

EXHIBIT NN

Case No. 1:14-cv-00857-TSC-DAR



AMERICAN
PSYCHOLOGICAL
ASSOCIATION

April 7, 2000

Ms. Jill Morningstar, Legislative Assistant
U. S. Senator Paul Wellstone
136 Hart Senate Office Building
Washington, DC 20510

Dear Jill:

We very much appreciate your seeking the endorsement of the American Psychological Association (APA) for the bill introduced by Senator Wellstone this week on Fairness and Accuracy in Student Testing. We wanted to get back to you quickly with an interim response about the Association's views.

As we mentioned on the phone earlier this week, APA is not in a position to endorse the bill at this time. APA policy on this issue is guided by *The Standards for Educational and Psychological Testing*, the 1999 version of which was developed over a six-year period with the American Educational Research Association and the National Council on Measurement in Education. While we are drafting a more comprehensive response for the Senator, we wanted to let you know the general areas of our concerns. First, we believe that the most fair and appropriate way to approach the problems Senator Wellstone seeks to address is to investigate high stakes decision-making in educational settings, how such decisions are informed, and the impact on educational outcomes. The critical issue is not tests, per se (provided they are valid, reliable instruments) but rather the instances wherein they and other measures of accountability (i.e. grades), are used inappropriately in making these decisions. Second, the bill appears to misstate the intent of both the *Standards* and the National Academy of Sciences (NAS) report by holding tests for tracking, retention and graduation to the same standard. The NAS study and the *Standards* treat those uses differently. Both documents treat graduation tests as certification tests, and do not prohibit assigning them determinative weight, but offer recommendations for implementing them fairly.

We understand the Senator wants to offer this amendment to the Elementary and Secondary Education Act soon and would appreciate APA's assistance in doing so. Regrettably, APA has not yet adopted a policy that goes beyond the *Standards* to address the issues of how high stakes decision-making in schools affects education and educational outcomes more broadly, particularly for certain groups of children. Accordingly, it would be impossible for us to offer you language that the Association could endorse within your timeframe. However, we would welcome the opportunity to work with you on your Plan B: an amendment to authorize specific evaluation research that would provide additional needed information about the impact of high stakes decision-making on the educational opportunities of different populations of children, and on the educational system more generally. The NAS report and other sources point to significant gaps in the data. We hope to provide specific draft language to you late next week.

Again, we appreciate your willingness to work with us. Senator Wellstone has been a champion of many of the issues most important to psychologists, and we are grateful to him.

Sincerely,

Ellen G. Garrison, Ph.D.
Director of Public Interest Policy

Patricia C. Kobor
Director of Science Policy

750 First Street, NE
Washington, DC 20002-4242
(202) 336-5500
(202) 336-6123 TDD

Web: www.apa.org



Please Recycle

106TH CONGRESS
2D SESSION

S. _____

IN THE SENATE OF THE UNITED STATES

Mr. WELLSTONE introduced the following bill; which was read twice and referred to the Committee on _____

A BILL

To provide for fairness and accuracy in student testing.

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 STUDENT TEST-**
4 **ING.**

5 (a) FINDINGS.—Congress finds the following:

6 (1) The use of large-scale achievement tests in
7 education has grown significantly in recent years.
8 States and local school districts have increasingly
9 used these tests in such contexts as raising student
10 academic standards to make high-stakes decisions
11 with important consequences for individual students,

1 such as tracking (assigning students to schools, pro-
 2 grams, or classes based on achievement level), pro-
 3 motion of students to the next grade, and gradua-
 4 tion of students from secondary school.

5 (2) The serious and often adverse consequences
 6 resulting from the sole or determinative reliance on
 7 large-scale tests have increasingly resulted in ques-
 8 tions and significant concerns by students, parents,
 9 teachers, and school administrators about how to en-
 10 sure that such tests are used appropriately and in
 11 a manner that is fair.

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

20 (4) In 1999, the National Academy of Sciences,
 21 through its National Research Council, completed its
 22 study and issued a report entitled 'High Stakes:
 23 Testing for Tracking, Promotion and Graduation'.
 24 Guided by principles of measurement validity, attri-
 25 bution of cause, and effectiveness of treatment, the

1 National Research Council made key findings for ap- ✓
 2 appropriate test use in educational settings, including
 3 the following:

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

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

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

1 mastery. In addition, instruction in core sub-
 2 jects typically has been and remains highly
 3 stratified. What teachers teach and what stu-
 4 dents learn vary widely by track, with those in
 5 lower tracks receiving far less than a world-
 6 class curriculum.

7 (D) Problems of test validity are greatest
 8 among young children, and there is a greater
 9 risk of error when such tests are employed to
 10 make significant decisions about children who
 11 are less than 8 years old or below grade 3, or
 12 about their schools. However, well-designed as-
 13 sessments may be useful in monitoring trends
 14 in the educational development of populations
 15 of students who have reached age 5.

16 (5) The National Research Council made the
 17 following recommendations:

18 (A) If parents, educators, public officials,
 19 and others who share responsibility for edu-
 20 cational outcomes are to discharge their respon-
 21 sibility effectively, they should have access to
 22 information about the nature and interpretation
 23 of tests and test scores. Such information
 24 should be made available to the public and
 25 should be incorporated into teacher education

1 and into educational programs for principals,
2 administrators, public officials, and others.

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

8 (C) High-stakes decisions such as tracking,
9 promotion, and graduation should not automati-
10 cally be made on the basis of a single test score
11 but should be buttressed by other relevant in-
12 formation about the student's knowledge and
13 skill, such as grades, teacher recommendations,
14 and extenuating circumstances.

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

19 (E) High-stakes testing programs should
20 routinely include a well-designed evaluation
21 component. Policymakers should monitor both
22 the intended and unintended consequences of
23 high-stake assessments on all students and on
24 significant subgroups of students, including mi-

Handwritten notes in a box: *rewrite report disaggregated*

Handwritten checkmark

Handwritten checkmark

1 norities, English-language learners, and stu-
2 dents with disabilities.

3 (6) These principles and findings of the Na-
4 tional Academy of Sciences are supported in signifi-
5 cant measure by the Standards for Educational and
6 Psychological Testing, adopted and approved in De-
7 cember of 1999, by the leading experts and profes-
8 sional organizations on testing, including the Amer-
9 ican Educational Research Association, American
10 Psychological Association, and the National Council
11 on Measurement in Education.

12 (b) TEST PERFORMANCE.—If performance on a
13 standardized test is considered as part of any decision
14 about the retention, graduation, tracking, or within-class
15 ability grouping of an individual student by a State edu-
16 cational agency or local educational agency that receives
17 funds under the Elementary and Secondary Education Act
18 of 1965, such test performance shall not be the sole deter-
19 minant in such decision and may be considered in making
20 such decision only if—

21 (1) the test meets professional standards of va-
22 lidity and reliability for the purpose for which the
23 test's results are being used, including the validity
24 and reliability of any cut score or performance
25 standard set or established for use on the test;



1 (2) the test allows its users to make score inter-
2 pretations in relation to a functional performance
3 level, as distinguished from those interpretations
4 that are made in relation to the performance of oth-
5 ers, is based on State or local content and perform-
6 ance standards, and is aligned with the curriculum
7 and classroom instruction;

8 (3) multiple measures of student achievement
9 are utilized, including grades and evaluations by
10 teachers, so that scores from large-scale assessments
11 are never the only source of information used nor
12 assigned determinative weight in making a high-
13 stakes decision about an individual student;

14 (4) students tested have been provided multiple
15 opportunities to demonstrate proficiency in the sub-
16 ject matter covered by the test;

17 (5) the test is administered in accordance with
18 the written guidance from the test developer or pub-
19 lisher;

20 (6) the State educational agency or local edu-
21 cational agency has evidence that the test is of ade-
22 quate technical quality for each purpose for which
23 the test is used;

24 (7) the State educational agency or local edu-
25 cational agency provides appropriate accommoda-

1 tions and alternate assessments for students with
2 disabilities that provide the students with a valid op-
3 portunity to show what they know and can do;

4 (8) the State educational agency or local edu-
5 cational agency provides appropriate accommoda-
6 tions for students with limited English proficiency,
7 including—

8 (A) if such a student is tested in English,
9 the student received academic instruction pri-
10 marily in English for at least 3 years prior to
11 the test, or if the student received instruction in
12 English for more than such 3 years, the local
13 educational agency determines that the student
14 has achieved sufficient English proficiency to
15 ensure that the test will accurately measure the
16 student's subject matter knowledge and skills;

17 (B) in the case of students with limited
18 English proficiency who have not been taught
19 primarily in English for 3 years prior to the
20 test, such students are assessed, to the greatest
21 extent practicable, in the language and form
22 most likely to yield accurate and reliable infor-
23 mation about what those students know and
24 can do; and

O:\KOS\KOS00.280

S.L.C.

1 (C) in the case of Spanish-speaking stu-
 2 dents with limited English proficiency, such
 3 students are assessed using tests developed and
 4 written in Spanish, if Spanish language tests
 5 are more likely than English language tests to
 6 yield accurate and reliable information on what
 7 those students know and can do; and

7
~~written~~
 written
 normed in
 Spanish?
 OTIS
 for gifted kids?

8 (9) the test is not used for a decision about
 9 promotion or placement in special education for a
 10 child below the age of 8 or grade 3.

11 (c) EVALUATIONS.—

12 (1) STATE EDUCATIONAL AGENCIES.—Each
 13 State educational agency that receives funds under
 14 the Elementary and Secondary Education Act of
 15 1965 and uses a standardized test as part of a high
 16 stakes decision described in subsection (b), shall con-
 17 duct a comprehensive evaluation of the impact of the
 18 test's use on students' education and educational
 19 outcomes, with particular consideration given to the
 20 impact on individual students and subgroups of stu-
 21 dents disaggregated by socioeconomic status, race,
 22 ethnicity, limited English proficiency, disability, and
 23 gender. The State educational agency shall make the
 24 results of the evaluation available to the public and
 25 shall provide clear and comprehensible information

1 about the nature, use, and interpretation of the test
2 and the scores the test generate.

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

21 (3) DEPARTMENT OF EDUCATION.—The Sec-
22 retary shall—

23 (A) conduct an evaluation similar to the
24 evaluation described in paragraph (1) among a

1 representative sample of States and local edu-
2 cational agencies;

3 (B) report the results of such evaluation to
4 Congress; and

5 (C) make the results of the evaluation
6 available to the public.

7 (d) DEFINITION OF STANDARDIZES TEST.—In this
8 section the term “standardized test” means a test that
9 is administered and scored under conditions uniform to
10 all students so that the test scores are comparable across
11 individuals: