

Do Discussion, Debate, and Simulations Boost NAEP Civics Performance?

By Kei Kawashima-Ginsberg
With support from the S.D. Bechtel, Jr. Foundation

April 2013

Ten years ago, the Civic Mission of Schools report ([Gibson & Levine, 2003](#)) clarified goals of civic education and identified six promising practices of civic education pedagogy. Three of these practices were measured on the National Assessment of Educational Progress (NAEP) Civics test in 2010: discussing current events, debating current issues (including controversies), and participating in simulations of democratic processes and procedures.

In this fact sheet, we explore who had access to the three recommended practices, whether these instructional practices were associated with higher NAEP scores, and whether the effects of these practices varied for different demographic groups.

The NAEP Civics assessment only measures certain kinds of knowledge: predominantly, abstract knowledge about perennial features of the US political system. The promising practices recommended by the Civic Mission of Schools have additional major objectives, such as learning about current events, and learning to deliberate and collaborate with other citizens. Furthermore, the NAEP's criteria for "proficiency" are somewhat arbitrary (see [this CIRCLE Fact Sheet](#) for more details). Nevertheless, it is important to know whether recommended teaching practices such as discussion, debate, and simulations boost NAEP scores. If they do, that is a strong argument for using these approaches. If they do not, educators may face something of a tradeoff between teaching the concepts tested on the NAEP and using interactive civics pedagogy.

At the 8th grade and 12th grade level, we found that White students and students from higher socioeconomic backgrounds received more of the promising practices. Exposure to these practices was associated with higher NAEP scores for all groups, although the recommended pedagogies did not compensate for gaps in NAEP scores. Indeed, more advantaged and White students seemed to derive more benefit than their less advantaged and ethnic minority peers when they received the same level of exposure to these pedagogies. Perhaps other aspects of their school and community environments

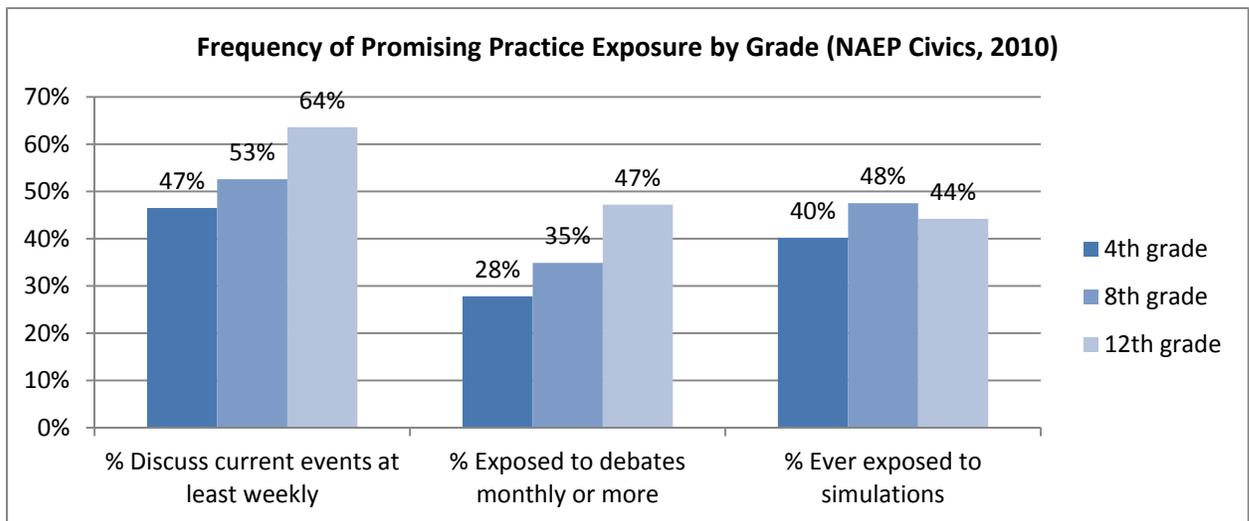
are more favorable to civic learning.

These patterns were not seen at the 4th grade level – in fact, the promising practices were associated with lower NAEP scores, raising questions that require further research. The 4th grade data may not be reliable, or elementary teachers (with very limited time for civics) may face a tradeoff between teaching civics information and teaching current events and civil dialogue.

How Frequently Are Students Exposed to the Recommended Practices?

First, we assessed how often students were exposed to the three promising practices captured by the NAEP data. (We could not investigate service-learning, extracurricular participation, or student voice in schools—the three remaining promising practices—because they were not measured by the NAEP.) We set the cut-off for regular exposure at different levels for each practice. For example, we did not expect teachers to set up simulation exercises on a daily or even weekly basis, but discussion of current events could occur more regularly. After examining the distribution of these experiences, we established a cut-off for each of these four practices as indicated in the following figure. For detailed distribution tables, please refer to the [supplement page](#).

Debates and current events discussions happened more frequently for older students. For simulations, there was no clear pattern related to age, but a little less than half of all students experienced simulations. From the data we have, it is not clear which specific types of simulation activities occurred at specific grades, but examples might include Model United Nations, simulated voting, or mock trials.



Who gets promising practices?

We examined whether exposure to each of the three promising practices varied by demographic backgrounds. At 12th grade, Hispanic students were overall less likely to experience current events discussions, debates, and simulations compared to other racial groups. Parental education and lunch-program eligibility (both indicators of a family's socioeconomic status) were highly predictive of exposure to high quality civic education (Table 1).

Table 1: Exposure to High Quality Civic Education Practices by Demographic Backgrounds, 12th grade

Practice	Current events (Weekly or more)	Debates (Monthly or more)	Simulation (Ever)
All 12th graders	63.6%	47.2%	44.2%
Boys	62.9%	47.6%	44.9%
Girls	64.2%	46.8%	43.5%
White students	64.7%	48.2%	44.9%
African American students	62.5%	48.0%	47.7%
Hispanic students	60.3%	42.9%	37.8%
Asian-Pacific Islander students	63.7%	47.2%	46.5%
Student whose parents have Less than High School Diploma (LTHS)	53.8%	41.6%	37.9%
Students whose parents have high school diploma	58.1%	40.4%	39.3%
Students whose parents have some college	64.8%	47.9%	43.0%
Students whose parents completed college	66.6%	50.2%	47.6%
Students who are eligible for free- or reduced-lunch program	58.5%	43.6%	40.9%
Students who are ineligible for lunch program	65.3%	48.8%	45.6%

Source: CIRCLE's analysis of the NAEP 2010 Civics Assessment restricted data, 12th grade (NCES, 2012). Sample includes the entire reporting sample, including ELL students and students with disabilities.

A generally similar pattern is evident in the 8th grade data (Table 2). Hispanic students and students who had less educated and/or affluent families were less likely to be exposed to all of the promising practices. For 8th graders, African American students are slightly more likely to report current events discussions and simulations compared to the overall 8th grade rates and compared to Hispanic and White students.

Table 2: Exposure to High Quality Civic Education Practices by Demographic Backgrounds, 8th grade

Practice	Current events (Weekly or more)	Debates (Monthly or more)	Simulation (Ever)
----------	---------------------------------	---------------------------	-------------------

All 8th graders	52.6%	34.9%	47.5%
Boys	51.7%	35.4%	47.1%
Girls	53.5%	34.4%	48.0%
White students	53.8%	34.9%	48.1%
African American students	56.5%	38.5%	53.1%
Hispanic students	45.4%	31.2%	40.6%
Asian-Pacific Islander students	52.6%	37.7%	53.0%
Student whose parents have Less than High School Diploma (LTHS)	48.7%	31.2%	41.2%
Students whose parents have high school diploma	49.9%	32.4%	40.8%
Students whose parents have some college	54.3%	36.5%	47.5%
Students whose parents completed college	53.7%	35.9%	51.1%
Students who are eligible for free- or reduced-lunch program	51.9%	34.0%	43.7%
Students who are ineligible for lunch program	53.1%	35.6%	50.2%

Source: CIRCLE's analysis of the NAEP 2010 Civics Assessment restricted data, 8th grade (NCES, 2012). Sample includes the entire reporting sample, including ELL students and students with disabilities.

Fourth-grade data show almost the opposite trend (Table 3). African American *and* Hispanic students were *more* likely to report experiencing some of the best-practice pedagogies than White and Asian/Pacific-Islander students. African American and Hispanic 4th graders were more likely to experience current events discussions and debates compared to their peers, and African American students were also slightly more likely to engage in simulation than other groups. Similarly, 4th graders who were eligible for the free or reduced-price lunch program were *more* likely to be exposed to debates and current event discussions than those who were not eligible. It is certainly possible that these trends are real, but questions could be raised about whether 4th graders in general are able to report pedagogical practices in civics accurately.

Table 3: Exposure to High Quality Civic Education Practices by Demographic Backgrounds, 4th Grade

Practice	Current events (Weekly or more)	Debates (Monthly or more)	Simulation (Ever)
All 4th graders	45.9%	29.4%	40.8%
Boys	44.5%	30.3%	40.7%

Girls	47.4%	28.6%	40.9%
White students	42.9%	27.0%	40.6%
African American students	54.7%	35.1%	43.4%
Hispanic students	47.5%	32.3%	40.8%
Asian-Pacific Islander students	43.4%	27.9%	37.1%
Students who are eligible for free- or reduced-lunch program	49.2%	33.6%	41.6%
Students who are ineligible for lunch program	43.2%	25.9%	40.0%

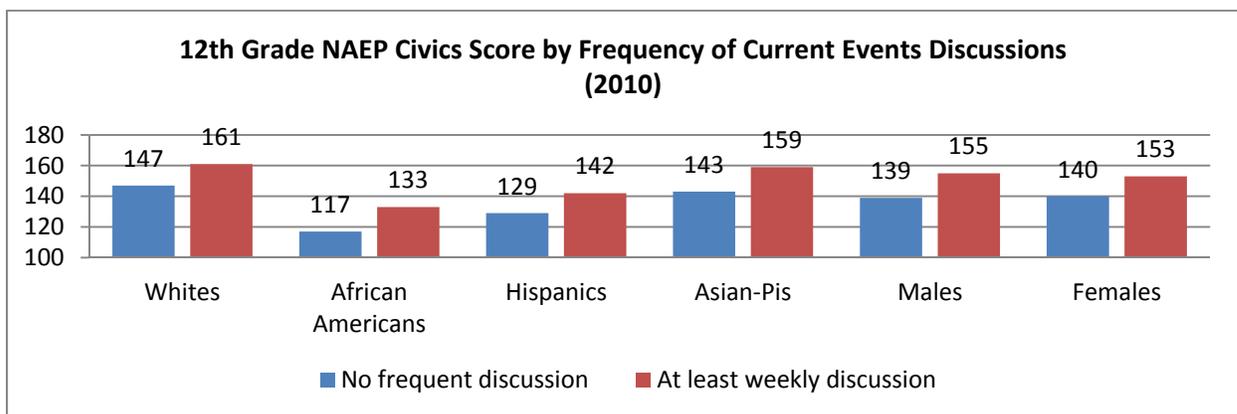
Source: CIRCLÉ's analysis of the NAEP 2010 Civics Assessment restricted data, 8th grade (NCES, 2012). Sample includes the entire reporting sample, including ELL students and students with disabilities.

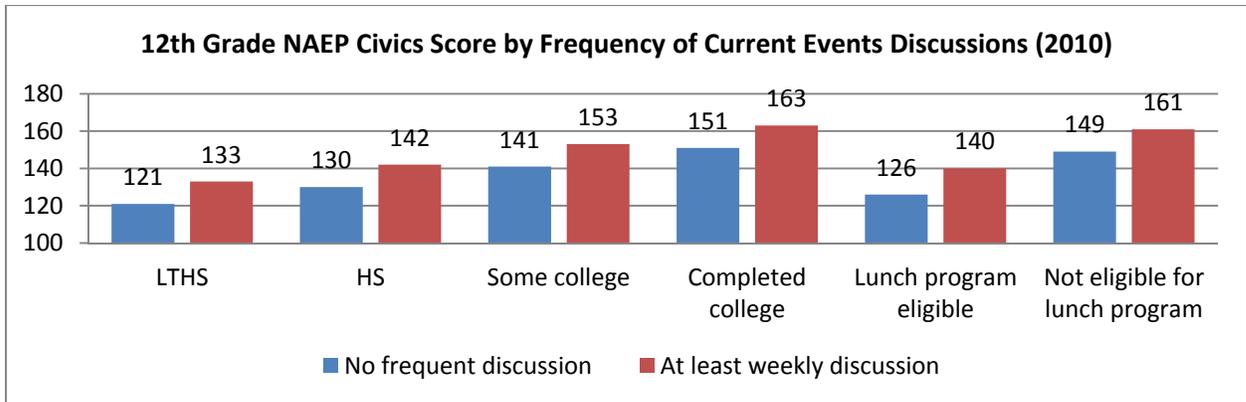
*Note: Parent education is not assessed among 4th graders.

Discussion of Current Events in Social Studies and Civics Scores

We identified students who were exposed to current events discussion on a regular basis: those who had experienced this feature at least once a week. We looked at the relationship between receiving the promising practices and NAEP Civics scores. It appears that current events discussion was most beneficial for the 12th graders, modest for 8th graders, and inconsistent for 4th graders. Importantly, for 12th graders, regular discussion of current events compensated for some of the gaps related to socioeconomic background.

For 12th graders, regular discussion of current events had a positive effect on males (16 point gain), females (13 point gain), White students (14 point gain), Black students (16 point gain), Hispanic students (13 point gain), and Asian students (16 point gain). These gains were all statistically significant. Current events discussion narrowed civic achievement gaps observed between students of different backgrounds in some cases. For example, students whose parents had high school diplomas but were exposed to regular discussion in school performed at the same level as students whose parents had some college experience but did not engage in discussion regularly (142 compared to 141).





We also calculated a group percentile rank and effect size for students who did or did not receive regular current event discussions. We found that the students who did receive current events discussions tended to place far ahead of the students who did not engage in frequent discussions in percentile ranking. For example, a typical African American student who did not engage in weekly discussion placed at 20th percentile (of all 12th graders), while an African American student who did engage in weekly discussion would have placed at the 35th percentile. The effect size ranged from “small” to “medium”¹ (See Table 5).

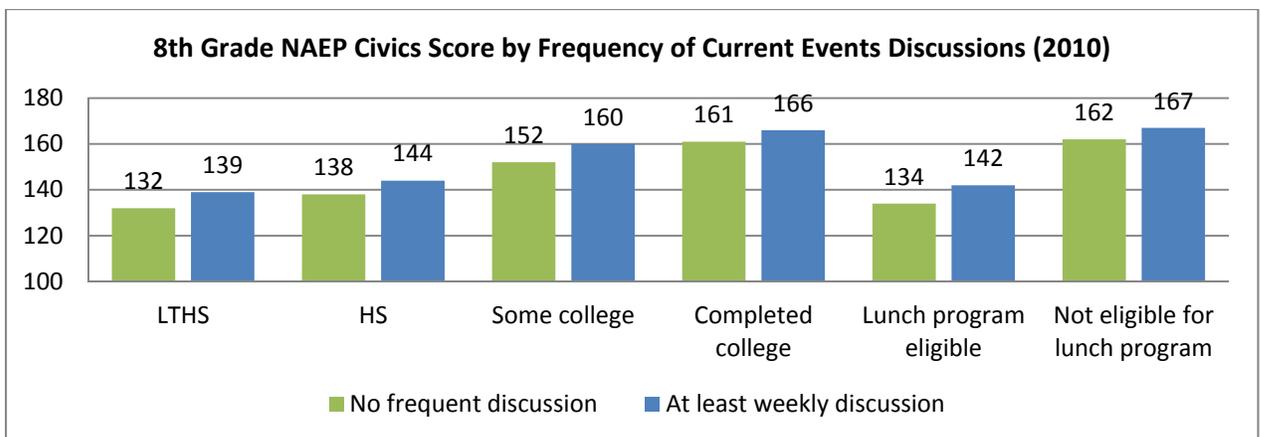
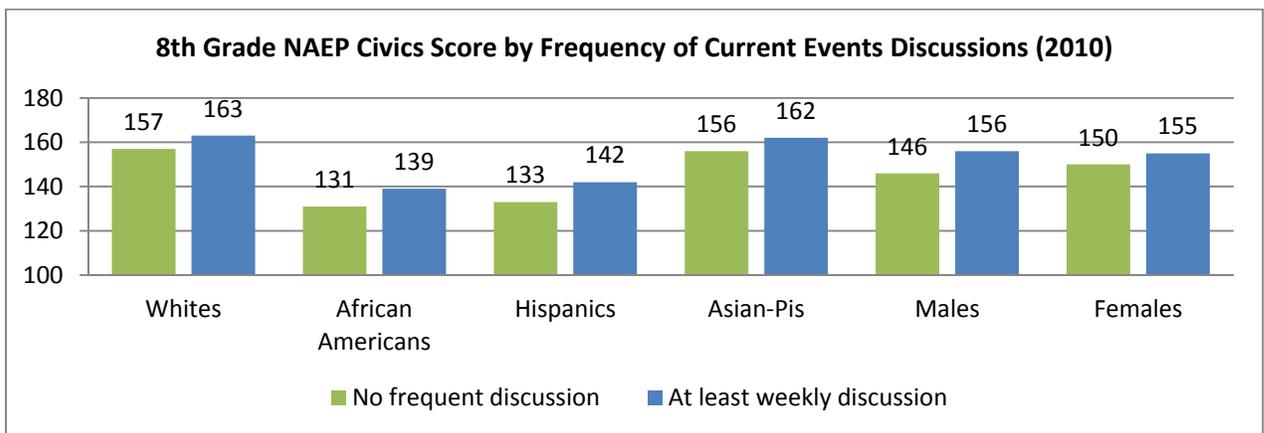
Table 5: Overall Group Percentile Rank and Effect size for Weekly Discussion of Current Events among 12th Graders

Group	Overall Percentile rank		Group Percentile Rank		Effect size within group
	Infrequent discussion	Frequent discussion	Infrequent discussion	Frequent discussion	
White students	50	66	39	56	.44
African American students	20	35	39	58	.47
Hispanic students	31	44	41	56	.37
Asian students	46	63	39	57	.44

¹ Effect size is a commonly accepted unit of measurement for the size of difference between two groups, such as a group that received a specific type of practice and a group that did not. Effect size is more useful than a statistical significance test in many cases because it is a metric for the size of the difference, rather than an indicator of whether there is any difference at all (i.e., significance test). It is calculated by dividing the group average score difference by pooled standard deviation. According to Cohen (1992, p. 157), an effect size of .20 or smaller is considered a “small effect,” an effect size of around .50 is considered a “medium effect,” and an effect size of .80 or larger is considered a “large effect.”

Male students	41	59	41	58	.43
Female students	42	57	41	56	.39
Parents with LTHS	24	32	39	56	.35
Parents with HS	28	45	38	56	.34
Parents with some college	38	56	46	58	.38
Parent with college degree	54	67	39	56	.38
Lunch program eligible	25	40	45	59	.40
Not lunch program eligible	51	64	46	58	.38

Eighth grade data revealed a similar but weaker effect of discussions. The biggest gain from regular discussion of current events was observed among males (10 points) but the effect was weaker for other groups such as females, students who parents have college degrees, and students who do not qualify for fee- or reduced lunch programs (5 points each).



The effect size and percentile ranking figures (shown in Table 6) indicate that the 8th graders who were engaged in regular current-events discussions did modestly better than the 8th graders in the same demographic group who did not have access to such opportunities. In all the demographic subgroups, the difference would be considered a “small” effect.

Table 6: Overall Group Percentile Rank and Effect size for Weekly Discussion of Current Events among 8th Graders

Group	Overall Percentile rank		Group Percentile Rank		Effect size within group
	Infrequent discussion	Frequent discussion	Infrequent discussion	Frequent discussion	
White students	57	64	46	54	.20
African American students	28	36	45	55	.24
Hispanic students	30	40	46	56	.26
Asian students	56	63	48	55	.18
Male students	44	56	49	54	.29
Female students	48	55	49	55	.15
Parents with LTHS	29	36	48	56	.22
Parents with HS	35	42	49	56	.19
Parents with some college	49	63	43	59	.41
Parent with college degree	61	67	49	55	.17
Lunch program eligible	31	40	48	57	.24
Not lunch program eligible	63	68	49	56	.17

For fourth graders, the effect of regular discussion was only significant among African American students and students who were eligible for a free- or reduced-lunch program, and the gains were modest. For all other groups, regular discussion of current events did not make a significant difference in the performance score. The effect size and percentile ranking findings show the same trend, which is that the students who were regularly engaged in current-events discussion performed very similarly to the students from the same demographic group who did not.

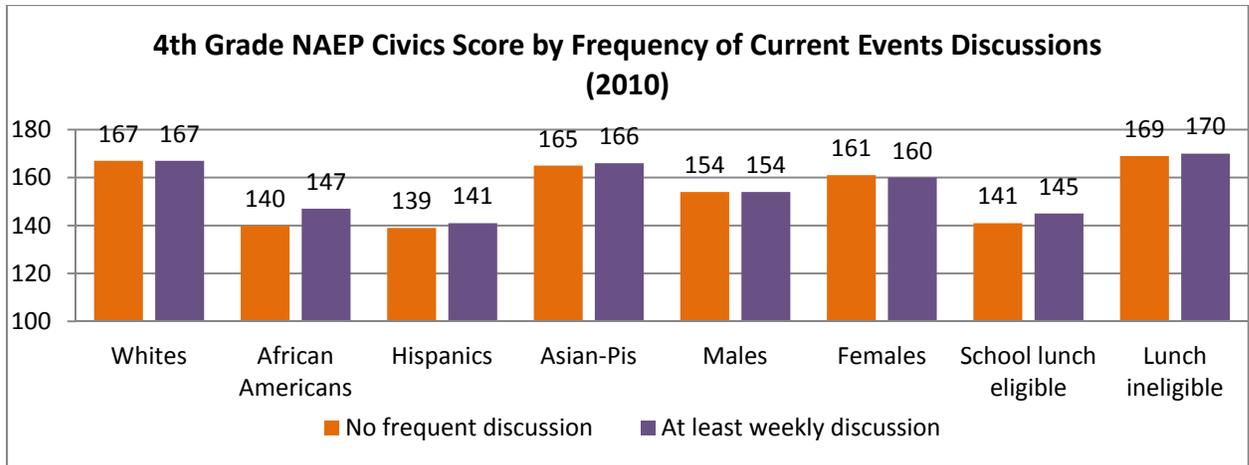


Table 7: Overall Group Percentile Rank and Effect size for Weekly Discussion of Current Events among 8th Graders

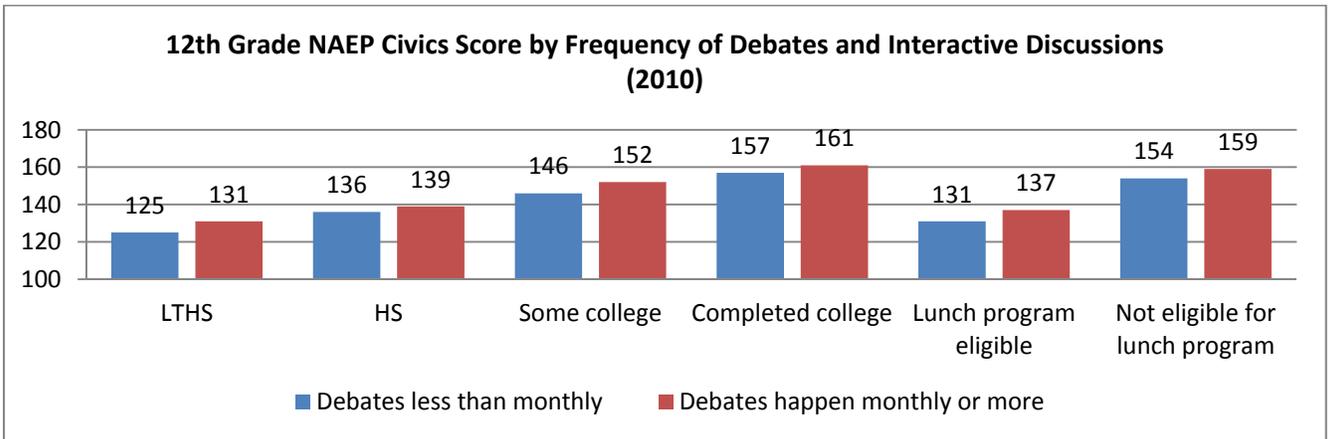
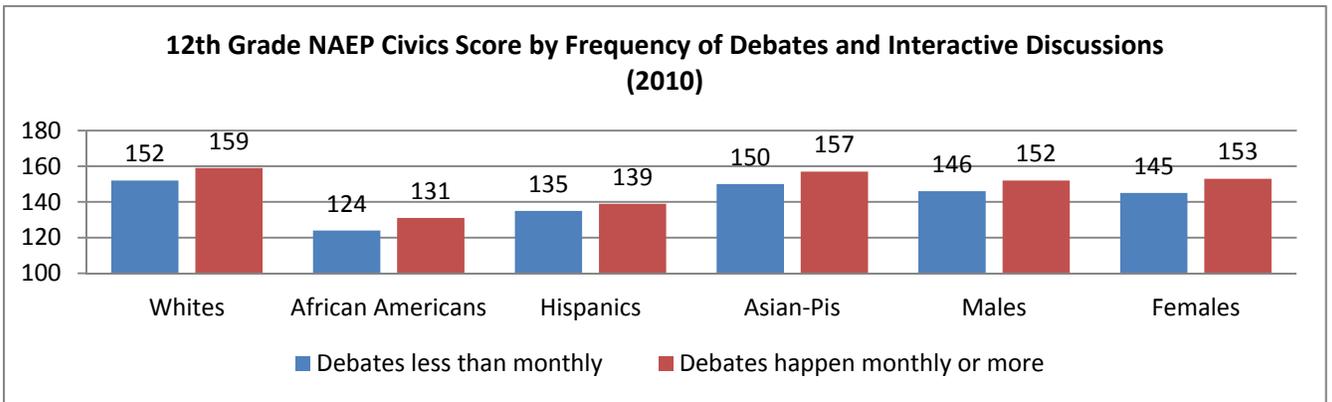
Group	Overall Percentile rank		Group Percentile Rank		Effect size within group
	Infrequent discussion	Frequent discussion	Infrequent discussion	Frequent discussion	
White students	62	62	50	50	0
African American students	30	38	46	55	.23
Hispanic students	29	31	49	51	.06
Asian students	60	61	51	52	.03
Male students	46	46	51	51	0
Female students	55	54	51	50	-.03
Lunch program eligible	31	35	47	53	.13
Not lunch program eligible	64	66	50	52	.04

Debates and Interactive Discussions and Civics Performance

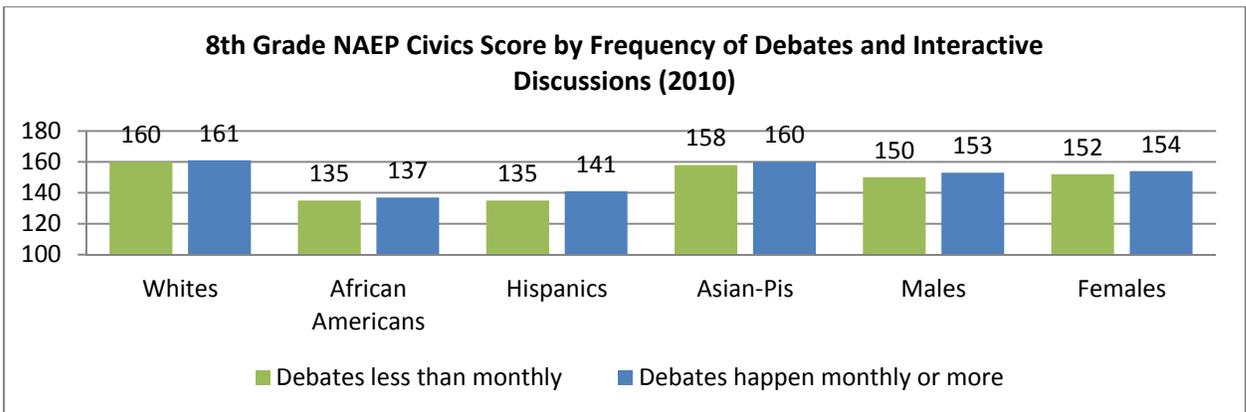
Consistent with the finding about current-events discussion, we found that the effect of debates and interactive discussion was strongest for 12th graders, but mixed for 8th and 4th graders. Within each grade, there were some variations by demographic group in the extent to which debates had positive effect.

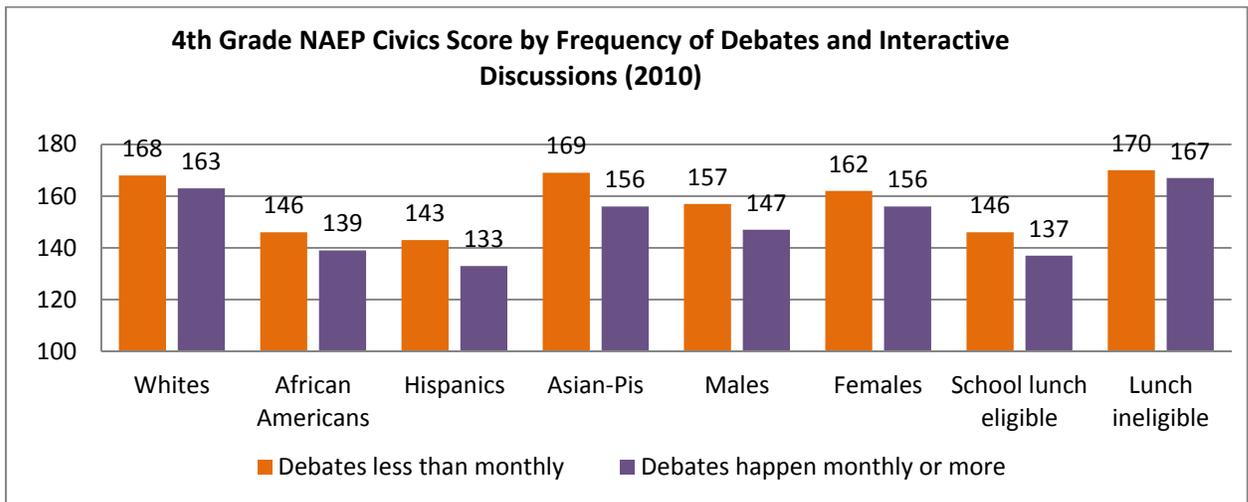
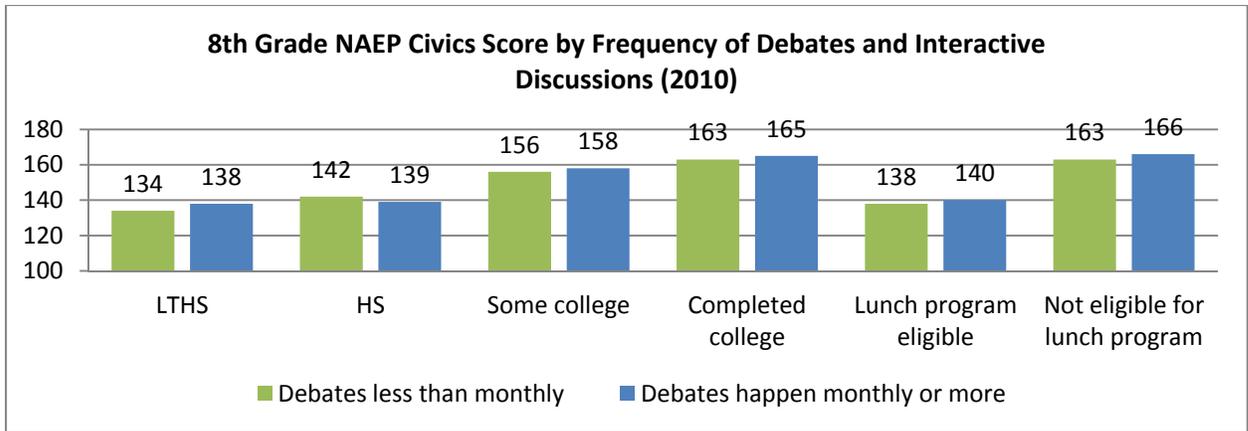
Debate had a modest but significant effect for 12th graders. Male 12th graders gained 6 points from this practice, and females gained 8 points. For racial groups, the findings were similar, but non-significant for Hispanic and Asian students, partly due to smaller

sample sizes. Hispanic students who were exposed to debates scored only 4 points higher than their peers who did not have this opportunity, a less significant gain than White and African American students (+7 points each).



For 8th graders, debates had no apparent effect. Across groups, the effect of debates was non-significant, with the exception of Hispanic students, who showed a 6-point gain when they had exposure to debates at least monthly.





For the 4th graders, the effects of debates were puzzling. The students who were exposed to debates regularly performed at lower levels than students who did not engage in debates, in virtually all subgroups of students. Interestingly, we found that students who were English Language Learners and students with disabilities were more likely to say that they engaged in debates frequently than those who do not fall into either of these categories. The apparent negative effect of debates was smaller when we excluded ELL students and students with disabilities, but the negative effect persisted.² The difference was quite large for most of these groups, and it raises a question about why debates do not seem to work for 4th graders, and what kind of debates are occurring in these classrooms. As noted above, it could be that reports of teaching practices are unreliable or invalid at the 4th grade level, but if the pattern is real, it deserves attention.

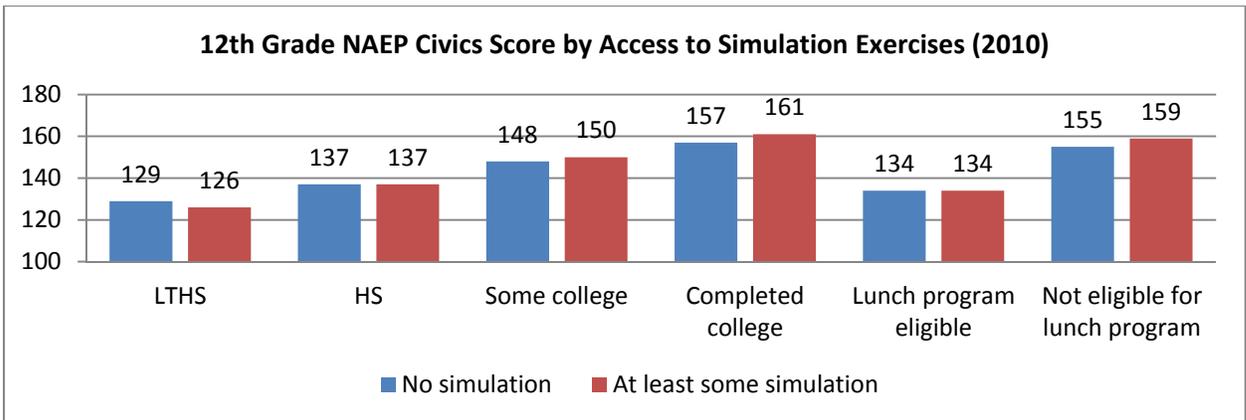
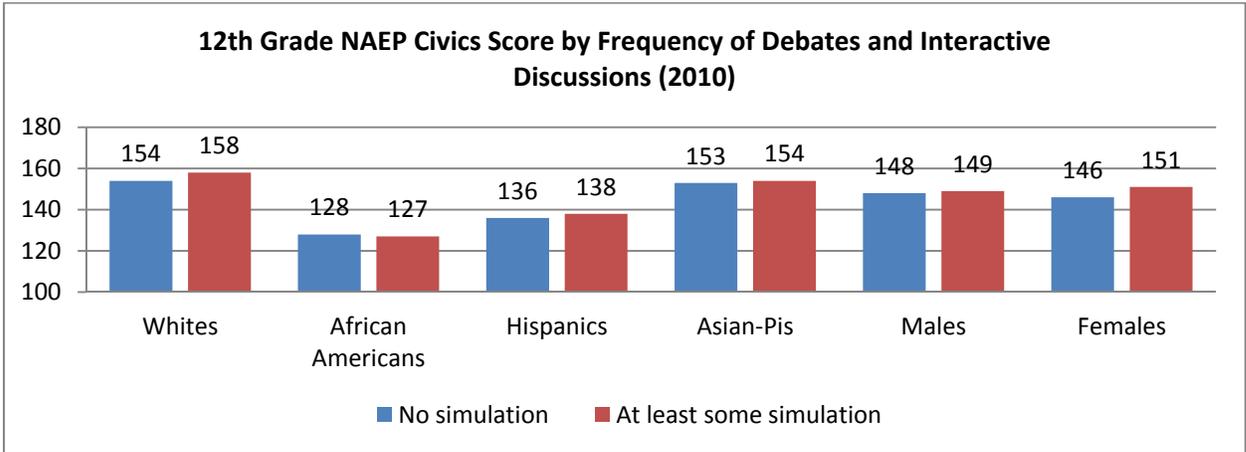
Simulations and Civics Score

We also tested whether having any opportunity to participate in a simulation affected Civics scores. Our criterion was that students had to be exposed to this type of opportunity at least once, based on the assumption that opportunities like this do not happen on a regular basis in most schools. Simulations were relatively rare across

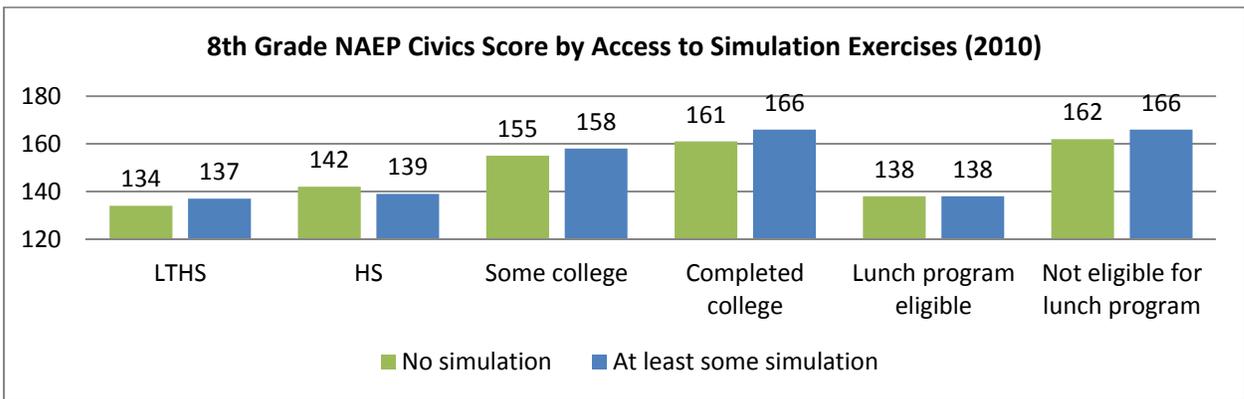
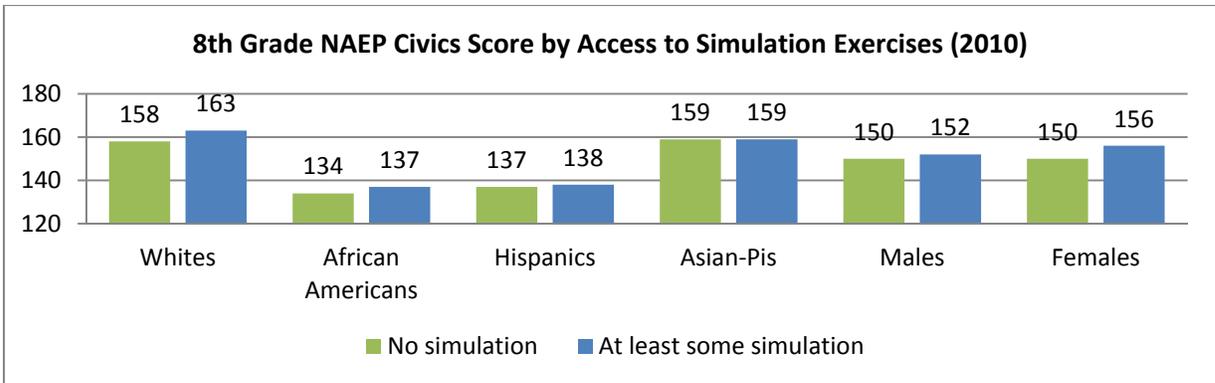
² Please contact the author for detailed figures.

grades—more than half of all students had never engaged in any simulation activities at all.

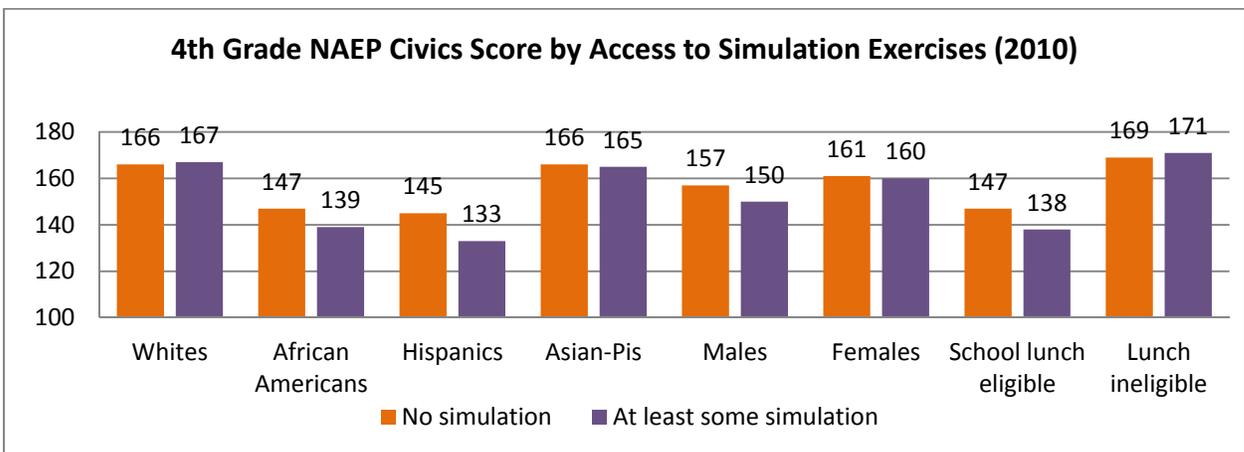
Among 12th graders, simulation activities were beneficial mostly for White and/or female students, and for students who were relatively advantaged (i.e., students with college-educated parents and those who are not eligible for school-lunch program). Simulations had no apparent benefit for racial minority students, males, and for students from less advantaged backgrounds.



The 8th grade data on simulations showed a similar pattern. Simulations were helpful for females, White students, and students whose parents had college degrees and/or students who were ineligible for school-lunch program. Simulations had no effect for other groups.



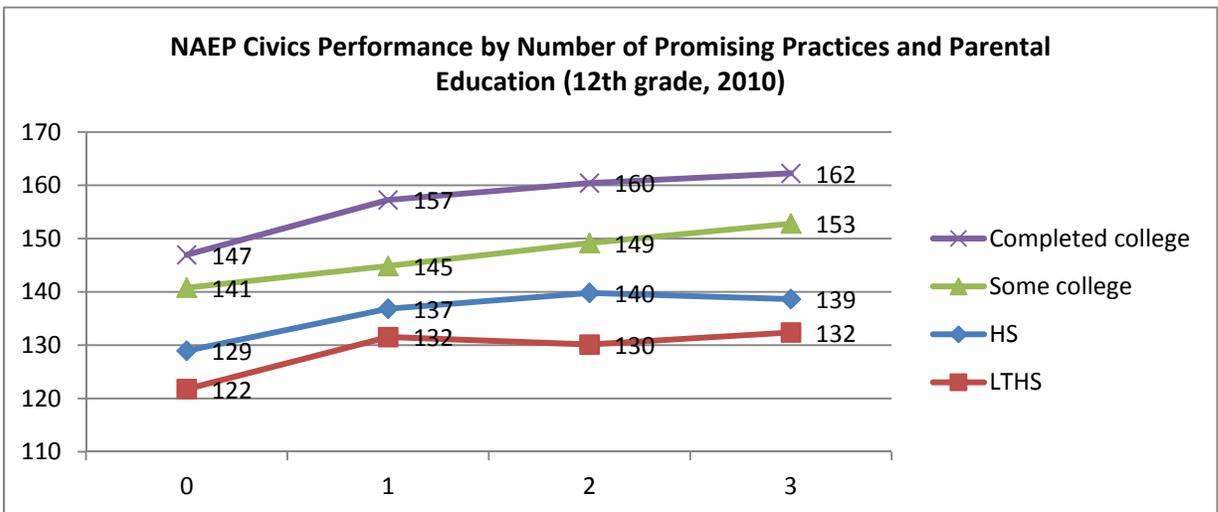
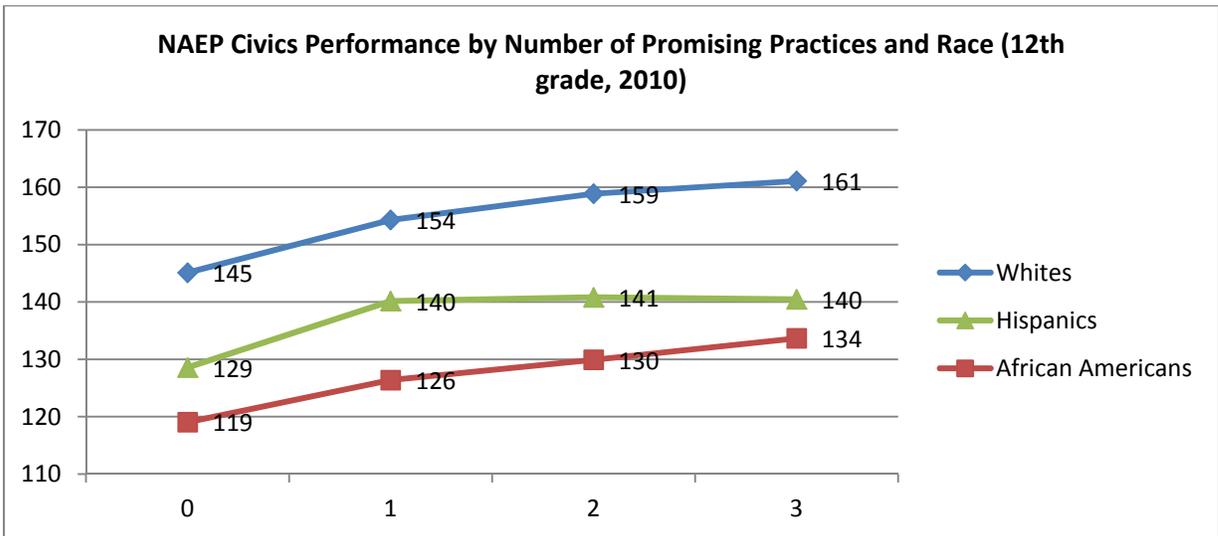
For 4th graders, simulations seemed to have no effect for some groups and negative effects for others. While there was no apparent effect on the performance of Whites, females, and Asian-Pacific Islanders and students who were not eligible for lunch program, it had a negative effect on African Americans, Hispanics, and students who were eligible for school lunch programs.



The Relationship between Number of Promising Practices and Performance

We then looked at whether the number of promising practices was related to NAEP Civics scores. There was a positive relationship between the number of promising practices and civics performance among 12th graders. However, there were some variations by groups. The figure below shows the relationship between the number of

promising practice and civics performance for 12th graders from different racial backgrounds. We found that the biggest gain was between no promising practices and at least one practice for all groups, but Hispanic students did not seem to benefit additionally from more than one practice, while White and African American students showed additive gains as the number of practices rose.³ We also found that students whose parents had less than high school or high school educations gained significantly between no practice and one practice, but their scores did not rise very much after that point. Furthermore, the average test scores for African American and Hispanic students who received all the promising practices still remained below the average score of the White students who did not get any of the promising practices. We found the same trend for students with different parental educational attainment levels. The students whose parents had no college experience performed below the level of students whose parents had any college experience and did not have access to *any* promising practices.



³ For more data on number of practices and gender, race, and socioeconomic background, please contact the author.

Unfortunately, exposure to promising practices did not minimize the race- and class-related achievement gaps in the NAEP Civics performance. Table 8 below shows the score gap between two groups at each level of best-practice exposure. While the overall trend is that all groups benefitted at least to some degree from promising practices, recommended pedagogies did not shrink the gap between the advantaged and disadvantaged students. In most cases, the gap expanded when students were exposed to more of the promising practices.

Table 8: Race and Class Gap Expands when Students Receive more Promising practices (NAEP Civics 12th grade data – numbers are differences in average test score)

	White-Black Gap	White-Hispanic Gap	College to LTHS gap	College to HS gap	Gap by free-or reduced-lunch eligibility
At No Promising practice	26	17	25	18	22
At 1 Promising practice	28	14	26	20	20
At 2 Promising practice	29	18	30	21	25
At 3 Promising practice	27	21	30	24	23
The gap expands by ...	1	4	5	6	1

Note: White-Asian gap is not significant at any practice level. Contact author for the 4th and 8th grade data.

Discussion and Implications

We analyzed the effect of three specific pedagogical approaches on the NAEP Civics Performance for 4th, 8th, and 12th graders. Some of the findings were promising, while others need more exploration. As an overall trend, 12th graders were likely to benefit from being exposed to all three types of practices, while findings were mixed, in some cases negative, for 4th graders.

There were some encouraging findings. Namely, the middle- and high-school students who are exposed to promising practices tended to perform better than their peers in the same demographic groups. It is important to note that, at least for 12th graders,

promising practices seem to have positive effect on NAEP Civics test performance across demographic groups. Additionally, we found that when students were exposed to more types of promising practices, they generally performed better on the NAEP Civics test.

On the other hand, there were some counterintuitive findings that need further exploration. First, the benefit of the particular pedagogical practices that we chose to explore was less clear for 8th and especially 4th graders. One explanation may be that the 4th graders did not answer the questions about what they had experienced in the same way as the 8th or 12th graders. However, it is also possible that spending time on the kind of experiential instructional techniques that we examined does sacrifice some concrete factual learning that would directly result in higher score on the NAEP Civics. As *The Nation's Report Card* and others have found, elementary school students generally receive far less instruction in Civics or Government than middle- and high-school students.⁴ Therefore, these experiential learning opportunities may in fact take some time away from learning about Bill of Rights or other facts related to Civics. However, it should also be emphasized that the NAEP Civics and other standardized tests are limited in that they do not assess students' civic skills, dispositions, or future intentions for civic participation. In other words, NAEP and standardized tests assess knowledge, just one aspect of civic competency. Therefore, it would be incorrect to assume that exposing elementary school students to experiential civic learning experiences would harm them. More than likely, these practices have other types of benefits, such as enhancing students' interest in the local government, political processes, or discussions that were not tested in the NAEP.

Another counter-intuitive finding was that students from different backgrounds did not respond in the same way when they were exposed to more types of recommended approaches. Unfortunately, at the highest dosage of promising practices, the civics score gap between students from less advantaged backgrounds and students from more advantaged backgrounds was larger, compared to the gap at the lowest dosage. There could be many reasons for this pattern, one of them being that students from advantaged backgrounds who receive these promising practices also are advantaged in other aspects of their educational lives, such as experiencing enriching extracurricular activities, exposure to formative experiences such as internships and community involvement, and access to adults who might instill civic knowledge outside of the classrooms (e.g., at dinner, as coaches, as mentors) compared to students from less advantaged backgrounds. Future research should incorporate how students' educational experiences, defined broadly, relate to civics test performance. Furthermore, future research could explore how the teachers' experiences, exposure to professional development, and expertise relate to the quality of instructions in promising practices, and ultimately, students' civics test performance.

⁴ Mark Hugo Lopez, Karlo Barrios Marcelo and Peter Levine, *Getting Narrower at the Base: The American Curriculum After NCLB*, CIRCLE monograph, December, 2008.