

107TH CONGRESS
1ST SESSION

S. 460

To provide for fairness and accuracy in high stakes educational decisions
for students.

IN THE SENATE OF THE UNITED STATES

MARCH 6, 2001

Mr. WELLSTONE introduced the following bill; which was read twice and
referred to the Committee on Health, Education, Labor, and Pensions

A BILL

To provide for fairness and accuracy in high stakes
educational decisions for students.

1 *Be it enacted by the Senate and House of Representa-*
2 *tives of the United States of America in Congress assembled,*

3 **SECTION 1. FAIRNESS AND ACCURACY IN HIGH STAKES**

4 **EDUCATIONAL DECISIONS FOR STUDENTS.**

5 (a) FINDINGS.—Congress makes the following find-
6 ings:

7 (1) The use of large-scale achievement tests in
8 education has grown significantly in recent years.
9 States and local school districts have increasingly
10 used these tests in such contexts as raising student

1 academic standards to make high-stakes decisions
2 with important consequences for individual students,
3 such as tracking (assigning students to schools, pro-
4 grams, or classes based on achievement level), pro-
5 motion of students to the next grade, and gradua-
6 tion of students from secondary school.

7 (2) The serious and often adverse consequences
8 resulting from the sole reliance on large-scale tests
9 have increasingly resulted in questions and signifi-
10 cant concerns by students, parents, teachers, and
11 school administrators about how to ensure that such
12 tests are used appropriately and in a manner that
13 is fair.

14 (3) In 1997, Congress directed the National
15 Academy of Sciences to “conduct a study and make
16 written recommendations on appropriate methods,
17 practices, and safeguards to ensure that, among
18 other things, . . . existing and new tests that are used
19 to assess student performance are not used in a dis-
20 criminatory manner or inappropriately for student
21 promotion, tracking, or graduation.”.

22 (4) In 1999, the National Academy of Sciences,
23 through its National Research Council, completed its
24 study and issued a report entitled “High Stakes:
25 Testing for Tracking, Promotion and Graduation”.

1 Guided by principles of measurement validity, attri-
2 bution of cause, and effectiveness of treatment, the
3 National Research Council made key findings for ap-
4 propriate test use in educational settings, including
5 the following:

6 (A) When tests are used in ways that meet
7 relevant psychometric, legal, and educational
8 standards, students' scores provide important
9 information, that combined with information
10 from other sources, can lead to decisions that
11 promote student learning and equality of oppor-
12 tunity.

13 (B) Tests are not perfect. Test questions
14 are a sample of possible questions that could be
15 asked in a given area. Moreover, a test score is
16 not an exact measure of a student's knowledge
17 or skills.

18 (C) To the extent that all students are ex-
19 pected to meet world-class standards, there is a
20 need to provide world-class curricula and in-
21 struction to all students. However, in most of
22 the Nation, much needs to be done before a
23 world-class curriculum and world-class instruc-
24 tion will be in place. At present, curriculum
25 does not usually place sufficient emphasis on

1 student understanding and application of con-
2 cepts, as opposed to memorization and skill
3 mastery. In addition, instruction in core sub-
4 jects typically has been and remains highly
5 stratified. What teachers teach and what stu-
6 dents learn vary widely by track, with those in
7 lower tracks receiving far less than a world-
8 class curriculum.

9 (D) It is a mistake to begin educational re-
10 form by introducing tests with high stakes for
11 individual students. If tests are to be used for
12 high stakes decisions about individual mastery,
13 such use should follow implementation of
14 changes in teaching and curriculum that ensure
15 that students have been taught the knowledge
16 and skills on which the students will be tested.

17 (E) Problems of test validity are greatest
18 among young children, and there is a greater
19 risk of error when such tests are employed to
20 make high stakes decisions about children who
21 are less than 8 years old or below grade 3, or
22 about their schools. However, well-designed as-
23 sessments may be useful in monitoring trends
24 in the educational development of populations
25 of students who have reached age 5.

1 (5) The National Research Council made the
2 following recommendations:

3 (A) If parents, educators, public officials,
4 and others who share responsibility for edu-
5 cational outcomes are to discharge their respon-
6 sibility effectively, they should have access to
7 information about the nature and interpretation
8 of tests and test scores. Such information
9 should be made available to the public and
10 should be incorporated into teacher education
11 and into educational programs for principals,
12 administrators, public officials, and others.

13 (B) A test may appropriately be used to
14 lead curricular reform, but it should not also be
15 used to make high-stakes decisions about indi-
16 vidual students until test users can show that
17 the test measures what they have been taught.

18 (C) High-stakes decisions such as tracking,
19 promotion, and graduation should not automati-
20 cally be made on the basis of a single test score
21 but should be buttressed by other relevant in-
22 formation about the student's knowledge and
23 skill, such as grades, teacher recommendations,
24 and extenuating circumstances.

1 (D) In general, large-scale assessments
 2 should not be used to make high-stakes deci-
 3 sions about students who are less than 8 years
 4 old or enrolled below grade 3.

5 (E) High-stakes testing programs should
 6 routinely include a well-designed evaluation
 7 component. Policymakers should monitor both
 8 the intended and unintended consequences of
 9 high-stake assessments on all students and on
 10 significant subgroups of students, including mi-
 11 norities, English-language learners, and stu-
 12 dents with disabilities.

13 (6) These principles and findings of the Na-
 14 tional Academy of Sciences are supported in signifi-
 15 cant measure by the Standards for Educational and
 16 Psychological Testing, adopted and approved in De-
 17 cember of 1999, by the leading experts and profes-
 18 sional organizations on testing, including the Amer-
 19 ican Educational Research Association, American
 20 Psychological Association, and the National Council
 21 on Measurement in Education.

22 (b) TEST PERFORMANCE.—If performance on a sin-
 23 gle large-scale test is considered as part of any decision
 24 about the retention, graduation, tracking, or within-class
 25 ability grouping of an individual student by a State edu-

1 cational agency or local educational agency that receives
 2 funds under the Elementary and Secondary Education Act
 3 of 1965, such test performance shall not be the sole cri-
 4 terion in such decision and may be considered in making
 5 such decision only if—

6 (1) the test, including any cut score or perform-
 7 ance standard set or established for use on the test,
 8 meets professional standards of validity and reli-
 9 ability for the purpose for which the test's results
 10 are being used;

11 (2) the test allows its users to make score inter-
 12 pretations in relation to a functional performance
 13 level, as distinguished from those interpretations
 14 that are made in relation to the performance of oth-
 15 ers;

16 (3) the test is based on State or local content
 17 and performance standards and is aligned with the
 18 curriculum and classroom instruction;

19 (4) the test follows implementation of changes
 20 in teaching and curriculum that ensure that stu-
 21 dents have been taught the knowledge and skills on
 22 which the students will be tested;

23 (5) multiple measures of student achievement,
 24 including grades and evaluations by teachers, are
 25 utilized to ensure that scores from the test are never

1 the only source of information used, nor the sole cri-
2 terion used, in making a high-stakes decision about
3 an individual student;

4 (6) students tested have been provided multiple
5 opportunities to demonstrate proficiency in the aca-
6 demic subject covered by the test;

7 (7) the test is administered in accordance with
8 the written guidance from the test developer or pub-
9 lisher;

10 (8) the State educational agency or local edu-
11 cational agency involved has evidence that the test
12 is of adequate technical quality for each purpose for
13 which the test is used;

14 (9) the State educational agency or local edu-
15 cational agency provides appropriate accommoda-
16 tions and alternate assessments for students with
17 disabilities that provide the students with a valid op-
18 portunity to show what the students know and can
19 do;

20 (10) the State educational agency or local edu-
21 cational agency provides appropriate accommoda-
22 tions and alternative assessments for students with
23 limited English proficiency (if the agency involved
24 determines that the students have not achieved suffi-
25 cient English proficiency to ensure that the test will

1 validly and reliably measure the subject matter
2 knowledge and skills of the students), including—

3 (A) the use of a test other than an
4 English-only test;

5 (B) the use of alternate assessments (con-
6 sisting of psychometrically equivalent tests in
7 the students' native language) in order to pro-
8 vide such students with a valid and reliable op-
9 portunity to demonstrate what the students
10 know and can do; and

11 (C) in a case in which the Secretary of
12 Education determines that more than 5 percent
13 of the students enrolled in kindergarten
14 through grade 12 in a State are members of a
15 single language minority group and are limited
16 English proficient—

17 (i) the assessment of the students in
18 that group using tests developed in the
19 language of that group, if the State or
20 local educational agency determines that
21 such tests are more likely than English-
22 only tests to yield accurate and reliable in-
23 formation regarding what those students
24 know and can do; or

1 (ii) if the language of the group is
 2 oral or unwritten or, in the case of Alaska
 3 Natives and other American Indians, if the
 4 predominant language of the group is his-
 5 torically unwritten, the furnishing of oral
 6 instructions, assistance, and other nec-
 7 essary information to such students relat-
 8 ing to the English-only test; and

9 (11) the test is not used for a decision about
 10 promotion or placement in special education for a
 11 child below the age of 8 or third grade.

12 (c) EVALUATIONS.—

13 (1) STATE EDUCATIONAL AGENCIES.—Each
 14 State educational agency that receives funds under
 15 the Elementary and Secondary Education Act of
 16 1965 and uses a large-scale test as part of a high
 17 stakes decision described in subsection (b), shall pe-
 18 riodically conduct a comprehensive evaluation of the
 19 impact of high stakes decisions on students' edu-
 20 cation and educational outcomes, with particular
 21 consideration given to the impact on individual stu-
 22 dents and subgroups of students disaggregated by
 23 socioeconomic status, race, ethnicity, limited English
 24 proficiency, disability, and gender. The State edu-
 25 cational agency shall make the results of the evalua-

1 tion available to the public and shall provide clear
2 and comprehensible information about the nature,
3 use, and interpretation of the test and the scores the
4 test generate.

5 (2) LOCAL EDUCATIONAL AGENCY.—Each local
6 educational agency that receives funds under the El-
7 ementary and Secondary Education Act of 1965,
8 uses a large-scale test as part of a high stakes deci-
9 sion described in subsection (b), and is located in a
10 State that does not conduct an evaluation under
11 paragraph (1), shall periodically conduct a com-
12 prehensive evaluation of the impact of high stakes
13 decisions on students' education and educational
14 outcomes, with particular consideration given to the
15 impact on individual students and subgroups of stu-
16 dents disaggregated by socioeconomic status, race,
17 ethnicity, limited English proficiency, disability, and
18 gender. The local educational agency shall make the
19 results of the evaluation available to the public and
20 shall provide clear and comprehensible information
21 about the nature, use, and interpretation of the test
22 and the scores the test generate.

23 (3) DEPARTMENT OF EDUCATION.—The Sec-
24 retary shall—

1 (A) conduct an evaluation similar to the
 2 evaluation described in paragraph (1) among a
 3 representative sample of States and local edu-
 4 cational agencies;

5 (B) report the results of such evaluation to
 6 Congress; and

7 (C) make the results of the evaluation
 8 available to the public.

9 (d) DEFINITIONS.—In this section:

10 (1) IN GENERAL.—The terms used in this sec-
 11 tion have the meanings given the terms in section
 12 14101 of the Elementary and Secondary Education
 13 Act of 1965.

14 (2) LARGE-SCALE TEST.—The term “large-scale
 15 test” means a test that is administered and scored
 16 under conditions uniform to all students so that the
 17 test scores are comparable across individuals.

18 (3) SOLE CRITERION.—The term “sole cri-
 19 terion” means the only one standard (such as a test
 20 score) used to make a judgment or a decision, in-
 21 cluding a step-wise decisionmaking procedure where
 22 students must reach or exceed one criterion (such as
 23 a cut score of a test) independent of or before other
 24 criteria can be considered.

