# Excellence for All: World Class Instructional Systems for Our Schools

## Setting College-Ready Qualification Scores for Chemistry & Coordinated Science

The *Excellence for All* initiative calls for students to demonstrate competence in five subject matter areas – mathematics, English language arts, history, the sciences and the arts -- to be eligible to claim a proficiency-based diploma as early as the close of their sophomore year in high school. Depending on the aligned instructional system their school has chosen, students will take end-of-course examinations in each of these subjects from either the University of Cambridge International General Certificate of Secondary Education (IGCSE) or ACT *QualityCore* systems. The National Center on Education and the Economy's (NCEE) Technical Advisory Committee has established the qualification scores for English and mathematics for both systems, but the states are responsible for setting the qualification scores in the other three subjects.

Ascertaining what qualification scores are associated with future college success is critical to the success of the initiative. At the request of the participating states, NCEE convened a Science Task Force to recommend qualification scores for both IGCSE and *QualityCore* Chemistry and IGCSE Coordinated Science.[1] State education agencies in Arizona, Kentucky and Mississippi and the Capitol Region Education Council in Connecticut were each invited to appoint members to the Task Force. Participants included high school science teachers, community college and university faculty, state education department curriculum professionals, and representatives of the private sector. The Task Force, which met previously in May and June 2012 to learn about the role of the end-of-course examinations in *Excellence for All,* and define qualification scores in biology, was reconvened with a somewhat revised membership (to reflect the change in emphasis from biology to chemistry) to recommend qualification scores for the three science exams noted above. At each of these sessions they were advised by Andrew Shouse, a leader in science education from the University of Washington, and Lloyd Bond, one of the nation's premier measurement experts and a consulting scholar at the Carnegie Foundation for the Advancement of Teaching.

The Task Force was charged with two specific goals. First, each qualification score should be an indicator of readiness to move forward in education, either to an upper division high school science course or to the initial credit-bearing course in a science program of study in an open enrollment college. Second, each measure is also construed as an approximation of basic scientific literacy that would be consistent with a definition of a scientifically literate citizen, signaling that when students pass the exam they are conversant with the fundamental concepts of science that may bear on their life experiences (at the ballot box, at the doctor's office, in the media, etc.). The qualification score is *not* intended to approximate readiness for a career in the sciences, nor for entrance into a science program at a selective post-secondary institution. The Task Force assumed that such college- or career-readiness paths would demand a higher benchmark than "literacy."

The two assessment systems use two different scales to report student proficiency. IGCSE scale scores range from G to A* with numerical equivalents that range from 20 to 100,[2] and

[1] Coordinated Science is a course in biology, chemistry and physics providing schools and students with the flexibility of studying the core concepts of these three subjects in two years.

[2] The numerical range for A* is 90-100, for A it is 80-89, for B it is 70-79, etc.

*QualityCore* scale scores range from 125 to 175. The Task Force had a limited body of student outcome data and correlated measures of college readiness to inform its recommendations. But it did have past examinations to inspect as well as scoring rubrics and exemplary models of student work for constructed response tasks that provided a window into the relationship between the quality of student knowledge and skills and exam scores.

In considering the evidence in hand the Task Force recognized that the qualification scores for these examinations needed to be seen as not just a mark for a single subject, but as one part of a larger qualification framework where students would also have to reach a satisfactory score in (i) another science course (most likely biology),[3] (ii) similar standards in history and the arts, and (iii) meet demanding standards in English and mathematics that had already been set by the NCEE Technical Advisory Committee (TAC). While the TAC placed special value on preventing false positives (i.e., ensuring that students not ready for credit-bearing courses in open enrollment colleges not be told they are), the Task Force placed special value on preventing false negatives (i.e., ensuring that students who could succeed in college were not misclassified and thus have their path to enrolling in credit-bearing community college courses after their sophomore year blocked). They also took the position that while English and math competence is essential for success in almost every college program, the same claim was more difficult to make for science, even as there was wide agreement that literacy in fundamental scientific principles and practices is essential to meet the shared objective that students leave high school ready to lead the life of an educated person. Taken together this line of reasoning led the Task Force to believe that the qualification scores for these three science examinations need not be as stringent as the scores for English and mathematics.

Key considerations for developing qualification score recommendations for these examinations were to be found in the answers to these questions:

- What knowledge, skills and dispositions are priorities for each exam?
- How do these priorities align with what is necessary for success in open-enrollment colleges, in upper division high school science courses and, most importantly, for scientific literacy (referred to among the Task Force as the educated person standard)?
- What indicators of readiness for success at open-enrollment colleges exist, what are their strengths and weaknesses, and how are these reflected in end-of-course exam grades in each of the two programs?

### *ACT QualityCore Chemistry*

The Task Force noted that one half of the ACT *QualityCore* Chemistry exam time and an even larger portion of the raw score (35 out of 59 points) is comprised of multiple-choice questions, with the remaining points and allotted time dedicated to three constructed response questions. Initial data analysis focused on identifying scores that are associated with future college success, using the ACT Science Benchmark, the ACT exam science score that predicts that in their first college science class there is a 50% chance of a student earning at least a B and a 75% chance of a student earning at least a C. With the benchmark set at 24 the Task Force first considered a scale score of 157 on the *QualityCore* Chemistry exam, as it is the scale score with the highest

[3] This criterion does not apply to the coordinated science course, which represents proficiency across three subject areas. It is not assumed that students who take this course would also take an additional science course in high school.

likelihood of translating into a 24 on the ACT even though the task was not to be bounded by college readiness. However, the historical experience with this exam reveals that only 5% of students taking the exam have performed at this level. The Task Force saw this as excessively stringent. Having no interest in creating a program only for elites, they decided to approach the qualification question from another angle.

This second approach involved using the Biology qualification score developed in June 2012 as a benchmark, reasoning that while the Biology and Chemistry scale scores are not tightly related, the qualification scores for each should represent roughly equivalent prospects for future success. The Biology qualification score is associated with an 80% chance of scoring an A or B in the next high school science course and a 25% chance of reaching the ACT science benchmark for college success. Looking to find Chemistry scores that mirrored the Biology scores and leaning on ACT research on *QualityCore* Chemistry, the Task Force found that a 145 Chemistry scale score was associated with an 80% likelihood of scoring an A or B in the next high school science course, and a Chemistry scale score of 148 to 149 was associated with a 25% chance of students meeting the ACT science benchmark. Between these possibilities they favored the result tied to the ACT benchmark more than the result tied to future high school performance on two grounds. First, high school performance data can be distorted by the fact that a more select group of students is usually taking the next science course. Second, linking this decision to college performance rather than high school performance was the more intellectually consistent stance to take.

In the *QualityCore* Chemistry exam, like its sister Biology exam, there are multiple ways for students to reach such scale scores as were considered in this process as points are awarded in a compensatory manner where weaknesses in one set of questions can be offset by strengths in another. A typical way for students to reach scale scores such as those noted above is to answer roughly two-thirds of the multiple-choice items correctly and earn roughly half of the available points on the constructed response tasks. After reviewing several of the multiple choice items and student work and scoring rubrics for several constructed response tasks, Task Force members concluded that student performances that could yield these kinds of scores were consistent with their own views of where an educated person standard ought to be set.

Consideration of these multiple factors led to a consensus recommendation of **148** as the qualification score for the ACT *QualityCore* Chemistry exam. Members of the Task Force were also influenced by the prevalence of technical language and an emphasis on rote memorization in several assessment items that they thought might artificially depress scores, so that some students earning a 148 actually understand chemistry at a level associated with a higher score – something more important in our definition of scientific literacy than knowledge of vocabulary.

***Cambridge IGCSE Chemistry and Coordinated Science***

The Task Force went through the same exercise for both Cambridge IGCSE Chemistry and Cambridge IGCSE Coordinated Science. The IGCSE courses do not have a formal benchmark set for each exam, but there is a general view across the Cambridge community that a score of C, or 60 in numerical terms, indicates that students are ready to attend open enrollment colleges or move on to A-level studies, the equivalent of Advance Placement and International Baccalaureate courses.

These exams have a multiple-choice section, a constructed response section and a coursework component that students fulfill during the course of the school year. The "assessment objectives" are defined explicitly and are built into the assessment rubrics provided for scorers, teachers and students. The Task Force was impressed by the demanding nature of the courses, as well as the alignment of the syllabus, the test instruments and the assessment rubrics.

The Task Force spent considerable time examining students' responses to test questions at the B and C achievement levels (as the "A" and "A*" level responses were clearly going to qualify and "D" level performances were likely to be inadequate so neither received much scrutiny). They also reviewed the examiners' comments on the student work. Discussions largely revolved around the quality of student thinking and writing produced for the test, in comparison to what is required for success in community colleges.

Each IGCSE exam allows a wide range of performance levels to allow the highest performing students to demonstrate their command of the subject matter (so at first glance each appears especially demanding). But this is coupled with rubrics that permit these instruments to recognize multiple levels of student competence. When translated into letter grades a Cambridge "B" is a tougher grade to earn than what we typically think is necessary to earn a B in U.S. schools and likewise for Cs and As.

After studying recent IGCSE Chemistry exams and associated examples of student work the Task Force came to a recommendation of ***60***, or baseline C. They came to the same conclusion with respect to the Coordinated Science exam, recommending a ***60*****,** or baseline C**.** Scores of 60-69 all fall under the definition of a Cambridge C, but the entry-level mark was seen as a solid performance for these purposes. A typical way a student might earn this score on the Chemistry exam is to get 55% of the multiple choice items right and earn 28% of the available constructed response points. On Coordinated Science a similar scenario would be 45% and 20%. While this latter standard may appear quite modest, when one considers the demands of the assessment tasks, it becomes apparent that an overall score of 60 represents a real accomplishment and is consistent with the notion of scientific literacy that animated these deliberations.

These findings were also supported by the prior decision of the Task Force to set the IGCSE Biology qualification score at 60 as these three IGCSE exams were seen as being roughly equivalent in terms of cognitive demand, plus there was no plausible argument that anyone thought had merit that might suggest that one of these science courses had greater value or importance than the other, save for the fact that Coordinated Science represents two years of study, not one. But this difference is accounted for by the fact that this longer course counts for two grades toward earning a new *Excellence for All* proficiency-based diploma.

Over the coming years the *Excellence for All* initiative, at the states' request, will gather additional evidence and suggest refinements to the qualification scores where appropriate. Future recommendations may take into consideration the performance of pilot school students on more advanced ACT and Cambridge courses, as well as how they do on college admissions exams as they advance from grade to grade and eventually on the grades they receive in college.

*January 2013*

Appendix: Task Force Members