, 40(1), 7–16. <https://doi.org/10.1111/emip.12415>
- Sireci, S. G., & Parker, P. (2006). Validity on trial: Psychometric and legal conceptualizations of validity. *Educational Measurement: Issues and Practice*, 25(3), 27–34. <https://doi.org/10.1111/j.1745-3992.2006.00065.x>
- Spar, D. L., & Burns, J. (2002). Hitting the wall: Nike and international labor practices. *Harvard Business School Case*, 9-700-047. <https://www.hbs.edu/faculty/Pages/item.aspx?num=26875>
- Spar, D. L., & La Mure, L. T. (2003). The power of activism: Assessing the impact of NGOs on global business. *California Management Review*, 45(3), 78–101. <https://doi.org/10.2307/41166177>
- Systemwide Academic Senate, University of California (2020). Report of the UC Academic Council's Standardized Testing Task Force (STTF). Retrieved from https://senate.universityofcalifornia.edu/_files/committees/sttf/sttf-report.pdf
- Zadek, S. (2004). The path to corporate responsibility. *Harvard Business Review*, 82, 125–132. <https://hbr.org/2004/12/the-path-to-corporate-responsibility>