

Spring 3-1-2006

## Evading the No Child Left Behing Act: State Strategies and Federal Complicity

Evan Stephenson

Follow this and additional works at: <https://digitalcommons.law.byu.edu/elj>



Part of the [Education Law Commons](#)

---

### Recommended Citation

Evan Stephenson, *Evading the No Child Left Behing Act: State Strategies and Federal Complicity*, 2006 BYU Educ. & L.J. 157 (2006).  
Available at: <https://digitalcommons.law.byu.edu/elj/vol2006/iss1/4>

.

This Article is brought to you for free and open access by BYU Law Digital Commons. It has been accepted for inclusion in Brigham Young University Education and Law Journal by an authorized editor of BYU Law Digital Commons. For more information, please contact [hunterlawlibrary@byu.edu](mailto:hunterlawlibrary@byu.edu).

# EVADING THE NO CHILD LEFT BEHIND ACT: STATE STRATEGIES AND FEDERAL COMPLICITY

*Evan Stephenson\**

## INTRODUCTION

The U.S. Department of Education (Education Department) has the legal duty to implement the three-part vision of the No Child Left Behind Act (NCLB or the Act).<sup>1</sup> Under NCLB, all public school students must progress toward 100 percent proficiency in math and reading within a twelve-year period ending in 2014.<sup>2</sup> Student proficiency is to be determined by uniform statewide tests.<sup>3</sup> These tests must be “challenging,”<sup>4</sup> meaning that they align with state educational standards.<sup>5</sup> Second, NCLB requires that student test improvement must progress evenly over time.<sup>6</sup> Third and finally, NCLB ensures that disadvantaged students may not be left behind under the Act’s accountability mandates.<sup>7</sup> In this respect, racial and ethnic minorities, limited English speakers, the poor, and disabled students must progress as the overall general students. To ensure that these student subgroups

---

\* Law Clerk, Hon. John M. Rogers, United States Court of Appeals for the Sixth Circuit, 2005–06. J.D., Univ. of Va., 2005. I would like to thank Terri Schwartzbeck, Jared Jacobs, Jeff Bennion, Tristen Stephenson, Brent Olson, Ed Stephenson, Jared Berg, Yvonne Stephenson, and Professor Charles J. Goetz for their insightful comments. Thank you also to Professor Jim Ryan, who oversaw and inspired this piece.

1. 20 U.S.C. §§ 6301–7941 (Supp. 2002); see *Chevron, U.S.A. Inc. v. Nat. Res. Def. Council, Inc.*, 467 U.S. 837, 842 (1984) (“If the intent of Congress is clear, that is the end of the matter; for the court, as well as the agency, must give effect to the unambiguously expressed intent of Congress.”).

2. See 20 U.S.C. § 6311(b)(1)(A)–(B) (requiring that all students be included in the state accountability system); *id.* at § 6311(b)(1)(F), (b)(1)(G)(iv), (b)(3)(A). See also Rod Paige, *Key Policy Letters Signed by the Education Secretary or Deputy Secretary*, <http://www.ed.gov/policy/elsec/guid/secletter/020724.html> (July 24, 2002).

3. 20 U.S.C. § 6311(b)(3)(A), (C).

4. *Id.* at §§ 6301, 6311(b)(1)(A).

5. *Id.* at § 6311(b)(1)(D)(ii).

6. *Id.* at § 6311(b)(2)(H)(i) (providing that intermediate goals for meeting NCLB’s requirements shall “increase in equal increments over the period covered by the State’s timeline”).

7. *Id.* at § 6311(b)(2)(C)(v)(1), (H)(aa)–(dd); 34 C.F.R. § 200.13.

keep pace, states must track their subgroups' test results separately.<sup>8</sup>

Notwithstanding its duty to implement this three-part vision, the Education Department has approved three devices that can each be used to evade a part of the Act's aims. These devices allow states to inflate their student proficiency statistics, thereby minimizing the likelihood of falling subject to the Act's harsh sanctions for inadequate test performance.<sup>9</sup> Furthermore, these devices lessen states' incentives to accomplish the same ends by allowing them to use the (more harmful) alternative of lowering educational standards.<sup>10</sup>

The first device allows states to evade NCLB's vision of evenly distributed progress over time by "backloading" their planned student proficiency gains.<sup>11</sup> This device, or *Balloon Schedule*, named for its similarity to "balloon mortgage" repayment schedules in which payments swell (i.e., balloon) in the later years,<sup>12</sup> allows a state to schedule the majority of student proficiency gains for the second half of NCLB's twelve-year timeline.<sup>13</sup> For example, Wyoming, Georgia, and thirteen other states have scheduled two-thirds of all student proficiency gains in the last four years of the twelve-year timeline.<sup>14</sup> Seven other states have scheduled the majority of such gains for the latter half of the timeline.<sup>15</sup>

The second device can be used by states to evade NCLB's mechanism for watching over disadvantaged subgroups. The Education Department permits states to exclude selected schools' subgroups from their statistical reports by electing to raise the minimum number of students that must belong to a subgroup in each school before it is tracked for accountability purposes (referred to below as *selection of minimum subgroup sizes*).<sup>16</sup> For instance, Missouri's minimum subgroup size for racial and ethnic minorities is thirty.<sup>17</sup> Thus, a Missouri school

---

8. *Id.*

9. *See infra* sec. II.A–II.D.

10. *See infra* sec. III.

11. *See infra* sec. I.C.–I.D.

12. A backloaded increase structure is "not unlike a balloon mortgage, leaving the heaviest lifting to those who will be in office long after the designers of that state's plan have departed the scene." Chester E. Finn, Jr. & Frederick M. Hess, *On Leaving No Child Behind*, 157 *Pub. Interest* 35, 43 (Fall 2004).

13. This definition is adapted from the Education Department's own description of the strategy. *See* U.S. Dept. Educ., *infra* n. 56 (noting that "Alaska and a number of other states, including Ohio and Arizona, created a trajectory that is more aggressive in the second half" of the timeline).

14. *See infra* tbl. 1 (derived from U.S. Dept. Educ., *Approved State Accountability Plans*, <http://www.ed.gov/admins/lead/account/stateplans03/index.html> (accessed Feb. 23, 2006)).

15. *See infra* tbl. 1.

16. *See infra* sec. I.F.1.

17. D. Kent King, *Missouri Consolidated State Application Accountability Workbook* 28,

with twenty-nine or fewer minority students will not be required to separately account for these students' test performance. Under some circumstances, perhaps that choice could be justifiable. Suppose, hypothetically, that Missouri wanted to strategically exclude certain special education subgroups from its statistical accountability reports. Missouri could do so by merely raising the minimum subgroup size for disabled students high enough that few schools would have enough disabled students to trigger separate reporting. For whatever reasons, Missouri *has* set a relatively high minimum subgroup size of fifty for the disabled.<sup>18</sup>

The third and final device allows states to avoid the Act's chief aim of 100 percent proficiency.<sup>19</sup> States can escape the 100 percent proficiency goal by calculating their statistics using *confidence intervals*.<sup>20</sup> A confidence interval is a statistical device that defines a "margin of error."<sup>21</sup> Schools may be considered NCLB-compliant when they reach state proficiency goals *minus the margin of error*.<sup>22</sup> Hypothetically, if a state's proficiency goal is 100 percent and its margin of error is eight percent, the state's schools may be considered compliant if they reach only ninety-two percent student proficiency. With such cushioning of their statistics, schools may never have to reach 100 percent proficiency, or any goal leading there.

This article examines the possible motives of states and the federal government in crafting these devices, their legality, and their probable effects on educational standards. Ultimately, this article furnishes evidence that pressure from NCLB to rapidly achieve 100 percent student proficiency encourages states to lower their educational standards, and that the Balloon Schedule, selection of minimum subgroup sizes, and confidence intervals have the benefit of releasing states from some of this pressure that was arguably intended by the Act.

Part I offers background on NCLB and its goals, and examines in detail three devices that states have used to evade the Act, with particular emphasis on the Balloon Schedule. Part II shows that states have adopted these devices to minimize their schools' exposure to NCLB's harsh sanctions for inadequate test performance. Part II further argues that the Education Department has approved the use of these devices in part to save the Act from its own unrealistic requirements and probably also as a

---

<http://www.ed.gov/admins/lead/account/stateplans03/mocsa.pdf> (updated July 5, 2005).

18. *Id.*

19. *See infra* sec. I.F.2.

20. *Id.*

21. *Id.*

22. *Id.*

means of preventing NCLB from being considered a failure. Part III demonstrates that the Balloon Schedule, selection of minimum subgroup sizes, and the use of confidence intervals, though at odds with NCLB's vision, reduce the pressure on states to define proficiency down by lowering their educational standards. To this extent, these devices are good for educational standards in America.

## I. THE BALLOON SCHEDULE AND OTHER DEVICES: DEFINITIONS AND ILLUSTRATIONS

### *A. NCLB's Timeline for 100 Percent Student Proficiency in Math and Reading*

In January 2002, President George W. Bush announced that as a result of the newly-passed *No Child Left Behind Act of 2001*,<sup>23</sup> "America's schools will be on a new path of reform, and a new path of results."<sup>24</sup> These new results are supposed to include an increase in the level of student academic achievement and elimination of the "achievement gap" between children of various backgrounds.<sup>25</sup> In crafting NCLB, lawmakers assumed that every child could score proficiently on tests<sup>26</sup> and that a key to student success is setting high expectations.<sup>27</sup>

23. 20 U.S.C. §§6301–7941.

24. Andrew Rudalevige, *The Politics of No Child Left Behind*, 3 Educ. Next 62, 63 (Fall 2003) (available at <http://www.educationnext.org/20034/pdf/62.pdf>).

25. See 20 U.S.C. § 6301 (Statement of Purpose); 34 C.F.R. § 200.13(a)(1)–(2) (emphasizing the necessity of meeting academic standards and closing the achievement gap); Alex Duran, *Factors to Consider When Evaluating School Accountability Results*, 34 J.L. & Educ. 73, 74 (2005); James E. Ryan, *The Perverse Incentives of the No Child Left Behind Act*, 79 N.Y.U. L. Rev. 932, 932 (2004).

26. See George W. Bush, *The Essential Work of Democracy*, 86 Phi Delta Kappan 114 (Oct. 2004) ("These reforms were entitled No Child Left Behind (NCLB) to reflect my belief that every child can learn. When expectations are high, America's children will rise to meet them."); George W. Bush, Speech, *Remarks to the National Urban League Conference* (Det., Mich., July 23, 2004), in 40 Wkly. Comp. Pres. Docs. 1365 (July 26, 2004) ("The philosophy of the No Child Left Behind Act says every child can learn. [W]e expect every child to learn and we expect you to show us whether or not every child is learning."); Michael Dobbs, *Former Math Teacher Recalculates No Child Left Behind Initiative*, Wash. Post A19 (Mar. 16, 2004) (quoting Assistant Education Secretary Raymond Simon saying the "only thing that is inflexible about [NCLB] is the idea that every child is capable of learning").

27. See Bush, *The Essential Work of Democracy*, *supra* n. 26; George W. Bush, Speech, *Remarks by the President in "Ask President Bush" Event* (Fond Du Lac, Wis., July 14, 2004), in White House Press Releases and Documents (available at 2004 WLNR 2541352) [hereinafter Bush, *Remarks*]; Jay Matthews, *To Educators, "No Child Goals Out of Reach"* Wash. Post A12 (Sept. 16, 2003) (quoting former U.S. Education Secretary Rod R. Paige).

NCLB ambitiously orders states to ensure that all students in affected schools test proficiently (with proficiency defined by each state) in math and reading within twelve years or by the school year ending in 2014.<sup>28</sup> Accordingly, every state has submitted an accountability workbook to the Education Department, delineating a schedule of the state's student proficiency goals for specific years.<sup>29</sup> The Department has provisionally approved all of these workbooks.<sup>30</sup>

NCLB requires states to set one annual measurable objective (AMO) applicable to each school, consisting of a percentage of students who must test proficiently.<sup>31</sup> AMOs are a component of adequate yearly progress (AYP), a phrase signifying whether a school meets all of NCLB's annual student improvement requirements.<sup>32</sup> For example, Virginia's plan requires seventy percent (the AMO) of its middle school students to test proficiently in reading on state tests in 2005, 2006, and 2007.<sup>33</sup> Seventy percent is thus a Virginia AMO for those years. If, hypothetically, only sixty-nine percent of students in a Virginia middle school test proficiently in 2007, then that school would fail to achieve this AMO and would also fail to make AYP for that year.

The AMO must also be *independently* achieved by the following four student subgroups in each school: the economically disadvantaged, limited English proficiency speakers, "major" racial or ethnic minorities, and the disabled.<sup>34</sup> If a school as a whole or *any* of its subgroups fails to achieve the AMO, the whole school fails to make AYP that year.<sup>35</sup>

28. 20 U.S.C. § 6311(b)(1)(A)-(B) (requiring that all students be included in the state accountability system); *Id.* at § 6311(b)(1)(F), (b)(1)(G)(iv), (b)(3)(A).

29. *See infra* n. 50 and accompanying text.

30. The Education Department provisionally approved all 50 state plans, including those laying out the Balloon Schedule. *See* U.S. Dept. of Educ., *Press Releases: President Bush, Secretary Paige Celebrate Approval of Every State Accountability Plan under No Child Left Behind*, <http://www.ed.gov/news/pressreleases/2003/06/06102003.html> (June 10, 2003); William J. Erpenbach et al., *Statewide Educational Accountability under No Child Left Behind 3* (Council of Chief St. Sch. Officers 2003) (available at <http://www.ccsso.org/publications/details.cfm?PublicationID=215>; *select* Download Free PDF(s) now) (observing that all the states' plans have been provisionally approved).

31. NCLB commands states to establish "statewide annual measurable objectives" towards total proficiency by 2014. 20 U.S.C. § 6311(b)(2)(G). These must be the "same for all schools and local education agencies in the state." 20 U.S.C. § 6311(b)(2)(G)(ii). They also must consist of a "single minimum percentage of students" who must score proficiently. 20 U.S.C. § 6311(b)(2)(G)(iii). *See also* 34 C.F.R. § 200.18.

32. *See* 20 U.S.C. § 6311(b)(2)(B)-(C) (setting forth the definition of "adequate yearly progress"); 34 C.F.R. § 200.14.

33. *See infra* tbl. 1.

34. 20 U.S.C. § 6311(b)(2)(C)(v)(I), (II)(aa)-(dd); 34 C.F.R. § 200.13.

35. 20 U.S.C. § 6311(b)(2)(C). This is true unless the Act's safe harbor provision applies, which is unimportant for this discussion. *See* 20 U.S.C. § 6311(b)(2)(I)(i) (safe harbor provision).

Hypothetically, if in 2007 only sixty-nine percent of a Virginia middle school's special education (i.e., disabled) students test proficiently in reading, then that school would fail to make AYP for that year—even if every other student in the school scored proficiently.<sup>36</sup>

As guidance for setting AMOs, the Act requires states to choose a “starting point” based on either the proficiency rate of the state's lowest performing subgroup or the proficiency rate of the school ranked at the state's twentieth percentile in student proficiency.<sup>37</sup> The difference between the starting point and 100 percent comprises the total percentage in student proficiency to be gained.<sup>38</sup> States must also determine the school years in which the percent proficient will *increase*.<sup>39</sup> Goals for increased proficiency are “intermediate goals.”<sup>40</sup> NCLB mandates that “[e]ach State shall establish intermediate goals for meeting the requirements, including measurable objectives . . . that shall increase in equal increments over the period covered by the State's timeline.”<sup>41</sup> The first such increase must occur no later than 2005, with later increases delayed no longer than three years.<sup>42</sup> Of course, AMOs and intermediate goals must place every school on a path toward 100 percent student proficiency by 2014.

### B. “Continuous and Substantial” Student Progress Forbids the Balloon Schedule

NCLB requires student improvement to be evenly distributed over time. The Act and regulations repeatedly emphasize that student progress must be “*continuous and substantial*,”<sup>43</sup> reflecting Congress's intent that states' timelines distribute proficiency gains evenly and not backload them. The intermediate goals of Arkansas and South Carolina for middle school reading are good examples of evenly scheduled proficiency

36. See e.g. Mark Goldberg, *Test Mess 2: Are We Doing Better a Year Later?*, 86 Phi Delta Kappan 389 (Jan. 2005) (“In North Carolina, more than 90% of schools are meeting the state growth goals, but only 47% of schools made AYP. Of those that failed, 283 missed the federal goal because one subgroup fell below its proficiency level.”).

37. See 20 U.S.C. § 6311(b)(2)(E) (“Each State, using data for the 2001–2002 school year, shall establish the starting point for measuring, under subparagraphs (G) and (H), the percentage of students meeting or exceeding the State's proficient level of academic achievement on the State assessments under paragraph (3) and pursuant to the timeline described in subparagraph (F).”). See also *id.* at § 6311(b)(2)(E)(i)–(ii); 34 C.F.R. § 200.16.

38. See *supra* nn. 32–37 and accompanying text.

39. See *infra* nn. 40–41, 46–47 and accompanying text.

40. 20 U.S.C. § 6311(b)(2)(H).

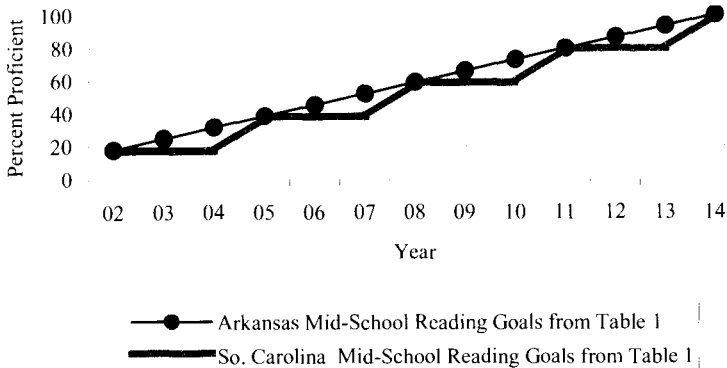
41. *Id.*

42. See *id.*

43. See 20 U.S.C. § 6311(b)(2)(C)(iii), (b)(2)(C)(v); 34 C.F.R. 200.13(b)(3).

growth.<sup>44</sup>

**Figure 1: Straight-Line and Stair-Step Schedules**



Arkansas uses an annually increasing straight-line schedule, and South Carolina opts for tri-annually increasing stair-steps.<sup>45</sup> Both schedules have trend lines that share the same slope and both distribute gains evenly over the timeline's first and second halves.

To specifically forbid the Balloon Schedule, Congress included language in the Act that limits states' flexibility in setting intermediate goals. Although states may schedule such goals annually, bi-annually, or tri-annually,<sup>46</sup> such goals must facilitate progress "over the period covered by the State's timeline" in "equal increments."<sup>47</sup> One state's 2003 NCLB workbook acknowledges that this statutory language expresses Congress's intent to forbid backloading of proficiency increases: "The Congressional intention [in] using [the 'equal increments'] language was to *ensure that no State waited until near the end of the timeline and then expected enormous, unrealistic growth in the last two or three years.*"<sup>48</sup> "Continuous and substantial" improvement facilitated by "equal" intermediate goals over the timeline, therefore, refers to even distribution of proficiency gains and prohibits the Balloon Schedule.

44. *Infra* tbl. 1.

45. Arkansas and South Carolina have the same starting point (eighteen percent) for intermediate/middle school reading. *See infra* tbl. 1.

46. *See* 20 U.S.C. § 6311(b)(2)(H).

47. *See* 20 U.S.C. § 6311(b)(2)(H)(i).

48. *See* Ill. St. Bd. Educ., *Accountability Workbook Originally Adopted in June 2003* 9, <http://www.ed.gov/admins/lead/account/stateplans03/ilcsa.pdf> (revised Aug. 23, 2005) (emphasis added).



*C. Many States Have Adopted the Balloon Schedule*

Instead of evenly distributing their proficiency gains as required by the Act,<sup>49</sup> many states have backloaded student proficiency increases. Table 1 on the next page shows the states' AMOs and intermediate goals for intermediate/middle school reading, as contained in the states' 2003 accountability workbooks.<sup>50</sup> Table 1 identifies, by italics, the many states that have arguably used the Balloon Schedule in their 2003 accountability workbooks. Five plans lacked sufficient AMO information for inclusion in the table. Of the forty-five remaining plans, twenty-two evidence the Balloon Schedule.

---

49. *See supra* sec. I.B.

50. Table 1 compiles data from the fifty states' 2003 approved accountability workbooks as posted on the Education Department website. *See* U.S. Dept. Educ., *Approved State Accountability Plans*, <http://www.ed.gov/admins/lead/account/stateplans03/index.html> (accessed Feb. 23, 2006).

		Table 1: 2003 State Accountability Workbooks Intermediate/Middle School Reading Proficiency Goals and Increases													
		Year, % Proficient													
		02-start	03	04	05	06	07	08	09	10	11	12	13	14	
1	Alabama*	--				x		x		x	x	x	x	x	
2	Alaska	64	64	64	70	70	70	76	76	76	82	88	94	100	
3	Arizona	31	31	31	43	43	43	54	54	54	66	77	89	100	
4	Arkansas	18	25	32	39	45	52	59	66	73	80	86	93	100	
5	California	14	14	14	24	24	24	35	46	57	68	78	89	100	
6	Colorado	75	75	75	81	81	81	87	87	87	94	94	94	100	
7	Connecticut	62	62	62	72	72	72	81	81	81	91	91	91	100	
8	Delaware*	--			x		x		x	x	x	x	x	x	
9	Florida	31	31	31	48	48	48	65	65	65	82	82	82	100	
10	Georgia	NA	60	60	67	67	67	73	73	73	80	87	93	100	
11	Hawaii	30	30	30	44	44	44	58	58	58	72	72	86	100	
12	Idaho	Not Enough Information													
13	Illinois	40	40	40	48	48	55	63	70	78	85	93	93	100	
14	Indiana*	--			x			x			x	x	x	x	
15	Iowa	61	61	61	68	68	68	74	74	74	81	87	94	100	
16	Kansas	51	51	51	63	63	63	76	76	76	88	88	88	100	
17	Kentucky	46	46	46	51	56	62	67	73	78	84	89	95	100	
18	Louisiana	?	37	37	47	47	47	58	58	58	68	79	89	100	
19	Maine	35	35	35	42	42	42	50	58	67	75	83	92	100	
20	Maryland*	--			x			x			x			x	
21	Mass.	71	71	76	76	81	81	85	85	90	90	95	95	100	
22	Michigan	31	31	31	43	43	43	54	54	54	66	77	89	100	
23	Minnesota	70	70	70	73	76	79	82	85	88	91	94	97	100	
24	Mississippi	30	30	30	48	48	48	65	65	65	83	83	83	100	
25	Missouri	18			39			59			80			100	
26	Montana	Not Enough Information													
27	Nebbraska	61	61	61	71	71	71	81	81	81	91	91	91	100	
28	Nevada	37	37	37	48	48	48	58	58	69	69	79	90	100	
29	New Hamp.	60	60	60	70	70	70	80	80	80	90	90	90	100	
30	New Jersey	58	58	58	66	66	66	76	76	76	87	87	87	100	
31	New Mexico*	--	x	x	x	x	x	x	x	x	x	x	x	x	
32	New York*	--			x	x	x	x	x	x	x	x	x	x	
33	N. Carolina	69	69	69	77	77	77	84	84	84	92	92	92	100	
34	N. Dakota	61	61	61	71	71	71	81	81	81	90	90	90	100	
35	Ohio*	--			x			x			x	x	x	x	
36	Oklahoma	Not Enough Information													
37	Oregon*	--			x			x			x	x	x	x	
38	Pennsylvania	45	45	45	54	54	54	63	63	63	72	81	91	100	
39	Rhode Island	68	68	68	73	73	73	79	79	79	84	89	95	100	
40	S. Carolina	18	18	18	38	38	38	59	59	59	79	79	79	100	
41	S. Dakota	65	65	65	71	71	71	77	77	77	83	88	94	100	
42	Tennessee	Not Enough Information													
43	Texas	47	47	47	54	54	60	60	67	73	80	87	93	100	
44	Utah	65	65	65	71	71	77	77	83	83	89	89	95	100	
45	Vermont*	--			x			x			x			x	
46	Virginia	61	61	61	70	70	70	80	80	80	90	90	90	100	
47	Washington	30	36	42	48	53	59	65	71	77	83	88	94	100	
48	W. Virginia	Not Enough Information													
49	Wisconsin	61	61	61	68	68	68	74	74	74	81	87	94	100	
50	Wyoming	35	35	35	45	45	45	56	56	56	67	78	89	100	

\*This state initially designated the years for improvement (marked as "x") without exact goals.

For example, from 2003 to 2008—the first half of the Act’s timeline—Georgia expects its middle-school reading proficiency rate in each school to increase from sixty percent to seventy-three percent, a thirteen percent jump. But from 2009 to 2014—the timeline’s second half—Georgia anticipates a much bigger increase from seventy-three percent to one hundred percent, a twenty-seven percent leap. Although Georgia’s six intermediate goals are all about seven percent<sup>51</sup> and are thus equal to each other, only two are scheduled for the first half of the timeline, while four are scheduled for the second half.<sup>52</sup> Rather than distributing its proficiency gains evenly over time, Georgia has scheduled about twice as much growth for the timeline’s final half, and has therefore backloaded its gains by *timing* intermediate goals *unevenly*.

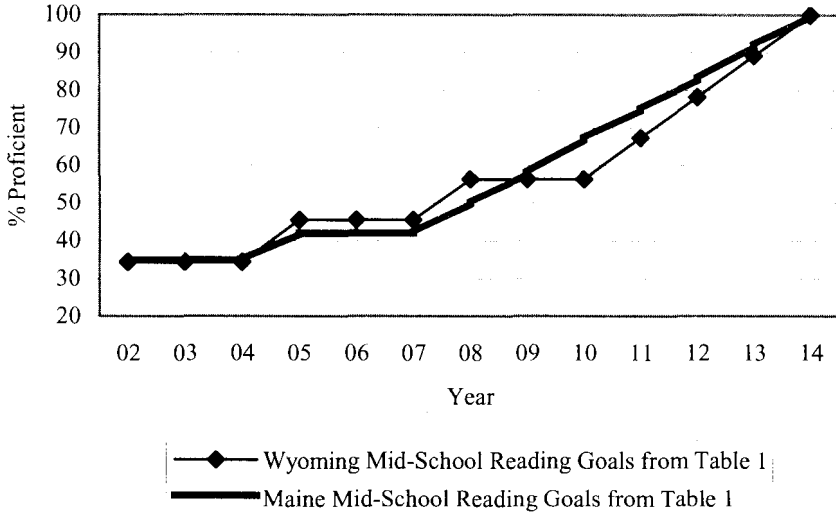
Figure 2 on the following page illustrates the Balloon Schedule as used by Wyoming and Maine.<sup>53</sup>

---

51. The total amount of Georgia’s combined proficiency increases—forty percent (i.e., the 100 percent final goal minus the sixty percent starting point)—divided by six (the number of intermediate goals selected by Georgia) equals six and two-thirds. Georgia rounded up to seven percent for all of its intermediate goals except the two scheduled for 2008 and 2013, which are six percent. Thus, four goals of seven percent (twenty-eight percent), plus two goals of six percent (twelve percent) equal forty percent. *Supra* tbl. 1.

52. *Supra* tbl. 1.

53. Wyoming and Maine share the same starting point (thirty-five percent) for intermediate/middle school reading. *Id.* Figure 2 illustrates data displayed in Table 1.

**Figure 2: Example Versions of the Balloon Schedule**

Wyoming’s and Maine’s proficiency goals, like those of Georgia, also employ the Balloon Schedule, putting the majority of the desired growth in the second half of the timeline. Wyoming’s plan sets six equal intermediate goals for 2005, 2008, 2011, 2012, 2013, and 2014. Maine’s schedule similarly opts for eight equal intermediate goals with most scheduled at the timeline’s end. These and any other such versions of the Balloon Schedule run contrary to the “continuous and substantial” and “equal increments” provisions of NCLB.<sup>54</sup>

*D. How States Obtained Approval for the Balloon Schedule:  
The Education Department’s Misreading of NCLB*

Despite the Act’s clarity in forbidding the Balloon Schedule, the Education Department has allowed states to use it.<sup>55</sup> How has the Department justified doing so? Quite simply, the Department approved the Balloon Schedule by misinterpreting the clear “equal increments” statutory language analyzed above. An Education Department press release states “[w]hile states have to ensure that their intermediate goals increase in equal increments over the NCLB timeline, states have great

54. See *supra* sec. I.B.

55. See *supra* n. 30 and accompanying text.

flexibility in determining *how often* their intermediate goals increase. States can raise their intermediate goals *every year or every two or three years.*"<sup>56</sup>

The key phrase in this press release is "how often." In the Department's view, states may backload proficiency gains, so long as their intermediate goals are quantitatively equal to each other. Increases may each equal ten percent, or 15 percent, or some other number. But states can backload gains by selecting how often such increases occur. Georgia, Wyoming, and Maine, as described above,<sup>57</sup> have each planned most of their equal increases for the timeline's second half<sup>58</sup>—which has the same effect as timing all increases consistently and making later increases larger. Interestingly, the Education Department's interpretation of the Act would not allow states to make later increases larger than earlier increases, but it does allow the functional equivalent of doing so by letting states plan a larger number of equal increases in later phases of the schedule. The Department does not read the phrase "over the NCLB timeline" to mean that increases should be equal over time and therefore does not regulate the timing or rate of progress. This places few, if any, constraints on when equal increases may take place, because it has allowed backloading. The Education Department's view then arguably contravenes NCLB's provisions. As such, the Balloon Schedule is inconsistent with any plausible interpretation of "continuous and substantial" progress in "equal increments"<sup>59</sup> over time.

#### *E. Did Congress Allow For an Accelerated Growth Interpretation of NCLB?*

Some Balloon Schedules state claims in their accountability workbooks that student proficiency will accelerate in later years, "after teachers are given time to align their instruction with academic content standards, after districts are given the opportunity to increase their capacity to support needed reforms, and after there is a highly qualified teacher in every . . . classroom."<sup>60</sup> Michigan anticipates acceleration in

---

56. U.S. Dept. Educ., *Charting the Course: States Decide Major Provisions under No Child Left Behind*, <http://www.ed.gov/news/pressreleases/2004/01/01142004.html#elements> (Jan. 14, 2004) (emphasis added). This same section of the press release pointed to some balloon strategies to show NCLB's "flexibility." *Id.*

57. *See supra* sec. I.C.

58. *See supra* tbl. 1.

59. *See supra* sec. I.B.

60. *See* Cal. Dept. Educ., *State of California Consolidated State Application Accountability Workbook* 30, <http://www.ed.gov/admins/lead/account/stateplans03/cacsa.pdf> (amended Aug. 23, 2005); *see also* Peter McWalters, *Rhode Island Department of Education Consolidated State Application Accountability Workbook* 30, <http://www.ed.gov/admins/lead/account/stateplans03/>

proficiency for these reasons and because it expects educational norms to change: “These shared norms and expectations require a significant investment in the knowledge and skills of teachers in low-performing schools and school districts before the most substantial improvement gains will be realized.”<sup>61</sup>

Some education research can be stretched to support the notion that schools’ proficiency growth can accelerate in the later years of an accountability system.<sup>62</sup> Teacher and school quality both exert some influence on student performance.<sup>63</sup> Thus, if these variables improve at an accelerated rate, as some Balloon Schedule states claim they will,<sup>64</sup> then it would be reasonable to expect student performance also to accelerate.<sup>65</sup> In these circumstances, backloading may be theoretically sound. Nevertheless, this theory of accelerated proficiency growth, even if accepted, is simply not allowed by the Act’s language, nor should it be unquestioningly accepted on its own merits.

As discussed above, Congressional intent is reflected in the Act’s language, which prohibits backloading.<sup>66</sup> The Senate conference members who explained the final version of NCLB to the Senate in December 2001 disparaged backloading as both prohibited by the Act and as unrealistic. When Senator Joseph Lieberman explained what the “conferees intend” regarding “this system of setting [the] progress bar and raising it in equal increments over a 12-year period,” he spelled out that “[i]t will further ensure that state plans outline realistic timelines for

ricsa.pdf (revised July 2004).

61. Mich. Dept. Educ., *Michigan Consolidated State Application Accountability Workbook* 28, <http://www.ed.gov/admins/lead/account/stateplans03/micsa.pdf> (amended June 2005).

62. See generally Harold Wenglinsky, *How Schools Matter: The Link between Teacher Classroom Practices and Student Academic Performance*, 10 *Educ. Policy Analysis Archives* 12 (Feb. 13, 2002), <http://epaa.asu.edu/epaa/v10n12/> (The abstract of the article observes that “the effects of classroom practices, when added to those of other teacher characteristics, are comparable in size to those of student background, suggesting that teachers can contribute as much to student learning as the students themselves.”). Therefore, if California, Rhode Island, and Michigan are correct that their overall school quality improvement (and related factors) will accelerate over time, see *supra* nn. 60–61 and accompanying text, student performance could conceivably accelerate also. See generally Linda Darling-Hammond, *Teacher Quality and Student Achievement: A Review of State Policy Evidence*, 8 *Educ. Policy Analysis Archives* 1 (Jan. 1, 2000), <http://cpaa.asu.edu/epaa/v8n1/>; Dan Goldhaber, *The Mystery of Good Teaching*, 2 *Educ. Next* 50, 52 (2002) (available at <http://www.educationnext.org/20021/50.html>) (finding that “high quality teachers raise student performance” and that teacher characteristics alone account for 8.5 percent of the variation in student achievement); Larry E. Suter, *Is Student Achievement Immutable? Evidence from International Studies on Schooling and Student Achievement*, 70 *Rev. Educ. Research* 529 (2000).

63. See *id.*

64. See *e.g. supra* nn. 60–61 and accompanying text.

65. See *supra* nn. 60–62 and accompanying text.

66. See *supra* sec. I.B.

getting to proficiency, and prohibits states from ‘backloading’ their expected proficiency gains in the out years.”<sup>67</sup> This statement suggests the Senate conference members, therefore, did not consider backloading to be “realistic.” In fact, the Congressional Record publishes no legislator’s views to the contrary.<sup>68</sup> Congress intended to forbid plans based on accelerated proficiency growth.

Second, this accelerated growth justification for the Balloon Schedule should not be unquestioningly accepted on its own merits. The education literature that can be stretched to support this acceleration theory at best bolsters the Balloon Schedule’s legitimacy only indirectly, if at all.<sup>69</sup> It is not clear, for instance, why accumulated improvements in school quality, teacher quality, curriculum alignment with content standards, and educational norms would accelerate rather than increase at a constant, linear rate.<sup>70</sup> Even if some schools could accelerate performance results, research by noted education scholar Richard Rothstein suggests that solutions to rapidly accelerated performance results are not generally applicable to all schools.<sup>71</sup>

### *F. Other Evasion Devices*

#### *1. Selection of Minimum Subgroup Sizes*

States have used other devices in addition to the Balloon Schedule to improve their NCLB statistics, including the selection of minimum subgroup sizes. More than a dozen states have excluded disadvantaged student subgroups from their accountability reports by raising the minimum number of students that must be in a subgroup before it is tracked for accountability purposes.<sup>72</sup> For example, one Maryland school

---

67. 147 Cong. Rec. S13403 (daily ed. Dec. 18, 2001) (statement of Sen. Lieberman) (emphasis added).

68. *See id.*

69. For examples of this literature, see *supra* n. 62 and accompanying text.

70. Although the literature cited *supra* in note 62 supports the notion that teacher and school quality affect student learning, they do not specifically state that student learning or teacher and school quality have any tendency (or capability) to improve *at an accelerated rate*. The claim that any of these can improve at an accelerated rate is derived not from the scholarly literature cited *supra* in note 62 but by certain states’ own accountability workbooks. *E.g.* *supra* nn. 60–61 and accompanying text.

71. *See* Richard Rothstein, *Class and Schools: Using Social, Economic, and Educational Reform to Close the Black-White Achievement Gap*, 72–83 (Econ. Policy Inst. 2004).

72. *See* Erpenbach et al., *supra* n. 30, at 34–35 (“Some States (e.g., Nebraska, Oklahoma, Ohio, Wisconsin) have established higher minimum ‘n’s’ for accountability determinations with SWDs [students with disabilities] subgroups.”); *id.* at 52 (noting that this strategy has a “long-term impact on the number of schools and districts identified for improvement”). For statutory provisions relevant to minimum subgroup sizes, *see* 20 U.S.C. § 6311(b)(2)(C)(v)(II)(dd), (b)(2)(1)(ii),

has already failed to make AYP “because its 10 special education students”<sup>73</sup> failed to test proficiently, while in the same year a Virginia school made AYP “despite its 24 special education students’ failing to make AYP. The difference? The minimal group size for reporting was five in Maryland and 50 in Virginia.”<sup>74</sup> As this example shows, disabled students can easily cause an entire school to fail AYP because their test results are counted separately. With so much riding on subgroups’ test performance, it is not surprising that states are exercising some strategic discretion in determining minimum subgroup sizes.

Although subjective motives cannot be determined with certainty, some states’ actions seem problematic in the light of observable data. In 2004, for example, Washington State raised the minimum subgroup size for disabled students and limited English proficiency speakers, but for no other subgroups, from thirty to forty.<sup>75</sup> As a result, any Washington school with thirty-nine or fewer disabled students will not separately be held accountable for the test performance of these students. Oddly, Washington waited until two years into the NCLB timeline to raise the minimum size for *only the disabled and limited English speakers*. New Jersey acted identically with respect to the disabled by increasing its minimum subgroup size for disabled students from twenty to thirty-five, also in the year 2004. Had this increase been in place in 2003, it would have dramatically cut the number of schools failing to make AYP.<sup>76</sup> In fact, one New Jersey newspaper estimates the number of such schools in the hundreds.<sup>77</sup> If Washington and New Jersey were indeed manipulating their minimum subgroup sizes, they arguably transgressed the Act’s commands.

In setting subgroup sizes, NCLB mandates that states consider only whether a given size effectuates the collection of reliable information and ensures student anonymity.<sup>78</sup> Outside of reliability or anonymity

(b)(3)(C)(xiii).

73. Gerald W. Bracey, *The 14th Bracey Report on the Condition of Public Education*, 86 Phi Delta Kappan, 149 (Oct. 2004).

74. *Id.*

75. David Wickert, *WASL Tweaks Bring Standards within Reach; Scoring Changes: Schools to Get Higher Marks, Even if Students Don’t*, Morn. News Trib. (Tacoma, Wash.) A1 (Oct. 21, 2004).

76. John Mooney, *U.S. Eases Rules on Test Results for Jersey’s Special Ed Students*, The Star-Ledger (Newark, N. J.) 40 (June 25, 2004).

77. Maia Davis, *U.S. Education Department Approves New Jersey’s Special Education Plan*, The Record (Bergen County, N.J.) A4 (June 24, 2004) (“An analysis by The Record (Bergen County, NJ), published last month, found that hundreds of schools deemed underperforming last year may now escape that label because they have fewer than thirty-five special education students.”).

78. See 20 U.S.C. § 6311(b)(2)(C)(v)(II)(dd), (b)(2)(I)(ii), (b)(3)(C)(xiii).



concerns, there is no obvious justification for setting a higher subgroup size for the disabled and for limited English speakers than for racial minorities or the poor.<sup>79</sup> Washington's and New Jersey's alteration of the minimum subgroup sizes *only* for the disabled and limited English speakers seems consistent with a hypothesis of strategic manipulation.

In 2004, many states, including Alaska, Arkansas, New Jersey, South Carolina, and Washington obtained Education Department approval of higher minimum subgroup sizes.<sup>80</sup> By granting to states the ability to raise minimum subgroup sizes—even in arguably suspicious cases like those of Washington and New Jersey—the Education Department has invited manipulation.

## 2. Confidence Intervals

To improve their NCLB statistics, a majority of states<sup>81</sup> have implemented another evasion device—the use of confidence intervals—which can dramatically increase a state's AYP numbers.<sup>82</sup> To cite an extreme example, one Utah newspaper claims that fifty-six percent of all Utah schools that made AYP in 2004 did so by using a confidence interval.<sup>83</sup>

To understand confidence intervals, imagine a typical news poll asking a sample group of randomly selected voters whom they will vote for in an upcoming election. The various polls, even if conducted in the same manner, will tend to disagree somewhat due to random variation in the sample groups.<sup>84</sup> But almost all the time, a repeated poll will have

---

79. A state might permissibly show that special education students' scores have higher variance and therefore a larger group of students is needed to make scores "reliable," if the state can also show that the Act's usage of "reliable" calls for a certain level of variance in test scores. For the argument that reliability may call for a certain level of variance in test scores, see generally Robert L. Linn & Carolyn Haug, *Stability of School-Building Accountability Scores and Gains*, 24 *Educ. Evaluation & Policy Analysis* 29 (2002).

80. See *supra* nn. 75–77 and accompanying text; Ctr. on Educ. Policy, *Rule Changes Could Help More Schools Meet Test Score Targets for the No Child Left Behind Act* 3–4, 9–10, <http://www.ctredpol.org/nclb/StateAccountabilityPlanAmendmentsReportOct2004.pdf> (Oct. 22, 2004). The District of Columbia, Puerto Rico, and some states appear interested in doing the same. See Diana Jean Schemo, *States' End Run Dilutes Burden for Special Ed*, N.Y. Times A1 (June 7, 2004).

81. For an overview of confidence interval use as found in states' accountability workbooks, see Erpenbach et al., *supra* n. 30, at 21–22. See also Ctr. on Educ. Policy, *supra* n. 80, at 3 ("About half the states had already included the use of confidence intervals in their original accountability plans. Since then, 12 states have either introduced the use of confidence intervals or changed the way they plan to use them to determine AYP.").

82. See *infra* nn. 149–154 and accompanying text.

83. Jennifer Toomer-Cook, *84% of Utah Schools Make the Grade on AYP Reports*, *Deseret Morning News* (Salt Lake City, Utah) B7 (Dec. 7, 2004).

84. For this paragraph, see Study Works Online, *Polling: Margin of Error*,

results that fall within a range, called the *margin of error*. The margin of error defines the *confidence interval*.<sup>85</sup> For example, if fifty percent of respondents in a poll say they will vote for Candidate A, and the margin of error is eight percent, then the repeated poll will almost always randomly register between fifty-eight and forty-two percent support for Candidate A. Fifty-eight to forty-two percent is the confidence interval.

States are using this statistical technique to cushion their schools' AYP results with a margin of error.<sup>86</sup> A recent *New York Times* report offered the following example of how a confidence interval may affect AYP results in practice: A state opts for a confidence level that results in a twenty-three percent margin of error for a small "sample,"<sup>87</sup> size of thirty minority students.<sup>88</sup> The potential effect of the confidence interval "is significant, if seemingly technical: For a class of 30 minority students at a school where 40 percent of each group must pass a given exam, the [confidence interval] cushion grants the school victory if only 17 percent, or 5 rather than 12 students, succeed."<sup>89</sup>

Suppose in the above example that this hypothetical school's subgroup achieves exactly seventeen percent proficiency (and falls twenty-three percent short of the AMO of forty percent). The justification for considering the school compliant is that its students' inadequate scores were the result of random variation in the "sample" of students.<sup>90</sup> The state may claim that another "sample" of students from the same population would have met the AMO.<sup>91</sup> Failure, then, was not the school's fault but the result of the random variation in the student

[http://www.studyworksonline.com/cda/content/article/0,,EXP545\\_NAV2-76\\_SAR542,00.shtml](http://www.studyworksonline.com/cda/content/article/0,,EXP545_NAV2-76_SAR542,00.shtml)  
(accessed Feb. 23, 2006).

85. See generally Mario F. Triola, *Elementary Statistics* 284–302 (6th ed., Addison-Wesley Publ. Co. 1995).

86. For an overview of how confidence intervals in the AYP context work, see generally ASR-CAS Jt. Study Group on Adequate Yearly Progress & Scott Marion et al., *Making Valid and Reliable Decisions in Determining Adequate Yearly Progress* 65–67 (Council of Chief St. Sch. Officers 2002) [hereinafter Marion et al.] (available at <http://www.ccsso.org/content/pdfs/AYPpaper.pdf>).

87. Quotes are put around the word "sample" because students are not samples in the NCLB system. See *infra* n. 90 and accompanying text.

88. See Schemo, *supra* n. 80.

89. *Id.*

90. Of course, students generally are not randomly assigned to schools and so cannot be considered a "random sample." Further, school populations within the Act's AYP framework are not really "samples" either because every student (not a mere sample of students) is supposed to be tested. See Schemo, *supra* n. 80 ("In addition, some statisticians . . . question the validity of using confidence intervals for this purpose. The cushions are most often used to allow for variations in statistical sampling, but schools are reporting on *actual students*, not *samples* of them." (emphasis added)).

91. See generally Linn & Haug, *supra* n. 79, at 29–36.

sample. If they rely on confidence intervals heavily enough, *schools using a confidence interval may never actually have to reach 100 percent student proficiency*, because they can always subtract the margin of error from 100 percent and be considered compliant.<sup>92</sup> This use of confidence intervals arguably compromises the Act's goals because the mandate of NCLB—that every public school student will test proficiently by 2014<sup>93</sup>—is rendered unnecessary. NCLB nowhere authorizes confidence intervals or anything like them.<sup>94</sup> Their use came about as a result of the Education Department's process of reviewing and approving state accountability workbooks.<sup>95</sup>

In short, the Balloon Schedule clearly violates NCLB. The selection of minimum subgroup sizes and the use of confidence intervals also arguably violate the Act.

## II. WHY HAVE STATES AND THE EDUCATION DEPARTMENT ADOPTED THE BALLOON SCHEDULE?

### *A. A Brief Overview of NCLB's Sanctions*

States using the Balloon Schedule help their schools avoid NCLB's severe sanctions for failing to make AYP. When a school falls short of AYP for two consecutive years, the school district must identify it for "school improvement."<sup>96</sup> Students in such a school receive the option to transfer to another school within the school district,<sup>97</sup> with transportation costs paid by the school district.<sup>98</sup> With the technical assistance of the

92. See e.g. Angela Pascopella, *Did the Tail Wag Dog? (Inside the Law: Analyzing, Debating and Explaining No Child Left Behind)* 41 Dist. Admin. 123 (Jan. 2005).

93. See *supra* n. 2 and accompanying text.

94. See Marion et al., *supra* n. 86, at 72 ("[T]he law does not mention confidence intervals at all."); E-mail from Terri Duggan Schwartzbeck, Policy Analyst, Am. Assn. of Sch. Adminstrs., to the author (June 28, 2004) (copy on file with author).

95. See *id.*

96. 20 U.S.C. § 6316(b)(1)(A) ("[A] local educational agency shall identify for school improvement any elementary school or secondary school served under this part that fails, for 2 consecutive years, to make adequate yearly progress as defined in the State's plan under section 6311(b)(2) of this title.")

97. *Id.* at § 6316(b)(1)(E)(i) ("In the case of a school identified for school improvement under this paragraph, the local educational agency shall, not later than the first day of the school year following such identification, provide all students enrolled in the school with the option to transfer to another public school served by the local educational agency.")

98. *Id.* at § 6316(b)(9) ("In any case described in paragraph (1)(E) for schools described in paragraphs (1)(A), (5), (7)(C)(i), and (8)(A), and subsection (e)(10)(C)(vii) of this section, the local educational agency shall provide, or shall pay for the provision of, transportation for the student to the public school the student attends.")

school district and after a possible review of the school's data,<sup>99</sup> the failing school must develop and implement a two-year plan to improve.<sup>100</sup> Failure to achieve AYP for three years triggers intensified school district assistance.<sup>101</sup>

If a school fails to meet AYP for a fourth consecutive year it may be subject to "corrective action."<sup>102</sup> Such action "might include replacing school personnel, instituting a new curriculum, extending the school year, . . . authorizing students to transfer to higher-performing schools[.]" or a reduction in funding to the target school.<sup>103</sup>

Finally, if "corrective action" does not cause the school to make AYP after another full year,<sup>104</sup> the school becomes subject to "restructuring."<sup>105</sup> The school district must pick an option from the law's restructuring menu, including:

- (1) Reopening the school as a public charter school;
- (2) Replacing school staff "relevant" to the school's failure;
- (3) Hiring a private company to run the school;
- (4) Turning the school over to the state government to run.<sup>106</sup>

The Act also contains a catchall phrase allowing the use of additional types of restructuring that would be effective.<sup>107</sup>

As a school passes from improvement to corrective action to restructuring, the sanctions for failure to achieve AYP grow "increasingly harsh."<sup>108</sup> It is not surprising that states would wish to help their schools avoid these sanctions by any means available, including the Balloon Schedule and other evasions of the Act.

99. *Id.* at § 6316(b)(2) ("Before identifying an elementary school or a secondary school for school improvement under paragraphs (1) or (5)(A), for corrective action under paragraph (7), or for restructuring under paragraph (8), the local educational agency shall provide the school with an opportunity to review the school-level data, including academic assessment data, on which the proposed identification is based.")

100. *Id.* at § 6316(b)(3).

101. *Id.* at § 6316(b)(5).

102. *Id.* at § 6316(b)(7) (defining corrective action).

103. Cory Shindel, Student Author, *One Standard Fits All? Defining Achievement Standards for Students with Cognitive Disabilities within the No Child Left Behind Act's Standardized Framework*, 12 J.L. & Policy 1025, 1044-45 (2004).

104. 20 U.S.C. § 6316(b)(8)(A).

105. *Id.*

106. *Id.* at § 6316(b)(8)(B).

107. *Id.*

108. Ryan, *supra* n. 25, at 933.

*B. NCLB's Unrealistic Goals Pressure States to Adopt Evasion Strategies*

The true underlying problem may be that NCLB's timeline and goals, though laudable in their intent, do not derive from sound research. Nearly all education experts see NCLB's 100 percent proficiency goal and its timeline as unrealistic<sup>109</sup> or even malignant.<sup>110</sup> To reach full proficiency by 2014, states must maintain significant proficiency growth for up to twelve years,<sup>111</sup> but there appears to be no "state that's sustained a growth of 5 percent a year over 5 years, let alone 12."<sup>112</sup> Indeed, "a large number of independent scholars have demonstrated that it is not possible to have 100 percent of students achieve a high standard."<sup>113</sup>

Studies simulating student performance based on past test results also paint a glum picture. "With few exceptions, the State simulation studies show that a high proportion of schools will likely not meet the new AYP requirements within two or three school years."<sup>114</sup> Even if states calculate AYP using statistical methods that eliminate smaller schools from their accountability reports, "almost all schools will end identified for improvement within five or six years."<sup>115</sup> Tellingly, education scholar Abigail Thernstrom, a prominent Republican and a George W. Bush appointee to the U.S. Commission on Civil Rights,<sup>116</sup>

109. See W. James Popham, Presentation, *Ruminations Regarding NCLB's Most Malignant Provision: Adequate Yearly Progress 2*, (D.C., July 28, 2004) (available at <http://www.cepdc.org/pubs/Forum28July2004/RuminationsReNCLB-AYP-ss-071488.pdf>) (referring to "the outlandish unrealism of the [NCLB's] adequate yearly progress (AYP) provisions," and specifying that "NCLB currently calls for schools to produce unrealistic increases in the test scores of successive cohorts of students."); Marion et al., *supra* n. 86, at 12 ("In a number of cases, the proficient level has been set so high [by the state] that it may be completely unrealistic to expect all students to reach that level by 2014").

110. Popham, *supra* n. 109, at 1 (referring in the title to NCLB's adequate yearly progress requirements as "[m]alignant"). See M. Hayes Mizell, *From Muck to Mountaintop*, 33 J.L. & Educ. 261, 261 (2004) ("A friend of mine who is a charter school administrator characterizes the law's application and compliance provisions as 'draconian and bordering on the sadomasochistic.'").

111. See *supra* tbl. 1.

112. Diana Jean Schemo, *Sidestepping of New School Standards Is Seen*, N.Y. Times A21 (Oct. 15, 2002).

113. William J. Mathis, *No Child Left Behind Act: What Will It Cost States?* 77 Spectrum: J. of St. Govt. 8 (Apr. 1, 2004). See Goldberg, *supra* n. 36 ("[T]he fact is that no large city, state, or country—other than mythical Lake Wobegon—has ever produced an entire population of students who are above average.").

114. Marion et al., *supra* n. 86, at 64.

115. *Id.* at 65.

116. See generally Manhattan Institute for Policy Research, *Abigail Thernstrom* [http://www.manhattan-institute.org/html/thernstrom\\_a.htm](http://www.manhattan-institute.org/html/thernstrom_a.htm) (accessed Feb. 23, 2006) (providing biographical data about Abigail Thernstrom).

dismisses the suggestion that all students can achieve a “Proficient” rating on her state’s tests—or even a “Basic” rating on the National Assessment of Educational Progress (NAEP), a national proficiency test.<sup>117</sup> The former she criticizes as “ludicrous,” the latter as “utopian.”<sup>118</sup>

So far, AYP statistics are not encouraging. Of the nation’s 91,400 public schools, approximately 26,000 (28.4 percent) failed to make AYP in the school year ending in 2003.<sup>119</sup> Nearly all states set their AMOs very low in 2003—at their starting points.<sup>120</sup>

One can only guess how quickly schools will fail to make AYP as intermediate goals come due. One study predicts that by 2014, ninety-nine percent of California public schools will have failed to make AYP.<sup>121</sup> The Connecticut Education Association estimates that “[m]ore than 90% of Connecticut elementary and middle schools won’t meet federal education standards in 10 years.”<sup>122</sup> Researchers forecast that by 2014 “nearly all schools in all states will fail under the law.”<sup>123</sup> Relying on similar research and facts, education researcher William J. Mathis voices doubt about NCLB’s core premise that every child *can* succeed in the sense contemplated by NCLB.<sup>124</sup> Even some NCLB supporters are pessimistic about schools’ chances for reaching full proficiency.<sup>125</sup>

117. Abigail Thernstrom, Speech, *Comments*, in *No Child Left Behind: What Will it Take?* 103, <http://www.edexcellence.net/doc/NCLBreport.pdf> (Feb. 2002). For an overview of the National Assessment of Educational Progress, see Natl. Ctr. for Educ. Statistics, *Overview: What is NAEP?* <http://nces.ed.gov/nationsreportcard/about> (updated Nov. 25, 2005).

118. Thernstrom, *supra* n. 117.

119. See Goldberg, *supra* n. 36 (“Approximately 26,000 of the nation’s 91,400 public schools failed to make AYP in the 2002–03 school year.”).

120. See *supra* tbl. 1.

121. Bracey, *supra* n. 73, at 139 (“[T]he projection is for 99% failure by the witching year of 2014, when 100% of students must be ‘proficient’.”).

122. Associated Press, *Study: 9 Out of 10 Schools Won’t Meet Federal Standards in a Decade*, Hartford Courant 1 (May 28, 2004).

123. Diana Jean Schemo, *Effort by Bush on Education Hits Obstacles*, N.Y. Times A1 (Aug. 18, 2004). See also Robert L. Linn, Presentation, *Rethinking the No Child Left Behind Accountability System* 3 (D.C., July 28, 2004) (available at <http://www.cep-dc.org/pubs/Forum28July2004>) (“[A]most all schools will fall short of the adequate yearly progress (AYP) targets within the next few years, unless major changes are made in the definition of AYP.”).

124. See Mathis, *supra* n. 113.

125. See e.g. Michael D. Casserly, Speech, *Comments*, in *No Child Left Behind: What Will it Take?* 71, <http://www.edexcellence.net/doc/NCLBreport.pdf> (Feb. 2002) (“Finally, [NCLB] has a strong accountability system that we also backed. . . . We expect to have a great deal of difficulty executing the bill’s AYP provisions. . . . [O]ur biggest challenge will be getting our instructional programs to do what this legislation envisions.”). The pro-NCLB Education Trust, an advocacy group, released a study in 2004 that emphasized progress in narrowing racial achievement gaps but admitted that “the pace of progress was generally insufficient to reach the goal of full proficiency by 2014.” Michael Dobbs, *Kerry Competes to Claim Issue of Reform*, Wash. Post A4 (Oct. 20, 2004) (summarizing findings from this study). See Educ. Trust, *Measured Progress: Achievement Rises*

Former Education Secretary Rod Paige, however, has dismissed such criticisms and tried to rebut them with positive anecdotes about a few schools that have experienced notable success.<sup>126</sup>

The impossibility of complying with NCLB on the merits, combined with the harshness of its sanctions, generates the pressure motivating states to help their schools evade the Act.

### *C. The Functions of the Balloon Schedule and Other Devices*

The Balloon Schedule helps states evade NCLB in two ways. First, backloading artificially lowers states' AMOs during the first half of the Act's timeline.<sup>127</sup> For example, if Georgia had applied South Carolina's intermediate goal structure<sup>128</sup> of even progress over the Act's timeline, Georgia would have scheduled a twenty percent increase in middle school reading proficiency by 2008, rather than a thirteen percent increase. (The Balloon Schedule thus gives Georgia middle schools a seven percent cushion.<sup>129</sup>) Due to this cushion, Georgia's middle schools with reading proficiency increases between thirteen percent and twenty percent will artificially<sup>130</sup> achieve AYP in reading in 2008.

Second, if states utilize a Balloon Schedule or similar tactics, they have less of an incentive to adopt other evasion strategies. For instance, the cushion Georgia schools gain by using the Balloon Schedule reduces pressure to further embellish student proficiency.

### *D. The Federal Government's Interest in Approving the Balloon Schedule*

The Balloon Schedule and the other evasion strategies inflate the number of schools making AYP, creating the appearance of school improvement and NCLB success, for which the federal government (especially the Education Department) may take credit. President George W. Bush and former Education Secretary Rod Paige both lauded NCLB's supposed "success" during the 2004 presidential election

---

*and*      *Gaps*      *Narrow,*      *but*      *Too*      *Slowly*      1-8,  
[http://www2.edtrust.org/edtrust/images/MeasuredProgress.doc .pdf](http://www2.edtrust.org/edtrust/images/MeasuredProgress.doc.pdf) (Oct. 2004).

126. See e.g. Rod Paige, Speech, *Remarks of Secretary Paige* ¶¶ 24-26 (The Educ. Writers Assoc. 2002 Annual Seminar, Apr. 27, 2002) (available at <http://www.ed.gov/news/speeches/2002/04/20020427.html>).

127. Such inflation is "artificial" to the extent that something other than genuine student improvement causes it, and such improvement is over and above what the state would accomplish while complying with the Act (i.e., by not using a Balloon Schedule).

128. See *supra* tbl. 1.

129. The seven percent cushion is relative to a stair-step intermediate goal structure.

130. On the meaning of "artificial" as used here, see *supra* note 127 and accompanying text.

season.<sup>131</sup> For example, former Secretary Paige's address at the 2004 Republican National Convention urged voters to reelect President Bush on the basis of NCLB's proven results: "Ladies and gentlemen, No Child Left Behind is working. . . . Only one candidate has created an education system worthy of a great nation: President George W. Bush."<sup>132</sup> This logic, however, cuts in both directions. Without ostensible successes in education, NCLB will look like a failure, and the federal government will absorb the blame. The Balloon Schedule and other evasion devices help ensure that NCLB redounds to the government's credit.<sup>133</sup>

President Bush and former Education Secretary Rod Paige have taken much credit for ostensible improvement in schools since the Act's passage.<sup>134</sup> Bush Administration officials typically have claimed success by saying that test scores have gone up.<sup>135</sup> Both President Bush and former Secretary Paige point to improved math scores of fourth- and eighth-graders from 2000 to 2003 on a nationwide proficiency test,<sup>136</sup> the NAEP.<sup>137</sup> Similarly, Paige's successor, Secretary Margaret Spellings, attributed gains in reading on the NAEP among nine-year-olds between 1999 and 2004 to NCLB.<sup>138</sup> On the other hand, there was no

131. See e.g. George W. Bush, Speech, *Remarks by the President at Victory 2004 Rally* (Oshkosh, Wis., Oct. 15, 2004), in White House Press Releases and Documents (available at 2004 WL 2558203) [hereinafter Bush, *Remarks at Victory Rally*] ("The No Child Left Behind Act is working."); Erik W. Robelen, *Paige: It's Not Too Early to Call School Law a Success*, 24 Educ. Week 23 (Oct. 6, 2004) (available at 2004 WL 63594841) ("Rod Paige is declaring the No Child Left Behind Act a success.").

132. Rod Paige, Speech, *Remarks by Education Secretary Rod Paige* ¶¶ 20, 27 (Republican National Convention, Aug. 31, 2004) [hereinafter Paige, *Republican National Convention*] (available at <http://www.washingtonpost.com/wp-dyn/articles/A50460-2004Aug31.html>).

133. Cf. Anne C. Lewis, *States Feel the Crunch of NCLB*, 86 Phi Delta Kappan 339 (Jan. 2005).

134. See *infra* nn. 131–158, and accompanying text.

135. See George W. Bush, Speech, *President's Remarks in New Mexico* ¶ 40 (Alamogordo, N.M., Oct. 24, 2004), in White House Press Releases & Documents (available at 2004 WLNR 3336295) ("We passed the No Child Left Behind Act, which is bringing high standards to our classrooms. . . . We're seeing great progress across this country. Math and reading scores are on the rise."); Paige, *Republican National Convention*, *supra* n. 132, at ¶¶ 20–21 ("No Child Left Behind is working. . . . All across America, test scores are rising; students are learning; the achievement gap is closing."). See also Bush, *Remarks at Victory Rally*, *supra* n. 131.

136. Bush, *The Essential Work of Democracy*, *supra* n. 26; Rod Paige, Speech, *Paige Kicks Off Annual Celebration of Historically Black Colleges and Universities* ¶ 18 (Sept. 13, 2004) (available at <http://www.ed.gov/news/speeches/2004/09/09132004.html>).

137. Recall that NAEP stands for "National Assessment of Educational Progress." See *supra* n. 117.

138. See U.S. Dept. Educ., *Spellings Hails New National Report Card Results*, <http://www.ed.gov/news/pressreleases/2005/07/07142005.html> (July 14, 2005) ("Today's Report Card is proof that No Child Left Behind is working—it is helping to raise the achievement of young students of every race and from every type of family background. . . . More than half of the progress in reading for 9-year olds during the Report Card's entire history has been made in the last five years. It is not a



statistically significant improvement in the average scaled NAEP reading scores of thirteen-year-olds and seventeen-year-olds over the same period (1999–2004).<sup>139</sup> Since NCLB's definition of AYP pivots not on the NAEP<sup>140</sup> but on state tests,<sup>141</sup> it is unclear whether NAEP proficiency gains are related to the Act. Nonetheless, any increases in NAEP scores are a good thing and may be at least partially attributable to implementation of the Act.

In other claims of success, the federal government has sought to take credit for some states' artificially<sup>142</sup> inflated progress. "Federal officials brag about the reduction in the number of schools failing to make adequate yearly progress (AYP) in 2004 in most states, forgetting to mention that changes in regulations are more responsible for the assumed improvement than any increase in the capacity of states and schools to significantly improve student learning."<sup>143</sup>

As the 2004 presidential election season grew more intense, the White House issued a detailed September press release publicizing NCLB's supposed successes.<sup>144</sup> Some of these claimed successes have undoubtedly been inflated by state evasion strategies.<sup>145</sup> For example, the press release noted that in Maryland the "percentage of African-American third graders scoring in the proficient range on state tests in reading increased 16 points in one year," and noted similar progress for Maryland's Hispanic fifth-graders.<sup>146</sup> President Bush similarly highlighted these figures.<sup>147</sup> Maryland, however, made its tests easier to

coincidence that progress accelerated so dramatically during this time period.").

139. See Natl. Ctr. Educ. Statistics, *National Trends in Reading by Average Scale Scores*, <http://nces.ed.gov/nationsreportcard/ltt/results2004/nat-reading-scalescore.asp> (last updated July 6, 2005). ("Thirteen-year-olds. The average score in 2004 was higher than the average score in 1971, but no difference from the average score in 1999 was found.") ("Seventeen-year-olds. There was no statistically significant difference between average scores in 1999 and 2004").

140. See Linn, *supra* n. 123 ("[T]he use of state-level NAEP results are not specified in the law . . .").

141. The President and former Secretary Paige also do not claim that this progress was fast enough to result in full proficiency by 2014. In fact, this progress is much too slow.

142. On the meaning of "artificial" as used here, see *supra* note 127 and accompanying text.

143. Lewis, *supra* n. 133.

144. See George W. Bush, *Education: The Promise of America*, § 3, <http://www.whitehouse.gov/news/releases/2004/09/20040926.html#3> (Sept. 26, 2004).

145. See *supra* n. 143 and accompanying text; *infra* nn. 148–158 and accompanying text. Perhaps some of these supposed successes present some genuine student improvement, and the Act may have facilitated some gains. Nevertheless, some of these supposed signs of success are artificial.

146. See *supra* n. 144.

147. See Bush, *The Essential Work of Democracy*, *supra* n. 26 ("In Maryland, the percentage of African American third-graders who are reading proficiently increased 16 percentage points in one year. The percentage of Hispanic fifth-graders achieving proficiency in math increased nearly 10 percentage points. And 25 schools exited school improvement status this year after meeting their performance objectives.").

pass for 2004,<sup>148</sup> and thus some of Maryland's improvement should be attributed to its easier tests.

The White House press release similarly pointed to North Carolina's increase in the percentage of schools achieving AYP, from forty-seven percent in 2003 to seventy percent in 2004.<sup>149</sup> At least some of these gains, however, derive from North Carolina's use of a confidence interval to calculate AYP. If North Carolina had used its confidence interval to calculate AYP in 2003, "about 200 more schools would have hit the [AYP] mark, raising the state's success rate from 47 percent to 57 percent."<sup>150</sup> The statistical legerdemain for 2004 has turned out to be stunningly higher than fifty-seven percent; a startlingly high seventy percent of North Carolina's schools achieved AYP, partially due to its use of a confidence interval.<sup>151</sup>

The press release also cites similar improvement in 2004 in Pennsylvania, a state that, like North Carolina, adopted a confidence interval in 2004.<sup>152</sup> Pennsylvania saw a nineteen percent increase in the proportion of schools making AYP in 2004 (from sixty-two to eighty-one percent).<sup>153</sup> But sixteen percent of this nineteen percent improvement occurred "because of the addition of a confidence interval."<sup>154</sup> In other words, eighty-four percent of increase in the proportion of Pennsylvania schools that made AYP in 2004 derived from the use of a confidence interval to calculate AYP. Pennsylvania has also adopted a Balloon Schedule<sup>155</sup> which artificially<sup>156</sup> lowers its early AMOs. The press release points to the AYP-related progress of eight other states.<sup>157</sup> But half of them (California, Delaware, Georgia, and Wisconsin) adopted the Balloon Schedule and therefore currently have artificially low goals.<sup>158</sup>

148. See *infra* nn. 187–188 and accompanying text (observing that Maryland's 2004 test scores may show illusory gains in proficiency because the state lowered its testing standards).

149. See *supra* n. 144.

150. Ann Doss Helms, *Many More Schools Expected to Meet Goals; But Those That Fail Must Let Students Transfer Out*, Charlotte Observer 1B (July 19, 2004).

151. See Emily S. Achenbaum, *Union County Schools Show Improvement*, Charlotte Observer 6U (July 22, 2004); Lynn Olson, *Data Show Schools Making Progress on Federal Goals*, 24 Educ. Week (Sept. 8, 2004) ("In North Carolina, 70 percent of schools met all federal AYP goals in 2003–04, up from 47 percent in 2002–03. That difference can be traced in part to the federal government's giving the state permission to use a 'confidence interval' . . .").

152. See Pa. Dept. Educ., *Adequate Yearly Progress 2003–04 Changes*, <http://www.pde.state.pa.us/pas/cwp/view.asp?a=3&q=104800> (accessed Feb. 13, 2005).

153. See Olson, *supra* n. 151.

154. *Id.* (quoting state director for assessment and accountability Carina Wong).

155. See *supra* tbl. 1.

156. On the meaning of "artificial" as used here, see *supra* note 127 and accompanying text.

157. See *supra* n. 144.

158. See *supra* table 1 for the AMO schedules of California, Delaware, Georgia, and Wisconsin. West Virginia's 2003 Accountability Workbook lacked sufficient information to

Although the Education Department has of necessity been modest in what it claims as success, taking even this kind of credit will be precarious when a larger share of schools fails to make AYP.<sup>159</sup> The federal government's interest in evasion strategies can be expected to grow stronger in the future when intermediate goals come due and further embellishment of AYP results is needed to prove that NCLB is "working."

### III. IN DEFENSE OF THE BALLOON SCHEDULE

#### *A. NCLB Gives States the Right to Define Proficiency Down*

NCLB's approach to state educational standards is out of tune with the rest of the Act. When it comes to the twelve-year timeline, the 100 percent proficiency goal, and its sanctions, the Act is rule-like. But regarding the substance of state educational standards, the Act imposes few or no constraints.<sup>160</sup> Noted education law expert Professor James E. Ryan has called this inconsistency "regulatory stringency and laxity,"<sup>161</sup> and has observed that it "could well prove disastrous. It will encourage states to lower their standards, make their tests easier, or lower the scores needed to be deemed proficient."<sup>162</sup>

NCLB includes two sets of educational standards: content standards and performance standards. "Content standards define the skills and knowledge that all students are expected to obtain and be able to demonstrate while performance standards define proficiency levels for skills and knowledge."<sup>163</sup> NCLB gives states the right to change, at any time, either or both sets of standards.<sup>164</sup> "Nothing in this part," says the Act, "shall prohibit a State from revising, consistent with this section,

---

categorize its AMO schedule as either a Balloon Schedule or not.

159. See *supra* sec. II.B (noting that most schools will eventually fail to make AYP).

160. See 20 U.S.C. § 6311(b)(1)(D) (providing that "Standards under this paragraph shall include—(i) challenging academic content standards in academic subjects that—(I) specify what children are expected to know and be able to do; (II) contain coherent and rigorous content; and (III) encourage the teaching of advanced skills; and (ii) challenging student academic achievement standards that—(I) are aligned with the State's academic content standards; (II) describe two levels of high achievement (proficient and advanced) that determine how well children are mastering the material in the State academic content standards; and (III) describe a third level of achievement (basic) to provide complete information about the progress of the lower-achieving children toward mastering the proficient and advanced levels of achievement"). This provision in the Act in no way limits states' authority to change or set their standards as they see fit.

161. Ryan, *supra* n. 25, at 944.

162. *Id.*

163. See Duran, *supra* n. 25, at 81.

164. See 20 U.S.C. § 6311(b)(1)(A)–(D).

any” educational standard relevant to NCLB.<sup>165</sup> The Act makes it illegal for the Education Department to require states to even *submit* their standards for review.<sup>166</sup>

If the Education Department wishes to prevent states from evading NCLB by defining proficiency down, it may not do so by diktat. Instead, the Department must induce states to voluntarily refrain from defining proficiency down—by allowing more effective and more benign evasions such as the Balloon Schedule, selection of minimum subgroup sizes, and the use of confidence intervals when determining AYP.

### *B. Effects of Defining Proficiency Down*

Unlike lowering educational standards, the Balloon Schedule, selection of minimum subgroup sizes, and confidence intervals have no obvious effect on educational standards. These three evasions have been created for the sake of NCLB’s accountability system.<sup>167</sup> As a result, they mainly manipulate AYP numbers that appear on federal accountability reports.

Defining proficiency down, on the other hand, is comparable to grade inflation and directly lowers educational standards.<sup>168</sup> Grade inflation occurs when average grades rise without a commensurate increase in student performance.<sup>169</sup> Defining proficiency down does exactly this: it increases the reward for the same level of performance by lowering the educational standard.<sup>170</sup>

165. *Id.* at § 6311(b)(1)(F) (“Nothing in this part shall prohibit a State from revising, consistent with this section, any standard adopted under this part before or after January 8, 2002.”).

166. *Id.* at § 6311(b)(1)(A) (providing that “a State shall not be required to submit such [academic content] standards to the [Education] Secretary”).

167. All of these devices are created by provisions in the Act. *See supra* secs. I.D–I.F.

168. *See infra* n. 169 and accompanying text.

169. *See* Robert Birnbaum, *Factors Related to University Grade Inflation*, 48 J. Higher Educ. 519, 522 (1977) (“In the context of grades, inflation can be viewed as a process in which a defined level of academic achievement results in a higher grade than awarded to that level of achievement in the past. In plain language, grade inflation implies that it is now ‘easier’ to get a high grade in a course than previously.”); Perry A. Zirkel, *Grade Inflation: Leadership Opportunity for Schools of Education?*, 101 Teachers College Rec. 247, 247 (1999) (“Grade inflation . . . is a rise in academic grades not accompanied by a commensurate increase in academic achievement.”).

170. It is worth noting that grade inflation does not always cause standards to drop—if grade expectations rise at the same rate that average grades do. Instead of educational standards dropping in such a case, the symbols representing the standards (and thus the exact same expectations) would merely *shift upward*. For example, if a college’s average grade level rose from “C” to “B,” employers’ expectations for graduates’ grades might also rise from “C” to “B.” The symbols representing the grades would have shifted upward, but employers’ rising expectations would leave the standard unchanged. Some employers and graduate schools today expect to see “high and relatively undifferentiated grades, and therefore rely on other criteria.” Henry Rosovsky & Matthew Hartley, *Evaluation and the Academy: Are We Doing the Right Thing? Grade Inflation and Letters*

When schools lower standards, they discourage higher levels of student learning.<sup>171</sup> The predictable result of defining proficiency down will be that students learn less and perform worse.<sup>172</sup> “[W]hen schools through inflated grades promote lowered standards . . . students routinely perform far below their capability, and they know it.”<sup>173</sup> Students “who are rewarded regardless of their performance eventually lose respect for their teachers and for the subject.”<sup>174</sup> Colorado may be a case in point. Since it defined proficiency down, Colorado’s fourth-grade reading scores on the NAEP<sup>175</sup> have dropped.<sup>176</sup>

The Act’s proponents understand this point well. According to the Education Trust, an advocacy group that supports NCLB, setting standards too low “ultimately stunt[s] the academic growth of our young people.”<sup>177</sup> President Bush has said of NCLB: “We’re challenging what I call the soft bigotry of low expectations. In other words, if you believe certain children can’t learn, they won’t. It’s just as simple as that. If you’ve got low expectations . . . you’ll achieve mediocre results.”<sup>178</sup> Former Education Secretary Rod Paige has similarly stated: “If expectations are high, then students will thrive. If expectations are low, then they will come to believe they are hopeless causes and they will surrender.”<sup>179</sup> One education researcher has even opined that the Act

*of Recommendation 12* (Acad. of Arts & Sciences, 2002). Of course, defining proficiency down for purposes of NCLB cannot fit into this relatively benign category of grade inflation because the “proficient” label’s value is fixed by law. Expectations cannot rise to match lowering standards unless NCLB is amended somehow. To be sure, states would want to define proficiency down only if it lowered standards, or doing so would not help evade NCLB.

171. See William Cole, *The Perils of Grade Inflation*, Chron. Higher Educ. B1 (Jan. 6, 1993) (“[B]y rewarding mediocrity, we discourage excellence.”).

172. See Clifford H. Edwards, *Grade Inflation: The Effects on Educational Quality and Personal Well-Being*, 120 Educ. 538 (Spring 2000) (“Grade inflation generally promotes lower academic standards and gives students a distorted view of their academic achievements and abilities.” (citations omitted)).

173. *Id.*

174. Gregory Stanley & Lawrence Baines, *No More Shopping for Grades at B-Mart: Re-establishing Grades as Indicators of Academic Performance*, 74 Clearing House 227, 228 (Mar. 2001).

175. See *supra* nn. 117, 137 and accompanying text (discussing the NAEP).

176. Susan Saulny, *State to State: Varied Ideas of ‘Proficient’*, N.Y. Times B8 (Jan. 19, 2005) (“Colorado’s proficiency rate fell to 37 percent on the national test, but that score was high enough to rank fifth in the nation.”). Of course, this fact *alone* does not prove that Colorado’s lowering of standards caused the entire drop, but the drop is consistent with the theory that defining proficiency down lowers student achievement.

177. Educ. Trust, *Stalled in Secondary: A Look at Student Achievement Since the No Child Left Behind Act 2*, <http://www2.edtrust.org/NR/rdonlyres/77670E50-188F-4AA8-8729-555115389E18/0/StalledInSecondary.pdf> (accessed Jan. 23, 2005).

178. Bush, *Remarks*, *supra* n. 27.

179. Matthews, *supra* n. 27.

might never have been needed but for rampant grade inflation.<sup>180</sup>

*C. Banning the Balloon Schedule Would Encourage States to Define Proficiency Down*

If the Balloon Schedule and its cousins were banned, states would be encouraged to inflate their proficiency results by defining proficiency down.<sup>181</sup> The “pressures to comply” with NCLB have already caused some states to “ease their standards for what it means to be ‘proficient’ in reading and math.”<sup>182</sup> For example, Colorado, Connecticut, Louisiana, Maryland, Texas, Utah, and Washington have made it easier for their students to test proficiently.<sup>183</sup> A Utah school curriculum director has said that she cannot be sure how much of her school’s statistical improvement in 2004 derived from the state’s lower “cut” scores.<sup>184</sup> Washington State estimates that four to nine percent more of its seventh-graders will test proficiently in reading and math in the coming year, as a result of a lower “cut” score on its test.<sup>185</sup> An assistant superintendent candidly remarked of this change: “The results are going to *look like* they have improved.”<sup>186</sup> Maryland also “set easier passing standards [for its tests] so that double the percentage of students were deemed ‘proficient’ under the new state tests than were deemed so under the old tests.”<sup>187</sup> According to one editorial writer’s assessment of Maryland’s higher scores in 2004, “‘improvements’ in test scores may be largely an

180. See John Stone, *Leaving Children Behind: Grade Inflation a Growing Problem*, Wash. Times A19 (Aug. 31, 2004).

181. See Marion et al., *supra* n. 86, at 13; Sam Dillon, *Playing the Standards Game*, N.Y. Times 2 (May 25, 2003); Sam Dillon, *States Are Relaxing Education Standards to Avoid Sanctions from Federal Law*, N.Y. Times A29 (May 22, 2003); David J. Hoff, *States Revise the Meaning of ‘Proficient’*, Educ. Week, 1 (Oct. 9, 2002); Erpenbach et al., *supra* n. 30, at 7; *No Child Act Leaves Many in Perplexity*, Lincoln J. Star (Lincoln, Neb.) (Dec. 3, 2003), (available at [http://www.journalstar.com/articles/2003/12/03/editorial\\_main\\_ljs/10041623.txt](http://www.journalstar.com/articles/2003/12/03/editorial_main_ljs/10041623.txt)) (quoting former U.S. Education Secretary Rod Paige: “Unfortunately, some states have lowered the bar of expectations to hide the performance of their schools.”); Gary W. Ritter & Christopher J. Lucas, *Puzzled States*, 3 Educ. Next 54, 60 (2003). See Matthew H. Joseph, . . . *And in Maryland*, Wash. Post B8 (Aug. 1, 2004); Linda Shaw, *Panel Lowers Bar for Passing Parts of WASL*, Seattle Times B2 (May 11, 2004).

182. Marion et al., *supra* n. 86, at 13 (internal quotation marks omitted).

183. See Joseph, *supra* n. 181 (Maryland); Ronnie Lynn, *2 Jordan Schools Fail Feds’ Test; District Must Pay to Transport Students Who Transfer*, Salt Lake Trib. C1 (Nov. 1, 2004); Ryan, *supra* n. 25, at 947–48; David Wickert, *WASL Tweaks Bring Standards within Reach; Scoring Changes: Schools to Get Higher Marks, Even if Students Don’t*, Morn. News Trib. (Tacoma, Wash.) A1 (Oct. 21, 2004).

184. Lynn, *supra* n. 183.

185. Wickert, *supra* n. 183.

186. *Id.* (emphasis added).

187. Joseph, *supra* n. 181.

illusion.”<sup>188</sup>

The states that have lowered their standards are relatively few in number—seven, to be exact.<sup>189</sup> Due to the Education Department’s approval of so many effective methods for inflating AYP numbers, states apparently do not feel a great need to lower their standards so far.

The relatively low number of states defining proficiency may also be partially due to the Act’s requirement that some students from every state take the NAEP. The Act provides that each state’s accountability workbook must contain assurances that, by the 2002–2003 school year, it will “participate in biennial State academic assessments of 4th and 8th grade reading and mathematics under” the NAEP.<sup>190</sup> Popularly known as “the Nation’s Report Card” and administered by a sub-unit of the U.S. Department of Education, the NAEP is the “the only nationally representative and continuing assessment of what America’s students know and can do in various subject areas.”<sup>191</sup>

The NAEP may help encourage states to keep standards high in the following manner: if they *lower* their standards, states may experience a drop in their students’ NAEP performance,<sup>192</sup> leading to embarrassment and perhaps public criticism of whatever leaders lowered the standards. Such consequences appear to be the *only* incentives supplied by the Act’s NAEP requirement, because AYP does not account for NAEP scores.<sup>193</sup> Improvements or drops in NAEP scores alone—holding all other variables constant—would have no effect on any school’s AYP status.<sup>194</sup> Thus, while it must be acknowledged that the NAEP does provide some incentive, however small, to keep standards high, the Act’s accountability system pulls strongly in the opposite direction—and successfully, if the examples of the seven states that have lowered

188. *Id.*

189. These states are: Colorado, Connecticut, Louisiana, Texas, Utah, Washington, and Maryland. *See supra* n. 183 and accompanying text. It should be noted that Missouri seems to be planning to lower its testing “cut” scores sometime in 2006. *See* Ltr. from Thomas R. Davis, Pres., Mo. Dept. of Elementary and Secondary Educ., to all Missouri Educators, (Feb. 18, 2004) (available at <http://www.dese.state.mo.us/stateboard/letters/2004/02182004.html>) (“The stated goal of NCLB is for all children to be proficient in reading and math by 2014. However, the annual testing that is required to measure our progress toward that goal will not be available until 2006. At that time all of Missouri’s definitions (such as ‘proficiency’) will have to be redone to create new cut scores on the annual tests.”). The Missouri Board of Education initially declined to lower “cut” scores. *See* Deann Smith, *State School Board Maintains Current Testing Standards*, *Kansas City Star* (Apr. 17, 2003) (available at <http://www.kansascity.com/mld/kansascity/news/local/5649270.htm?1c>).

190. 20 U.S.C. § 6311(c)(2).

191. Natl. Ctr. for Educ. Statistics, *supra* n. 117.

192. For one possible example of this, *see e.g. supra* n. 176 and accompanying text.

193. The Act nowhere incorporates the NAEP into its definition of AYP. *See* 20 U.S.C. § 6311(2)(b)(1)(B)–(C) (setting forth the definition of “adequate yearly progress”).

194. *See id.*

standards are any indication.<sup>195</sup> Indeed, as pressure from the Act's accountability system builds, it is easy to imagine states preferring some shame and embarrassment from lower NAEP scores to rising numbers of schools being subject to improvement, corrective action, and restructuring.<sup>196</sup>

Rather than lower education standards, states have thus far much preferred to employ the evasion devices blessed by the Education Department: the Balloon Schedule, selection of minimum subgroup sizes, and confidence intervals. Recall that twenty-two states opted for the Balloon Schedule in their 2003 accountability workbooks.<sup>197</sup> A majority of states use a confidence interval.<sup>198</sup> Over a dozen states have arguably manipulated minimum subgroup sizes, with more considering this tactic.<sup>199</sup> But only two of the states that have lowered educational standards (Louisiana and Texas) also use the Balloon Schedule.<sup>200</sup> It is also telling that, of the thirteen states whose governors announced in early 2005 their plans to raise educational standards, ten (seventy-seven percent) are Balloon Schedule states.<sup>201</sup> In all, about half of all states are Balloon Schedule states.<sup>202</sup>

The pressure to define proficiency down, if these other strategies were eliminated, could be enormous. For schools in Balloon Schedule states, AMOs would immediately jump if the Balloon Schedule were banned. Witness Georgia, a state where middle school reading AMOs for 2006 could soar above currently prescribed levels by as much as seven percent.<sup>203</sup> In North Carolina, Pennsylvania, and other states, confidence intervals have aided hundreds of schools in making AYP.<sup>204</sup> Likewise, without selection of subgroup sizes many schools would fail to make

195. *See supra* n. 189 and accompanying text.

196. For an overview of sanctions such as school improvement, corrective action, and restructuring, *see supra* section II.A.

197. *See supra* tbl. 1.

198. *See supra* n. 81 and accompanying text.

199. *See supra* n. 72 and accompanying text.

200. *See supra* tbl. 1.

201. Robert Pear, *Governors of 13 States Plan to Raise Standards in High Schools*, N.Y. Times A13 (Feb. 28, 2005) ("The 13 states are Arkansas, Georgia, Indiana, Kentucky, Louisiana, Massachusetts, Michigan, New Jersey, Ohio, Oregon, Pennsylvania, Rhode Island and Texas. Other states are expected to join the coalition in the next few weeks.") Of these, only Arkansas, Kentucky, and Massachusetts are not Balloon Schedule states. *See supra* tbl. 1. The lists of states cited so far stand out as bizarre in this respect: two of the states whose governors are now calling for higher standards (Louisiana and Texas) are Balloon Schedule states *and* have already lowered their standards.

202. *See supra* tbl. 1.

203. *See supra* n. 129 and accompanying text.

204. *See supra* nn. 150–154 and accompanying text.



AYP due to their special education students.<sup>205</sup> Any crackdown on these strategies would likely contribute to an increase in the number of states lowering educational standards beyond its currently modest number of seven.<sup>206</sup>

### CONCLUSION

In an effort to save the No Child Left Behind Act from its own unrealistic requirements, the Education Department has allowed states to heavily backload planned student proficiency gains, to change minimum subgroup sizes in ways that exclude disadvantaged students from accountability reports, and to use confidence intervals to calculate AYP. Although allowing these evasions may violate the Act, their use is arguably the right thing to do because it reduces the incentive for states to do something even worse: harmfully define educational proficiency down.

---

205. See e.g. Karen Hill, *Making the Grade: Special Education: Acceptable Score Hinges on Progress of Disabled If Too Few Pass, Whole School Fails*, Atlanta J. Const. D6 (Aug. 1, 2004) (“Although failure in any one of 14 subgroups would flunk an entire school, it is the ‘disabled’ subgroup that shoved many schools in metro Atlanta off the list.”).

206. See Ryan, *supra* n. 25, at 114–16 (observing that states have four options for avoiding failure under NCLB: (1) raise proficiency to 100 percent on the merits; (2) use an evasion strategy, such as the Balloon Schedule; (3) opt-out of Title I funds; and (4) lower educational standards).