EXCERPT FROM
GETTING READY FOR COLLEGE, CAREERS AND THE COMMON CORE:
WHAT EVER EDUCATOR NEEDS TO KNOW
Dr. David Conley
To be published in Fall 2013 by Jossey-Bass, San Francisco

CHAPTER 10:
FROM AN ASSESSMENT SYSTEM TO A SYSTEM OF ASSESSMENTS

This chapter makes the case for the use of multiple assessment measures and a wider array of data sources for determinations of student college and career readiness and for a range of other purposes as well. A multidimensional system of assessments within a state yields data that can be used at all levels, from the statehouse to the schoolhouse. Such a system will be more consistent with the model of college and career readiness presented in earlier chapters, which emphasizes the wide range of aspects that need to be developed in students for them to be ready to succeed in college and careers.

More complex and varied data can yield improved insight into the degree of overall student readiness. Students can also be afforded alternative means to demonstrate readiness. A grade point average can still generate useful information in combination with other sources, although its role is becoming somewhat limited as more students have GPAs that approach or exceed 4.0. An increasing number of states administer college entrance examinations to all students. The ACT, with its EPAS system of EXPLORE/PLAN/ACT, and College Board, through its Pathway system consisting of ReadiStep, the PSAT, and the SAT, provide an overall judgment in relation to a score cut point and nationally normed comparative information on student readiness for college and careers. An increasing number of students complete college-level work in high school through Advanced Placement or International Baccalaureate tests or from college courses taken while in high school through concurrent or dual enrollment programs. A few states have instituted other testing requirements as well. In short, a range of information is already being collected, but most of it is not yet being integrated and reported in complementary ways that create a valid and complete picture of college and career readiness.

Results from the consortia assessments are now being added to the mix, including performance tasks that require more complex and extensive writing and other skills, such as problem solving. This information can offer additional insight into readiness. Evidence strongly suggests that skills such as being able to write a research paper independently are highly indicative of potential preparedness for a wide range of postsecondary programs. Knowing more about how well students write research papers is an example of additional evidence that is highly useful when making readiness determinations and offers students an alternative way to demonstrate readiness. Valuing this type of more complex information that a student work product generates also clearly signals to students the importance of developing the academic skills necessary to produce quality products.

Many schools throughout the country already espouse a philosophy of learning that can be measured only through more complex performance assessments. Schools such as the Aspire and Envision charters ask students to complete semester and year-long projects that require integrating a broad range of skills and knowledge. The New York Performance Standards Consortium consists of schools that build instruction around performance tasks. The Coalition of Essential Schools has a long tradition stretching back to the mid-1980s of juried reviews of student work products and projects. Deborah Meier and her colleagues at Central Park East Secondary School, for example, expected students to produce highly complex and challenging graduation portfolios. Many other schools have long histories of using performance tasks. In addition, a number of states have instituted requirements for culminating projects or other types of performance demonstrations as a required element for high school graduation. Massachusetts, Washington State and Oregon, for example, allowed students to submit collections of evidence in place of scores on high school exit exams. In other words, complex assessment is not entirely new or novel, and many schools throughout the nation can already serve as laboratories or demonstration sites for those interested in implementing tasks in their schools and for states looking to adopt multiple routes for students to demonstrate necessary knowledge and skills.

As new tests are implemented and current state tests are reviewed and revised, the opportunity exists to move away from an assessment system composed of often overlapping, redundant, or disconnected scores or reports and toward a system of assessments model that yields information that addresses state accountability requirements and also provides students, teachers, schools, and postsecondary education institutions with performance data to inform a range of decisions that leads to ongoing improvement. Rarely is it possible to achieve this goal with a single test or even multiple measures when such measures overlap with one another on what they test.

While the Common Core State Standards do not specify everything that is necessary for postsecondary readiness and success, they do include many standards that are cognitively complex and critical for success in college and careers. The consortia assessment developers readily admit that their tests do not capture every factor necessary for college and career readiness because they test only English and mathematics. However, even in these two subject areas, consortia assessments will not be able to assess some Common Core standards. Following are some examples of concepts contained in the standards, explicitly or implicitly, that are not well addressed by the consortia assessments in their current configuration and that require assessment beyond what the consortia exams will be able to offer:

- Conducting research and synthesizing information
- Developing and evaluating claims
- Reading critically and analyzing complex tasks
- Communicating ideas through writing, speaking, and responding
- Planning, evaluating, and refining solution strategies
- Designing and using mathematical models
- Explaining, justifying, and critiquing mathematical reasoning

A number of states that want to provide students, teachers, schools, and postsecondary programs with the information they need to determine college and career readiness levels for students are considering a system of assessment approach. For example, the Innovation Lab Network, sponsored by the Council of Chief State School Officers (CCSSO), is a group of ten states committed to a series of educational reforms to support, supplement, and at times go beyond the Common Core State Standards. One of their initiatives is the "Consortium-Plus" model of assessment in which states commit to develop additional measures beyond the common core. Kentucky is perhaps the best example—as of the spring of 2013—of a state that has taken steps toward developing a more comprehensive set of measures that moves the state toward a system of assessments model.

While some may disagree with specific elements of the Kentucky model or question the overall amount of assessment that is going on in the state, the important point is that the model has been developed to meet the needs of the Kentucky educational system. The idea behind a system of assessments approach is that each state that adopts such a strategy will likely have a different mix of measures based on state goals, priorities, and traditions. Some states might opt for more classroom-based assessment, demonstrations, and performance tasks aligned with a competency grading system, while others might add data points and measures that would yield high-quality information for teacher evaluation and instructional improvement, along with other important purposes. The system of assessments approach allows each state to craft the set of measures necessary for its schools to demonstrate performance in relation to key state goals while addressing school-level needs and contributing to college and career readiness determinations. Table 10.1 shows Kentucky's system of assessments.

TABLE 10.1 SYSTEM OF ASSESSMENTS
FOR THE STATE OF KENTUCKY

Kentucky System of Assessments	Grade Level	Purpose of Exams
Assessing Comprehension and Communication in English State-to-State (ACCESS)	K-12	English language proficiency
ACT	10-11	College readiness
ACT QualityCore	10-12	End of course
ACT Work Keys (Applied Math, Locating Information, and Reading for Information)	12	Career readiness
Armed Services Vocational Aptitude Battery (ASVAB)	10-12	Career readiness
COMPASS	12	College placement assessment

EXPLORE	8	College readiness assessment
Industry Certificate Assessments	9-12	Career readiness
Kentucky Occupational Skills Standards Assessment (KOSSA)	10-12	Career readiness
Kentucky Performance Rating for Educational Progress (K-PREP)	3-8	Kentucky Core Academic Standards
Kentucky Online Testing (KYOTE)	12	College readiness
PLAN	10	College readiness

A carefully designed system of assessments takes into account the varied needs of all the constituents who use assessment data: teachers, principals, superintendents, and boards of education; college and university officials and administrators in proprietary training programs; state education department staff, legislators, and governors; staff at the US Department of Education and in Congress; members of education advocacy groups; parents; the business community; and many others. The system of assessments approach collates information from enough different sources to address a wider range of needs than does rating schools as succeeding or failing based on one set of test scores. The system does so in a way that results in a more holistic picture of students, schools, and educational systems that does not waste or duplicate information or effort but also does not use a single source of data inappropriately. This method also allows a wider range of students to demonstrate what they know and can do by means of additional options and methods.

For example, using test scores to determine where individual students fall on a performance level system (e.g., approaches, meets, exceeds) requires a great deal of attention to the stakes associated with the decisions. As the stakes get higher, more information is needed. For example, if scores are going to be used to make determinations about graduation, remediation, placement, financial aid, or admissions, more than a single test score should be factored into the equation. Additional sources of information on the knowledge and skills associated with readiness and success help reduce classification errors, which occur when a student who is ready is labeled not ready or vice versa. A system of assessments can provide more valid and reliable information for a variety of purposes, including high stakes decisions about individual students.

A system of assessments can derive information on college and career readiness from a wide range of potential sources. For example, tests or end-of-course exams in subject areas other than reading, writing, and math yield better insight into breadth of knowledge. Surveys of student learning skills and techniques reveal whether students are becoming lifelong learners. Inventories of student knowledge about the college selection and application process, financial aid, and matriculation demonstrate the degree to which students are learning about the transition process.

Performance Tasks: Key Element in a System of Assessments

One additional source of information is the performance task, which comes in a wide variety of shapes and sizes. Administered in the classroom, these tasks can vary dramatically in length and complexity. They can take anywhere from a period to a semester to complete and have the potential to yield information that standardized tests do not. Performance tasks are not necessarily the same as a teacher-generated assignment. Task quality is generally more tightly controlled by identifying relevant content, specifying the conditions for administration, designating the scaffolding or support available to students, standardizing the scoring process, and controlling how results are reported. Additional quality control can be achieved through the use of moderation processes in which selected tasks are rescored by someone other than the teacher and the moderation scores are then used to help teachers score more reliably in the future.

Information from classroom-based performances is used more frequently than standardized tests in a number of educational systems around the world. For example, the states of Queensland and Victoria in Australia and schools in the International Baccalaureate Diploma Programme have invested in developing educator skills as classroom-based performance assessors. They devote resources to training teachers how to develop quality assessments at the classroom level, then expect teachers to incorporate complex assignments and projects into their courses, and then assess student work appropriately and accurately. It is not uncommon for students to write essays or give presentations and have them scored by their teachers or members of expert panels who use scoring guides. Much of this assessment occurs at the end of courses as well as at the culmination of studies at designated levels.

This type of assessment can take place because educators have acquired over time very similar ideas and explanations of what adequate performance on these papers and tasks looks like. In nations as varied as the Netherlands and Singapore, teachers share mental models of student performance that allow them to administer performance tasks with consistency and to judge student performance reliably. These mental models are developed from the earliest stages of teacher preparation and induction and are reinforced by high-quality in-course assessments and grading practices that also use scoring guides closely aligned with teachers' mental models of adequate student performance. In this fashion, teachers can score performance tasks efficiently and effectively.

Range of Performance Task Types

The types of performance tasks or measures that are useful in a system of assessments can cover a wide span (figure 10.1). At one end are simple assignments that can be completed in a portion of a class period. Exhibit 10.1 and 10.2 contain examples of tasks that simply require students to write several paragraphs. Other short performance tasks may ask students to take information that is given and interpret it, reorganize data and draw diagrams that explain relationships, or classify and categorize objects or data into like and unlike groups. Even simple tasks can assess knowledge and skills that cannot be gauged well with multiple-choice items. Teachers often devise these types of tasks

themselves, pull them from curriculum materials, or access them online. They are generally closely tied to the content at hand and require only modest extrapolation and application of terms, ideas, and concepts being learned in class. An example of this type of task might be one in which students are asked to write a new ending to a story using a different literary style. A math-oriented in-class, single-period tasks might provide students with a list of prices at local gas stations and ask them to determine which was the best value from a cost-benefit perspective, taking into account variables such as time to travel to the station, gas costs of driving there, proximity to other stops that must be made while getting gas, and other salient factors.

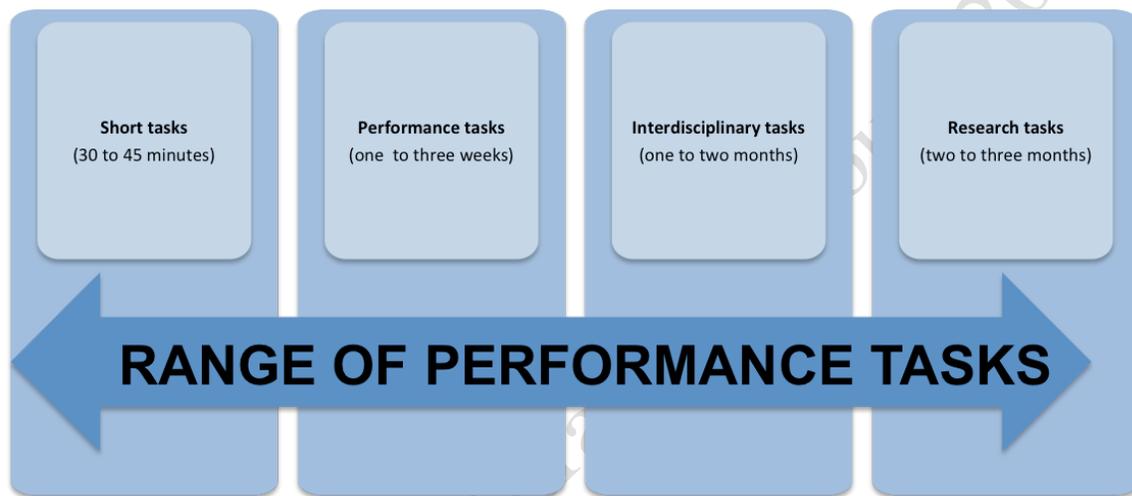**FIGURE 10.1 CONTINUUM OF ASSESSMENT FOR DEEPER LEARNING**

| Short tasks (30 to 45 minutes) | Performance tasks (one to three weeks) | Interdisciplinary tasks (one to two months) | Research tasks (two to three months) |

RANGE OF PERFORMANCE TASKS

EXHIBIT 10.1 SHORT PERFORMANCE TASK

Grade 10 Prose Constructed Response

Use what you have learned from reading "Daedalus and Icarus" by Ovid and "To a Friend Whose Work Has Come to Triumph" by Anne Sexton to write an essay that provides an analysis of how Sexton transforms "Daedalus and Icarus."
As a starting point, you may want to consider what is emphasized, absent, or different in the two texts, but feel free to develop your own focus for analysis.
Develop your essay by providing textual evidence from both texts. Be sure to follow the conventions of standard English.

Grade 10 Prose Constructed Response—Sample #1 from Literary Analysis Task."*Parcconline.org*. Partnership for Assessment of Readiness for College and Careers, n.d. Web. 18 June 2013. http://www.parcconline.org/samples/english-language-artsliteracy/grade-10-prose-constructed-response%E2%80%94sample-1-literary-analysis

EXHIBIT 10.2 – SHORT PERFORMANCE TASK

Grade 6 Writing Task – Cell Phones

Read the text and complete the task that follows it.

Cell Phones in School—Yes or No?

Cell phones are convenient and fun to have. However, there are arguments about whether or not they belong in schools. Parents, students, and teachers all have different points of view. Some say that to forbid them completely is to ignore some of the educational advantages of having cell phones in the classroom. On the other hand, cell phones can interrupt classroom activities and some uses are definitely unacceptable. Parents, students, and teachers need to think carefully about the effects of having cell phones in school.

Some of the reasons to support cell phones in school are as follows:
• Students can take pictures of class projects to e-mail or show to parents.
• Students can text-message missed assignments to friends that are absent.
• Many cell phones have calculators or Internet access that could be used for assignments.
• If students are slow to copy notes from the board, they can take pictures of the missed notes and view them later.
• During study halls, students can listen to music through cell phones.
• Parents can get in touch with their children and know where they are at all times.
• Students can contact parents in case of emergencies.

Some of the reasons to forbid cell phones in school are as follows:
• Students might send test answers to friends or use the Internet to cheat during an exam.
• Students might record teachers or other students without their knowledge. No one wants to be recorded without giving consent.
• Cell phones can interrupt classroom activities.
• Cell phones can be used to text during class as a way of passing notes and wasting time.

Based on what you read in the text, do you think cell phones should be allowed in schools? Using the lists provided in the text, write a paragraph arguing why your position is more reasonable than the opposing position.

Next along the continuum of performance tasks are those that require at least some out-of-class work. These are incrementally more complicated because the teacher has to verify that all the work produced is the student's own. For example, students might be required to access information from US Census databases to answer specific questions about local conditions. The task could measure ELA knowledge and skills, math knowledge and skills, or a combination of the two. Part of the requirement would be a

draft and then a final version with edits and revisions. Tasks of this type are increasingly available from both commercial sources and online task banks. Although some are teacher developed and not necessarily reviewed for their content validity or other psychometric properties, a growing number of such tasks have been carefully designed and thoroughly vetted to ensure they measure what they purport to measure and can be scored reliably. Exhibit 10.3 presents an example of a ThinkReady task that requires students to work in class and out of class to collect information and conduct research on a topic of interest and produce multiple drafts.

EXHIBIT 10.3 ONE TO THREE-WEEK PERFORMANCE TASK WITH IN-CLASS AND OUT-OF-CLASS WORK

A Modest Solution: Writing Satire about Current Events

Students write a satire in the form of Jonathan Swift 's essay, "A Modest Proposal." Students first research a real-world problem that interests them to learn about its causes and possible solutions. For their proposal, they create a persona who offers an overt, absurd solution while also communicating a covert, but real, solution proposed by the student.

In another performance task example, students in middle school math might be asked to use information about traffic volumes and flow to identify the best routes to take to get to various destinations and then to make recommendations on how to improve traffic flow overall or where to site a new hospital so that it is accessible but not in an area of high congestion. The first part might be completed in class individually, while the second might require additional work outside class followed by group work in class. Scoring might include a component score for correct use of mathematics, a second for problem-solving techniques, and a third for thoroughness of proposed solution.

A third example is a type of task that is longer in duration and may take up to a third of a semester. This is really best described as a project. Often it is the student who defines the focus of the project and is responsible for organizing the task and locating all the necessary information to complete it. The student may be expected to follow a particular outline or to address a range of requirements in the process of completing the project. The project may be judged by the teacher alone or in conjunction with someone who has expertise in the subject area in which the project is focused. Exhibit 10.4 contains an example of an interdisciplinary task from Envision Schools requiring students to integrate knowledge from several subject areas. For this type of project, a student or team of students undertakes an investigation of some sort, such as locally sourced foods. The investigation requires them to conduct research on a number of topics:
- Where food they eat comes from
- What proportion of the price represents transportation
- How dependent they are on other parts of the country for their food and what would happen to local food supplies if the national transportation system were disrupted for a week or a month

- What choices they could make if they wished to eat more locally produced food, what the economic implications of doing so would be, and whether doing so could cause unintended economic consequence in other parts of the country

---

EXHIBIT 10.4 ONE- TO TWO-MONTH INTERDISCIPLINARY TASK

Disaster in the Gulf Project

In response to the April 2010 BP Deepwater Horizon oil drilling rig accident, seniors at Envision Schools explore effects and impact through Disaster in the Gulf, an interdisciplinary project:
• *AP Government*: Produce a research paper about our government 's role in responding to such a disaster, including the role of federal agencies and our national emergency management system.
• *World Literature*: Write a three- to four-minute speech using rhetorical skills and deliver the speech at a simulated congressional hearing.
• *AP Environmental Science*: Explore the environmental impact of the oil spill. Consider different methodologies of cleaning the affected areas, along with the social, economic, and environmental impact of the oil and cleaning.
• *Advanced Visual Arts*: Create sculptures and other art forms from petroleum-based materials.

This project takes place over nine weeks and is reviewed by the subject-area teacher using a rubric from the College Success Student Performance Assessment System.

The project is presented to the class and scored by the teacher using a guide that included ratings of the use of mathematics and economics content knowledge; the quality of argumentation; the appropriateness of information sources cited and referenced; the quality and logic of the conclusions reached; and overall precision, accuracy, and attention to detail.

Finally, a fourth type of performance assessment is really more like a culminating examination in which skills acquired over multiple years are applied to a large, complex problem. The prompt may be standardized and developed externally, as in the case of the example in exhibit 10.5 , or may be generated at the school site or by the teacher, in the case of a culminating project undertaken to meet graduation requirements. This type of task takes the form of a paper that is scored with a common scoring guide. This task has elements of the interdisciplinary task but is distinguished primarily by the longer period of time over which it takes place (up to a semester), combined with the higher level of rigor that comes from the external scoring. The extended essay reflects the types of assessments used in many other countries to gain greater insight into student thinking skills and content understanding and mastery.

Students are supported throughout the process with advice and guidance from a supervisor (usually a teacher at the school). Performance tasks can generate insight into other aspects of student learning skills and strategies. For example, teachers can report on student ability to sustain effort when confronted with difficult tasks; to manage time to complete complex, multistep assignments; and to work with others to improve both individual and group performance. This evidence of readiness for postsecondary educational opportunities and career pathways can be used in combination with scores on English and math tests to allow students with a lower score in, say, math, to compensate with strong evidence of effective study habits, good collaborative skills, and the ability to seek help from instructors. This more varied information can come from performance tasks, where teachers observe the learning skills, techniques, and strategies students employ. Scoring guides can rate these types of learning skills along with content knowledge. Such performance task scores can be used to identify students with postsecondary potential who may be struggling academically but respond well to performance tasks as a means of expressing their learning skills, thus showing their greater potential for success in a postsecondary program that requires the ability to learn independently and seek help when needed.

EXHIBIT 10.4 FOUR-MONTH RESEARCH TASK

The extended essay is an independent, self-directed piece of research that takes place over the course of a semester, culminating in a four-thousand-word paper. As a required component, it provides:
• Practical preparation for the kinds of undergraduate research required at tertiary level
• An opportunity for students to engage in an in-depth study of a topic of interest within a chosen subject.

Emphasis is placed on the research process:
• Formulating an appropriate research question
• Engaging in a personal exploration of the topic
• Communicating ideas
• Developing an argument

Participation in this process develops the capacity to:
• Analyze
• Synthesize
• Evaluate knowledge

How a System of Assessments Addresses a Wider Range of Standards

The Common Core State Standards include many areas that are crucial for postsecondary readiness and success but cannot easily be assessed with the types of methods being used by the consortia assessments (the list near the start of this chapter contains examples). These standards and others are candidates for the range of performance tasks just described. It's easy to think of extended activities that would cause students to demonstrate the more complex cognitive skill sets these standards are attempting to

develop. This could be accomplished through structured tasks that require the application of content knowledge and problem-solving skills to a defined problem, such as choosing among several proposed plans to improve local economic development, or they could be much more open ended, such as requiring students to identify an issue or topic, explain their choice, and then indicate which standards the project will address. Students would be challenged to formulate the problem and then present their interpretation and potential solution.

This type of more active engagement with content not only helps students understand what they have been taught and to integrate their knowledge; it also develops cognitive skills and learning strategies that help them to be successful with new content. When the focus of schooling is on one test, no matter how good that test is, classroom instruction and learning naturally gravitate toward what is on the test. Aligning instruction with a test is not automatically a bad thing, but it can become problematic when the test does not or cannot address all of the standards that are supposed to frame overall curriculum and instruction. Widening the range of assessment types and integrating assessment more fully into classroom practice helps ensure that the full range of standards, including the complex and intellectually demanding ones, is properly valued and evaluated.

A system-of-assessments approach opens the door to a much wider array of measurement instruments and approaches and of opportunities for students to become involved in and take ownership of the work they are doing. Currently states limit their assessment options because almost all measures are viewed through the lens of accountability purposes and the technical requirements of high-stakes testing. These requirements generally end up emphasizing reliability over validity. In other words, instruments that yield the same results over repeated administrations are valued more highly, even if what the instrument measures is not necessarily the most accurate representation of the learning or behavior that is of interest in the first place. Reliability is important because people tend to find unreliable measures to be unfair. However, validity is also important, and it is strengthened by using appropriate means to measure what is important.

A system of assessments would be necessary and useful even if the consortia assessments tested all of the Common Core State Standards perfectly because it would allow more students to demonstrate their knowledge and skills in different ways. However, they do not and cannot capture information from some of the most important standards and other key college and career readiness variables. Therefore, states that wish to ensure that students, teachers, schools, and postsecondary programs have the right kinds of evidence necessary to determine college and career readiness will want to consider moving toward a system of assessments. Such an approach can open the door to student profiles that generate more comprehensive portraits of student performance across a wider range of relevant factors and variables and that enable more students to meet standard. This helps reduce the probability of lowering standards due to pressure that results from many students doing poorly on one test or measure.