Beyond Breadth-Speed-Test: Toward Deeper Knowing and Engagement in an Advanced Placement Course

Walter C. Parker
Jane Lo
Angeline Jude Yeo
Sheila W. Valencia
Diem Nguyen
Robert D. Abbott
Susan B. Nolen
John D. Bransford
Nancy J. Vye
University of Washington

We report a mixed-methods design experiment that aims to achieve deeper learning in a breadth-oriented, college-preparatory course—AP U.S. Government and Politics. The study was conducted with 289 students in 12 classrooms across four schools and in an "excellence for all" context of expanding enrollments in AP courses. Contributions include its investigation of a model of deeper learning, development of a test to assess it, and fusion of project-based learning with a traditional curriculum. Findings suggest that a course of quasi-repetitive projects can lead to higher scores on the AP test but a floor effect on the assessment of deeper learning. Implications are drawn for assessing deeper learning and helping students adapt to shifts in the grammar of schooling.

Advanced high school courses take many forms around the world but often lag behind contemporary research on how people learn and what learning is. Broad coverage combined with a fast pace and a high-stakes summative exam typically count for "rigor." This breadth-speed-test formula, while efficient for some purposes, may exacerbate the problem it attempts to solve; it may drain advanced courses of the kinds of intellectual work they require if they are to be considered "advanced." Here we report findings from the second year of an iterative, mixed-methods design

experiment that aims to deepen students' learning in a widely enrolled advanced high school course in the United States: Advanced Placement U.S. Government and Politics.

The U.S. government and politics course, whether or not the Advanced Placement (AP) version, is a staple in the American high school curriculum. After a temporary decline during the 1970s, its enrollment returned to high levels in the 1980s, and approximately three-quarters of high school students take the course as of the latest count (Niemi & Smith, 2001). The course has two central concepts: *politics* (the processes of getting and using political power) and *government* (the product of politics). The chief variable in different forms of the course is whether the course has more of an active-citizenship, experiential goal or more of a political science, academic-learning goal. There is overlap, but the emphases are distinct. The former is more likely to engage students in civic action, perhaps using a module such as the popular *Project Citizen* (Center for Civic Education, 2010). Students identify a public policy problem in their community, develop

WALTER C. PARKER is professor of social studies education and, by courtesy, political science at the University of Washington, Seattle, 122 Miller Hall, Seattle WA 98195-3600 (e-mail: *denver@uw.edu*). His research focuses on school-based civic education in national and global contexts.

JANE LO is a doctoral student and research assistant in the College of Education at the University of Washington, Seattle. Her research focuses on the civic development of youth.

ANGELINE JUDE YEO is a doctoral candidate in the College of Education at the University of Washington, Seattle. Her research focuses on the influence of contexts and identities on students' historical thinking.

SHEILA W. VALENCIA is professor of language, literacy, and culture at the University of Washington, Seattle. Her research focuses on reading instruction, policy, and teacher development.

DIEM NGUYEN is research scientist and director of the Knowledge in Action Project at the College of Education, University of Washington, Seattle. Her research focuses on equity, diversity, and the quality of education in urban public schools.

ROBERT D. ABBOTT is professor of educational statistics at the University of Washington, Seattle. His research interests include the analysis of longitudinal data and the use of latent variables in educational research.

SUSAN B. NOLEN is professor of learning sciences and human development at the University of Washington. Her research focuses on the development of motivation and identity in learning contexts and their relationship to engagement and learning.

JOHN D. BRANSFORD is the Shauna C. Larson Professor of the Learning Sciences in the College of Education at the University of Washington, Seattle. He conducts research on learning, problem solving, transfer, assessment, and roles for technology in education.

NANCY J. VYE is principal research scientist in the College of Education at University of Washington, Seattle. Her research focuses on choice, agency, and challenge-based learning in K–12 and workplace settings.

a solution, and then plan a course of political action to effect change. The latter emphasis involves less doing and more knowing—less participation and more intensive study of the structures and functions of U.S. government, including its constitutional underpinnings, the *Federalist Papers*, landmark Supreme Court cases, federalism, civil rights and liberties, interest groups, elections, and so forth. The AP U.S. Government and Politics course (APGOV) falls squarely into the latter approach. According to the College Entrance Examination Board's (2010) course description,

> This course includes both the study of general concepts used to inter-
> pret U.S. government and politics and the analysis of specific exam-
> ples. It also requires familiarity with the various institutions, groups,
> beliefs, and ideas that constitute U.S. government and politics. (p. 6)

We value both poles in this tension—both participatory and academic civic learning. In the design-based research (DBR) reported here, we attempted to join them in a hybrid version of the course and to do so in the breadth-oriented, exam-based context of APGOV. We developed collab-oratively—with teachers, curriculum and instruction scholars, learning scien-tists, and political scientists—a kind of project-based learning (PBL) that puts students in authentic civic roles as legislators, cabinet secretaries, interest group members, judges, journalists, lobbyists, and citizens. In this way, students were engaged in simulated political action, but they also had to understand the roles they were playing and their interdependencies and contexts. The aim was for students to "experience" government and politics while also studying them intensively, learning to act and acting to learn in tandem, thus combining the two modal approaches.

This article is organized as follows. We begin with a problem statement that includes a brief on the AP program, including its rapid growth and "democratization" (Lacey, 2010, p. 34) in recent years along with increasing concern that student learning in these courses is superficial and unadaptive (i.e., insufficiently transferable and generative). Next, we present the con-ceptual framework, followed by a methods section, which includes our mea-sure of deeper learning and adaptive transfer, and then the course design principles. Methods are followed by Findings in two parts: quantitative find-ings from the two tests and then qualitative findings from end-of-year inter-views with students as well as from their responses to the deep learning measure. We conclude with a discussion and implications for extending the design experiment into a third year and for introducing it to classrooms in poverty-impacted urban schools.

Problem

Honors, Advanced Placement, the International Baccalaureate (IB), and Dual-Placement are familiar names of advanced programs in U.S. high

schools. There are more, many of them home-grown rather than affiliated with standardized programs (Sadler, Sonnert, Tai, & Klopfenstein, 2010), and some of these are in the nation's most prestigious prep schools. Increasingly in public schools, however, AP has become the norm for what often are called "rigorous college-prep courses." It has become the main brand and largest program of advanced high school coursework in the United States. Also, its quality is widely (though certainly not universally) regarded; indeed, it often is called in popular media the "gold standard" of the American high school (e.g., Associated Press, 2012; Hess, 2012; Mathews, 2009).

There are more than 30 AP courses, ranging from Calculus to World History. Of these, U.S. Government and Politics typically ranks fourth in student participation, just behind English Language, English Literature, and U.S. History. The courses are developed by committees of scholars and AP teachers working together with assessment specialists from the College Board, the association that develops and markets AP and other tests, such as the SAT. Some courses are one semester in length, others one year, but all end with a high-stakes summative test. Students who receive a passing score (a 3 or higher of 5 points possible) often have an edge in college applications, and some colleges allow these students to bypass the introductory course and proceed directly to advanced courses.[1]

The present study is grounded in two problems. One pertains to the remarkable growth of AP in public schools in recent decades and the subsequent gap between increased enrollment and success in these courses. The other pertains to AP lagging behind contemporary research on learning. Regarding the first, AP was developed in the 1950s for high-achieving students at elite prep schools but soon expanded to comprehensive public high schools, often as the upper curriculum track (Schneider, 2011). The number of AP test takers nearly doubled between 1950 and 1980 and then tripled between 1990 and 2000. According to the College Entrance Examination Board (2011), the number of seniors leaving high school having taken an AP test doubled between 2001 and 2010. As the number of students taking AP courses increases, so does the ethnic and socioeconomic diversity of students in AP. There was a 13% jump in Latino and African American test takers between 2008 and 2009 alone. There are two prominent explanations for these increases, one direct and one indirect. The direct explanation is that the College Entrance Examination Board and the U.S. Department of Education (2000) together mounted an initiative in 2000 urging high schools in the country to offer at least 10 advanced courses by 2010. The Department of Education then incentivized this expansion, subsidizing exam fees for low-income students and professional development for teachers (see Klopfenstein, 2004; Wakelyn, 2009). This helped further to push AP from its more elite private and suburban public enclaves to urban school districts where there are greater numbers of students generally and students of color

and lower income particularly. According to the College Entrance Examination Board and the U.S. Department of Education report, "The greatest challenge is to increase the numbers of poor and minority students taking Advanced Placement . . . and other demanding courses" (p. 1).

The second explanation is broader and helps to explain why this policy was able to gain traction. According to Schneider's (2011) historical analysis, an "excellence for all" trend gained momentum in the late 1980s and brought social-justice school reformers into an alliance with social-efficiency reformers. An historic tension was resolved, somewhat, between advocates of curriculum tracking for "excellence" and advocates of curriculum de-tracking for "equity." Schneider writes, "With advocates in government, non-profit organizations, and philanthropic foundations, these new activists won significant policy victories" (p. 3). These include No Child Left Behind under President Bush and Race to the Top under President Obama. In both initiatives, "excellence" was to be "democratized" (Lacey, 2010, p. 34). Many schools today, particularly in urban districts, are expanding AP participation by lowering or removing entrance requirements to the courses, such as prior achievement, and encouraging all or many more students to tackle them (Associated Press, 2012; College Entrance Examination Board, 2011; Sadler et al., 2010). The thrust of the discourse is that all students, not just elite private and suburban public students, should have access to the "rigor" of the "gold standard." For the problem of unequal access to excellent curriculum, "AP for all" is presented as a solution.

Increased enrollments have been accompanied, however, by increased numbers of students who fail the end-of-course AP test (Dougherty & Mellor, 2010; Lichten, 2007). The democratization of *enrollment* in these courses has not meant the democratization of student *success* in these courses. "Facile slogans that Advanced Placement is for everyone," according to Dougherty and Mellor (2010), "do not relieve educators of the responsibility to prepare students in the earlier grades" (p. 220). Stung by the high failure rates and the burst of criticism from teachers and scholars for open-enrollment policies, the College Entrance Examination Board adopted in 2011 a more cautious stance, staying well shy of unlimited access:

> The AP Program encourages educators to make equitable access a guiding principle for their AP courses by giving *all willing and academically prepared* [italics added] students the opportunity to succeed in rigorous, college-level opportunities. (p. 8)

If more students, whether or not they are "willing and academically prepared," are to gain access to advanced courses, then we prefer that the quality of the courses be improved, not assumed. Moreover, we want traditionally underserved students who are now increasingly gaining access to AP to be supported and successful, not only admitted. Increasing AP enrollments

allows policymakers to gesture toward equity by raising curriculum standards, reciting the "excellence for all" mantra, and providing equal opportunity for students to enroll in the classes; but, raising standards and granting access says little about another key: equal opportunity to learn (Au & Valencia, 2010).

We turn now to the second problem. Regardless of who is included in the courses and exams, there is increasing concern that AP lags behind contemporary research on how people learn and current conceptions of what learning is (National Research Council, 2000, 2002). These are contested issues, of course, and the subjects of much theory, practice, and empirical research; and, they are felt intensely and debated often in relation to AP. Despite AP's strengths, its Achilles' heel may lie in the gap between a curriculum and a course. AP courses are notorious for curriculum scope, such that there can be too much curriculum for a time-bound course—a grand stuffing of topics into a space too small to contain them meaningfully. This can result in rushed coverage and superficial learning, a "pancake course" that is "a mile wide and an inch deep," which does not make a satisfying conception of intellectual rigor nor realize the potential of advanced high school coursework. Certainly, there is room for improvement, which the College Entrance Examination Board itself recognizes.[2]

In 2002, a National Research Council report recommended that AP courses be redesigned to reduce accelerated coverage and to better reflect what is now known about *how* students learn and what it means to *learn*. According to that report, "The inclusion of too much accelerated content can prevent students from achieving the primary goal of advanced study: deep conceptual understanding of the content and unifying concepts of a discipline" (p. 1). "Well-designed programs," by contrast, "help students develop skills of inquiry, analysis, and problem solving so that they become superior learners" (p. 12). Accordingly, the goals of the design experiment reported here were (a) deeper conceptual learning and capacity for adaptive transfer; (b) same or higher scores on the AP test; (c) greater engagement, with appeal and success for a wider array of AP students; and (d) a course architecture that is sustainable and scalable by design.

Conceptual Framework

We focus this research on interlocking theoretical issues. First, how can deeper understanding be achieved in advanced high school courses and how can PBL contribute to it? Accelerated learning is only weakly related to deeper learning, we assume, because conceptually it refers to the pace at which topics are covered rather than the extent or nature of a learner's understanding. A second issue is what counts as learning. This issue imbricates and, to a degree, anchors the depth/breadth problem in curriculum and instruction. Have students who passed the breadth-oriented AP test

really learned anything in a meaningful way? Intuitively we might say "yes, of course," but does broad surface knowledge count? At issue is what learning means. For example, has something been learned if it cannot be retrieved or applied flexibly in a transfer task? A third issue, then, concerns the assessment of deep and adaptive learning. A core part of this study is developing a test that asks students to mobilize their knowledge in a transfer situation rather than only to recall it on the AP test. A fourth issue concerns instructional sequencing. What is the optimal timing and placement of reading and lecturing in project-based teaching and learning? Must some book learning, for example, precede the more experiential project activity—a field trip or simulation—so as to facilitate sense-making during the activity? Or should reading and listening *follow* the experience, motivated now by a need to understand what happened? This leads to a fifth issue concerning the way students experience a forced shift in the way they are accustomed to "doing school," which connects to a sixth issue that has already been broached: how to join experiential or participatory civic learning with the academic study of government and politics.

PBL is regarded generally as an ambitious form of instruction whether in kindergarten, a high school classroom, or medical school (Barron & Darling-Hammond, 2008; Ravitz, 2009; Walker & Leary, 2009). As we will show in this section, our approach to PBL has both typical and unique dimensions. Projects are at the center of our design; they are the "main course, not dessert" (Larmer & Mergendoller, 2010). However, we aimed also at deeper learning: specifically, learning for adaptive transfer. Toward that end, we unite the projects under a master course question and use a learning cycles approach (National Research Council, 2000): Students revisit the master course question as they move through subsequent project cycles, thereby giving the course a single main focus alongside its many secondary and tertiary topics. In this way, depth and breadth of learning are orchestrated in a continuous dialectic. The master course question gives a quasi-repetitive quality to the projects because while each project is unique in focus (congress, judiciary, elections, etc.), each is aimed at the same overarching question.

While PBL can take numerous forms, generally it interrupts the well-established classroom routine in which the teacher gives lectures, homework readings from the textbook, and quizzes and tests. This routine has deep cultural roots, a long institutional history, and a comfortable place in the popular imagination; it is a custom or, as Tyack and Cuban (1995) wrote, a "grammar" of schooling (p. 85). Goodlad (1984) found across more than 1,000 classrooms an "extraordinary sameness . . . . : repeated lecturing, questioning, monitoring, and quizzing" (p. 249; cf. Stodolsky, Ferguson, & Wimpelberg, 1981). We believe this grammar applies especially to AP courses. The anecdotal record—assembled from conversations with teachers and principals and from attending summer AP institutes where teachers

study released test items and learn from experienced AP teachers how to teach for increased pass rates—robustly supports it, as does examination of released test items over the past 10 years and textbooks written for the course. Yet, the anecdotal record does not substitute for an empirical warrant, and the empirical warrant is thin. We conjecture that the presence of the high-stakes AP exam and the accelerated, test-prep, coverage-oriented culture of the AP classroom (National Research Council, 2002; Sadler et al., 2010) combine to intensify rather than attenuate the established grammar.

Narrowing this characterization to civic education specifically, we know from the most recent National Assessment of Educational Progress report (National Center for Education Statistics, 2011) that only 18% of 12th-grade government students reported taking part in more experiential classroom practices such as role-playing, mock trials, and moot courts more than once a month, and 56% of the students reported never taking part in role-playing. Only 20% of students reported working on projects more than once a week. Seventy-one percent reported never writing letters to register their opinions or to help solve problems. Other research indicates that these more active classroom practices are distributed unequally on the basis of social class, ethnicity, and college-going plans. Kahne and Middaugh (2010) found in a study of California high school seniors that African American and Latino students and students who did not expect to go on to college were less likely than others to have had participation-oriented government courses generally and less likely to report participating in simulations in particular.

The goal of PBL is not only experiential and authentic learning but meaningful learning rather than simple recollection—a contrast articulated by numerous scholars near the beginning of the cognitive revolution in psychology (Bruner, 1990) and more recently in the National Research Council's (2000) *How People Learn*. Meaningful learning is more satisfying to learners because they feel that they understand, but also because it is applicable and actionable: They can *use* it as a basis for solving new problems. In other words, PBL appears to support transfer (Bransford & Schwartz, 2000). For example, a student does not remember only that the concept of checks and balances in government means that the judiciary needs to be independent of the executive and the legislature, but *why* this is so, *what* is at stake, *how* the judiciary checks executive power, and *how* it could be applied in an established democracy like ours or a developing democracy in the Middle East. Achieving this kind of preparation for lifelong learning is in some sense the ultimate goal of educational research and practice—a confounding but worthy pursuit.

To give material expression to meaningful learning, a standard feature of PBL is student investigations that result in authentic products. These products are termed "authentic" because they are like those found outside school

in workplaces, laboratories, legislatures, studios, and so forth. Learning activities involve some degree of student initiative, collaborative work, problem solving, argumentation, revision, and sometimes, deep exploration of the concepts and principles of a discipline. This has been summarized as *inquiry-based learning* (Darling-Hammond et al., 2008) because intellectual investigations are front and center. Collaborative, realistic, challenging investigations are often missing in the conventional grammar of teaching and learning, including traditional Advanced Placement courses.

We are attempting to contribute to the theory and practice of PBL by developing and testing a particular kind of PBL that is suited to advanced high school coursework and aimed at our four goals: (a) deeper conceptual learning of the course content, (b) same or higher scores on the AP test than students in a traditionally taught course, (c) greater engagement and appeal for a wider array of students, and (d) a course that is sustainable and scalable by design. Furthermore, we aim to situate this contribution within a broader one: to interrupt the conventional meaning of "rigorous teaching and learning" as vast coverage at a fast pace followed by a high-stakes exam. In so doing, we hope to demonstrate that deep and adaptive learning is not incompatible with advanced high school coursework.

We now detail the course's five design principles:

1. rigorous projects as the spine of the course;
2. quasi-repetitive project cycles where each builds on the other, yielding deeper understanding;
3. engagement that creates a need to know;
4. teachers as co-designers;
5. a course that can scale (migrate).

Project-Centered Design

In our approach to PBL, students work both collaboratively and alone to develop knowledge and skills through an extended inquiry that is structured around complex, authentic challenges. Project work requires heightened communication—ample public discourse as students work to interpret texts and problems and make decisions about the best course of action (Parker, 2006). Through project activities and the recurring phases of project anticipation, execution, and reflection, students have multiple opportunities to try out their current levels of understanding, revise them, and in this way deepen them. Also, we aimed to create a course experience in which challenging projects were at the center of the course, providing its spine, not the appendages; or using the other popular metaphor, the "main course, not dessert." The point of such metaphors, and perhaps a reason for their proliferation, is to communicate an important distinction—an inversion of the typical course grammar in which projects, if any, are both treated and

located as add-ons or end-of-course capstones—valuable activities done *after* reading and remembering and *after* background information has been acquired. Instead, projects are the centerpieces of the course.

Depth Through Quasi-Repetitive Learning Cycles

Quasi-repetitive activity cycles (Bransford et al., 2006; National Research Council, 2000), or what our teacher collaborators dubbed "looping," means that students have opportunities to revisit questions, ideas, and problems throughout the course. Looping, we reasoned, is key to deepening (elaborating, differentiating, and integrating; Parker et al., 2011) their evolving understanding of the course topics. Expertise in any domain, from podiatry to public policy, generally grows with the right sort of repeated practice—with "trying again" under novel conditions and with feedback. Central to this design principle is that the projects are united by a course "master question." As students move through the different projects, they revisit (loop back on) the master question and "try again" to generate a response, reflecting on what they have gleaned from the prior project cycles and the project cycle at hand. Here is inquiry-based learning—an intellectual investigation (Darling-Hammond et al., 2008)—but stretched and looped throughout the entire course. By unifying the projects, the master question gives the course one big topic amid myriad smaller topics. Also, the question is authentic; that is, as relevant outside school as it is to the course subject matter.

Engagement First

In their paper "A Time for Telling," Schwartz and Bransford (1998) explored when to use texts and lectures within the total repertoire of instructional methods. They concluded that there is a readiness for learning from textbook readings or lectures after some understanding has been generated in other ways. A third design principle, therefore, was that engagement in project work (e.g., being assigned to the role of a legislator with the task of setting up an office and advancing a legislative agenda) would typically precede telling (e.g., a PowerPoint lecture or homework reading on how Congress interacts with other institutions of national government). The purpose of this sequencing is to create a readiness (ideally, an eagerness) for telling so that the information students gain from it, whether through textbook reading or listening to a lecture, is needed for making progress on the project and constructing a suitable understanding. The telling has somewhere to go because there is already something going on—a context has been created. Students are already engaged in an action arena to which the telling can be of service; the telling serves to explain, clarify, and elaborate what is going on in the project work. "When telling occurs without readiness," Schwartz and Bransford conclude, "the primary recourse for students is to treat the new information as ends to be memorized rather than as tools

to help them perceive and think" (p. 477). This is a central reason why our team chose PBL as the basic architecture for the course. Managed well, it constrains the piling on of background information prior to project work, presumably so that students will "know enough" to participate in the project. That sequence, reversed here, is a persistent and deeply rooted grammar of schooling and one that can prevent students from ever getting to transfer and to *try*, again, their understandings over an extended sequence of quasi-repetitive trials.

Together, the first three principles compose the learning theory we are testing in this study. The conceptual framework of design-based research has to include, additionally, the practicalities of the research-and-development process; hence, the next two principles shape the working relationship of the researchers and practitioners as well as the approach taken to scaling (broadening, migrating) this work.

Teachers as Co-Designers

Brown (1992) concluded that if classrooms are to be transformed from "academic work factories to learning environments that encourage reflective practice among students, teachers, and researchers" (p. 174), then experimentation on complex classroom interventions must be done as a collaborative undertaking among teachers, researchers, and school administrators. Penuel, Fishman, Cheng, and Sabelli (2011) emphasized the same point when they concluded that design-based research, particularly design-based *implementation* research, is committed to "using research to solve practical problems," which requires the research to be plainly and directly "practice centered" (p. 332). Accordingly, our teacher collaborators became designers and curriculum makers—continually working with researchers to integrate AP content with a set of projects selected and adapted by the team.

Scalability

Our aim was not a "hothouse" experiment that would display what is possible but improbable. We were not satisfied merely to show that complex interventions could establish a change in practice. Instead, we worked with an eye to scalability, or what Brown (1992) aptly called "migration." She wrote: Researchers "must operate always under the constraint that an effective intervention should be able to migrate from our experimental classroom to average classrooms operated by and for average students and teachers, supported by realistic technological and personal support" (p. 143). We aimed for a design that other educators could adapt and who in turn could further the research and development in their own environments, thereby widening the community of teachers and researchers working to deepen learning in challenging high school courses. A caveat is in order, however. By embracing migration or scalability as a design principle, we do not

want to suggest that this course design can be "taken to scale," to use a popular phrase that often overestimates an intervention's suitability for additional settings. Rather, as will be explained in the Course Design section in the following, we aimed to create a course that would lend itself to adoption and adaptation in settings where there was both interest in and a supportive infrastructure for research and development.

As will be evident in the following sections, these final two principles are not cursory. They shaped the selection and revision of projects, compelled us to build student voice into the iterative design work, and generated theoretical categories such as the *two-worlds problem*.

Methods

This was a mixed-methods quasi-experimental design within the DBR approach. An innovation was designed—collaboratively by teachers, curriculum specialists, disciplinary scholars, and researchers—and then iteratively implemented, tested, and refined in real classroom settings, not a laboratory (Anderson & Shattuck, 2012; Brown, 1992; Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003; Penuel et al., 2011). Both quantitative and qualitative data were gathered and analyzed. Brown's (1992) seminal work on design experiments clarified the theoretical and methodological challenges researchers face when they are planning and revising as they go—when they are "simultaneously involved in designing under conditions of continuous flux" (p. 152). Such improvisation, while the exception in controlled laboratory experiments, is the norm in design experiments. As Collins, Joseph, and Bielaczyc (2004) write, "Because design experiments are set in learning environments, there are many variables that cannot be controlled" (p. 18). This limited control is not to be interpreted as a deficit, for the design tradition emphasizes the importance of learning during design so that the designers can intervene deliberately in the situation at hand, brimming with emergent and unpredictable properties, to reach a desired goal (Cobb et al., 2003).

There are four subsections: Research Design, Quantitative Measures, Qualitative Data and Analysis, and Course Design. Our design experiment is now in its third year, and here we report findings from Year 2. Findings from Year 1 (Parker et al., 2011), while favorable to the PBL approach, resulted in modifications to the course design, which are discussed later under Course Design.

Research Design

The PBL-AP approach was implemented in a suburban school district in the Pacific Northwest of the United States. Like many suburban districts, it had an adequate resource base and the institutional stability to accommodate the upheavals of innovation and also to fund a social studies curriculum

coordinator with whom we could collaborate along with teachers. Pursuant to our Goal 3 and pertinent to the "democratization" or "excellence for all" trend, the district had received national attention for encouraging AP participation for nearly all students, regardless of their preparation for it.

In both years, the PBL-AP approach was implemented in two of the district's five high schools that were eager to increase AP enrollments. Because we aim to develop, iteratively, a course that will be fruitful for the more diverse population of students now enrolling in AP courses, we selected the two high schools in the district with the greatest contrast in terms of students' race/ethnicity, socioeconomic status, and the school's historical achievement. One (School A) was a high-achieving school and the other (School B) a moderately achieving school. The latter previously had fewer students taking or passing the AP test. It also had three times the number of students receiving free or reduced-price meals and three times the number of Hispanic students. This afforded a within-district contrast relevant to Goal 3 discussed earlier: greater engagement, with appeal and success for a wider array of AP students. In the findings section that follows, we distinguish between the results of these two schools in contrast to moderately achieving comparison Schools C and D, located in a neighboring state.[3]

The research consisted of two comparisons with a total of 289 students, 12 classes, and five teachers. Each comparison used a nonrandomized intervention design in which outcomes from PBL-AP classes were compared to those for traditionally taught AP classes. Comparison 1 compared the results for students in three PBL-AP classes at the high-achieving School A to the results for students in AP classes in six traditionally taught AP classes in another, moderately achieving school district (Schools C and D). Comparison 2 compared the results for students enrolled in three PBL-AP classes at the moderately achieving school (School B) to the same traditionally taught comparison classes in the second district. For comparability, all classes were yearlong AP U.S. Government and Politics, whether PBL or traditional. Class sizes ranged from 24 to 37, with a mean class size of 31. Within each school, all of the yearlong AP U.S. Government and Politics courses were of the same type (i.e., PBL-AP or traditional AP). Table 1 provides demographic details for each school and Table 2 displays the numbers of students, classes, and teachers.

We used hierarchical linear (random coefficient) modeling (HLM) for the statistical analyses. For the AP scores, HLM models were used to compare the results taking into account students' nesting into classrooms, treatment condition, and students' prior achievement, including grade point average (GPA) and mean score on prior AP tests taken. For the scores on the Complex Scenario Test (CST; explained below), the models took into account students nesting into classrooms, treatment condition, and student's prior GPA as the measure of prior achievement on classroom-based assessments. These models took the form shown in Table 3. Table 3 follows

Table 1
Demographic Characteristics of the Schools (%)

School	Reading/English Language Arts Test Performance	Free/ Reduced Lunch	White	Black	Asian/ Filipino	Hispanic	Other
School A	92.7[a]	11.6	61.8	2.6	25.6	3.4	6.6
School B	78.8[a]	32.7	52.4	4.5	20.6	14.0	8.5
School C	77[b]	49.8	50.3	1.7	16.5	31.0	0.5
School D	74[b]	47.8	57.4	1.4	14.6	25.1	1.5

[a]State average = 78.9%.
[b]State average = 74%.

Table 2
Total Number of Students, Classes, and Teachers by School and Research Condition

School	Number of Students	Research Condition	Classes	Teachers
High-achieving School A	86	Experimental project-based learning (PBL) AP course	3	1
Moderate achieving School B	89	Experimental PBL AP course	3	2
Moderate achieving Schools C and D	114	Traditional AP course (comparison)	6	2
Total students	289		12	5

Raudenbush and Bryk's (2002) notation. Posttest Outcome$_{ij}$ is the outcome for student i in classroom j, γ_{00} is the overall mean of the outcome, γ_{01} is the additional effect of the PBL class (the traditional class is coded 0) on class j, γ_{10} is the regression of student i in classroom j posttest on the student's mean of prior AP tests, and γ_{20} is the regression of the posttest on the student's prior GPA. Both the student variables and the treatment contrast (PBL vs. traditional class) are fixed effects in the model. The focus is on γ_{01}: the contribution of the contrast of PBL versus traditional to the mean of classroom j.

The overarching research question was this: Can a rigorous form of project-based learning be applied to high school AP U.S. Government and Politics course and improve student learning? There were three more specific questions.

Research Question 1: Can we create a PBL-AP course in which students do as well or better on the AP test than students in a traditionally taught AP course?

Table 3
Hierarchical Linear Models

Level 1 model
$\quad$ Posttest Outcome$_{ij}$ = β_{0j} + β_{1j} (prior AP) + β_{2j}(prior GPA) + r_{ij}

Level 2 model
$\quad$ β_{0j} = γ_{00} + γ_{01} (Traditional vs PBL) + u_{0j}
$\quad$ β_{1j} = γ_{10}
$\quad$ β_{2j} = γ_{20}

Research Question 2: Will students demonstrate a deeper level of knowledge as assessed by a complex scenario test?

Research Question 3: Will students report greater engagement in their learning experience, both inside and outside the classroom?

In this article, we report quantitative findings on Questions 1 and 2, after which we address Question 3 with a portion of our qualitative data. There, we report on interviews conducted with students in the experimental course to learn their summative views on the course, thereby qualifying the experimental findings with students' own descriptions of their response to the course while also giving them a voice in the redesign of the course for the following year. Doing this is faithful to DBR methodology, with its aim to improve practice through iterative modifications of the intervention.

Quantitative Measures

To address the first two research questions, we administered the following measures: the College Board-administered AP test and the Complex Scenario Test, which was created by the research team.

College Board-Administered AP Test

The AP U.S. Government and Politics exam was officially administered near the end of the course in May.

Complex Scenario Test

The CST is an in-class, open-ended, paper-and-pencil assessment of applied knowledge or adaptive transfer. It uses a real-world problem of politics and government to assess students' learning in the course. Whereas the AP test primarily measures students' ability to identify and describe the structures and functions of government and change in them over time, the CST aims to assess how well students can apply that knowledge in a particular scenario, which they are encountering for the first time, in which their charge is to formulate a plan for intelligent political action on a controversial

issue that is currently in the news and heating up. The scenario in Year 2 involved an indigent person's right to a fair trial. Features of the CST include the following:

- Places students in the role of adviser to a congressperson or an interest group.
- In the scenario, students must mobilize knowledge from across the course to draft an action plan for their client.
- Students are told their client will want to decide for her/him/itself; therefore, students need to justify their recommendation.
- The scenario centers on a controversial issue, one the students have not encountered in the course and are not expected to be expert in (facts are provided within the scenario; e.g., in an actual news article about the issue).
- The topic is from current news headlines but students are told, "While some of the facts and materials of the case are made up, others are real, notably . . . ."

Neither the PBL teachers nor the traditional course teachers had knowledge of the CST, and our teacher-collaborators were not involved in its development. They were unable, therefore, to aim instruction at it or to engage students in practice for it.

Qualitative Data and Analysis

A sample of students in each of the six PBL-AP classrooms was interviewed individually three times during the experimental course. Then, following the administration of the AP and CST tests in late May, each of the six PBL-AP classes was interviewed in a large-group, fishbowl format. We had two objectives: (a) to learn from students' experiences in the course so that we could gauge its appeal and effectiveness for them and (b) to afford them a voice in making adjustments to the course design for the project's third year. The latter was a top priority for our teacher-collaborators. Furthermore, a primary goal of DBR, as discussed by Cobb et al. (2003), "is to improve the initial design" (p. 11). Another primary goal, as discussed by Penuel et al. (2011), is theory building. The latter includes problem redefinition and the refinement of central categories—in our case PBL itself and deep conceptual learning (adaptive transfer). The qualitative data presented here were gathered and analyzed to serve both objectives and both goals.

The fishbowl interviews were recorded and transcribed. Transcripts were then read iteratively by three members of the research team, and themes were generated that were pertinent to the study's questions. Using rotating inner circles of speakers, students were asked five questions:

1. What are you taking away from this course experience that's of value to you personally—that has helped you or served your needs in some way?
2. Looking back, how well would you say the course has achieved its intended purpose: enabling AP students to learn about U.S. government and politics in a meaningful and deep way?

3. Please comment on any specific components or features of the course and how well they worked. How well did the parts hang together? Was the whole greater than the sum of the parts?
4. How could we make this course more successful for kids like you?
5. If you had known what the course was going to be, would you have signed up? If changes were made along the lines suggested here today?

In this article, we address qualitative data from the fishbowl interviews at the moderate achieving experimental high school where the mixture of "veteran" and "newcomer" AP students and the range of students' prior achievement were greatest. A "two-worlds problem"[4] emerged as a prominent theme. In our analysis, we use Bourdieu's (1977) concept of habitus along with the idea of learning to "do school" (Pope, 2003). Following this, we turn to an analysis of students' written responses on the CST. There, we peer into one student's depth of understanding on three criteria: differentiation, elaboration, and integration. But first we turn to the experimental PBL-AP course design.

Course Design

The course was designed according to the four goals and five principles presented earlier. As indicated under Design Principle 5, scalability, we aimed to adopt and adapt projects from other practice communities rather than creating original projects. By so doing, we could rely on projects that not only were readily available elsewhere, thereby supporting scaling, but that had "matured" thanks to being vetted and revised in other practice arenas. For example, the project called "111th Congress," was based on a simulation called LegSim (http://faculty.washington.edu/jwilker/353/WhatisLegSim.pdf) developed by political scientist John Wilkerson, a Congress scholar. Another project cycle centered on the popular "moot court" simulations of Supreme Court hearings (www.landmarkcases.org). Another project is "A Government for Xlandia." The Buck Institute of Education (www.bie.org) developed this simulation and then improved it by integrating feedback from teachers who used the project in their classrooms. As it turns out, our teachers decided against using this particular simulation again in Year 2. In its place, they created their own project, "Town Hall." The reason for the replacement was that a project was needed that would directly target, early in the course, the concept *federalism* as it is laid out in the U.S. Constitution, debated in contemporary politics, and assessed on the AP test. The focus of "Xlandia" was broader and, some teachers believed, more appropriate for a *comparative* government course. The replacement project opened the course in Year 2. As this change exemplifies, Design Principles 4 and 5 functioned together: Scalability often had to be negotiated with teachers as co-designers (see Penuel et al., 2011). Accordingly, if the teachers felt that an imported project would not work

for their students and school context, we made a revision even if it meant that the teaching-research team needed to create a project from scratch.

What follows are thumbnail descriptions of the six projects cycles that were the spine of the course. Each involves role taking in simulated political activity. Student engagement in these roles and activities began immediately with a project's commencement, not after background information had been established (the "engagement first" principle), thereby provoking a need to learn more through reading and lecture. These projects were unified under a master course question that was the result of a deliberation among the teachers, researchers, curriculum coordinator, and political scientists on the team. That question in both Years 1 and 2 was "What is the proper role of government in a democracy?" It is this question, recall, to which students return in each project, quasi-repetitively "looping" their evolving understandings. Also, note that this question is an authentic one that can be found at the heart of contemporary political debates and elections in the United States.

1. *Town Hall*. Students are citizens participating in town hall meetings where they deliberate policy alternatives in four issue domains: economy, environment, foreign policy, and social welfare.
2. *Government in Action*. Students are new House members in the U.S. Congress interacting with other branches of government in the same four issue areas.
3. *Congress 111 (LegSim)*. Students are legislators in the U.S. Congress, writing and working to pass legislation on matters of personal and shared concern, consistent with the interests of their district.
4. *Election 2010*. Students are campaign consultants planning and executing strategies for victory on election day.
5. *Supreme Court*. Students are justices of the Supreme Court, petitioners, or respondents in landmark cases.
6. *Taking It Global* (after AP test). Students are members of the U.N. Human Rights Commission deliberating proposals for U.N. intervention in hot spots around the world.

Quantitative Findings

In this first of two findings sections, we present the results related to the first two research questions: (1) Can we create a PBL-AP course in which students do as well or better on the AP test than students in a traditionally taught AP course? (2) Will students demonstrate a deeper level of knowledge as assessed by a complex scenario test? Recall that one PBL school was historically high achieving and the other was historically moderate achieving. Also, recall that the comparison group was moderate achieving and that statistical analyses were adjusted for students' prior achievement.

Table 4
**Means and Standard Deviations of AP Test Scores and
Percentage of Students With High Pass**

	Project-Based Learning AP High-Achieving School A	Project-Based Learning AP Moderate Achieving School B	Traditional AP Moderate Achieving Schools C and D
Average score (*SD*)	3.37 (1.27)	2.33 (1.15)	2.03 (1.07)
Students with high pass (4–5)	47.7%	16.8%	5.7%
Number of students	86	89	87

AP Test Results

Regarding Question 1, the results indicate that it is possible to get same or higher scores on the AP test with a PBL course. Table 4 displays the means and standard deviations of the AP scores of the students in the three groups. The table also shows the percentage of students with a "high pass" (score of 4 or 5) on the AP test. PBL-AP students scored significantly higher on the AP test than the traditionally taught AP students in both the moderate achieving PBL-AP school, $t(7) = 3.12$, $p = .018$, effect size (ES) = 0.25, and, with a greater effect size, the high-achieving PBL-AP school, $t(7) = 6.73$, $p = .001$, ES = 0.78. Also, because some colleges assign college credit for high scores (4–5) on the AP test, we compared these "high pass" rates. Significantly more students at the PBL Schools A (47.7%), $t(7) = 4.52$, $p = .001$, and B (16.8%), $t(7) = 2.66$, $p = .033$, achieved a high pass than traditional students at Schools C and D (5.7%).[5]

We note that these are conservative conclusions with respect to learning differences in comparisons of the two groups of moderate achieving students. In the "traditional AP" group, many more students opted out of taking the AP test than in the PBL classrooms. Almost all (98%) of the PBL students in School B took the AP test, whereas only 73% took the AP test in the comparison group (Schools C and D). It may be that comparison students who opted out felt less prepared for the AP test and would have scored poorly. At least, this was the conjecture of their teachers.

CST Results

As noted earlier, the AP test primarily measures students' ability to identify and describe the structures and functions of government and change in them over time. Regarding Research Question 2, we designed the CST to measure how well students can *apply* what they have learned to a novel and authentic problem. Accordingly, students are given a real-world controversial issue and asked to formulate a plan for well-informed political action.

Students' written answers to the CST were scored on four dimensions, and the quality of the student's response related to each dimension was assigned a score ranging from 1 to 6 (6 = highest quality response). The dimensions were:

1. Overall Quality—Gives a high quality response overall.
2. Task and Client—Directs advice to the *particular* congressperson or citizen group (appropriate to Congressperson X or Citizen Group Y).
3. Influencing Public Policy—Gives an informed political process account (using political process concepts and vocabulary).
4. Grasping Controversial Issues—Analyzes the public policy issue at stake and what makes it controversial

Scoring was blind with respect to student group. It involved making two passes through a response. Initially, scorers judged the response on Dimensions 2, 3, and 4, giving a score for each. Then, scorers judged the response holistically and gave an Overall Quality score. The percentage agreement of two independent raters on these dimensions were Task and Client (90%), Influencing Public Policy (90%), Grasping Controversial Issues (82%), and Overall rating (90%). When the two raters did not agree within 1 point, the rubric was applied by an independent third rater and the mean rating was assigned to the paper.

Table 5 displays the means and standard deviations for both high- and moderate achieving PBL students in comparison to the traditional students on the four dimensions of the CST. Again, HLM models similar to those in Table 3 were modeled (but only student GPA was used as a student-level predictor because we considered it a better control of prior achievement on results of the CST). And again the focus is on γ_{01}, which is the contribution of the contrast of PBL versus traditional to the mean of classroom *j*. PBL students in the high-achieving School A scored significantly higher on three of the four dimensions of the CST as compared with students in the traditional courses; Overall score: $t(7) = 2.63$, $p = .034$, ES = .50; Task and Client: $t(7) = 3.80$, $p = .008$, ES = .71; Influencing Public Policy: $t(7) = 3.16$, $p = .017$, ES = .83; and Grasping Controversial Issues: $t(7) = 2.22$, $p = .06$, ES = .40. These findings suggest that these PBL students more deeply understood the AP content to the point that they were able to apply it in a novel situation to solve a complex problem.

Meanwhile, students in the moderate achieving PBL School B did not perform significantly differently than students in the traditional courses: Overall score: $t(7) = 1.14$, $p = .291$; Task and Client: $t(7) = 2.05$, $p = .080$; Influencing Public Policy: $t(7) = 1.18$, $p = .276$; and Grasping Controversial Issues: $t(7) = 1.58$, $p = .159$. The lack of difference between PBL students from the moderately achieving school and the comparison students (also

<div align="center">

Table 5
Complex Scenario Test: Means and Standard Deviations of Scores

</div>

	Project-Based Learning AP High-Achieving School A	Project-Based Learning AP Moderate-Achieving School B	Traditional AP Moderate-Achieving Schools C and D
Overall:			
Average score	2.23	1.97	1.88
SD	(0.70)	(0.78)	(0.60)
Task and Client:			
Average score	2.45	2.19	2.00
SD	(0.86)	(0.80)	(0.74)
Influence Policy:			
Average score	2.15	1.79	1.78
SD	(0.74)	(0.71)	(0.60)
Controversial Issues:			
Average score	2.41	2.31	2.14
SD	(0.79)	(0.89)	(0.67)
Number of students	82	77	114

in moderately achieving schools) is, we surmise, probably the result of a "floor effect": Both groups scored low on the CST—on average a low 2.0 on a scale of 1 to 6. This test requires proficient reading and writing skills and is given within a relatively short time limit. In other words, it is so difficult that variance was limited—many students did poorly in both the intervention and comparison groups. This finding supports our goal as we go forward (see the following) of creating reading, writing, and other scaffolds that can help less prepared students succeed. It also supports current efforts to lower the difficulty of the CST test in such a way that a greater range of scores is achieved, as is the case with the AP test.

Qualitative Findings

In this second findings section, we turn to our group interview data and then to a sample of higher scoring CST responses. Our goal in examining these data is theory building with an eye toward design improvement in the following year per DBR conventions (Penuel et al., 2011). First, we focus on data gathered in end-of-year group interviews with students in the PBL classes at the moderate achieving high school. Of the two treatment schools, this is the one where the mixture of "veteran" and "newcomer" AP students was greatest, as was the range of students' prior achievement. We wanted to gauge the course's appeal and effectiveness for them while giving them

a voice in revising the course for the project's third year. (The semi-structured interview protocol was given earlier.)

"You Learn Two Different Skill Sets"

The fishbowl interviews at the end of the prior year had revealed considerable student frustration with the "engagement first" design principle. Many students were accustomed to "doing school" (Pope, 2003) in accordance with a school grammar in which they had long been immersed: The teacher gives lectures, homework readings from the textbook, and tests; and when a teacher uses projects, they are not the main course but a side dish that comes after a sufficient foundation of "background information" has been laid down, or as dessert that comes after the AP test has been administered. Inverting that sequence means that there may be greater readiness for learning from telling *after* some initial understanding and need-to-know has been generated by actual engagement in the project activities. This is not simply "just-in-time-teaching," but creating the structural conditions for it. Many students in the first year of this design experiment found it frustrating, arguing that without sufficient background information "we don't know what we're doing" (Parker et al., 2011, p. 552). These students wrestled with how course components could best be sequenced for learning, and many preferred that the course stick to the traditional grammar: First lay a foundation of prior knowledge—some sort of "government for dummies" or "floaties," as they put it—so that the project work feels more stable and grounded, with less floundering in the deep end of the pool without supports. The "engagement first" or "time for telling" (Schwartz & Bransford, 1998) concept does not mean an absence of support, but its implementation in Year 1 felt that way to a good number of students.

Accordingly, in the second year our teacher-collaborators tried to orient students to a new way of doing school. Of course, it was easier for teachers to do this in the second year because they themselves had experienced the curriculum in the first year. As a result, students in the Year 2 fishbowls reported greater comfort with the "engagement first" design. For example, "I like the six project cycles, and that I knew every time we started a project cycle what the basic layout of it would be. . . . It's kind of a schedule. So I would know basically what we were going to be doing." And from another student: "[The project cycles] were all alike; the basics of it were similar. . . . It was all like we kind of knew what to expect and what we would be doing."

But as one problem recedes, others can be seen more clearly. In Year 2 what emerged was what we call the two-worlds problem. We refer to the perception of numerous students that there was a kind of firewall between AP and PBL. This perception was acute for several of the AP veterans especially. The problem was already apparent in Year 1, as stated by this AP veteran:

> Okay, this is where I had my biggest struggle with the course. . . . I thought I was taking an AP course so that I can get AP credit for college because that was what I wanted. I have taken all these other AP classes, and it's designed to enable you to do well on the test, and in this course I found myself doing a lot of work that was not preparing me directly for the test. It may have been interesting and, you know, engaging, and, you know, you learn stuff, you learn a lot of different skills; but you might not have been directly learning about things that was on the test. (Parker et al., 2011, p. 554–555)

This student evaluated the course not on how or what he was learning but on how efficiently it taught to the test. He had taken numerous AP classes, which he regarded generally (and happily) as test-preparation classes and, consequently, struggled with the PBL approach. Moreover, he had normalized the approach of past AP classes such that he regarded the new approach as the disruption of a tradition. He was "pulled up short"—a term used by Kerdeman (2003) to capture the experience of being caught off guard, where an individual is forced to confront the limitations of his or her routine knowledge or know-how (p. 296).

Let's turn to a Year 2 student at the moderate achieving high school. Like the student just quoted, she also is an AP veteran, and she too was pulled up short by the disruption of a habit that she had established "throughout the years," as she put it. But she also is trying to make sense of the new pedagogy:

> [S9:25:23]: Throughout the years, like, every AP history I have taken has had the same structure. And so that initial discomfort, where we didn't really know what was going on, [was] because we weren't forewarned. I mean we were told that this is project-based, but I don't think any of us really knew what it meant. . . . There was a lot of times when none of us felt like we were doing anything, like, extremely significant. It's just kind of like let's go to government and who knows what we'd do today. . . . I like the course like now, but during a lot of those days I felt like the class was pointless because I don't see any direct results from what we are doing. Because I think we were all so used to, like, AP World, AP U.S., or like AP Psych where you get a textbook and you know when each chapter was going on and how many chapters you have before the AP test. . . . It's a lot of time just communicating with groups and just doing kind of obscure things rather than just listening to lecture and stuff like that.

This student was comparing the course's PBL approach and, especially, its "engagement first" principle to the straightforward efficiency of test-prep teaching and learning. So completely did she grasp the purpose and pedagogy of traditional AP that by contrast, PBL seemed "obscure" or indirect and even "pointless." (Recall that students are assigned to roles and engaged

in simulations before learning the requisite information and concepts; ambiguity is, therefore, inherent in the design.)

This next student, also an AP veteran, is accustomed to taking classes without regard to personal interest or motivation. She was not as put off by the "engagement first" dimension of the course, but she resisted its effectiveness; it was causing her to learn more than she wanted or believed she needed. Bolstering this view is her conception of college to which she believes this course does not align: College courses are not project-based, so why is this one?

> [T9:25:14]: [I]f you're passionate about the subject, then you've got to want to do more in class and focus on what you are doing in class because that will really help you figure out what you want to do, if you want to be politically active. But for me . . . it was just a class to take, it's not a passion; it's not a subject that I ever want to learn. It's just a class to take for me and a test I want to pass. . . . That's why I read the textbook stuff. And maybe the projects were more frustrating for me because it was like, why can't we just do a PowerPoint, take notes, and go home like a college course would be? I just didn't feel it was super helpful to do the projects. But I see how somebody who really wants to do something in government would enjoy the projects and take that as experience they can use in their lives.

But consider these comments from students who were AP newcomers. For them, the issue wasn't the contrast between AP and PBL pedagogy—a distinction they didn't possess—but instead between textbook-based learning and PBL—a distinction they did possess. The two-worlds problem was manifest now in the use of curriculum materials. Reflecting on the AP test, which he had taken shortly before this interview, this student says it was based on the textbook whereas the projects were not.

> [T2:14:27]: I think you get two different things out of the projects and out of the textbook learning. I mean, I know the class tried to fuse them together in a way, but I think people could go the whole year and do fine [on the projects] without opening their textbook once just because they weren't really as connected. . . . You learn two different skill sets, I think, from the book and then from the projects.

The second and third students agree but add their own inflections:

> [T10: 20:07]: I think the pressuring part about the difference between the textbook and the project is that the AP test is based off the textbook and you can't separate a class completely from the textbook that the test is based on. So I feel like if you totally eliminate the textbook, you still learn the same amount but you won't learn the correct thing that the test will want you to know.

> [T11:22:07]: It kind of depends on like which way they [fellow students] want to be successful in the class. If they were just looking to do well on the AP test, then they want to read the textbook. I mean, this class is really good like teaching how to learn the ways that you learn in real life 'cause you are not going to learn life skills through a textbook. You are going to learn to work with other people on issues you care about. So if you just try to, like, learn about what the class is teaching you, then you don't need the textbook. I know that I didn't read the textbook because I didn't care about the AP test and I'm just fine with it.

For these students taken together, our PBL-AP approach interrupted a school *habitus* (Bourdieu, 1977) that had been long in formation. According to Bourdieu (1977), a habitus is an embodied system of socially acquired predispositions that enables social actors to "size up" a situation and produce an appropriate response or improvisation. A particular, built-up way of doing school generally or doing AP specifically is a habitus, and a habitus is by definition both sensible and invisible to the persons who embody it. When students find themselves in a situation where old "rules of the game" are no longer in sync with the circumstances, a remodeling of the habitus might result but more likely is a sense of dissonance or rupture—of being pulled up short and finding oneself unable to make sense of the situation.

Recall that student S9 reported that she felt lost and purposeless for much of the course. However, in the end, she and others reported that the PBL approach had indeed helped them perform well on the AP test. Unlike in test-prep teaching and learning where, perhaps, students can better monitor their own progress, our PBL approach seems to take students by surprise:

> [S9:27:08]: I think this is one of the most helpful classes as far as AP test-wise. But this whole year, I just didn't really know how this would help me at all.

> [S3:34:13]: [A]t the end of the year [when] you start reviewing everything, that's when I started to finally understand everything and that helped me out in the AP test.

> [V6:00:47:] I think it's . . . a lot like learning as you go along. And you might not know or realize that you are necessarily learning things. . . . Like for me when we were looking at the practice questions on the AP test, I didn't realize that I registered all that stuff in my mind, but I knew a lot of them.

In summary, these fishbowl interviews indicated that many students—both AP veterans and newcomers—experienced a two-worlds problem in this engagement first course. We note with interest that the disjuncture was related to the curriculum track they had experienced in the past: AP

veterans finding it unlike other AP courses and AP newcomers finding a disconnect between project activity and textbook reading. The differences threw some students off balance, forcing them to make adjustments to their school habitus. Many students reported that in hindsight, however, the new practices did help prepare them for the AP test.

"Take It to a Higher Power"

But what do these students know "deeply" about U.S. government and politics? This presents a different interpretive challenge; accordingly, we shift now to students' written responses to the CST. This assessment was our attempt to measure depth and adaptability of students' knowledge by asking them (a) to apply what they know (b) in a particular scenario where (c) they take a role as a political advisor and (d) must produce a plan for intelligent political action (e) on a complex political controversy that is currently in the headlines and heating up. The scenario in Year 2 involved the right of an indigent person, Mr. Weis, to a fair trial in a case that could involve, upon conviction, the death penalty. A divided Georgia Supreme Court had ruled in 2006 that his prosecution could proceed despite Mr. Weis's complaint that he was not receiving a fair trial. The state had appointed a public defender team after denying the defendant's request for a legal team with experience in death penalty cases. The team subsequently asked to withdraw from the case, saying that it did not have the time, funds, or qualifications to pursue a death penalty case.

The CST informed students that an interest group called "Travesty of Justice" had formed in response to the case and that they (the students) had been hired to advise this group. Students were directed to use knowledge from across the course and, using no texts or notes, to draft an action plan that would advance Travesty's goals. The prompt included two primary documents—a description of the case from the *New York Times* and a statement by a Travesty member that had been posted on a social networking site. The prompt ended, "Please write your response in the form of a memo to Travesty of Justice beginning on the next page." Students' own views on the matter were not at issue; rather, their knowledge of government and politics would enable them to (here again are the four dimensions on which responses were scored) give advice that is not general but tailored to the goals and interests of this particular interest group, an informed political process account (e.g., knowledge of "lower" and "higher" courts in the judicial system while also distinguishing between state and federal courts and between the judiciary and other branches), an understanding of why the public policy issue at stake is controversial, and a high-quality response overall. Students were told that Travesty of Justice will want to decide for itself what to do; therefore, students need to justify their recommendation and make clear the reasons and system requirements for it.

Quantitative results using HLM models were given earlier. Recall that the scores generally were low on the CST in both the treatment and comparison conditions. Here, we examine the 10 highest scoring responses with respect to three criteria of deep conceptual learning: differentiation (the understanding is complex, multifaceted), integration (it is coherent, connected), and elaboration (it is detailed, nuanced). Because theory development is a primary purpose of design-based research (Penuel et al., 2011), we chose high-scoring responses for this analysis so that we could examine the theoretical categories at work in student responses (rather than simply noting their absence). Our interest is in refining this model of deeper learning.

In terms of the first criterion, each of the 10 responses indicated the ability to distinguish the central constitutional issue at hand and to consider the roles played by the three branches of government and linkage institutions (e.g., media and interest groups). For example, each response recognized the role the U.S. Supreme Court would play based on the divided ruling of the Georgia Supreme Court. Travesty would need to "take it to a higher power" to resolve the conflict, one student wrote. In the excerpt that follows, a student refers to the sixth amendment guarantee of "a speedy and public trial, by an impartial jury of the State and district wherein the crime shall have been committed," and then to national versus state courts, and then to the state legislature and media:

> Constitutionally, in the Bill of Rights everyone is ensured a fair and speedy trial. The Weis case has been neither. . . . Those who would support our efforts of legal representation reform for the poor would be interest groups, anti-death penalty groups, and lawmakers/elected representatives whose constituents have a high poverty rate. . . . Our target venues should be the U.S. Supreme Court [and] the Georgia State legislature. We would have a high possibility of getting a *writ of certiorari* accepted by the Supreme Court and [have] them [overturn] the Georgia Supreme Court's decision in this case because this case raises questions of Constitutionality. . . . The decision to proceed in the Weis case [without] competent legal representation goes totally against the Constitution . . . it would be a good idea to use the media (for ex: local news) to do a human interest piece on Weis's plight.

While one of the highest scoring, this response averaged only 3.5 points out of 6 across the four dimensions. Still, the response displays some differentiation in identifying the constitutional issue at stake (e.g., not Amendment 8's cruel and unusual punishment nor Amendment 1's freedom to petition the government, but Amendment 6's speedy and fair trial with the assistance of counsel) and also by its reference to the political entities involved and the division of powers in the U.S. federal system.

With respect to content integration—pulling it all together—the prompt asked students to explore the interaction of governmental institutions in

relation to the constitutional issue. The student quoted previously suggested that the strategic use of the media would "enrage the people of Georgia to get them to start demanding their state or local [representatives] to take action, which can be very effective." The student also advised Travesty to petition other interest groups and lawyers to "help write and support a *writ of certiorari* as well as submit *amicus* briefs if the case is taken" for review by the U.S. Supreme Court. Responses like this one indicate that the higher scoring students understood the necessity of coordinated institutional action in order to advance Travesty's agenda and that they could mobilize knowledge of the separate institutions in the context of this particular scenario.

These top scorers struggled the most with elaboration. Even the best responses were thin and vague as to the action plan. In general, there are few details and little nuance in the recommended actions for Travesty in relation to government and linkage institutions. Mentioning the writ of certiorari, as does the aforementioned student, does not indicate whether the student understands the process of petitioning a case to the Supreme Court, only that such an appeal is needed and that this is what it is called. A more elaborate treatment could indicate that Travesty, an interest group, cannot petition for a writ of certiorari as this needs to be done by the petitioner, Mr. Weis. Travesty could advocate and fund this process, however. More important, the student needs to state the constitutional grounds on which the petition would likely be granted. Instead, the action plan from this student was simply that Travesty should affiliate with other interest groups, getting them to help write amicus (friends of the Court) briefs. While rightly mentioning the state of Georgia's legislature, rather than the U.S. Congress, as the site of needed legislation, the student did not detail how Travesty could influence the passing of legislation besides "hire lobbyists," which is unelaborated and undifferentiated.

We cannot, however, reasonably attribute the lack of elaboration to students' lack of knowledge or inability to apply it to the scenario. Students may have felt they had done quite enough to provide a skeletal response; deep learning, after all, does not enjoy a secure place in the grammar of schooling, and students know that a skeletal response will typically get them a passing grade (Bereiter & Scardamalia, 1987). Furthermore, they may not have understood the meaning of a "political action plan" nor understood how to frame their response as a "memo" to Travesty, and then how to frame a memo for a particular client. Recall that neither the PBL nor the traditional classes had practice with this type of prompt. Furthermore, the CST was administered the day after the AP test and just before high school graduation. Students may have been too depleted (or too excited) for still more testing. And unlike the AP test, the CST is not a high-stakes exam. Students may have felt no need to put forth their best

effort as the test did not "count" either toward a course grade or a college application.

Discussion

Students in the PBL-AP schools and classes scored significantly higher on the AP test than students in the comparison schools and classes both in comparison 1, involving a high-achieving school, and in comparison 2, involving a moderate achieving school. In both analyses, comparison classrooms were in a moderate achieving school, and HLM models were used to account for initial differences in prior achievement across classrooms. The results are encouraging, we believe, especially when we remember that any substantial interference with test-prep pedagogy could sacrifice pass rates on the AP exam. (This desire for high pass rates combined with fear of the reverse probably is the elemental motivation for the penchant toward test-prep pedagogy in AP classrooms across the country; Sadler et al., 2010.) Comparison 2 displays a more remarkable finding: Students in the PBL-AP condition outperformed similar peers in traditional courses on the AP test. Meanwhile, on the alternative Complex Scenario performance assessment, students in the high-achieving school outperformed comparison students while students in the moderate achieving school performed the same. What is puzzling about the latter finding is that these moderate achieving students outperformed their peers on the breadth test but not on the depth test; *both* groups performed poorly on the CST. We suspect that the literacy task demands of this assessment were insurmountable for many of these students, undermining the measure's ability to assess what they had learned. There could be other problems as well as mentioned previously; for example, both groups' lack of familiarity with a test prompt of this sort: to give politically savvy advice on a novel problem. (Recall that there was no practice for the CST, unlike for the AP test.) As we go forward in this design experiment, refining and validating the CST is of paramount importance. We have already begun with think-aloud interviews of a purposive sample of students as they read the CST prompt, and we will incorporate these findings in the next design and administration of the CST. (A preliminary finding of interest is that students generally had difficulty not in decoding the CST but in bringing domain knowledge to bear in comprehending it.)

A two-worlds problem emerged in the qualitative findings as a source of (some) student frustration. This problem is a fertile one for educators aiming to infuse deep, experiential, and adaptive learning into advanced high school coursework. Such coursework is commonly inscribed in a breadth-speed-test grammar. An impending broad, high-stakes exam has a singular way of focusing students' and teachers' attention, and the orientation toward meaningful learning is felt by some students as mere annoyance. This problem can be framed as an advantage of working in the AP environment, for it

brings the depth-breadth tension into sharp relief in a way that neither PBL nor AP can alone. Furthermore, by working in the AP environment PBL finds itself joined to the constraints of the contemporary content standards movement, the college-readiness craze, and the "excellence for all" rhetoric—none of which has been associated with meaningful learning per se but rather with the testing-and-accountability movement. This confluence of disparate forces makes a fertile problem space for research and development on PBL.

Strengthening PBL with quasi-repetitive project cycling and sustained inquiry on a master course question appears, from the focus group interview data, to have intensified rather than mitigated the difference between learning for understanding and learning for test performance. For some students, this breach manifested as something like "projects versus textbook reading," the latter containing the tested vocabulary and the former being of uncertain purpose and value. But even these students admitted, once the AP test had come and gone, that "somehow" the projects had prepared them well for the test.

The ambiguity that attends engagement first learning was bewildering for some students at the same time others were finding it refreshing and enjoyable. We cannot say whether the frustration experienced by some students aids or undermines success on the two tests. Surely some frustration is inherent in rigorous learning. Adaptive experts are typically frustrated more often than novices because not only are they working on challenging problems but they are persevering, thereby prolonging the frustration. Adaptive experts are not accustomed "to treat new information as ends to be memorized" (Schwartz & Bransford, 1998, p. 477) as some veteran AP students may be; rather, they are inclined to treat new information "as tools to help them perceive and think." The engagement first design put students repeatedly in real-world political roles as legislators, justices, campaign aids, journalists, and so forth before they fully understood those roles or their relationship to one another in the broader framework of U.S. government and politics, thus creating a need to know. They continually had to "size up," read about, ask the teacher about, and communicate with other students about the overall situation. Routine expertise developed in prior, traditional AP course experiences was no match for the learning environment in which they now found themselves. As Martin, Pierson, Rivale, and Diller (2007) concluded, in challenge-based pedagogies "many students initially attempt to use their efficient learning strategies and find that they are inadequate" (p. 150; also Hatano & Inagaki, 1986). Joining terms from Tyack and Cuban (1995) and Bourdieu (1977), we can say that a profound alteration of the classroom grammar—such as PBL in an AP course—can rupture a habitus.

Still, we are not inclined to treat student frustration as an inevitable or desirable feature of rigorous learning nor as a necessary corrective for an entrenched test-prep habitus. Instead, we embrace our third goal: greater

student engagement with appeal and success for a wider array of students. This is why in the second year reported here the teachers and researchers aimed, with considerable success, to better orient students to this particular form of PBL pedagogy, with projects as the main course (not the occasional side dish or dessert) and engagement first as modus operandi.

Our learning cycles approach to project-based learning was a means not an end. The central aim was deep and meaningful understandings that are available and useful in future, novel problem settings—adaptive transfer. The CST presented students with such a scenario, asked them to size it up, and then to provide well-informed political advice to political actors who sought to influence its outcome. This kind of understanding is complex. The Constitution is not simply "the law of the land," but an assembly of principles, structures, and mandates for government and politics. Such an understanding is also coherent—integrated—rather than a heap of unrelated pieces. And it is elaborate, too: shades of gray are perceived and articulated, such as petitioning the court for a writ of certiorari and then, assuming it is granted, writing an amicus brief on behalf of the petitioner. Our analysis of the higher scoring responses to the CST leaves us cautiously confident that these criteria are worth pursuing in subsequent trials. We are aware that the links between an achieved understanding (even a complex one that is differentiated, integrated, and elaborated) and the ability and likelihood to transfer it flexibly to a novel scenario are empirically uncertain.[6]

A closely related aim was that students would "experience" government and politics while studying them intensively thanks to steady project engagement and role-taking—learning to act and acting to learn in tandem—and in this way blending the two types of civic education that are modal in this country and often opposed. Action-oriented or participatory approaches to civic learning have long been contrasted with more passive or "banking" (Freire, 1970, p. 72) approaches that give short shrift to problem posing, problem solving, and, in Dewey's (1916/1985) famous phrase, "conjoint communicated experience" (p. 93). Using simulations in every project cycle and putting students in the shoes of an array of actors engaged in political behavior, and doing so before rather than after so-called background information has been established: This was our strategy for meaningful learning and for blending knowledge and action—political action fused with academic knowledge of the structures and functions of government.

Conclusion

The principal contribution of this work is its fourfold attempt to interrupt a contemporary discourse of "rigor" defined as broad coverage at an accelerated pace followed by a high-stakes summative exam. This breadth-speed-test formula, we believe, emphasizes fast, superficial learning at the expense of meaningful learning, which is problematic because meaningful learning

appears to increase the likelihood of adaptive transfer. "The generalization potential for learning," according to Goldstone and Day (2012), "is just as important a facet of efficiency (as pace), even though far more research on assessment is needed to develop adequate measures of generalization potential" (p. 151). Developing such an assessment is a core feature of this design experiment.

We are attempting through iterative design-and-implementation research to theorize deep civic learning and to help students develop civic knowledge that is available for flexible use and additional learning in novel civic problems. We are attempting also to develop a measure to assess it, conduct systematic inquiry on an intervention that seeks to achieve it in actual classroom settings, and fuse project-based experiential learning with a traditional structures-and-functions-of-government curriculum. By undertaking this work on the AP platform, we are able to study and revise our model within a set of challenging real-world conditions. Among these is the "excellence for all" discourse that is broadening the population of students participating in AP. While we applaud this emphasis on equity of access to AP courses, we want the courses to be worthy of the new students; and, we want these students to be successful. Organizing the course around a set of active, experiential project cycles was the core of our approach to helping more students succeed.

We believe that gearing advanced high school courses on government and politics toward depth of understanding, engagement, and adaptive expertise rather than bare-bones test prep actually matters. It matters not only for student learning but also for democratic institutions such as independent judiciary, checks and balances, and equality under the law. These institutions require agents—citizens—for their maintenance and invigoration. The AP platform places severe limits on this aim by packing too much curriculum into a single course and then capping it with a high-stakes, breadth-oriented exam. In this design experiment, we approached the depth/breadth tension with an articulation of project-based learning and curriculum coverage; we attempted to mobilize PBL on behalf of both coverage and in-depth learning. Coverage was addressed by teaching to the broad sweep of the College Board's six topics and preparing students to take the impending AP exam. In-depth learning and course appeal for the wider array of students now enrolling in it were orchestrated through quasi-repetitive project cycling and an engagement first policy that puts telling and reading at the service of a situated need to know. Sustained inquiry on a course master question united the projects and helped assure a regular "try again," recursive cycling.

We close by emphasizing two limitations of the study with regards to generalizability and scalability. First, the results of a small quasi-experimental study do not solve the breadth-speed-test problem of advanced high school coursework. Nor do they solve the AP test failure rates among underserved

students who increasingly are encouraged to enroll in these courses. While generalizability is not a goal of design experiments, effective design-based research should "be able to migrate from our experimental classroom" to other classrooms (Brown, 1992, p. 143), as noted earlier. It remains to be seen whether our PBL-AP approach can "migrate" successfully to other settings, but we did design the course with that purpose in mind (Design Principle 5). In the first 2 years, the study involved a relatively advantaged, though still relatively diverse, suburban student population and a single AP subject. In Year 3, to be reported later, there were two expansions of the work. First, we migrated PBL-AP U.S. Government and Politics to high schools in three poverty-impacted urban school districts. Second, at the invitation of the suburban district we applied the same design principles to the development of a second course: AP Environmental Science. To ensure that students have the opportunity to learn in these courses, rather than merely to enroll in them (Goal 3), we are engaging our teacher-collaborators in the urban districts in the development of scaffolds to support learners in reading AP textbooks and collaborating productively with peers.

Second, we recognize that in numerous but short-lived successful initiatives "the successes were among students taught by the early adopters" (Darling-Hammond et al., 2008, p. 15). In the first 2 years of our design experiment, including Year 2 reported here, the teachers were, indeed, the early adopters who had co-designed the course with the researchers. Consequently, the successful migration to other courses in the AP stable, combined with the migration from suburban to urban schools, is far from assured. We believe it will need continually to involve the new teachers as collaborators—as adaptors rather than adopters. They best know their students, their courses, and their school contexts. Relying on their agency and on-the-ground experience will be key.

Notes

[1]This is a rapidly shifting landscape. See Labaree (2012), Schneider (2011), and Sadler, Sonnert, Tai, and Klopfenstein (2010).

[2]The College Board is undertaking course revisions that aim at more meaningful learning in Advanced Placement (AP) courses. See "What's Changing in AP" at http://advancesinap.collegeboard.org/.

[3]While comparison schools are not common in design-based research (see Brown, 1992; Penuel, Fishman, Cheng, & Sabelli, 2011), we believe there is value to the comparative data, especially in a context like this one where passing the high-stakes exam is of paramount importance and the risk of reducing pass rates looms as a decisive threat to innovation. If an alternative pedagogy is used, stakeholders want to know, "What were

their pass rates in comparison to students in traditional courses?" In conversations with school administrators and AP teachers, most recently at the national AP conference in summer 2012, this is normally the first question put to us.

[4]This problem is not to be confused with another two-worlds problem, the gap between clinical and academic learning settings in teacher education discussed by Feiman-Nemser and Buchman (1985).

[5]Effect sizes (ES) were calculated using Cohen's *d* on the adjusted means. Additional tables displaying hierarchical linear modeling (HLM) results are available on request from the first author.

[6]See the special issue of *Educational Psychologist* edited by Goldstone and Day (2012) on "New Conceptualizations of Transfer of Learning."

References

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, *41*(1), 16–25.

Associated Press. (2012, May 8). Advanced Placement surges as tool for schools raising standards. *Education Week*. Retrieved from http://www.edweek.org/ew.

Au, K. H., & Valencia, S. W. (2010). Fulfilling the potential of standards-based education: Promising policy principles. *Language Arts*, *87*, 373–380.

Barron, B., & Darling-Hammond, L. (2008). Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning. In L. Darling-Hammond, B. Barron, P. D. Pearson, A. H. Schoenfeld, E. K. Stage, T. D. Zimmerman, G. N. Cervetti, & J. L. Tilson (Eds.), *Powerful learning: What we know about teaching for understanding* (pp. 11–70). San Francisco, CA: Jossey-Bass.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ: Lawrence Erlbaum Associates

Bourdieu, P. (1977). *Outline of a theory of practice*. Cambridge, UK: Cambridge University Press.

Bransford, J. D., & Schwartz, D. L. (2000). Rethinking transfer: A simple proposal with multiple implications. In A. Iran-Nejad & P. D. Pearson (Eds.), *Review of research in education* (*Vol. 24*, pp. 61–100). Washington, DC: American Educational Research Association.

Bransford, J. D., Vye, N. J., Stevens, R., Kuhl, P., Schwartz, D., Bell, P., . . . & Sabelli, N. (2006). Learning theories and education: Toward a decade of synergy. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 209–244). Mahwah, NJ: Erlbaum.

Brown, A. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, *2*, 141–178.

Bruner, J. S. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.

Center for Civic Education. (2010). *Project Citizen*. Calabasas, CA: Author.

Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, *32*(1), 9–13.

College Entrance Examination Board. (2010). *Government and politics: United States, comparative, course description*. New York, NY: Author.

College Entrance Examination Board. (2011). *The 7th annual AP report to the nation*. New York, NY: Author.

College Entrance Examination Board & Department of Education. (2000). *Dispelling the culture of mediocrity: Expanding AP*. New York, NY, & Washington, DC: Authors.

Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of Learning Sciences*, *13*(1), 15–42.

Darling-Hammond, L., Barron, B., Pearson, P. D., Schoenfeld, A. H., Stage, E. K., Zimmerman, T. D., . . . & Tilson, J. L. (2008). *Powerful learning: What we know about teaching for understanding*. San Francisco, CA: Jossey-Bass.

Dewey, J. (1985). *Democracy and education* (Vol. 9). Carbondale, IL: Southern Illinois University Press. (Original work published 1916)

Dougherty, C., & Mellor, L. T. (2010). Preparing students for Advanced Placement. In P. M. Sadler, G. Sonnert, R. H. Tai, & K. Klopfenstein (Eds.), *AP: A critical examination of the Advanced Placement program* (pp. 219–232). Cambridge, MA: Harvard Education Press.

Feiman-Nemser, S., & Buchman, M. (1985). Pitfalls of experience in teacher preparation. *Teachers College Record*, *87*(1), 53–65.

Freire, P. (1970). *Pedagogy of the oppressed*. New York, NY: Seabury.

Goldstone, R. L., & Day, S. B. (2012). Introduction to "New conceptualizations of transfer of learning." *Educational Psychologist*, *47*, 149–152.

Goodlad, J. I. (1984). *A place called school*. New York, NY: McGraw-Hill.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York, NY: Freeman.

Hess, R. (2012). *Common Core architect and new College Board president David Coleman* (interview). Retrieved from http://blogs.edweek.org/edweek/rick_hess_straight_up/2012/06/straight_up_conversation_common_core_architect_and_new_college_board_president_david_coleman.html?intc=es

Kahne, J., & Middaugh, E. (2010). High quality civic education: What is it and who gets it? In W. C. Parker (Ed.), *Social studies today: Research and practice* (pp. 141–150). New York, NY: Routledge.

Kerdeman, D. (2003). Pulled up short: Challenging self-understanding as a focus of teaching and learning. *Journal of Philosophy of Education*, *37*, 293–308.

Klopfenstein, K. (2004). The Advanced Placement expansion of the 1990s: How did traditionally underserved students fare? *Education Policy Analysis Archives*, *12*(68). Retrieved from epaa.asu.edu/epaa/v12n68/

Labaree, D. F. (2012). School syndrome: Understanding the USA's magical belief that schooling can somehow improve society, promote access, and preserve advantage. *Journal of Curriculum Studies*, *44*, 143–163.

Lacey, T. (2010). Access, rigor, and revenue in the history of the Advanced Placement program. In P. M. Sadler, G. Sonnert, R. H. Tai, & K. Klopfenstein (Eds.), *AP: A critical examination of the Advanced Placement program* (pp. 17–48). Cambridge, MA: Harvard Education Press.

Larmer, J., & Mergendoller, J. R. (2010). *The main course, not dessert*. Retrieved from http://www.bie.org/tools/freebies/main_course_not_dessert/

Lichten, W. (2007). Equity and excellence in the College Board Advanced Placement program. *Teachers College Record*. Retrieved from www.tcrecord.org

Martin, T., Pierson, J., Rivale, S. R., & Diller, K. (2007). The role of generating ideas in challenge-based instruction in promoting adaptivity. In W. Aung (Ed.), *Innovations 2007: World innovations in engineering education and research* (pp. 147–160). Arlington, VA: International Network for Engineering Education and Research.

Mathews, J. (2009, August 28). Will Advance Placement replace the SAT? *The Washington Post*. Retrieved from http://voices.washingtonpost.com/class-struggle/2009/08/will_advanced_placement_replac.html

National Center for Education Statistics. (2011). *The nation's report card: Civics 2010.* Washington, DC: Institute of Education Sciences, U.S. Department of Education.

National Research Council. (2000). *How people learn: Brain, mind, experience, and school* (expanded ed.). Washington, DC: National Academy Press.

National Research Council. (2002). *Learning and understanding: Improving advanced study of mathematics and science in U.S. high schools, Executive Summary.* Washington, DC: National Academy Press.

Niemi, R. G., & Smith, J. (2001). Enrollments in high school government classes: Are we short-changing both citizenship and political science training? *PS: Political Science and Politics*, *34*, 281–287.

Parker, W. C. (2006). Public discourses in schools: Purposes, problems, possibilities. *Educational Researcher*, *35*(8), 11–18.

Parker, W. C., Mosborg, S., Bransford, J. D., Vye, N. J., Wilkerson, J., & Abbott, R. (2011). Rethinking advanced high school coursework: Tackling the depth/breadth tension in the AP US Government and Politics course. *Journal of Curriculum Studies*, *43*, 533–559.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, *40*, 331–337.

Pope, M. D. C. (2003). *Doing school.* New Haven, CT: Yale University Press.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Ravitz, J. (2009). Summarizing findings and looking ahead to a new generation of PBL research. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 4–11.

Sadler, P. M., Sonnert, G., Tai, R. H., & Klopfenstein, K. (Eds.). (2010). *AP: A critical examination of the Advanced Placement program.* Cambridge, MA: Harvard Education Press.

Schneider, J. (2011). *Excellence for all.* Nashville, TN: Vanderbilt University Press.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and instruction*, *16*, 475–522.

Stodolsky, S. S., Ferguson, T. L., & Wimpelberg, K. (1981). The recitation persists, but what does it look like? *Journal of Curriculum Studies*, *13*, 121–130.

Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform.* Cambridge, MA: Harvard University Press.

Wakelyn, D. (2009). *Raising rigor, getting results: Lessons learned from AP expansion.* Washington, DC: National Governor's Association.

Walker, A., & Leary, H. (2009). A problem based learning meta analysis: Differences across problem types, implementation types, disciplines, and assessment levels. *Interdisciplinary Journal of Problem-Based Learning*, *3*(1), 12–43.