J. CURRICULUM STUDIES, 2011, VOL. 43, NO. 4, 533–559



# Rethinking advanced high school coursework: tackling the depth/breadth tension in the AP *US Government and Politics* course

WALTER PARKER, SUSAN MOSBORG, JOHN BRANSFORD, NANCY VYE, JOHN WILKERSON and ROBERT ABBOTT

This paper reports a design experiment that attempted to strike a balance between coverage and learning in an exam-oriented, college-preparatory, high school course—Advanced Placement (AP) *US Government and Politics*. Theoretically, the study provides a conceptual framework for penetrating the depth/breadth tension in such courses, which are known for coverage and perhaps 'rigour', but lag behind contemporary research on how people learn and what learning is. Methodologically, the paper details a mixed-methods study of an alternative approach to AP coursework, conducted with 314 students across three high schools. First-year findings indicate that a course of semi-repetitive, content-rich project cycles can lead to same or higher scores on the AP exam along with deeper conceptual learning, but that attention is needed to a collateral problem: orienting students to a new kind of coursework.

Keywords: high school curriculum; advanced placement; project based learning; government and politics; innovative pedagogy

Advanced Placement (AP) courses in the US are functionally similar to the A-levels in the UK and former members of the Commonwealth (e.g.

*Walter C. Parker* (corresponding author) is Professor of Social Studies Education and (by courtesy) Political Science at the University of Washington, Seattle. His research focuses on civic education, K-12, in national and global contexts.

*Susan Mosborg* is a former Research Scientist at the University of Washington, Seattle. Her interests include curriculum design practices, learner identities, and education system innovation.

*John D. Bransford* is the Shauna C. Larson Professor of the Learning Sciences 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 current research interests focus on choice, agency, and challenge-based learning in K-12 and workplace settings.

*John Wilkerson* is Associate Professor of Political Science at the University of Washington, Seattle, and Director of the Center for American Politics and Public Policy. He is the inventor of LegSim, an on-line virtual legislature used in high school and college courses, and a research specialist in American legislative politics.

*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.

*Journal of Curriculum Studies* ISSN 0022-0272 print/ISSN 1366-5839 online © 2011 Taylor & Francis
http://www.informaworld.com
DOI: 10.1080/00220272.2011.584561

Singapore and India), influencing students' college admissions and largely defining advanced high school coursework. The number of students enrolling in AP courses in the US is increasing as more school districts remove entrance requirements and encourage all students to tackle 'rigourous' courses, a decision often framed in terms of access, equity, and college readiness (Riley 2005, Schneider 2009, Wakelyn 2009, Sadler *et al.* 2010). AP courses are considered by many to be among the best to be found in the American high school. Indeed, they are often touted as 'the gold standard' of the American high school curriculum (Mathews 2009: 8), and are looked upon as one piece of the puzzle from which to develop a 'common core' system of high-quality standards and assessments across states (National Governors Association 2009). Yet many worry that AP lags behind research on how people learn and what learning is (National Research Council 2002). The chance to rethink AP courses, thus, represents an important opportunity for curriculum study and innovation in education today.

AP courses and exams are developed collaboratively by scholars and teachers working with the College Board (the association that develops and markets AP and other tests). These courses are not the product of a single teacher working 'behind the classroom door' (Goodlad and Klein 1970); rather, each AP course—the course description and the exam—is the product of a deliberation among disciplinary scholars, high school teachers, and test developers, which opens it to public scrutiny and debate during and after development. AP courses and exams have this important public advantage—a degree of transparency. Course descriptions, sample exams, and scoring criteria are available for public inspection at www.collegeboard.com. Their achilles' heel, however, lies in the gap between a curriculum and a course. In AP, there is often too much curriculum for the time-bound course—a great stuffing of arguably important topics into a space temporally too small and pedagogically too meagre to contain them meaningfully. AP courses are notorious for coverage (breadth, scope), and famous for 'rigour', although the latter typically goes unquestioned and undefined in popular media (e.g. Long 2010, Mellon 2010).

In 2002, the National Research Council, which functions under the auspices of the National Academy of Sciences, recommended that AP courses be redesigned to reduce coverage and better reflect what is now known about how students 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' (National Research Council 2002: 1). 'Well-designed programmes', in contrast, 'help students develop skills of inquiry, analysis, and problem-solving so that they become superior learners' (p. 12).

Here we report a mixed-methods study of an alternative approach to AP coursework that aimed to build students' transferable conceptual understanding and inquiry skills, as advocated by the National Research Council and others. This was a design experiment involving the course *AP US Government and Politics*, which is one of the most-taken AP exams (College Board 2010). We investigated whether a re-designed AP course

using principles drawn from *How People Learn* (National Research Council 2000) and project-based learning (Barron *et al.* 1998, Darling Hammond *et al.* 2008, Ravitz 2009) would improve student learning and engagement when compared to a traditionally-taught AP course. The project's goals were (1) same or higher scores on the AP test, (2) deeper conceptual learning and greater capacity for adaptive reasoning, (3) greater engagement, with appeal and success for a wider array of AP students, and (4) a course that is sustainable and scaleable by design.

Our approach entailed, centrally, treating the depth/breadth problem not as a contradiction or dichotomy but as a tension or contrary—a key distinction in the semiotic study of oppositions (Greimas 1987). Contradictions cannot logically occupy the same space; tensions can, although not without difficulty and problems of balance—hence, 'tension'. Oppositions such as depth vs breadth, constructivism vs behaviourism, friend vs enemy, and so forth are thereby loosened from the dichotomous negation of *versus* to the various possibilities and hybridities of *and* or *with*. Depth *versus* breadth is opened up to co-ordinations of depth *and* breadth or depth *with* breadth. Such a framework is helpful in designing coursework in the AP arena where breadth is extensive in both the published course outline and the impending AP exam, which together can encourage 'test-prep' or 'coverage' teaching and learning over pedagogical alternatives more in line with current research on what learning is and how people learn.

The depth/breadth tension is the subject of lively debate (Sizer 1984, Hirsch 1988, 2001, Newmann 1988, Wineburg 1997). Researchers' conceptions of depth and breadth have ranged from the amount of instructional time spent on a topic (Sadler and Tai 2001, Schwartz *et al.* 2008) to the nature of that instruction and the kinds of learning that result (National Research Council 2000, Darling-Hammond *et al.* 2008). Schwartz *et al.*, for example, found that students in high school science courses that included at least one topic covered for a month or longer earned higher grades in college science. Darling-Hammond *et al.*, reviewing a number of studies, concluded that students who are engaged in collaborative inquiries involving sustained constructive work along with regular formative assessments are more likely to achieve complex and applicable understandings than when they are engaged in tasks and assessments that emphasize memorization of a broad swath of information. A complex and applicable understanding is 'deep', we stipulate, when it is integrated (coherent), differentiated (multifaceted and elaborated), and flexible (adaptable); consequently it supports future learning in novel situations (Hatano and Inagaki 1986, Bransford and Schwartz 2000). A deep understanding is typically both domain-dependent and domain-specific. It is domain-dependent because as expertise develops, the student knows, in an increasingly integrated, differentiated, and flexible way, something in particular, not things generally; and it is domain-specific because the student-expert's knowledge of, say, an AP Government concept such as *limited government* differs not only substantively but also epistemologically from one's knowledge of an AP Environmental Science concept such as *energy conversion*.

We proceed as follows. First we describe principles driving the redesigned course that aimed to co-ordinate depth and breadth in such a way that students would achieve, in comparison to a control group, the same or higher scores on the AP exam *and* deeper understanding. Next we present our methodology followed by the quantitative findings and then qualitative data from end-of-course interviews with students. The latter capture students' responses to the experimental course on two dimensions of the depth/breadth tension: the extent to which they felt prepared for the broad AP exam and their reflections on how best to co-ordinate the course's component parts: projects, lectures and readings. We conclude with a discussion of both the quantitative and qualitative findings in light of subsequent developments in the experimental course.



## Course design

We begin with some background on the AP programme in order to situate our course design in recent developments in the social context of AP and in contemporary principles of learning. In many US states, AP enrolment is 'skyrocketing' (Long 2010: 1). The number of exam takers nearly tripled between 1990 and 2000 and there was a 13% jump in Latino and African American exam takers just between 2008 and 2009. Increasingly, open enrolment policies, where any student is allowed to take AP, are being implemented in both suburban and urban school systems. Typically the problem frame is unequal access to quality curriculum, and 'AP for all' is presented as a solution. As AP moves from being a haven for elite students, elite private academies are increasingly dropping it (Schneider 2009). Nonetheless, the total number of exam takers is growing—and the number of students who fail it is growing too. An AP official reflects, 'I don't know an educator who wouldn't think it's a good tradeoff to take the risk and give more courses *that we know have been good for the few*' (quoted in Lewin 2010: 4, emphasis added).

While the issues here are manifold, our primary concern in this work is *the quality of the learning experience* to which students in greater proportion are gaining access. We aim to foster deep, transferable understandings and engaging learning experiences alongside good scores on the AP test. This is a tall order. However, if more students are to gain access to the course in the name of equity, which we generally support, then we prefer that its quality, including its 'rigour', be improved, not assumed. In this way we aim to create both a stronger course and a more appealing course for a wider array of students than normally participates in AP while at the same time increasing the long-term learning benefit of taking these courses for each student. Accordingly, we return to the 2002 recommendation of the National Research Council with which we began: AP courses need to be redesigned to reflect what we know about how students learn and to build students' transferable, conceptual understanding and inquiry skills. The report concluded: 'Although AP programmes ... currently are not well aligned with learning principles, they can be revised

with this research in mind. The resulting transformations are likely to make the programmes more successful in enhancing deep conceptual learning and make them more accessible to additional students' (p. 9). These were our principal aims as we designed a 1-year, project-based AP *US Government and Politics* course.

Why this course? First and foremost, this was a design experiment, and so we were sensitive to the wishes of the teachers and school district administrators with whom we collaborated. They requested that we begin with this course. Second, this is one of the most enrolled of the nearly 40 courses offered by AP. Indeed, the government course, whether AP or not, remains a staple of the US high school curriculum. Third, the course's content is important to society. Democracies (specifically, republics or constitutional democracies) are rare historically and inherently fragile. 'To ask why democracy does not exist at a particular time or in a particular country is on the face of it a distorting question', writes historian Robert Wiebe (1995: 9–10). 'Its absence does not compel explanation'. Accordingly, an AP *US Government and Politics* course can be singled out from others in the AP galaxy as one with unique historical gravitas. Democratic modes of association are not given by nature; on this the historical record could not be clearer. Rather, they are built, and much of the construction work is done by people who share an understanding of what kind of polity they are trying to create. These people are not born already grasping the difficult political principles of limited government, civil rights and liberties, toleration, and equality before the law. These are social, moral, and cognitive achievements (Parker 2003). Gearing an AP course on government and politics toward depth of understanding and adaptive expertise rather than bare-bones test prep actually matters if institutions such as representative government, an independent judiciary, and freedom of inquiry and expression are to be rejuvenated in each new generation.

Achieving same or higher scores on the AP exam was by itself an ambitious goal, as there is a vast quantity of material to 'cover', and any deviation from a straight-ahead, focused, test-prep, coverage-style teaching and learning regime could sacrifice exam scores. Indeed, with such a broad array of tested material, the traditional mode of test-prep instruction may be an efficient way to achieve this goal. However, in addition to same or higher scores, we had the further goal of deepening students' understanding, and for this we had to attend to learning principles. We began with careful readings of the College Board's (2008) *Course Description* for US Government & Politics. It states that the course

> will give students an analytical perspective on government and politics in the United States. This course includes both the study of general concepts used to interpret US government and politics and the analysis of specific examples. It also requires familiarity with the various institutions, groups, beliefs, and ideas that constitute US government and politics. While there is no single approach that an *AP United States Government and Politics* course must follow, students should become acquainted with the variety of theoretical perspectives and explanations for various behaviours and outcomes. Certain topics are usually covered in all college courses. (p. 5)

Six course topics are listed. We give them here with the percentages of multiple-choice questions devoted to each on the AP test. There is also a free-response portion of the test which addresses 'some combination' of the same six topics.

(1) Constitutional underpinnings (5–15%);
(2) Political beliefs and behaviours (10–20%);
(3) Political parties, interest groups, and mass media (10–20%);
(4) Institutions of national government: Congress, presidency, bureaucracy, federal courts (35–45%);
(5) Public policy (5–15%); and
(6) Civil rights and civil liberties (5–15%).

Our approach to helping students learn these topics both broadly and deeply was to adapt project-based learning (PBL) to the AP environment using a learning cycle approach. PBL is a broad and often unspecified umbrella term for a wide range of pedagogies. In this design experiment, we specified it with these five principles, which guided our course design:

rigourous projects as the spine of the course,
quasi-repetitive project cycles (looping),
engagement first,
teachers as co-designers, and
an eye for scaleability.

*Projects as the spine of the course*

In PBL, students, working collaboratively and sometimes alone, learn knowledge and skills through an extended inquiry process structured around complex, authentic challenges and carefully designed tasks and products (Darling-Hammond *et al.* 2008, Ravitz 2009). When PBL is well done, students have multiple opportunities to bridge their own prior, informal, and local knowledge with disciplinary, formal, and broadly vetted knowledge, reflect on their progress, and revise in order to 'work smarter' as the course proceeds. Project work requires heightened communication—lots of public talk as students collaborate to interpret texts and problems and make decisions (Parker 2010). Through project activities—and the recurring phases of project anticipation, execution, and reflection—students have multiple meaningful opportunities to try out their current levels of understanding, revise them, and in this way deepen them. This meant inverting the typical course organization where projects, if any, are treated as special add-ons or end-of-course capstones—valuable activities done *after* reading and remembering has been done, *after* 'background' information has been acquired (Paek *et al.* 2005). Instead, we aimed to create a course experience where challenging projects provided the spine of the course, not the appendages; that is, the entrée, not the dessert; the main show, not the side show: the core of the teaching and learning regime.

*Depth through looping*

Quasi-repetitive project cycles (Bransford *et al.* 2006) or what our teachers dubbed 'looping' means that students have opportunities to revisit questions, ideas, and problems. This, we reasoned, is a key to deepening (complicating, differentiating, and integrating) their evolving understandings of the core AP topics. Expertise in any domain, from playing baseball to making public policy, appears generally to grow with the right sort of practice—with trying again under somewhat different conditions. The course had five projects, each conceived as a knowledge-in-action learning cycle where students alternate between two modes—learning to act and acting to learn. 'Learning to act' here is when students are in traditional AP mode (textbook/lecture/test prep) and 'acting to learn' is when they are engaged in complex projects with real-world goals. The anticipation of one mode helps motivate and drive the other (National Research Council 2000). Moreover, the project cycles are united by a course 'master question'—for this course, *What is the proper role of government in a democracy?* As students move through the five project cycles, they repeatedly respond to (loop back to) the master question and 'try again' to generate a response, reflecting on what they gleaned from the prior project cycles and the project cycle at hand. Through this looping, we conjectured, knowing and acting would deepen in tandem (Bransford *et al.* 2006).

*Engagement first*

Schwartz and Bransford (1998: 476), in a paper entitled 'A Time for Telling', explored '*when* to use texts, lectures, and explanations within the total repertoire of instructional methods', and concluded that there is a 'readiness' for learning from textbook readings or lectures *after* some understanding has been generated in other ways. A central principle in our design experiment, 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) would typically precede 'telling' (e.g. how Congress interacts with other institutions of national government). The purpose of this sequencing is to create a readiness for telling so that the information students gain by it, whether through reading a textbook chapter or listening to a lecture, is *needed* for making progress on the project and constructing a suitable understanding. In this way the telling has 'somewhere to go' because there is already something going on—students are already engaged in an action arena in which the telling can be of service; that is, the telling serves to explain and elaborate what is going on in the project work. 'When telling occurs without readiness', Schwartz and Bransford (1998: 477) 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'. This is a key reason why our team chose PBL as the basic architecture for the course, for it reverses the piling on of

what is commonly called 'background information' *prior* to project work, presumably so that students will 'know enough' to participate in the project. That sequence, inverted here, is a persistent 'grammar' of schooling (Tyack and Cuban 1995: 85) and one that can prevent students from ever getting to *use*, and thereby to transform, information in action.

### *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 design experimentation on complex classroom interventions are inevitably collaborative undertakings. This allows the design to be 'constantly revised based on experience' (Collins *et al.* 2004: 18; also Bransford *et al.* 2010). This course design put our teacher collaborators in the position of being curriculum makers—continually working to integrate AP (centred on the College Board's Course Description) with a set of carefully selected projects (see 'Scaleability' below). To do this, teachers (and all of us learning to integrate these two) needed to create a course flow that loops effectively from one project cycle to the next in pursuit of success on the AP exam plus deep learning. Teachers and researchers needed to weave together AP content and routines with project content and routines so that deep knowledge and engagement could build across the course.

### *Scaleability*

Per our fourth goal, our aim was not a 'hot house' experiment that would display what is possible but improbable. 'Complex interventions in classroom practice', Brown's (1992: 141) moniker for design experiments, are not satisfied with establishing an existence proof (verification that a change in practice is possible) but are done with an eye to scale. 'This is intervention research designed to inform practice', she wrote (p. 143); accordingly, 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'. We worked, therefore, toward a design that could be adapted by others who could, in turn, further the design experiment in other circumstances, thus widening the community of teachers and researchers working to deepen learning in exam-based courses. Therefore, we attempted to create no projects from scratch but to adapt the best of the best projects from other practice communities. The first project, for example, was *A Government for Xlandia*, developed by the Buck Institute of Education (www.bie.org). Another, called 'Congress 111', was *LegSim* developed in the political science department of the University of Washington (www.legsim.org). Another was the popular 'moot court'—simulations of Supreme Court hearings (www.landmarkcases.org). By so doing, we were able to rely on

projects that not only are readily available elsewhere but that already are somewhat mature in that they have been through the (re)development process, vetted and revised in numerous practice arenas.

Four of the five projects cycles were simulations that involved role-taking, and one (Making a Difference) involved no role-taking. The five project cycles were:

(1) *A Government for Xlandia*. Students are members of a UN task force advising a new nation just emerging from a long dictatorship about the various forms and features of constitutional democracy.
(2) *Making a Difference*. Students interact with government and community leaders to propose public policy and action that will improve society.
(3) *Congress 111 (LegSim)*. Students are legislators in the US Congress, participating in politics, strategy, and social change.
(4) *Election 2008*. Students are party strategists advising candidates in the November congressional election.
(5) *Supreme Court*. Students are justices of the Supreme Court or attorneys arguing before the Court.

Daily classroom life in the course involved orchestrating these projects in accordance with the design principles and alongside textbook readings and lectures, and, always on the horizon, the impending AP exam. The 'engagement first' principle meant that readings and lectures were positioned inside project work. In *Congress 111*, for example, students immediately took the role of legislators setting up their offices and committees, and determining and advancing their legislative agenda; meanwhile, they read and listened to lectures and saw videos about how Congress works in relation to the other institutions of national government, political parties and beliefs, interest groups, and the Constitution (see the list of six topics). One day of *Congress 111* might feature legislative committee work, the next day a lecture or preparation for a floor debate, and the next day a mid-unit assessment of student learning. Homework consisted of reading, planning, and reviewing as well as working collaboratively at the project's website at www.legsim.org. A few students in each classroom were designated videographers and would use Flipcams to interview classmates and film committee meetings and other legislative events. Eventually, a culminating performance activity—a floor debate with an elected Speaker presiding—completed the project. An adult expert (e.g., a lawyer or legislator) was invited to play a role in the culminating performance. This elevated the authenticity of the project while affording students feedback on which aspects of their performance rang true or not to the expert's knowledge and experience.

## Methodology

This was a design experiment—an *experiment* in the sense that a particular curricular and instructional treatment is tested but a *design* experiment in

the sense that an innovation is designed and iteratively tested and modified in a real educational setting, not a laboratory (Brown 1992, Stevens *et al.* 2005). 'Design' here is an adjective modifying 'experiment', but it also has a verb sense, for the design work is ongoing and under revision at all times. Brown (1992), whose seminal work coalesced the design experiment concept and method while creating the complex classroom intervention known as 'reciprocal teaching' (Brown and Palinscar 1987), noted the theoretical and methodological challenges researchers face when they are (re)designing as they go—when they are 'simultaneously involved in designing under conditions of continuous flux' (Brown 1992: 152):

> Components are rarely isolatable, the whole really is more than the sum of its parts. The learning effects are not even simple interactions, but highly interdependent outcomes of a complex social and cognitive intervention. And this presents a methodological headache for traditional psychology, allergic as it is to multiply confounded experiments. (p. 166)

A teachers' strike occurred at the beginning of the school year, for example, just when the first project cycle was to be implemented. This caused an immediate revision and loss of instructional time. At the same time, one of the teachers of the experimental course took maternity leave and was replaced by a teacher who was not only new to AP but new to teaching. Improvisation is the exception in controlled laboratory experiments but the norm in design experiments. Collins *et al.* (2004: 18) put it plainly: 'Because design experiments are set in learning environments, there are many variables that cannot be controlled'. This is not always a disadvantage, for the design tradition as applied to education emphasizes the importance of learning *during* design so that the designers (here, the team of teachers, curriculum specialists, and researchers) 'can intervene purposefully in the situation at hand', with its emergent and unpredictable properties, 'to bring about a desired state or reach a desired goal' (Bransford *et al.* 2010: 841–842).

Our research was conducted in three high schools in a suburban school district in the western region of the US during the 2008–2009 school year. The general research question was, Can rigorous project-based learning be applied to high school AP courses and improve student learning? Specifically, (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? (3) Will students report greater engagement in their learning experience, both inside and outside the classroom?

In this article, we report the 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 at the end of the year to learn their summative views on the course, thereby qualifying the quantitative experimental findings with students' own descriptions of their engagement in the course, while also giving them a voice in the re-design of the course for the following

year. This is faithful to design-experiment methodology, with its aim to improve practice through iterative modifications of the intervention.

### *Research design*

The research consisted of two studies. Each study used a non-randomized intervention design in which outcomes from PBL-AP classes were compared to those for traditionally-taught AP classes. Study 1 compared the results for students in four PBL-AP classes at a high-achieving school (School A) with the results for students in four traditionally-taught AP classes at a second high-achieving school (School C). Study 2 compared the results for students enrolled in four PBL-AP classes at a moderate-achieving school (School B) to the four classes that were in the high-achieving control school (School C). For comparability, all classes were year-long AP *US Government and Politics*. Class sizes ranged from 23–32, with a mean class size of 28. Within each school, all of the year-long AP *US Government and Politics* courses were of the same type (i.e. PBL-AP or Traditional AP). In total, 314 high school students took part in the research, in 12 classes at the three high schools (eight PBL-AP classes in two schools and four Traditional AP classes at School C). Table 1 displays the numbers of students, classes, and teachers in each research group.

To control for students' prior achievement across schools, we used hierarchical linear modelling (HLM) for the statistical analyses of results. The models took into account students' prior achievement, including the student's GPA and scores on the PSAT, the Washington Assessment of Student Learning-Reading (WASL), and prior AP tests.

### *School selection process*

The district's recommendation for the intervention school was a high-achieving school (School A)—one of the two highest achieving of the five high schools in the district. We also selected a school where we could test the intervention with a less advantaged student population. This was a

**Table 1. Number of students, classes, and teachers by school and research condition.**

| | | Total number of students | | Classes | Teachers |
|---|---|---|---|---|---|
| Study 1 | High achieving school A | 106 | PBL AP course | 4 | 2 |
| | High achieving school C | 106 | Traditional AP course | 4 | 1 |
| Study 2 | Moderate achieving school B | 102 | PBL AP course | 4 | 2 |
| | total | 314 | | 12 | 5 |

moderate-achieving high school (School B). Of the schools in the district already with an AP *US Government and Politics* course, this was the school whose student population was least advantaged compared to our first intervention school (although certainly not disadvantaged by national standards). As the control school, we chose the other of the two highest-achieving schools (School C). School C's student population was demographically most like the high-achieving intervention school, and its students in previous years performed similarly on the AP *US Government and Politics* exam. It, too, has a reputation for strong academics and strong AP test performance.

### *Teacher selection process*

Due to the small number of teachers at each school, we could not freely select teachers for participation in either the PBL-AP or control sites. However, all five teachers (PBL and Traditional course) were well-credentialed and deemed effective by the district's social studies curriculum director. Each teacher had a Master's in Teaching earned from a university teacher education programme in the past 6 years. Each also had a bachelor's degree in political science or a related field and 1–8 years classroom teaching experience. (The Traditional course teacher had the most years teaching experience.) Only one of the PBL-AP teachers had taught year-long AP *US Government and Politics* before, and this teacher, along with the district's social studies curriculum director, helped to instigate what eventually became this collaborative design experiment. In contrast, one intervention teacher at the high-achieving school and one at the moderate-achieving school were entirely new to teaching this course. Both were interested in the project, however, and were welcomed aboard. In this way, the intervention teachers (with one exception) were relatively less experienced with teaching and/or with this particular AP course, and one could argue that this relative lack of experience makes for a conservative test of the effects of our PBL-AP design.

### *Measures*

To address our three research questions, measures of student learning and engagement were collected at various points during the 2008–2009 school year. In this report, we discuss the results from the following mea-

**Table 2. Student learning measures.**

| | Pre- | Post- |
|---|---|---|
| AP Test<br>*created by the College Board* | | √ |
| Complex Scenario Test<br>*a deep learning assessment created by the research team* | √ | √ |

sures: the College Board-administered AP test and the Complex Scenario Deep Learning Assessment, which we created.

Table 2 lists the measures and the schedule by which they were administered (at present and/or at post-test).

*College board-administered AP test*

The AP *US Government and Politics* exam was officially administered in May 2009.

*Complex scenario test*

The Complex Scenario Deep Learning Assessment was created by the research team. It used a simulated 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 Complex Scenario Test aims to assess how well students can *apply* that knowledge in a particular scenario where their charge is to monitor and influence public policy—specifically, to formulate a plan for intelligent political action on a controversial issue that is currently in the news and heating up. Key features of the Complex Scenario Test include:

> Places students in the role of adviser to a congressperson or a grassroots citizen group in a particular scenario.
>
> In the scenario, students must mobilize knowledge from across the project cycles to draft an action plan for their client.
>
> Students are told their client will want to decide for her/him/itself; therefore, as advisors, students need to justify their recommendations.
>
> The scenario centres on a controversial issue, one the students are not expected to be expert in (facts are provided within the scenario; for example, 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 (e.g. news article)'.

Whereas the AP Test is given once, at course end, the Complex Scenario Test was given as a pre-test/post-test, once during Week 3 of the Course and again during Week 36 or 37. The pre-test and post-test were on the same controversial issue; however, we modified the post-test scenario to represent a different stage in the political process. The pre-test was a 1-day (50-minute class period) paper and pencil assessment. The post-test took 2 days. Day 1 of the post-test (a 50-minute class period) consisted of the paper and pencil task. Then, for learning purposes, day 2 consisted of teamwork: sharing and further developing individuals' responses from Day 1, culminating in a 5-minute presentation that each team spoke into an audio recorder (such that the teams were presenting simultaneously). The results presented here are for the paper-and-pencil

portion of the test only—that portion of the test amenable to pre/post analysis at the individual student level.

### *Qualitative data and analysis*

A sample of students in each of the eight PBL-AP classrooms were interviewed individually three times during the experimental course. Then, following the administration of the two post-tests, each of the eight PBL-AP classes was interviewed in large-group, fishbowl format in late May. In this article, we address data from the fishbowl interviews. Our objective was to learn from students' experiences in the course so that we could gauge its effectiveness for them while affording them a voice in making adjustments to the course design for the project's second year. Using rotating inner circles of speakers, with opportunities for students in the outer group to contribute on any question, 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 US 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?

The group interviews were recorded and transcribed. Transcripts were then read iteratively and themes were generated that were pertinent to the course design. The relationship between frustration and learning, in the context of a high-stakes, exam-based course, emerged as a prominent theme. Following the presentation of the qualitative findings in the next section, we turn to this theme where we tease out students' thinking about how to coordinate the projects, lectures, readings, and impending AP test along with their thinking about the 'engagement first' course design principle.

## Quantitative findings

In this first findings section, we present the key findings from studies 1 and 2, respectively, as they relate to the first two of our guiding research questions: (1) Can we create a PBL-AP course in which students do as well or better on the AP test compared to students in a traditionally-taught AP course? (2) Will students taking the PBL-AP course demonstrate a deeper level of knowledge as assessed by a complex scenario test?

**Table 3. AP Test scores by research condition.**

| | PBL AP course School A (study 1) (N=103) | | PBL AP course School B (study 2) (N=100) | | Traditional AP course School C (studies 1 & 2) (N=66) | | PBL AP course Schools A & B combined (N=203) | |
|---|---|---|---|---|---|---|---|---|
| score of 5 | 29 | 28.2% | 8 | 8.0% | 4 | 6.1% | 37 | 18.2% |
| score of 4 | 18 | 17.5% | 10 | 10.0% | 9 | 13.6% | 28 | 13.8% |
| score of 3 | 31 | 30.1% | 21 | 20.0% | 21 | 31.8% | 52 | 25.6% |
| score of 2 | 21 | 20.4% | 36 | 36.0% | 19 | 28.8% | 57 | 28.1% |
| score of 1 | 4 | 3.9% | 25 | 25.0% | 13 | 19.7% | 29 | 14.4% |
| score 3 or higher | 78 | 75.7% | 39 | 38.0% | 34 | 51.5% | 117 | 57.6% |
| score 4 or higher | 47 | 45.6% | 18 | 18.0% | 13 | 19.7% | 65 | 32.0% |
| Mean score | 3.46* | | 2.40 | | 2.58 | | 2.94* | |
| *SD* | *1.21* | | *1.20* | | *1.14* | | *1.31* | |

*$p < .05$ (comparison between PBL AP course mean and traditional AP course mean).

*Results of study 1*

Study 1 compared the results for PBL-AP and Traditional-AP students from the two high-achieving schools (A and C on the tables). These *schools* had been selected for the study because they matched on prior achievement and student SES. To determine if the intervention and control *students* in our study samples were also similar, we compared their prior achievement (as defined earlier) and found no significant difference ($p < 0.05$). The results for students in four traditionally-taught AP classes were then compared with the results for students in four PBL-AP classes. HLM (random coefficient) models were used to compare the results, taking into account the students' nesting within classrooms, treatment condition, and students' prior achievement.

Regarding question 1, the results indicate that it is possible to get higher scores on the AP test with a PBL course. Table 3 shows that PBL students scored significantly higher ($p < 0.05$) on the AP test than the traditionally-taught AP students. Also, because many colleges assign college credit for AP scores of 3 or more, we compared 'passing' scores of 3 or more on the AP test. The results showed that more PBL students at school A (75.7%) achieved a passing score on the test than Traditional students at school C (51.5%).

Regarding question 2, we designed the Complex Scenario Test to measure students' deeper understanding of course content. 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; the Complex Scenario Test looks at how well students can *apply* what they learned to a novel problem. Students are given a real-world controversial issue and asked to formulate a plan for well-informed, smart political action. Students' written answers to the Complex Scenario Test were scored on four dimensions, and the quality of the student's answer related to each dimension was assigned a score ranging from 1–6 (6 = highest quality answer). 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 terminology).
(4) *Grasping Controversial Issues*: Analyses the public policy issue at stake and what makes it controversial.

Scoring involved making two passes through the 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 (88%), Influencing Public Policy (84%), Grasping Controversial Issues (81%), and Overall Quality rating (93%). When the two raters did not agree within one

point, the paper was scored by an independent third rater. In all cases the mean rating was assigned to the response. Only seven papers of 264 required a fourth rater.

Table 4 shows average scores for PBL and Traditional students on the four dimensions of the test. PBL students scored significantly higher ($p < 0.05$) on all four dimensions of the Complex Scenario Test as compared with students in the Traditional course. In this analysis, scores were adjusted for prior achievement using WASL-Reading scores and pre-test scores on the Complex Scenario Test. These findings suggest that 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.



### *Results of study 2*

Study 2 also compared the results for students enrolled in AP classes in two high schools, but this set of four PBL classes was in a *school* that historically had fewer students taking or passing the AP test as compared to the four classes in the control school (school B). We also looked at the entering characteristics of *students* in the PBL and traditionally-taught courses. As expected, based on the two schools' historical data, the entering PBL students were significantly lower on prior achievement ($p < 0.05$) than the entering Traditional course students. The results for students in the four traditionally-taught AP classes (School C) were compared with the results for students in four PBL classes. Again, HLM (random coefficient) models were used to compare the results taking into account the students' nesting within classrooms, treatment condition, and student-level prior achievement.

Regarding research question 1, in Study 2 the PBL students performed as well on the AP test (mean score = 2.4) as the traditionally-taught students from the higher-achieving school (mean score = 2.6). There was not a significant difference ($p < 0.05$) in the scores of the two groups. Goal 1, recall, was 'same or higher scores'. Also, more students in the PBL school passed the AP test in the year that the research was conducted (2008–2009) than in previous years at the school when the course was not project-based. As for passing scores of 3 or higher, the results showed that fewer PBL students (38%) achieved a passing score on the test when compared to Traditional students (51.5%) (see table 3).

Regarding question 2, as discussed earlier students' written answers to the Complex Scenario Test were scored on the test's four dimensions. Table 4 shows average scores for PBL and Traditional students on each dimension. PBL students scored significantly higher on all four ($p < 0.05$) dimensions of the Complex Scenario Test as compared with students in the Traditional course. As in Study 1, scores were adjusted for prior achievement using WASL-Reading scores and pre-test scores on the Complex Scenario Test. These findings suggest that PBL students more deeply understood the AP content; they were able to apply it to solve a complex problem.

**Table 4. Complex Scenario Test scores by research condition.**

| | PBL AP Course School A (study 1) (N=95) | PBL AP Course School B (study 2) (N=78) | Traditional AP Course School C (studies 1 & 2) (N=91) | PBL AP Course Schools A & B Combined (N=173) |
|---|---|---|---|---|
| Measure | Mean (SD) | Mean *(SD)* | Mean *(SD)* | Mean *(SD)* |
| Overall | 2.34* *(0.91)* | 2.07* *(0.83)* | 1.61 *(0.75)* | 2.22* *(0.88)* |
| Task & client | 2.42* *(0.97)* | 2.15* *(0.84)* | 1.62 *(0.76)* | 2.30* *(0.92)* |
| Influencing public policy | 2.20* *(0.82)* | 1.92* *(0.82)* | 1.53 *(0.70)* | 2.07* *(0.83)* |
| Grasping controversial issues | 2.60* *(0.83)* | 2.13* *(0.77)* | 1.76 *(0.77)* | 2.38* *(0.88)* |

*$p < .05$ (comparison between PBL AP course mean and traditional AP course mean).

## Student reflections: frustration and achievement

We turn now to our interview data in order to explore the adjustments students needed to make to the PBL-AP design. We wanted to understand students' responses to the alternative course and their advice for its future designs. Especially for experienced AP students, the alternative course design was definitely not 'business as usual', we knew, and their reflections were of particular interest as the year was ending and the research team began to consider modifications of the course for the following year.

Students had been interviewed individually three times during the experimental course and then in whole-class focus groups in late May, after the AP exam. In this space, we address data from the latter because they present the opportunity to peer into students' public, summative thinking about the course.

In this first section we focus on students' reflections on the course design, particularly as regards a problem central to the depth/breadth tension: As noted earlier, the impending AP exam's sweeping breadth of tested topics produces in traditionally-taught AP courses a great deal of telling—a great laying-in of information from both textbook reading and listening to lectures. The experimental course design attempted to orchestrate this information-giving with the circumstances in which students *needed* that information as a *tool* in their project work. We were aware that this would require some students to adjust how they 'do school', particularly AP veterans but also newcomers who may be accustomed to a routine pedagogical diet of telling-and-testing or projects as side-shows rather than as the spine of the course. This routine, to which some students in these interviews acknowledged they have been 'programmed' ([B2: 3:51]: 'A text-based based ... curriculum: that's sort of what we're programmed to use'), was, by design, replaced in the experimental course with our adaptation of project-based learning described earlier.

We are struck especially by two facets of students' reflections on the PBL-AP course. One is their pedagogical thinking about how to coordinate the central components of the course—the textbook readings, lectures, projects, and the AP exam. Closely related is their attention to the 'time for telling' problem (Schwartz and Bransford 1998). Being 'thrown', as one student put it, into the projects without benefit of much prior knowledge or 'background information' was, as displayed in some of the comments that follow, at times frustrating; students sometimes felt that they were floundering. Compared to traditional AP courses, this project-centred approach put students on a new, or as one student said, 'unorthodox', landscape.

### *Orchestration of course components*

We begin with a student who shares her views on the relationship of textbook readings, lectures, and project work. Each is important to her, and

she explains that project work represents 'doing', which needs to be informed by 'background', that is, knowing '*what* we're doing'.

> Lectures actually are really important, and for the unit that we didn't have lectures in, I felt sort of lost. ... So, I think it's really important to strike a balance between the readings, which we are going to have to do because we can't learn or do everything in class; the lectures, so we do know what we're doing; and the projects, which will help us incorporate it. But you can't miss out on one of those, because if you have projects and no lectures, then we're going to be just doing stuff just to get the project done, because we won't really understand it ... [A1:00:35:14.04]

Then, she offers a hypothesis about the effective sequencing of these components: Perhaps project work should *precede* the background information. Here she echoes one of our course design principles, Engagement First, which holds that there is greater readiness for learning from telling *after* some understanding and need-to-now have been generated by actual engagement with the project.

> Another thing though, background is really good, but if you don't know what you're learning it *for*, so maybe introduce a project, *then* learn the background.

If you don't know 'what you're learning it *for*', that is, how and in what circumstances you'll be *using* this information, there may be no readiness for it and no ability to use it as a means to other ends—and thus no way to push beyond a shallow understanding. This is precisely why our course design placed engagement before telling.

In this next excerpt, a second student agrees that each component is valuable, but he wants a reverse sequence.

> Well personally, I like both components of how Mr. ___ went about teaching this course, where we had the projects as well as his lectures. I think both of them are very valuable. I also really enjoy his lectures and I think they're very, very helpful. I think it would have been more helpful if we had the lectures first, or he would teach us some more information and then give us the projects, instead of just giving us these projects and giving us very little guidance and hoping we'll go out and figure out what to do and how to do it. I prefer to work on something if I have a basis of knowledge to start from and then from there I can build on it, which I think is much more efficient than just trying to figure things out and not really knowing what I'm doing or whether what I'm doing is correct. [B1: 00:17:42.11]

A third and fourth student agree with this sequencing model, pointing out that beginning with a project—diving into the water rather than only sticking a toe in—can leave students floundering. Asked to sink or swim, they may sink.

> This course kind of assumes that we had a basic knowledge of the different components of government, but I think for that there was ... almost too much of an assumption. We didn't know as much as you thought we did. [B1: 00:20:19.23]

> I completely agree ... I would have liked it if we were to start off with kind of like 'politics for dummies' or 'government for dummies' type of deal, just to build off of that, because we started off so quickly that instantaneously I was lost. [B1: 00:20:19.23]

Agreeing, a fifth student takes up the swimming pool metaphor and offers a theory about how students like her think and learn.

> Being pushed into the government pool without any floaties is kind of difficult because you ... it's like a process: you learn things and then you go and test it out, like in our life, that's the way things work in our mind. Like, we start from baby steps. We've gotten (almost; they are seniors) to the point where we are thrown off into college and do our own thing, based on the knowledge we have. Here, (however) it's like we are born and (then) we are thrown into college [laughing]. [B2: 17:27]

Another student agrees, but uses a different metaphor from the television show, *CSI*:

> It would be great to start a class with like, for example, if it was some CSI course, there is a murder, that is the course of the year, solve the murder kind of deal. You know [laughing], but in this case it was make a difference or write your project or create a government for Xlandia. But without any of the steps, I don't know how to fingerprint a dead body, same kind of deal; I don't know how to start writing a project. [B2: 17:27]

These students are wrestling with how course components can best be sequenced for learning, and all but the first would prefer that the course somehow lay a foundation of prior knowledge—some sort of 'government for dummies' or 'floaties'—so that the project work feels more stable and grounded, with less floundering and less feeling like 'I don't really know what I'm doing or whether what I'm doing is correct'.

### *Preparation for the AP test*

Recall, however, that students in the PBL-AP classes performed as well as (Study 2) or significantly better than (Study 1) students in the Traditional-AP classes on the AP test, and significantly better (both studies) on the complex scenario test of deep and applicable knowledge. This next set of excerpts may help to explain why. Here, five students reflect on the course vis à vis their preparation for and success on the AP exam. The course was 'unorthodox', the first student concludes.

> The whole way we did it was so unorthodox. We are so used to textbooks, learning what's supposed to be on the AP Exam. Just going in a linear fashion, whereas (in this class) you *felt* like you weren't learning anything; like, you *were* learning something, but it was different than other AP courses and that was pretty interesting. [C3: 00:07:20.28]

If the telling-and-testing routine *felt* like learning, this young man seems to be saying, then the alternative project-centred routine did

not. It was beyond the routine to which he was accustomed. Yet, he *was* learning.

Another student attributes this to both the 'hands-on' project work and the inter-personal dynamic of collaborative group-work, reflecting Cohen's (1994: 3) observation that students working in groups 'become actively involved with their work and are held there by the action of the group'.

> So I think this class did a lot better in preparing me for the AP test than other APs I've done. Part of it is the group-work that encouraged me to actually do my work, even though I still didn't have the best track record. Part of it, I think, is the hands on learning rather than just listening to lectures and book learning the whole time, at least for people like me, was a huge improvement (over) the AP's I've taken. When I went in to take this one, I felt best about this and I came out feeling best, and I knew this stuff better than I've known any other AP, and I feel this just really worked. [A1:00:33:19.09]

The next two students echo a course design principle: 'When telling occurs without readiness, 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' (Schwartz and Bransford 1998: 477).

> So, I felt like learning it in this kind of setting where we're actually doing projects and stuff was definitely a lot more benefitting than any other class where you just read out of a textbook. Because with all the other AP classes, we learn mainly just to get the credit on the AP test, and then after we get that credit, it's almost like goodbye. But whereas we have the experience of actually trying to do this, I think I'm definitely going to remember it because I've actually done it before. Maybe not actually 'done' it, but we did it in a *simulation*, and that was, I think, the most beneficial thing in actually learning it. Because just reading out of a textbook, you memorize all of this content, but when you don't need it anymore, you never actually use it, only just to write it on a piece of paper afterwards, just to make sure you remember it, but this is actually going through the process. You actually get to understand what all goes into politics and you become a lot more involved or at least a lot more aware of how things work so you can make sense of things while other people may say things about government, you can actually understand why government works the way it does. [A1: 00:36:49.01]

> I feel it taught us more, like project applications for what we learned. So a lot of the (AP) Chemistry and English (courses) seemed to go towards the test. We have a lot of work to spend towards this, but these skills I can actually use towards other things ... I really like that because I feel like this is actually useful stuff that matters as opposed to just being memorization for silly stuff that only matters once. [A1: 00:41:08.21]

But another student, also an AP veteran, judges the course not on how or what he is learning but solely on how efficiently it teaches to the test.

> 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 they might not have been directly learning about things that was on the test. [B2: 9:49]

## Discussion

Contemporary research suggests that learning well depends on (a) students' ability and inclination to reflect on their learning; (b) how and when they are encouraged to use, revise, and deepen prior knowledge in tasks that ask them to construct new knowledge; (c) scaffolding to elicit performances beyond what could be achieved without it; and (d) classroom and curriculum organization (National Research Council 2000). Advanced high school courses, like those leading to the A-Levels and AP exams, place extraordinary demands and constraints on these conditions by packing so much curriculum into a single, time-bound course, and then capping it with a high-stakes, breadth-oriented exam. We approached the depth/breadth tension with a particular co-ordination of project-based learning and curriculum coverage; that is, we tried to mobilize PBL on behalf of both coverage and learning. Coverage was addressed by teaching to the broad sweep of the College Board's six topics and by focusing students' attention on the impending AP exam. Meanwhile, we orchestrated in-depth learning through quasi-repetitive project cycling and an engagement-first policy that routinely put 'telling' at the service of a situated need-to-know. Sustained inquiry on a course master question united the projects and helped to assure a regular 'revise again', iterative cycling in the course design. Collaborative work, including role-taking and public discourse, was a central feature.

Students in the experimental course scored, on average, as well as (in study 2) or better than (study 1) students in the control classes on the AP exam ($p < 0.05$), and both scored better than control classes on an alternative, complex-scenario test of deep conceptual learning ($p < 0.05$). However, the qualitative data presented above show that it wasn't always smooth sailing. A good number of students experienced frustration with the 'engagement first' design or the way that principle was implemented in the first year. We cannot say for certain what caused the frustration or whether the frustration aided or undermined success on the two exams. We suspect, however, that some of the frustration was due to students—especially AP veterans—having been acclimated or, as some said, 'programmed' to a particular way of 'doing school': hearing lectures and reading texts, memorizing information, and then taking a test. In other words they had become accustomed 'to treat the new information as ends to be memorized rather than as tools to help them perceive and think' (Schwartz and Bransford 1998: 477). The adaptive learning, engagement-first design, in contrast, put them repeatedly in real-world political roles as legislators, justices, campaign aids, and so forth before they fully understood those roles or their relationship to one another in the broader framework of US government and politics, thus creating a need-to-know. They continually were having to figure out, 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 adaptive learning environment in which they now found themselves. 'In adaptive learning situations, such as challenge- or problem-based instruction, many students initially attempt to use their efficient learning strategies and find that they are inadequate' (Martin *et al.* 2007: 150).

However, we are not inclined to explain away students' feelings of floundering and frustration as either a necessary aspect of powerful learning nor as a needed corrective to routine A-Level or AP 'programming'. It is true that adaptive experts need to be able to weather ambiguity and persevere through messy problems, and equally true that rigourous learning can be enjoyable, as anyone knows who experiences 'flow' when 'lost' in a good challenge (Hatano and Inagaki 1986, Csíkszentmihályi 1996). However, the Engagement First design principle does *not* mean that students should be 'thrown into the deep end of the pool', and research on 'time for telling' (Schwartz and Bransford 1998) does *not* recommend it. Accordingly, we want to reduce the frustration that accompanied PBL for some students in the re-designed course. Because some amount of floundering and ambiguity is inherent in authentic project work, we don't want to reduce it to a negligible level, but we do want to make PBL-AP more enjoyable for more students.

As a consequence of listening to these student voices, the second iteration of the course in year 2 of the project was designed not only to improve the coherence and connectedness of the several projects but also to reduce student frustration. Preliminary data analysis suggests that this goal is being achieved. The teachers in year 2 have accumulated experience and, with it, a better sense of how the course hangs together—experience and knowledge that was impossible in the first iteration of the course, such as how projects connect to one another in actual practice, the routine phases of any project, and how to preview and review project work. Accordingly, they are better able to communicate this to students, providing advanced organizers and assisting students along the way—more and better 'floaties', if you will—as required by the Engagement First design principle. Simply put, teachers and researchers learned from year 1, and we are seeing this in year 2. This orientation work appears to be providing some of the assistance students need when navigating challenging project work. We are seeing in interview data that students in year 2 are experiencing less frustration while, we anticipate, still accomplishing the two primary goals: same or higher scores on the AP exam plus deeper conceptual learning.

We believe the design-experiment methodology was fruitful. Of course, it was 'a compromise' (Stevens *et al.* 2005: 140), for it was located in 'the ongoing dilemma of doing basic research in educational settings while at the same time supporting and improving educational practices for the students and teachers participating in the experiment'. Improvisations and modifications were common, and while they confounded the experiment they were done for the sake of improving learning outcomes in these classroom settings as well as caring for the well-being of students. Generalizability is not primarily the goal of design experiments as it is in pure

experiments; still, an effective design experiment should, as Brown (1992: 143) wrote, 'be able to migrate from our experimental classroom' to other classrooms. These are related but distinct goals. It remains to be seen whether our alternative AP design can migrate to other settings and courses, but we did design the course with that aim (goal 4). We intend to continue to modify the design in subsequent years of our work as we 'migrate' to other courses (*AP Environmental Science*) and to other settings (under-served urban schools). Additionally, we are eager to further develop the validity and reliability of our alternative measure, the Complex Scenario Test. Is the definition of deep learning (complex and applicable) with which we began this article adequate? Does this test measure it precisely and usefully? We are undertaking additional work on the alternative measure as we write this, and these questions are guiding it.

We believe the major contribution of this study, although limited in these ways, is to have begun to develop an approach to co-ordinating depth and breadth of learning in high-stakes, exam-based courses. It is an approach that may be relevant not only to other AP courses but also to similar, advanced high school courses in other countries. This is because it retains the high-stakes, breadth-oriented test but joins it with a course design that features active project-based learning at its 'spine' alongside quasi-repetitive cycling and sustained inquiry on a well-chosen master course question. Furthermore, the high-stakes, breadth-oriented test is partnered with (not replaced by) a measure of complex and applicable learning, which the course design is intended to achieve in tandem with a broad grasp of the main course topics. Perhaps, then, we have demonstrated how a reasonably high-quality, breadth- and exam-oriented curriculum, one that enjoys considerable public confidence, can interact with real-world projects that aim for a kind of learning and knowing that are deep and transferable. Now that international comparisons of test scores are the fashion, and the AP footprint is expanding dramatically in the US, we hope other researcher–teacher teams will tackle the *learning* issues inherent in this activity while making advanced courses more engaging and successful for a wider array of students.

## Acknowledgements

The authors are grateful to the George Lucas Education Foundation, the Bill and Melinda Gates Foundation, and the National Science Foundation grant (NSF-SLC# 0354453) for their support of this work, and to the Bellevue (WA) School District, including social studies curriculum co-ordinator John Brill and the teachers with whom we developed the alternative design described here—Amber Graeber, Adrienne Curtis Dickinson, Katherine Piper, and Donald B. Myers. The ideas expressed in this article are not necessarily those of our sponsoring partners.

## References

Barron, B., Schwartz, D. L., Vye, N. L., Moore, A., Petrosino, A., Zech, L. and Bransford, J. D. (1998) Doing with understanding: lessons from research on problem- and project-based learning. *Journal of Learning Sciences*, 7(3–4), 271–311.

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

Bransford, J. D., Mosborg, S., Copland, M. A., Honig, M. I., Nelson, H. G., Gawel, D., Phillips, R. S. and Vye, N. L. (2010) Adaptive people and adaptive systems: Issues of learning and design. In Hargreaves, A. Lieberman, A. Fullan, M. Hopkins, D. (eds), *Second international handbook of educational change* (Dordrecht, The Netherlands: Springer), 825–856.

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

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

Brown, A. L. and Palinscar, A. S. (1987) Reciprocal teaching of comprehension strategies: A natural history. In Day, J. D.Borkowski, J. G. (eds), *Intelligence and exceptionality: New directions* (Norwood, NJ: Ablex), 81–132.

Cohen, E. G. (1994) *Designing groupwork* (2nd ed.) (New York: Teachers College Press).

College Entrance Examination Board (2003) *A brief history of the Advanced Placement program* (New York: College Entrance Examination Board).

College Entrance Examination Board (2008) *US Government and Politics, course description* (New York: College Entrance Examination Board).

College Entrance Examination Board (2010) *The 6th annual AP report to the nation* (New York: College Entrance Examination Board).

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

Csíkszentmihályi, M. (1996) *Creativity: Flow and the psychology of discovery and invention* (New York: Harper Perennial).

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

Goodlad, J. I. and Klein, M. F. (1970) *Behind the classroom door* (Worthington, OH: C. A. Jones).

Greimas, A. J. (1987) *On meaning: Selected writings in semiotic theory*, P. J. Perron and F. H. Collins, Trans. (Minneapolis: University of Minnesota Press).

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

Hirsch, E. D., Jr. (1988) Brief response to Newmann. *Social Education*, 52(6), 436–438.

Hirsch, E. D., Jr. (2001) Seeking breadth and depth in the curriculum. *Educational Leadership*, 59(2), 22–25.

Lewin, T. (2010) Expansion of A. P. tests also brings more failures [Electronic Version]. *The New York Times*. Available online at: www.nytimes.com, accessed 11 February 2010.

Long, K. (12 February 2010) Rigorous college-prep classes skyrocketing in Washington state. *The Seattle Times*. Available online at: www.seattletimes.com, accessed 15 February 2010.

Martin, T., Pierson, J., Rivale, S. R., Vye, N. L., Bransford, J. D. and 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* (Arlington, VA: International Network for Engineering Education and Research), 147–160.

Mathews, J. (28 August 2009). Will Advance Placement replace the SAT? *The Washington Post*. Available online at: http://voices.washingtonpost.com/class-struggle/2009/08/will_advanced_placement_replac.html, accessed 13 February 2010.

Mellon, K. (25 March 2010) HISD considers mandatory policy to ensure courses are rigorous enough. *The Houston Chronicle*. Available online at: www.chron.com, accessed 27 March 2010.

National Governors Association (1 June 2009) Forty-nine states and territories join Common Core Standards Initiative. [Online news release]. Available online at: http://www.nga.org/portal/site/nga/menuitem.6c9a8a9ebc6ae07eee28aca9501010a0/?vgnextoid=263a584a61c91210VgnVCM1000005e00100aRCRD, accessed 1 April 2010.

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).

Newmann, F. M. (1988a) Can depth replace coverage in the high school curriculum? *Phi Delta Kappan*, 69(5), 345–348.

Newmann, F. M. (1988b) Another view of cultural literacy: Go for depth. *Social Education*, 52(6), 432–436.

Paek, P. L., Ponte, E., Sigel, I., Braun, H. and Powers, D. (2005) *A portrait of Advanced Placement teachers' practices* (Research Report No. 2005-7) (New York: College Board).

Parker, W. C. (2003) *Teaching democracy: Unity and diversity in public life* (New York: Teachers College Press).

Parker, W. C. (2010) Listening to strangers: Classroom discussion in democratic education. *Teachers College Record*, 112(11), 2815–2832.

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.

Riley, M. (2005) *AP as the 'common curriculum'*. Available online at: http://apcentral.collegeboard.com/apc/public/features/18762.html, accessed 21 February 2010.

Sadler, P. M. and Tai, R. H. (2001) Success in introductory college physics: The role of high school preparation. *Science Education*, 85, 111–136.

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

Schneider, J. (2009) Privilege, equity, and the Advanced Placement program: Tug of war. *Journal of Curriculum Studies*, 41(6), 813–831.

Schwartz, D. L. and Bransford, J. D. (1998) A time for telling. *Cognition and Instruction*, 16 (4), 475–522.

Schwartz, M. S., Sadler, P. M., Sonnert, G. and Tai, R. H. (2008) Depth versus breadth: How content coverage in high school science courses relates to later success in college science coursework [Electronic Version]. *Science Education*, published online in Wiley InterScience (www.interscience.wiley.com). Available online at: http://www3.interscience.wiley.com/journal/121580319/abstract, accessed June 24, 2009.

Sizer, T. R. (1984) *Horace's compromise: The dilemma of the American high schools* (Boston: Houghton Mifflin).

Stevens, R., Wineburg, S., Herrenkohl, L. R. and Bell, P. (2005) Comparative understanding of school subjects: Past, present, and future. *Review of Educational Research*, 75(2), 125–157.

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

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

Wiebe, R. H. (1995) *Self rule: A cultural history of American democracy* (Chicago: The University of Chicago Press).

Wineburg, S. (1997) Beyond 'breadth and depth': Subject matter knowledge and assessment. *Theory into Practice*, 36(4), 255–261.