Making the Most of Teacher Team Time to Support Equitable Assessment

Daniel P. Moore[a] & William R. Penuel [b]

Stanford University, Stanford, CA[a]; University of Colorado Boulder, Boulder, CO[b]

Summary

Assessment can help promote equity for racially minoritized students in our schools. The path to equity, though, does not lie with the common practice of using teacher team time to pore over test score data. Instead, it depends on teacher teams developing high-quality materials and preparing for relevant, meaningful instruction.

We recommend four evidence-informed strategies for making the most of teacher team time:

1. Choose and adapt more meaningful performance-based assessments.
2. Follow protocols that focus attention on evidence-informed practices for equity.
3. Use instructional materials that support culturally sustaining pedagogies.
4. Establish grading practices that focus on student strengths and feedback that supports growth.

Taken together, these recommendations will allow school leaders and principals to use assessment to promote equity.

Introduction

Many school leaders across the country have sought to use assessment as a driver for promoting more equitable outcomes for racially minoritized students in schools. Key tools include end-of-year tests and benchmark assessments, with the logic that by spotlighting inequities through test scores, those inequities can then be addressed. However, too often these scores are viewed as a valuable end in themselves, leading to practices like excessive test prep or school team protocols for poring over test data, rather than analyzing how to better educate our children.[1] Educators in such teams frequently attribute low scores to the motivations or abilities of students— often biased by factors like race, disability status, or home language— rather than to factors that educators can influence like teaching, student-teacher relationships, and school culture.[2] In addition, teams rarely turn to evidence for how to achieve better outcomes for racially minoritized students.

For assessment to promote equity, we need not only different kinds of assessment and grading practices, we need additional materials and new ways of using data and evidence to support changes to teaching and learning. That means choosing and adapting assessments that are consequential to students themselves; following protocols that draw attention to how practice influences the experiences of students; making use of culturally sustaining pedagogies; and establishing grading practices that focus on student strengths and feedback that supports growth.

The Problems with Data Teams

One of the most common practices schools do to improve equity through assessment is to create data teams to look at students' end-of-year test scores and identify gaps in performance between different racial groups. For example, a school with mostly Black students might compare its state exam scores with the average scores for white students in the state.[3] It sounds logical: if you want to promote equity, you need to close achievement gaps, and to do that, you need to see the data. This logic leads schools to devote weeks of instructional time to preparing students for spring testing season, and special notices are sent home to parents to ensure their students get enough rest and come to school well-fed on testing day.

Schools and districts invest millions each year in interim and benchmark exams that resemble standardized tests to gauge progress toward standards [4] and now, after the pandemic, to measure "learning loss." Special data teams, as well as teacher work groups or professional learning communities (PLCs), spend many meetings each year poring over performance on state tests and common assessments, identifying students as in need of remediation. A common protocol involves sorting students into high, middle, and low categories and identifying "bubble" students— those closest to moving from the middle to high group— to target for intervention.[5]

The trouble is that these investments of money and educators' time don't work. Talk of achievement gaps frequently leads to a form of "gap gazing," where explanations for racially minoritized students' lower test scores are attributed to deficits in the children themselves.[6] Data team meetings do sometimes have teachers leave with an action plan to adjust instruction. However, there is often a missing step of looking at what research says about how to better serve those students to whom we owe an education debt in our society, such as the Black, Latinx, and Native American students in our schools.

The assessments that are used don't adequately represent the demands of today's standards, either. Standardized multiple-choice tests do not provide opportunities for students to demonstrate their abilities to reason with evidence, model phenomena with mathematics and science ideas, or engage in complex literary interpretation. Additionally, data rarely provide clues about what to do next, leaving educators to double down on their existing strategies— which may have contributed to the gaps they identified in the first place. A range of well-designed studies, in fact, suggest data driven decision making of this kind doesn't work overall, and it doesn't promote equity.[7]

The solution to test pressure, it turns out, isn't more testing or obsessing over test score data in team meetings. There is no substitute for spending team time developing high-quality materials and preparing for instruction that connects to racially minoritized students' everyday lives. Teachers in school teams should focus on assessment that helps students see meaning in what they are learning in school and how to make sense of it within the broader sociopolitical contexts that impact them.

What We Can Do: Four Evidence-Informed Strategies

School leaders can use teacher teams effectively to better support equitable assessment and learning. There is evidence that this is possible,[8] but the way teams use their time needs to be informed by the research base. We recommend four strategies for schools to try: (1) using more meaningful performance-based assessments; (2) following protocols that focus attention on evidence-informed practices for promoting equity; (3) using instructional materials that support culturally sustaining pedagogies; and (4) establishing grading practices that focus on student strengths and feedback that supports growth.

Use Meaningful Performance-Based Assessments

Most contemporary standards demand that students gain a grasp of core disciplinary practices, such as explanation and argumentation.[9] But traditional standardized tests that rely on multiple-choice items give few opportunities for students to engage in those practices in a meaningful way. Such assessments effectively reduce and narrow a discipline[10] and they are hardly ever collaborative, even though most of these disciplinary practices require students to be able to think with others.

The kinds of assessments we need instead are called performance-based assessments. These assessments empower students to show what they know by creating a product for a real audience.[11] For students from racially minoritized communities, it is particularly valuable for promoting equity when performance-based assessments create a context for building relationships with families and communities,[12] and when they are grounded in a theory of learning that honors the cultural repertoires that students bring to the classroom with them.[13]

The assessments themselves can present contexts that make visible the sociopolitical realities students face in their lives. Jennifer Randall, a leading scholar in anti-racist assessment, offers the example of a math assessment focused on students' ability to model a number line by determining how many red beans are needed to feed Black Lives Matter protesters on their way to a march focused on immigrant rights.[14] A science assessment could ask students to analyze data to determine the impact of a border wall on the survival of jaguars whose habitats stretch across the US-Mexico border and propose solutions that consider the rights of both human beings and other species, as well.[15] These smaller, individual assessment tasks fit within a larger system of curriculum, teaching, and assessment in which students study sociopolitical topics in a deeper and ongoing way.

Rather than looking at rows and rows of scores on a spreadsheet, teacher teams should make use of their shared time by looking at student work samples. Looking at actual student work opens up conversations that go beyond "who got it" and "who didn't" and toward what teaching practices will support learning.[16] By looking at real student materials, teachers can look for strengths in student thinking to build from and focus on the specific qualities evident in students' work.[17]

There are three key considerations for making the most of looking at student work together. First, it requires a deep engagement with the content and how students might approach it. It is critical to bring the perspective that every student makes sense to themselves, and to seek to understand what a student is communicating. Second, teachers need to disrupt tendencies to evaluate student work against a standard of whiteness. They can do this by making explicit their criteria for student work, co-constructing rubrics with students, and interrogating the beliefs that underlie their criteria.[18] Third, teacher teams must center protocols that help focus attention on practice and evidence, rather than on student deficits. We describe these protocols below.

Follow Protocols that Focus Attention on Evidence-Informed Practices

Structured protocols can help teachers avoid the common pitfalls of data teams. For example, effective protocols for looking at student work invite teachers to look at samples from the same assignment and to interrogate both the conditions under which the work took place and their own criteria for evaluating the work.[19] Focusing on conditions and criteria is important because it can surface tacit differences in teachers' expectations of students, as well as tacit differences in teaching practices.[20] In such discussions, it is valuable to be honest and critical of practice, something that requires vulnerability on the part of educators.[21] In addition, teacher teams can integrate the study of research-informed strategies in the context of their own practice.[22] Such study can lead teachers to adopt new strategies rather than further relying on current practices that may be fostering inequitable impacts on student learning.

An example of a protocol focused on practice and equity is the Student Experience Improvement Cycle (SEIC), which was designed to enhance research and evidence use for equity in science education. The SEIC is based on a Plan-Do-Study-Act cycle.[23] In this cycle, teachers first adapt evidence-informed strategies for equitable small group and whole class knowledge building activities. Then, they gather and interpret student experience data to assess whether the strategy helped improve the overall student experience, as well as differences in experience associated with students' race and gender.

The SEIC strategy is novel in three ways. First, it integrates research-based summaries, ideas, strategies, and coaching into an inquiry cycle that also uses experience data. Second, it uses a system for collecting and visualizing student experience data (the Science SEET system, https://www.scienceseet.com/). Third, it incorporates a structured protocol for analyzing data, discussing affective responses to data, and considering how practice may help explain data patterns. Evidence from two recent pilot tests of the SEIC (n = 29, 25) shows its potential for enhancing the degree to which students experience science instruction as relevant to them, particularly Black students.[24]

Use Instructional Materials that Support Culturally Sustaining Pedagogies

Another pathway to more equitable assessment practice is through instructional materials that support culturally sustaining and justice-focused pedagogies. These pedagogies aim to center in the curriculum the literacies, practices, and struggles for justice of communities marginalized by systemic inequality.[25] They can be supported by open instructional materials that leave room for adaptation and improvisation, so as to connect disciplinary learning to students' interests and the concerns of their communities.[26] Importantly, such materials are a powerful route to facilitating better assessment practice, because materials can embed guidance about feedback as well as strategies for what to do next to address learning needs. Such materials need to pursue ambitious goals for learning and hold up a mirror to students, so they can see themselves in assessment opportunities that also open up new worlds.[27]

There are several ways that culturally sustaining instructional materials can become productive focal points for teacher teams. Building such materials with colleagues can facilitate a shared vision for how teaching and learning support justice projects of racially minoritized communities.[28] Studying actual instructional materials can help improve student learning outcomes.[29] Adapting and making use of culturally responsible assessments can support students in building more equitable classroom cultures and in meeting more ambitious learning goals.[30] Teacher teams can also co-construct liberatory assessments that center students' identities, cultural knowledge, and efforts to heal from systemic oppression, or that directly confront inequality, systemic racism, and white supremacy.[31] Examples include anti-racist writing tasks and projects that assess students' understanding of the knowledge and skills needed to be an effective activist.[32] Such assessments have restorative and transformative potential.

One example of a school district that has created a small-scale system of assessment that supports culture-based pedagogies is Denver Public Schools. With their partners in the inquiryHub research-practice partnership, they have built a multicomponent system that includes curriculum materials, embedded assessments, "transfer" assessments, and sustained professional learning opportunities.[33] This system aims to minimize the harms of high-stakes testing by centering curriculum-based assessment grounded in cultural perspectives, and through extensive, ongoing professional development.

Use Grading that Focuses on Student Strengths and Feedback to Support Growth

Because of the ubiquity of grading and its impact on students, eliminating harmful grading practices provides a powerful way to reduce inequity. Practices such as grading for effort and participation, which rely on teachers' subjective judgments, can be eliminated, as studies indicate that white teachers consistently rate Black students lower in terms of participation than other students.[34] Eliminating such practices can reduce the impact of teacher bias on students' grades. Dropping other problematic grading practices, like assigning grades for homework completion or averaging grades without considering student growth, can reduce disparities in performance.[35] Other alternatives to traditional grading practices include emphasizing qualitative summaries of student accomplishment over numbers, as well as providing students with opportunities to revise work to improve grades.[36]

Feedback that focuses on student growth and a path to deepen thinking can also promote equity. Though feedback is often thought of as a neutral, objective process for correcting mistakes, the nature of feedback in fact establishes a particular kind of relationship between teachers and students.[37] A productive relationship is one where students feel a sense of authentic care from the teacher— that is, a concern for their intellectual well-being coupled with high expectations.[38] Further, effective feedback is grounded in what we know about how conceptual change happens: by building from what students already know, rather than seeking to directly eliminate "problematic" ideas.[39]

Time in teacher teams can be spent reviewing grading practices, eliminating harmful ones, and exploring how to give feedback that builds authentic relationships of care between students and teachers. Teachers can spend team time engaged in book study or looking at evidence-informed strategies that focus on how grading can support equity. These practices can add benefit to looking at student work together by focusing attention on the criteria educators use to evaluate it. This work can also help a team develop shared approaches to grading and feedback grounded in care and compassion for students.

Summary of Key Conclusions and Recommendations

To summarize, here is what we know:
- The use of data-driven decision making alone does not result in reduced disparities in educational outcomes.
- Performance-based assessments give students a chance to engage in consequential assessment.
- Instructional materials that support using culturally sustaining pedagogies provide a basis for meaningful assessment.
- Grading and feedback that focus on student strengths can promote equity.

Consequently, we recommend four evidence-informed strategies for making the most of teacher team time:
1. Choose and adapt more meaningful performance-based assessments.
2. Follow protocols that focus attention on evidence-informed practices for equity.
3. Use instructional materials that support culturally sustaining pedagogies.
4. Establish grading practices that focus on student strengths and feedback that supports growth.

In particular, we recommend that school leaders and teachers eliminate:
- extensive discussions of interim, benchmark, and end-of-year test scores;
- use of assessments that limit opportunities for students to show their capabilities for engaging in complex practices;
- one-size-fits-all curricula and assessments that do not center students' cultures;
- grading for participation and effort; and
- assigning grades for homework completion.

Meanwhile, we recommend that school leaders and teachers do support:
- the use of performance-based assessment;
- grounding assessment discussion in instructional materials that help students see themselves mirrored within activities;
- allowing students to revise assignments to improve grades; and
- providing feedback that builds relationships of care between teachers and students.

Taken together, these recommendations will allow school leaders and principals to use assessment to promote equity.

A Case for Urgent Action

It is urgent to act on this issue now. As society and schools attempt to recover from the pandemic, there is an opportunity for the education field to "imagine our world anew"[40] and engage in a "hard re-set."[41] But we are already seeing efforts to return to "normal" and to reembrace the pre-pandemic status quo of schooling. Inaction— in this case, the continued use of school data teams to narrowly fixate on standardized test score data and deficitize children, particularly those from marginalized groups— will only function to maintain a status quo that reinforces inequality and white supremacy. We therefore urge action amongst school leaders, principals, and teachers to make the most of team time by supporting equitable assessment practices and to implement the research-based recommendations we have outlined.

We acknowledge that implementing the strategies in this brief is likely to be challenging, because policies and barriers at the system-level encourage the very practices we advocate moving away from. Accountability policies that demand widespread testing and attach sanctions to schools for low performance put pressure on principals to emphasize looking at test data. Alternative systems around the globe feature a diminished role for tests, which could ease the pressure to rely on test scores as the sole indicator of school quality.[42] In addition, states and districts can re-allocate assessment funding to prioritize opportunities to interpret student responses to more open-ended tasks.[43] Coordinated investments in leader learning are also important, so that principals and coaches systemwide better understand the value of looking at student work together as a leverage point for equitable change.

Acknowledgment

Thank you to Beth Vinson from Denver Public Schools for her thoughtful contributions to this brief.

References

1. Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. American Educational Research Journal, 42(2), 231-268. https://doi.org/10.3102/00028312042002231

2. Bertrand, M., & Marsh, J. A. (2021). How data-driven reform can drive deficit thinking. Phi Delta Kappan, 102(8), 35-39. https://doi.org/10.1177/00317217211013936

3. Bertrand, M., & Marsh, J. A. (2015). Teachers' sensemaking of data and implications for equity. American Educational Research Journal, 52(5), 861-893. https://doi.org/10.3102/0002831215599251

    Mandinach, E., & Schildkamp, K. (2021). The complexity of data-based decision making: An introduction to the special issue. Studies in Educational Evaluation, 69, 1-3. https://doi.org/10.1016/j.stueduc.2020.100906

4. Lazarín, M. (2014). Testing overload in America's schools. Center for American Progress.

5. Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas accountability system. American Educational Research Journal, 42(2), 231-268. https://doi.org/10.3102/00028312042002231

6. Bertrand, M., & Marsh, J. A. (2021). How data-driven reform can drive deficit thinking. Phi Delta Kappan, 102(8), 35-39. https://doi.org/10.1177/00317217211013936

    Gutiérrez, R. (2008). A "gap-gazing" fetish in mathematics education? Problematizing research on the achievement gap. Journal of Research in Mathematics Education, 39(4), 357-364. https://doi.org/10.5951/jresematheduc.39.4.0357

7. Penuel, W. R., & Shepard, L. A. (2016). Assessment and teaching. In D. H. Gitomer & C. A. Bell (Eds.), Handbook of Research on Teaching (pp. 787-851). AERA.

8. Ryoo, J., Goode, J., & Margolis, J. (2015). It takes a village: Supporting inquiry- and equity-oriented computer science pedagogy through a professional learning community. Computer Science Education, 25(4), 351- 370. https://doi.org/10.1080/08993408.2015.1130952

9. Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English language arts and mathematics. Educational Researcher, 42(4), 223-233. https://doi.org/10.3102/0013189X13480524

10. Koretz, D. (2017). The testing charade. In The Testing Charade. University of Chicago Press.

11. Darling-Hammond, L. (1994). Performance-based assessments and educational equity. Harvard Educational Review, 64(1), 5-31. https://doi.org/10.17763/haer.64.1.j57n353226536276

    Guha, R., Wagner, T., Darling-Hammond, L., Taylor, T., & Curtis, D. (2018). The promise of performance assessments: Innovations in high school learning and higher education admissions. Learning Policy Institute.

12. Kang, H., & Furtak, E. M. (2021). Learning theory, classroom assessment, and equity. Educational Measurement: Issues and Practice, 40(3), 73-82. https://doi.org/10.1111/emip.12423

13. Penuel, W. R. (2021). Possible futures for equitable educational assessment. Phi Delta Kappan, 103(4), 54-57. https://doi.org/10.1177/00317217211065829

    Taylor, C. S., & Nolen, S. B. (2022). Culturally and socially responsible assessment: Theory, research, and practice Teachers College Press.

14. Randall, J. (2021). "Color-neutral" Is not a thing: Redefining construct definition and representation through a justice-oriented critical antiracist lens. Educational Measurement: Issues and Practice, 40(4), 82-90.

15. Penuel, W. R., Carlson, C. E., Wachowski, S., Freitas, D., Alvarez, A., Hamilton, R., Schild, A., & Raza, A. (2022). Steps to designing justice-focused assessments. Institute for Science + Math Education, University of Washington.

16. Horn, I.S., & Little, J. W. (2010). Attending to problems of practice: Routines and resources for professional learning in teachers' workplace interactions. American Educational Research Journal, 47(1), 181-217.

17 Blythe, T., Allen, D., & Powell, B. S. (1999). Looking together at students' work: A companion guide to assessing student learning. Teachers College Press.

Minstrell, J. (1992). Facets of students' knowledge and relevant instruction. In F. Duit, F. Goldberg, & H. Niedderer (Eds.), Research in physics learning: Theoretical issues and empirical studies (pp. 110-128). IPN.

18. Inoue, A. B. (2015). Antiracist writing ecologies: Teaching and assessing writing for a socially just culture. Parlor Press.

19. Blythe, T., Allen, D., & Powell, B. S. (1999). Looking together at students' work: A companion guide to assessing student learning. Teachers College Press.

Dempsey, K., Beesley, A. D., Fazendeiro Clark, T., & , & Tweed, A. (2015). Authentic student work samples support formative assessment in middle school. In C. Suurtamm & A. R. McDue (Eds.), Annual perspective in math education 2015: Assessment to enhance learning and teaching (pp. 157-169). National Council of Teachers of Mathematics.

20. Goodwin, B., & Hein, H. (2016). Research says: Looking at student work yields insights. Educational Leadership, 73(7), 79-80.

21. MacDonald, E. (2011). When nice won't suce: Honest discourse is key to shifting school culture. Journal of Staff Development, 32(3), 45-51.

22. Dearman, C. C., & Alber, S. R. (2005). The changing face of education: Teachers cope with challenges faced by collaboration and reflective study. The Reading Teacher, 68(7), 634-640. https://doi.org/10.1598/RT.58.7.4

23. Bryk, A. S., Gomez, L. M., Grunow, A., & LeMahieu, P. G. (2015). Learning to improve: How America's schools can get better at getting better. Harvard University Press.

24. Raza, A., Penuel, W. R., & Salinas del Val, Y. (2022, March). Supporting science teachers in using student experience data to support more equitable participation in science classrooms NARST Annual Conference, Vancouver, CA.

25. Morales-Doyle, D. (2017). Justice-centered science pedagogy: A catalyst for academic achievement and social transformation. Science Education, 101, 1034-1060.

Paris, D. (2012). Culturally sustaining pedagogy: A needed change in stance, terminology, and practice. Educational Researcher 41(3), 93-97.

26. Girard, L., Wiley, K., DeBarger, A. H., Bichler, S., Bradford, A., & Linn, M. C. (2021). Self-directed science learning during COVID-19 and beyond. Journal of Science Education and Technology, 31, 258-271. https://doi.org/10.1007/s10956-021-09953-w

27. Gutiérrez, R. (2007). Context matters: Equity, success, and the future of mathematics education. In T. Lamberg & L. R. Wiest (Eds.), Proceedings of the 29th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education (pp. 1-18). University of Nevada.

28. Segura, D., Morales-Doyle, D., Nelson, S., Levingston, A., & Canales, K. (2021). Sustaining community-school relationships around shared visions of climate justice and science teaching. Connected Science Learning, 3(5). https://www.nsta.org/connected-science-learning/connected-science-learning-september-october2021/sustaining-community

29. Gonzalez, K., Lynch, K., & Hill, H. C. (2022). A meta-analysis of the experimental evidence linking STEM classroom interventions to teacher knowledge, classroom instruction, and student achievement. (EdWorkingPaper: 22-515). Annenberg Institute at Brown University. https://doi.org/10.26300/d9kc-4264

References

30. Penuel, W. R., DeBarger, A. H., Boscardin, C. K., Moorthy, S., Beauvineau, Y., Kennedy, C., & Allison, K. (2017). Investigating science curriculum adaptation as a strategy to improve teaching and learning. Science Education, 101(1), 66-98. https://doi.org/10.1002/sce.21249

31. Au, W. (2017). Can we test for liberation? Moving from retributive to restorative and transformative assessment in schools. Critical Education, 8(13)

32. Au, W. (2022). Unequal by design: High-stakes testing and the standardization of inequality. Taylor & Francis.

33. Penuel, W. R., & Watkins, D. A. (2019). Assessment to promote equity and epistemic justice: A use-case of a research-practice partnership in science education. Annals of the American Academy of Political and Social Science, 683, 201-216. https://doi.org/10.1177/0002716219843249

Shepard, L. A., Penuel, W. R., & Pellegrino, J. W. (2018). Using learning and motivation theories to coherently link formative assessment, grading practices, and large-scale assessment. Educational Measurement: Issues and Practice, 37(1), 21-34. https://doi.org/10.1111/emip.12189

34. Downey, D. B., & Pribesh, S. (2004). When race matters: Teachers' evaluations of students' classroom behavior. Sociology of Education, 77(4), 267-282. https://doi.org/10.1177/003804070407700401

35. Feldman, J. (2018). Grading for equity: What it is, Why it matters, and how it can transform schools and classrooms. SAGE.

36. Zerwin, S. M. (2020). Point-less: An English teacher's guide to more meaningful grading. Heinemann.

37. Vossoughi, S., Nzinga, K., Berry, A., Irvine, F., Mayorga, C., & Gashaw, M. (2021). Writing as a social act: The feedback relation as a context for political and ethical becoming. Research in the Teaching of English, 56(2), 200-222.

38. Valenzuela, A. (1999). Subtractive schooling: U.S.-Mexican youth and the politics of caring. SUNY Press.

39. Smith, J. P., III, diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. The Journal of the Learning Sciences, 3(2), 115-163. https://doi.org/10.1207/s15327809jls0302_1

40. Roy, A. (2020). Azadi: Freedom. fascism. fiction. Haymarket Press.

41. Ladson-Billings, G. (2021). I'm here for the hard re-set: Post pandemic pedagogy to preserve our culture. Equity & Excellence in Education, 54(1), 68–78. https://doi.org/10.1080/10665684.2020.1863883

42. Allen, R. (2012). Developing the enabling context for school-based assessment in Queensland, Australia. The World Bank.

43. Klenowski, V. (2011). Assessment for learning in the accountability era: Queensland, Australia. Studies in Educational Evaluation, 37(1), 78-83. https://doi.org/10.1016/j.stueduc.2011.03.003