

Requirements for Measuring Adequate Yearly Progress¹

CRESST Co-Director Robert L. Linn



Robert L. Linn

Robert L. Linn is Distinguished Professor of Education at the University of Colorado at Boulder and co-director of the National Center for Research on Evaluation, Standards, and Student Testing. He is the current (2002-2003) president of the American Educational Research Association and former chairperson of the National Research Council's Board on Testing and Assessment (1997-2000).

Final regulations were recently published for the No Child Left Behind Act of 2001 ([NCLB] 2002) addressing important issues of adequate yearly progress (AYP). The following federal requirements are now certain:

- Students in Grades 3 through 8 must make adequate yearly progress in mathematics and reading/language arts.
- Subgroups must make adequate yearly progress, including students who are economically disadvantaged, students from major racial and ethnic groups, students with disabilities, and students with limited English proficiency.
- States must specify a timeline so that all students will achieve at the proficient level or higher by the 2013-2014 school year.
- The AYP starting point must be defined based on 2001-2002 state test results, using the higher of "(1) the percentage in the State of proficient students in the lowest achieving subgroup of students . . . or (2) the percentage of proficient students in the school that represents 20 percent of the State's total enrollment among all schools . . ." (NCLB, 2002). The regulations prescribe the method for identifying this school.

Based on comments by U.S. Secretary of Education Rod Paige (2002; see also Olson, 2002), states and school districts were hopeful that more flexibility would be

provided for how states could define and report adequate yearly progress. However, the final AYP-relevant regulations remain very similar to the draft regulations (Department of Education . . . Proposed Rules, 2002; Department of Education . . . Rules and Regulations, 2002), leaving most states with little choice but to conform their existing accountability programs to the new, often very different federal requirements.

In this brief, we review some basic AYP issues and expand our NCLB accountability research. See the spring 2002 *CRESST Line* (Baker, Linn, & Herman, 2002) for NCLB and the purposes of assessment, and the fall 2002 *CRESST Line* (Linn, Baker, & Herman, 2002) for our recommendations on minimum group sizes to meet the new NCLB requirements.

A Simple AYP Example

Figure 1 provides a representation of adequate yearly progress under NCLB. The starting points established from the 2001-2002 state assessments in mathematics and reading/language arts together with the 2013-2014 target of 100% proficient determine the annual measurable objectives. If the starting points were, say, 52% proficient or above in reading/language arts and 40% proficient or above in mathematics (Figure 1), then the annual measurable objective would increase by 4%² a year for reading/language arts and 5%³ for mathematics. Intermediate checkpoints between 2001-2002 and 2013-2014 are to be established, with the first checkpoint to occur within, at most, 2 years and subsequent checkpoints to occur no more than 3 years apart.

The checkpoint targets for the example above would be a minimum of 60%⁴ proficient or above in reading/language arts and 50%⁵ proficient or above in mathematics in 2003-2004. Subsequent checkpoints in 2006-2007, 2009-2010, 2012-2013, and 2013-2014 would have targets for percent proficient or above of 72%, 84%, 96%, and 100%, respectively, for reading/language arts and 65%, 80%, 95%, and 100%, respectively, for mathematics. The model is straightforward and the requirements are rigorous.⁶

School and District Requirements to Meet AYP: Key Points and Challenges

For a school or district to meet the AYP target at any checkpoint, all students and each student subgroup must meet or exceed the state’s annual measurable objective. Ninety-five percent of each subgroup’s students must take the assessments on which the state AYP is based.

A subgroup exception is allowed if (a) the number of students in that subgroup scoring below proficient is reduced by at least 10% from the prior year, and (b) the subgroup made progress on one or more other state indicators.⁷

The exception has sometimes been referred to as the “safe harbor” provision. Let’s say that a school’s goal for a checkpoint year is 50% proficient or above for the school and each subgroup. The school and all subgroups make this goal, except the special education subgroup. Only 30% of special education students score proficient or above this year. The school apparently fails to meet its goal. Under the safe harbor provision, the school could make its goal

Example of the Safe Harbor Provision	
Year 2002-2003	78% not proficient or above
Year 2003-2004	70% not proficient or above
$(78 - 70) / 78 =$	10.2% reduction in students not proficient or above

if the percent of special education students below the proficient level has “decreased” from 78% in the previous year to 70% this year. If the subgroup also improves on another indicator, this subgroup then meets the safe harbor exception rule.

These requirements are laudable for the attention they bring to students who have lagged behind in the past. However, the requirements set extremely stringent standards that are highly challenging and possibly quite unrealistic. As Kane (2002) pointed out, the simple process of adding subgroups reduces the probability that a school will meet its goal, even if all other factors remain equal.

NCLB requires subgroup reporting for those groups

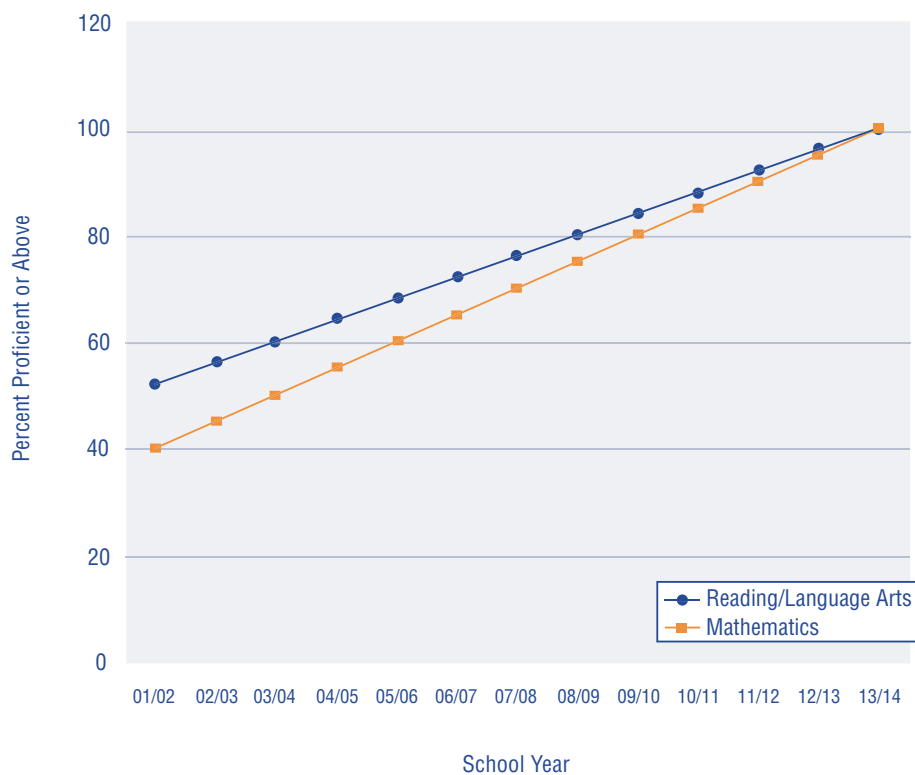


Figure 1. Adequate yearly progress in two subject areas.

with significant numbers of students to provide valid results. However, NCLB does not define what a minimum size should be. Since statistical reliability increases as the number of students within a category increases, states will need to establish a minimum number for any category. See Linn et al. (2002) for CRESST’s recommendation. We expect substantial variation in the minimum subgroup size set by different states.

NCLB allows states to average scores across grade levels but not between topics. Using a minimum subgroup size of 25 students, Table 1 demonstrates that averaging across grade levels for this real California school would add two subgroups, Asian and White, and be just short of enough students to add Pacific Islanders. In addition to adding more subgroups, the number of students in each subgroup increases, thereby improving reliability. For example, aggregating Grades 3 through 5 for African American students increases the total *N* to 120 students, compared to 34 students for third grade and 40 students for fifth grade. Thus, the case for aggregating across grades is strong.

Variability in Stringency of Progress Requirements for Different Schools

Although the measurable objective for schools is uniform at any checkpoint, the stringency of the target varies

Table 1
 Number of Students in Subgroups:
 Elementary School in California
 2001-2002 School Year

Grade	Asian	White	African American	Pacific Islander	Filipino	Hispanic	Male	Female
3	8	15	34	8	4	188	131	126
4	10	14	46	6	4	182	153	109
5	13	15	40	9	7	175	133	128
Total	31	44	120	23	15	545	417	363

greatly by school. Suppose that the 2001-2002 state starting point for mathematics is 40% proficient or above based on the school that is at the 20th percentile as allowed by NCLB. Assuming that schools are reasonably similar in size, that would imply that more than 40% of the students achieved at the proficient level or higher in roughly 4 out of 5 schools statewide. The uniform statewide AYP objective would be 50% proficient or above in mathematics in 2003-2004 for all schools.

For School A, which in 2001-2002 had only 30% of its students scoring at the proficient level or higher, an increase in percent proficient or above of 20% would be required over the 2-year interval to 2003-2004 (Figure 2). The increase for School B, which started out with 45% proficient or above, would be only 5% to meet the AYP target in 2003-2004. School C, where 75% of the students scored at the proficient level or higher in 2001-2002, could actually experience a reduction in the percent proficient or above of 20% and still be above the AYP objective in 2003-2004. Thus, NCLB has stringent standards for increases in the percent proficient or above for schools that start off with a low percentage of students meeting the standard, but is quite permissive for initially high-scoring schools, at least in the early years. School C, for example, could coast along with the same 75% proficient or above for several years. Not until 2009-2010, when the target would have reached 80%, would School C have to start

showing progress to avoid being placed in the “needs improvement” category.

Summary

The NCLB adequate yearly progress requirements represent enormous, if not overwhelming, challenges to schools, districts, and states. Lower performing schools, oftentimes in urban areas, have greater distances to go than higher achieving schools and also tend to have a larger number of subgroups, thereby increasing the probability that at least one subgroup will not make adequate yearly progress. Additionally, lower achieving schools must begin to make immediate progress, whereas higher performing schools will not have consequences for at least several years in the future. School-level results between states may vary significantly based on the mini-

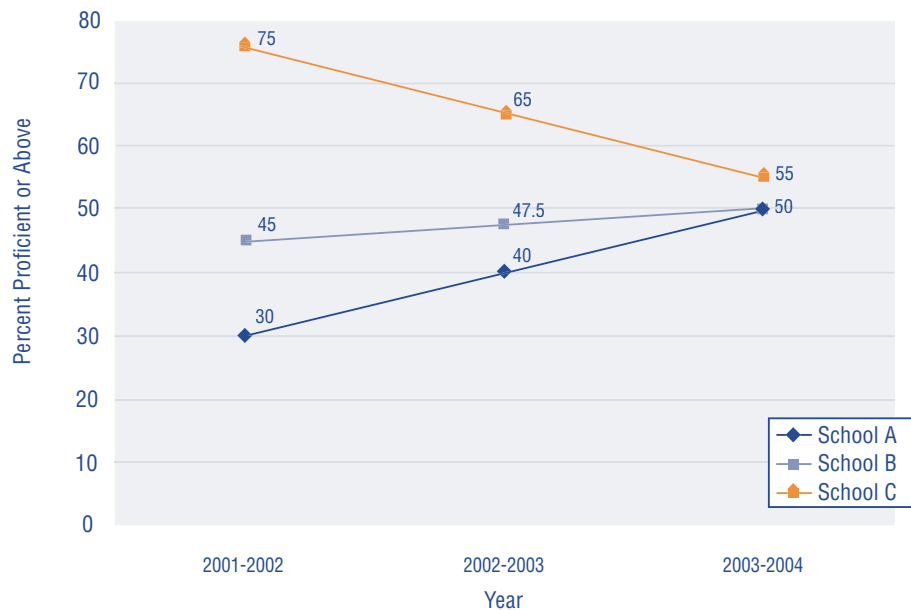


Figure 2. Adequate yearly progress for three schools.

mum subgroup size. Other challenges include the major differences between state accountability systems and the very rapid turnover of state tests. CRESST is committed to continuing our research into these topics and providing advice to policymakers and state and local education agencies. Our winter 2003 *CRESST Line* issue will further explore the differential challenges between low- and high-achieving schools, as well as other important NCLB accountability topics.

References

- Baker, E. L., Linn, R. L., & Herman, J. L. (2002, Spring). From the directors. *The CRESST Line*, 1-3. (Newsletter of the National Center for Research on Evaluation, Standards, and Student Testing [CRESST], University of California, Los Angeles)
- Department of Education. Title I—Improving the academic achievement of the disadvantaged: Proposed rules, 67 Fed. Reg. 50,986 (Aug. 6, 2002) (to be codified at 34 C.F.R. pt. 200).
- Department of Education. Title I—Improving the academic achievement of the disadvantaged: Rules and regulations, 67 Fed. Reg. 71,710 (Dec. 2, 2002) (to be codified at 34 C.F.R. pt. 200).
- Kane, T. (2002, September). *Racial subgroup rules in school accountability systems*. Paper presented at the 2002 annual CRESST conference, "Research Goes to School: Assessment, Accountability, and Improvement," University of California, Los Angeles.
- Linn, R. L., Baker, E. L., & Herman, J. L. (2002, Fall). Minimum group size for measuring adequate yearly progress. *The CRESST Line*, 1, 4-5. (Newsletter of the National Center for Research on Evaluation, Standards, and Student Testing [CRESST], University of California, Los Angeles)
- No Child Left Behind Act of 2001, Pub. L. No. 107-110, 115 Stat. 1425 (2002).
- Olson, L. (2002, November 27). States anxious for federal guidance on yearly progress. *Education Week*, pp. 1, 14, 15.
- Paige, R. (2002). Dear Colleague [Letter signed by the Secretary of Education, July 24]. Retrieved December 20, 2002, from www.ed.gov/News/Letters/020724.html

¹ I thank Richard Hill for valuable comments on an earlier draft of this brief. Any remaining shortcomings are, of course, the responsibility of the author.

² $(100 - 52)/12 = 4\%$

³ $(100 - 40)/12 = 5\%$

⁴ $52\% + 2(4\%) = 60\%$

⁵ $40\% + 2(5\%) = 50\%$

⁶ We have used equal intervals and uniform rates of change in the example for simplicity. It should be noted that the U.S. Department of Education does not require the use of equal intervals across the timeline so long as increments are substantial and the goal of 100% proficient is achieved by 2014.

⁷ The second requirement of making progress on an additional indicator applies to all subgroups, not just subgroups that qualify for the exception.

The work reported herein was supported under the Educational Research and Development Centers Program, PR/Award Number R305B960002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education. The findings and opinions expressed in this publication do not reflect the positions or policies of the National Institute on Student Achievement, Curriculum, and Assessment, the Office of Educational Research and Improvement, or the U.S. Department of Education.

GRADUATE SCHOOL OF
**EDUCATION &
INFORMATION
STUDIES**

UCLA Center for the Study of Evaluation
CSE/CRESST
GSE&IS BLDG MAILBOX 951522
LOS ANGELES CA 90095-1522
ADDRESS SERVICE REQUESTED

PRESORTED
STANDARD
U.S. POSTAGE
PAID
PASADENA, CA
PERMIT NO. 1132