RECOMMENDATIONS

Many issues will need to be addressed in order to bring about the fundamental changes in assessment practice necessary to promote and value deeper learning. The recommendations offered here are meant to serve as a starting point for a process that likely will unfold over many years, perhaps even decades. The question is: Can policymakers sustain their attention to this issue long enough to enact the policies necessary to bring about necessary changes? For that matter, can educators follow through with new programs and practices that turn policy goals into reality? And will the secondary and postsecondary systems be able to cooperate in creating systems of assessments and focusing instruction on deeper learning?

I believe that if we are to move toward these goals, education policymakers will need to:

1. **Define college and career readiness comprehensively.**
   States need clear definitions of college and career readiness that highlight the full range of knowledge, skills, and dispositions that research shows to be critical to students’ success beyond high school (including not only key content knowledge but also cognitive strategies, learning skills and techniques, and knowledge and skills related to the transition to college and the workforce).

2. **Take a hard look at the pros and cons of current state accountability systems.**
   If they agree that college and career readiness entails far more than just a narrow set of academic skills and knowledge, then policymakers should ask themselves how well—or poorly—existing state and district assessments measure the full range of things that matter to students’ long-term success. Further, policymakers should take stock of the real-world impacts that the existing assessment models have had on teaching and learning. For well over a decade, proponents of high-stakes testing have asserted that the prevailing model of accountability creates strong incentives for teachers and schools to improve. However, high-stakes testing is past due for an assessment of its own. State leaders should ask themselves: Are the existing tests, and their use in evaluating teacher and school performance, truly having the desired impact? In reality, what changes in instruction do teachers make in response to summative results and their use in evaluating their, and their schools’, performance? How much time and money is currently devoted to such tests, and what might be the opportunity costs? That is, to what extent could high-stakes testing be crowding out other, more useful ways of assessing student progress?

3. **Support the development of new assessments of deeper learning.**
   Across the country, many efforts are now underway to create assessments that address a wide range of knowledge and skills, going well beyond reading and mathematics, and these efforts need to be encouraged and nurtured. However, several key

Can policymakers sustain their attention to this issue long enough to enact the policies necessary to bring about necessary changes? For that matter, can educators follow through with new programs and practices that turn policy goals into reality?
problems will need to be resolved if assessments of deeper learning are to be scalable, reliable, and useful enough to justify their expense. In particular, when it comes to measures that require students to report on their own progress—or that require teachers to rate students in some way—measures will have to be developed by which to triangulate these reports against other data sources, in order to ensure a reasonable level of consistency. Further, it will be extremely important to institute safeguards to protect students’ privacy and ensure that this sort of information is not used inappropriately. And, finally, policymakers and educators will have to be careful to distinguish between assessment tools that are meant to serve low-stakes, formative purposes—generating information that can be used to improve teaching and learning—and those that can fairly be used as the basis for summative judgments about students’ learning or teachers’ performance.

4. **Learn from past efforts to build statewide performance assessment systems.** States’ pioneering efforts to develop performance assessments in the 1990s and early 2000s yielded a wealth of lessons that can inform current attempts to expand assessment beyond a limited set of tests. Most important is the need to proceed slowly at first, in order to develop systems by which to manage the sometimes-complex mechanics of collecting, analyzing, reporting, and using these types of richer information. Educators, especially, must have sufficient time to learn how to work with new assessments, not only how to score them but how to teach to them successfully.

5. **Take greater advantage of advances in information technology.** Many of the challenges that confronted states 25 years ago, when they first adopted performance assessment systems, can be addressed today through the use of vastly more sophisticated technology for information storage and retrieval. Online storage is plentiful and cheap, and it is far easier to move data electronically now than it was then. The technological literacy level of educators is higher, as are the capabilities of postsecondary institutions to receive information electronically. If districts and states take advantage of this new capacity to manage complex data in useful and user-friendly ways, they should find it much easier than in past decades to store student data in digital portfolios and access that information to meet the needs of audiences such as educators, admission officers, parents, students themselves, and perhaps potential employers.

6. **Adapt federal education policy to allow greater flexibility in the types of data that can be used to demonstrate student learning and growth.** The U.S. Department of Education’s waiver process has introduced some flexibility with respect to the measures of student learning that states—and, in at least one case, a consortium of school districts—can use to meet federal accountability requirements. However, any reauthorization of the Elementary and Secondary Education Act and its NCLB provisions should go much further to encourage the use of multiple forms of assessment and to make clear to states that such models can pass federal muster.

7. **Consider using the National Assessment of Educational Progress as a baseline measure of student problem-solving capabilities.** The design of NAEP, particularly the fact that not all test-takers are asked to complete the entire battery of NAEP items, allows it to include fairly complex and time-intensive tasks. This design characteristic can be used both to field-test more complex performance items as well as to generate a better national metric of student problem-solving skills in the areas NAEP assesses. Having a baseline that is consistent across states can help determine which states are making the most progress with their statewide systems of assessment of deeper learning. PISA, too, could be used in this fashion, but the implementation challenges would be much greater than building upon NAEP’s existing infrastructure.

8. **Build a strong base of support for a comprehensive system of assessments.** The process of developing a more complex system of assessments must not exclude any major group of stakeholders. Teachers in particular need to be centrally involved in designing, scoring, and determining how data from rich assessments of student learning will be used. State policymakers, too, have a compelling interest in finding ways to make sure that those assessments are both valid and reliable. And postsecondary and business leaders must have a seat at the table, as well, if they will be expected to make use of any new sources of information about students’ college and career readiness.

9. **Determine the professional learning, curriculum, and resource needs of educators.** Currently, few states do much, if anything, to gauge schools’ capacity to provide meaningful opportunities for professional learning. And as a result, most schools are unable to help their teachers acquire new skills. In order to implement any
new assessments successfully, it will be absolutely critical to determine—early on in the process—what resources will be necessary to ensure that all teachers are assessment literate, can use the information generated by multiple sources of assessment, are capable of developing assignments that lead to deeper learning, and can teach the full range of content and skills that prepare students to succeed in college and careers. It is worth noting that few state education departments or intermediate service agencies currently have the capacity to offer the level of guidance and support most schools, particularly those in smaller districts, need to undertake the type of professional learning program necessary to implement and use a system of assessments approach to instructional improvement.

10. Look for ways to improve the Common Core State Standards and related assessments so that they become better measures of deeper learning. This may be a tall order at a time when Common Core implementation is undergoing a rocky period. However, the surest way to undermine the credibility of the standards and the assessments would be to refuse to improve them in response to feedback from the field. Such a stance would only lead educators to view them as just another mandate to be complied with, rather than as a source of professional guidance and growth. Already, the standards are almost five years old, and it is past time to begin the lengthy process of designing and initiating a careful and systematic review process. Similarly, even though PARRC and SBAC are only just now completing their field testing, their designers must continue to seek out criticism, keep a close eye on their rollout, communicate more frankly and vocally the limitations of these assessments, while simultaneously suggesting ways to get at the various aspects of college and career readiness that these assessments currently overlook.

Ideally, the educational assessment system of the future will be analogous to a thorough, high-quality medical diagnostic procedure, rather than the cursory check-up described at the beginning of this paper. Educators and students alike will have at their disposal far more sophisticated and targeted tools to determine where they are succeeding, to show where they are falling short, and to point in the direction of how and what to improve. They will receive rich, accurate information about the cause of any learning problems, and not just the symptoms or the effects. Policymakers will understand that improved educational practice, just like improved health, is rarely achieved by compelling people to follow uniform practices or using data to threaten them but, rather, by creating the right mix of incentives and supports that motivate and reward desired actions, and that help all educational stakeholders to understand which outcomes are in their mutual best interests.

Research and experience make it clear that educational systems that can foster deeper learning among students must incorporate assessments that honor and embody these goals. New systems of assessment, connected to appropriate resources, learning opportunities, and productive visions of accountability, comprise a critical foundation for enabling students to meet the challenges that face them throughout their education and careers in the 21st century.