Adoption Level
August 4, 2010

RESOLUTION TO ESTABLISH SCORE STANDARDS FOR THE NEW JERSEY BIOLOGY COMPETENCY TEST

WHEREAS, the goal of public schools is to provide all students with a thorough and efficient education as defined by the *Core Curriculum Content Standards* so they may function politically, economically, and socially in our democratic society; and

WHEREAS, the New Jersey Biology Competency Test is provided for assessing student progress toward mastery of essential knowledge and skills; and

WHEREAS, N.J.A.C. 6A:8-4.1(b) requires State Board of Education approval of student performance levels for statewide assessments as specified for district certification; and

WHEREAS, the established levels of proficiency are partially proficient, proficient and advanced proficient; and

WHEREAS, the established scale score standards for the New Jersey Biology Competency Test are 200 for the proficient level and 250 for the advanced proficient level; and

WHEREAS, the total raw scores points possible are 78; 24 of which are aligned to the performance assessment; 54 of which are aligned to selected response items; and

WHEREAS, the corresponding raw cut scores recommended by the Commissioner of Education for the New Jersey Biology Competency Test are as follows:

41 for the proficient level; 58 for the advanced proficient level; now, therefore be it

RESOLVED, the cut scores recommended by the Commissioner of Education for the New Jersey Biology Competency Test shall apply to the 2010 administration and be the basis for reporting scores for future administrations, until such time as the Board shall adopt new performance standards for this assessment.

_____	_____
Willa Spicer	Arcelio Aponte
Acting Commissioner of Education	President
New Jersey State Board of Education	New Jersey State Board of Education

New Jersey Biology Competency Test
Standard Setting Executive Summary

Michael B. Bunch
Measurement Incorporated

The New Jersey Biology Competency Test (Biology Test) is a recent addition to the assessment program for the state. It is the first of a planned series of end-of-course tests for high school subjects and was administered for the first time in the spring of 2008 on a no-fault basis; i.e., there were no sanctions for students who did not perform at a satisfactory level on the test in 2008.

The New Jersey Department of Education (NJDOE) conducted standard setting for the Biology Test. Educators from throughout the state of New Jersey participated in this three-day meeting. Staff of Measurement Incorporated (MI), the contractor for the New Jersey High School Proficiency Assessment (NJ HSPA), facilitated the meeting, along with NJDOE staff.

The main goals of the meeting were to (1) describe and delineate the thresholds of performance, for each subject area, that are indicative of *Partially Proficient, Proficient,* and *Advanced Proficient* performance, and (2) establish recommended cut scores for the test that differentiate *Partially Proficient* from *Proficient* and *Proficient* from *Advanced Proficient* performance (i.e., two cut scores to yield three performance levels). These recommendations are designed to help inform the New Jersey State Board of Education (Board) as it completes its task of establishing performance standards for all statewide assessments.

MI staff met with representatives of the NJDOE and 27 educator-panelists from around the state to recommend performance standards. The panelists, nominated by district superintendents, were chosen specifically to represent the demographics of educators throughout the state. A profile of the 27 panelists is provided. Following an introduction to the process, an overview of the test development process by NJDOE staff, and specific standard-setting techniques by Dr. Michael Bunch (Senior Vice President of MI), the panelists divided into two independent groups (14 in one group, 13 in the other), each led by a facilitator from MI. The purpose of the two-group approach was to test for and neutralize any facilitator effects; i.e., influence on the final outcomes by the person facilitating the group. Panelists in both groups received a thorough orientation and practice exercises to prepare them for their task. MI staff provided additional information to panelists as they proceeded through three rounds of setting standards, discussing decisions, and settling on final recommendations.

In accordance with a plan approved by the New Jersey Technical Advisory Committee (TAC), a group of nationally recognized experts in testing, MI staff employed the bookmark standard-setting procedure. This procedure is the most widely used standard-setting procedure for statewide assessments and is thoroughly documented in the plan submitted to the NJ TAC, as well as in a recent book by Gregory Cizek and Michael Bunch. Dr. Bunch provided the orientation to the bookmark procedure. In this procedure, panelists review all test items in a specially formatted test booklet that places the easiest item on page one, the most difficult item on the final page, and all items in between ordered by difficulty. Panelists place a bookmark at

the point in the test booklet where they believe the probability of a minimally *Proficient* or minimally *Advanced Proficient* student would begin to have less than a two-thirds chance of answering correctly. These page numbers are then mathematically translated into ability estimates, which are translated into raw scores. The average of the panelists' ability estimates is then translated into a raw score which becomes the cut score. The procedure is more fully described in the body of the report.

NJDOE convened biology educators from across the state to review and revise performance level descriptors (PLDs) for *Partially Proficient, Proficient,* and *Advanced Proficient*. These PLDs were to provide the guidance standard-setting panelists would need in order to complete their tasks.

Throughout the standard-setting meeting, an external evaluator observed the process and took notes. The evaluator has prepared a separate report and will submit it for review by the NJDOE and the TAC in response to a TAC recommendation. The evaluator, Dr. Robert Lissitz, is well known in the educational assessment field, having served for many years as the head of the educational measurement department at a major university. He also participated in the review of standard setting for the NJ HSPA Science Test in 2007.

Results and Discussion

Panelists considered each test in three rounds. During Round 1, each panelist placed two bookmarks, one for *Proficient* and one for *Advanced Proficient*. MI staff analyzed the data for Round 1 and led discussions of the results. Panelists reviewed the results of Round 1 and then repeated the process of placing bookmarks in Round 2. After Round 2, MI staff again analyzed the data and presented results to the panelists, along with score distributions showing percentages of students who would be classified as *Partially Proficient*, *Proficient,* and *Advanced Proficient* on the basis of the Round 2 cut scores. After discussion of these results, panelists once again placed two bookmarks in Round 3. These bookmarks defined the final cut scores (averaged over all panelists) to be forwarded to the NJDOE.

Final recommended performance standards are reported in Table ES-1. The table includes the total number of points possible on the test. The Proficient and Advanced Proficient Cut Score column includes both the raw score mean and the mean expressed in terms of a percentage of the Total Points Possible. The final column in Table ES-1 shows the total number of points possible for the test.

Table ES-1
Final Cut Score Recommendations

Proficient Cut Score		**Advanced Proficient Cut Score**		**Total Points Possible**
Raw Score	% Correct	Raw Score	% Correct	
41	53	58	74	78

The recommended *Proficient* cut score is 41 out of 78 points, or 53 percent of the total score possible. This outcome is consistent with the goal of more rigorous standards, reflected in performance level descriptors developed by NJDOE with the assistance of New Jersey educators.

The impact of these cut scores on New Jersey students is summarized in Table ES-2. The final column of the table shows the total percentage of students whose scores would place them in the *Proficient* or *Advanced Proficient* category.

Table ES-2
Percentages of Students Classified at Each Level

Group	Number Tested	% Partially Proficient	% Proficient	% Advanced Proficient	% Proficient or Above
All Students*	58,099	44.9	41.4	13.6	55.1
By Race/Sex					
Asian	4,199	18.4	45.8	35.8	81.6
African American	6,988	65.6	30.8	3.6	34.4
Hispanic	8,037	59.8	35.3	4.9	40.2
White	57,376	25.8	54.5	19.7	74.2
Female	24,916	39.4	46.0	14.6	60.6
Male	23,231	35.7	46.9	17.4	64.3
By Status					
LEP	1,545	90.5	8.9	.6	9.5
Special Education	8,249	78.6	18.9	2.5	21.4

*NB: All analyses are based on "tested" flag (valid and attempted test scores only). No Braille or LP test takers are included due to their small sample size.
The analyses were also conducted using the equating sample which amounts to a minimum of 30% of the total population.
Total students tested in 2010: **104,954**

Measurement Incorporated believes and can document that the standard setting process for the Biology Competency Test was sound, both in conception and execution, representative of the highest standards in contemporary educational measurement, and representative of standards operating among state assessment systems nationwide; that the participants, New Jersey teachers, found it to be so; and that, as New Jersey's assessment vendor, with wide experience implementing assessment programs in other states, MI stands behind the validity of the Biology test standard setting results and the process which produced them, and is prepared to assist the NJDOE in communicating this validity to stakeholders and federal peer reviewers.

2010 New Jersey Biology Competency Test Performance by Subgroup

DFG	n (2010)	2010			
		PP	P	A	P+A
A	5141	74.4	23.6	2.1	25.6
B	4252	56.8	36.6	6.6	43.2
CD	4652	48.1	43.8	8.0	51.9
DE	6621	40.6	49.3	10.0	59.4
FG	6223	32.8	54.1	13.1	67.2
GH	8542	26.1	54.3	19.6	73.9
I	8141	16.7	53.4	29.9	83.3
J	2049	8.5	45.5	46.0	91.5
O	66	92.4	7.6	0.0	7.6
R	373	80.4	19.0	0.5	19.6
V	2159	38.0	42.7	19.3	62.0

Gender	n (2010)	2010			
		PP	P	A	P+A
Male	23231	35.7	46.9	17.4	64.3
Female	24916	39.4	46.0	14.7	60.6

Ethnicity	n (2010)	2010			
		PP	P	A	P+A
Asian	4199	18.4	45.8	35.8	81.6
African American	6988	65.6	30.8	3.6	34.4
Hispanic	8037	59.8	35.3	4.9	40.2
Native American	67	55.2	35.8	9.0	44.8
Hawaiian/Islander	114	28.1	54.4	17.5	71.9
White	27403	25.8	54.5	19.7	74.2

	n (2010)	2010			
		PP	P	A	P+A
SE	8249	78.6	18.9	2.5	21.4
LEP	1545	90.5	8.9	0.6	9.5
REGULAR	48237	37.7	46.4	16.0	62.3
TOTAL	58099	44.9	41.4	13.6	55.1

PP = Partially Proficient
P = Proficient
A = Advanced Proficient
P+A= Proficient and above