EXHIBIT NN

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

Pat Kober  
Director of Science Policy

Web: www.apa.org  

American Psychological Association
A BILL

To provide for fairness and accuracy in student testing.

Be it enacted by the Senate and House of Representatives of the United States of America in Congress assembled,

SECTION 1. FAIRNESS AND ACCURACY IN STUDENT TESTING.

(a) FINDINGS.—Congress finds the following:

(1) The use of large-scale achievement tests in education has grown significantly in recent years. States and local school districts have increasingly used these tests in such contexts as raising student academic standards to make high-stakes decisions with important consequences for individual students,
such as tracking (assigning students to schools, programs, or classes based on achievement level), promotion of students to the next grade, and graduation of students from secondary school.

(2) The serious and often adverse consequences resulting from the sole or determinative reliance on large-scale tests have increasingly resulted in questions and significant concerns by students, parents, teachers, and school administrators about how to ensure that such tests are used appropriately and in a manner that is fair.

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

(4) In 1999, the National Academy of Sciences, through its National Research Council, completed its study and issued a report entitled 'High Stakes: Testing for Tracking, Promotion and Graduation'. Guided by principles of measurement validity, attribution of cause, and effectiveness of treatment, the
National Research Council made key findings for appropriate test use in educational settings, including the following:

(A) When tests are used in ways that meet relevant psychometric, legal, and educational standards, students' scores provide important information, that combined with information from other sources, can lead to decisions that promote student learning and equality of opportunity.

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

(C) To the extent that all students are expected to meet world-class standards, there is a need to provide world-class curricula and instruction to all students. However, in most of the Nation, much needs to be done before a world-class curriculum and world-class instruction will be in place. At present, curriculum does not usually place sufficient emphasis on student understanding and application of concepts, as opposed to memorization and skill.
mastery. In addition, instruction in core subjects typically has been and remains highly stratified. What teachers teach and what students learn vary widely by track, with those in lower tracks receiving far less than a world-class curriculum.

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

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

(A) If parents, educators, public officials, and others who share responsibility for educational outcomes are to discharge their responsibility effectively, they should have access to information about the nature and interpretation of tests and test scores. Such information should be made available to the public and should be incorporated into teacher education
and into educational programs for principals, administrators, public officials, and others.

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

(C) High-stakes decisions such as tracking, promotion, and graduation should not automatically be made on the basis of a single test score but should be buttressed by other relevant information about the student's knowledge and skill, such as grades, teacher recommendations, and extenuating circumstances.

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

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

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

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

(1) the test meets professional standards of va-
"lidity and reliability for the purpose for which the
test's results are being used, including the validity
and reliability of any cut score or performance
standard set or established for use on the test;
(2) the test allows its users to make score interpretations in relation to a functional performance level, as distinguished from those interpretations that are made in relation to the performance of others, is based on State or local content and performance standards, and is aligned with the curriculum and classroom instruction;

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

(4) students tested have been provided multiple opportunities to demonstrate proficiency in the subject matter covered by the test;

(5) the test is administered in accordance with the written guidance from the test developer or publisher;

(6) the State educational agency or local educational agency has evidence that the test is of adequate technical quality for each purpose for which the test is used;

(7) the State educational agency or local educational agency provides appropriate accommoda-
tions and alternate assessments for students with
disabilities that provide the students with a valid op-
portunity to show what they know and can do;

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

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

(B) in the case of students with limited
English proficiency who have not been taught
primarily in English for 3 years prior to the
test, such students are assessed, to the greatest
extent practicable, in the language and form
most likely to yield accurate and reliable infor-
mation about what those students know and
can do; and
(C) in the case of Spanish-speaking students with limited English proficiency, such students are assessed using tests developed and written in Spanish, if Spanish language tests are more likely than English language tests to yield accurate and reliable information on what those students know and can do; and

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

(c) EVALUATIONS.—

(1) STATE EDUCATIONAL AGENCIES.—Each State educational agency that receives funds under the Elementary and Secondary Education Act of 1965 and uses a standardized test as part of a high stakes decision described in subsection (b), shall conduct a comprehensive evaluation of the impact of the test's use on students' education and educational outcomes, with particular consideration given to the impact on individual students and subgroups of students disaggregated by socioeconomic status, race, ethnicity, limited English proficiency, disability, and gender. The State educational agency shall make the results of the evaluation available to the public and shall provide clear and comprehensible information
about the nature, use, and interpretation of the test
and the scores the test generate.

(2) LOCAL EDUCATIONAL AGENCY.—Each local
educational agency that receives funds under the El-
ementary and Secondary Education Act of 1965,
uses a standardized test as part of a high stakes de-
cision described in subsection (b), and is located in
a State that does not conduct an evaluation under
paragraph (1), shall conduct a comprehensive eval-
uation of the impact of the test’s use on students’
education and educational outcomes, with particular
consideration given to the impact on individual stu-
dents and subgroups of students disaggregated by
socioeconomic status, race, ethnicity, limited English
proficiency, disability, and gender. The local edu-
cational agency shall make the results of the evalu-
ation available to the public and shall provide clear
and comprehensible information about the nature,
use, and interpretation of the test and the scores the
test generate.

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

(A) conduct an evaluation similar to the
evaluation described in paragraph (1) among a
representative sample of States and local educational agencies;

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

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

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