Running Head: Understanding the Student Experience in High Schools

Understanding the Student Experience in High School: Differences by School Value-Added Rankings

Marisa Cannata

Vanderbilt University

April 2013

This research was conducted with funding from the Institute of Education Sciences (R305C10023). The opinions expressed in this report are those of the authors and do not necessarily represent the views of the sponsor.

Abstract

Research on school reform and implementation has found that past efforts to scale up interventions in schools often results in little change in the core work of teaching and learning (Elmore, 1996). That is, there are often few changes in the work students do with teachers. Thus it is important for researchers and school reformers to know how students experience the school. Learning about the student experience is particularly important given the Center's emphasis on essential components of effective schools such as personalized learning connections, quality of instruction, and culture of learning. This paper uses data collected from shadowing high school students for a full day. The overarching research question guiding this paper is: What is the student experience in high schools and how does it vary by school effectiveness? Data come from the National Center on Scaling Up Effective Schools intensive case study research in eight high schools across two large urban districts in Florida and Texas.

Introduction

Three decades ago, American public high schools were indicted as large, impersonal bureaucracies where teachers make an implicit bargain with students to not expect too much of them in exchange for compliance (Powell, Farrar, & Cohen, 1985; Sedlack, Wheeler, Pullin, & Cusick, 1986; Sizer, 1984). Instruction was criticized as teacher-centered and curriculum varied widely for students in the same school based upon their chosen course trajectory. In the intervening years, the United States has seen numerous reforms targeted at changing high schools to increase standards, focus on college and career readiness, equalize opportunities, and create more engaging and personalized learning environments for students.

Despite these numerous reforms, national and international comparisons of student achievement indicate that underperformance in high school is a persistent problem, even as gains have been made in the elementary grads (Rampey, Dion, & Donahue, 2009). Research on school reform and implementation has found that past efforts to scale up interventions in schools often results in little change in the core work of teaching and learning (Elmore, 1996). That is, there are often few changes in the work students do with teachers.

Yet most of this research relies on data collected from teachers and other adults in schools; there is relatively little that uses data from students themselves to capture how students experience high school. Recent research has highlighted the importance of capturing the student voice and experience, particularly in high schools (Mitra, 2004; Wallace & Wildy, 2004; Wilson & Corbett, 1999). For example, the Gates Foundation Measures of Effective Teaching study found that students can distinguish between teachers based on effectiveness (Gates Foundation, 2010), suggesting that student perspectives can be valuable in understanding effective school practices. This paper explores how students experience high school and how these experiences vary by school effectiveness. Through exploring this question, the paper will draw

on data obtained shadowing students throughout their regular school day to answer such questions as: How do students spend their time in high school? With whom do students interact during the school day? How do student experiences vary by their educational track and school effectiveness?

Framework

Research on practices in effective schools usually focuses on the actions, perceptions, and beliefs of teachers and leaders in the schools. Yet schools are filled with students and it is students whose educational, social, and emotional lives are shaped by what happens in schools (Bidwell, 2006). As valuable as it is to understand schools from the teacher and leader perspective, the picture is incomplete without also viewing school from the perspective of students (Wallace & Wildly, 2004). Further, incorporating the student perspective into school reform efforts can improve youth development and achievement outcomes (Applebee, Langer, Nystrand & Gamoran, 2003; Mitra, 2004). For example, knowing the extent to which schools establish personalized learning environments for students where they are actively engaged in challenging academic work requires understanding how students experience schools.

Learning about the student experience is particularly important given the Center's emphasis on essential components of effective schools such as personalized learning connections, quality of instruction, and culture of learning. Personalized learning connections are the ways in which students in a school have a connection or sense of belonging to the school as a whole, as well as meaningful, positive connections with teachers and students in the school (McLaughlin, 1994; Lee & Smith, 1999). When schools create a sense of personalization for students, students feel connected to the school as a whole and other individuals within the school, while lack of personalization can lead to alienation (Nasir, Jones, & McLaughlin, 2011, Hallinan, 2008, Crosnoe, Johnson, & Elder, 2004). Despite the widespread recognition of the importance of personalized learning connections, little is known about the mechanisms by which students form positive and personal relationships with teachers and fellow students, highlighting the need for additional research on the opportunities students have to interact and connect with others in

the school.

Students who feel personally connected to the school do exhibit increased cognitive engagement with academic tasks and have supportive relationships with adults in school (Fredericks, et al., 2004; Walker & Greene, 2009). Just as personalized learning connections exist along a continuum, so too does student engagement that ranges along a continuum from the most teacher-directed to the most student-initiated where an engaged student might participate autonomously (Birch & Ladd, 1997; Buhs & Ladd, n.d.; Nystrand & Gamoran, 1991). The extreme end of cognitive engagement can be characterized as experiencing flow in which a person is so totally involved in an activity that they lose awareness of time and space captures the extreme end of the spectrum of behavioral engagement (Csikzentmihalyi & Csikzentmihalyi, n.d.). Student engagement is also characterized as having multiple dimensions, where cognitive or intellectual engagement relates to student investment in learning and a desire to exceed requirements and be challenged (Fredricks, Blumenfeld, & Paris, 2004). Cognitive engagement is also demonstrated by the ways in which students invest in the instructional tasks laid out by teachers and concentration on those tasks (Shernoff, Csikzentmihalyi, Schneider, & Shernoff, 2003; Yazzie-Mintz & McCormick, 2012). Cognitively engaged students exhibit self-regulation or the ability to be strategic about how they study. Behavioral or procedural engagement, on the other hand, refers to students demonstrating the specific behavioral markers of complying with basic expectations set out by the teacher (Birch & Ladd, 1997; Finn, Pannozzo, & Voelkl, 1995; Nystrand & Gamoran, 1991; Skinner & Belmont, 1993).

Students' opportunities for more cognitive and active engagement, however, depends on the types of expectations set by their teachers. For example, when teachers provide more time and structure for classroom discussion, effort and engagement are enhanced (Kelly & Turner, 2009; Nystrand, 1997). Student learning and engagement can also be enhanced through collaborative grouping strategies (Boaler & Staples, 2008). Opportunities for student success are also associated with tracking. Tracking students by ability is linked to increasing inequality in student outcomes as students in advanced tracks make

larger gains than those in non-advanced tracks (Carbanaro & Gamoran, 2005; Gamoran, 2009; Oakes, 2005). Further, it is the instructional structures used in classes in different tracks that lead to this inequality, as students in advanced tracks are more exposed to discussion-based and more engaging instructional strategies (Applebee, Langer, Nystrand, & Gamoran, 2003; Gamoran, 2004; Oakes, 2005; Watanabe, 2008).

Data and Methods

Data come from 8 high schools in two large urban districts, one in Florida and one in Texas. In each district, value-added data were used to identify two high schools that have relatively higher value-added scores in reading/language arts, math, and science and two high schools that have relatively lower value-added scores in these subjects. The schools serve primarily low-income (free and reduced price lunch eligible) and racial minority students, reflecting the populations of these urban districts. The data for this paper came from a larger study about practices that differentiate the higher and lower value-added high schools in these districts (for information on other larger study, see Cannata, Taylor Haynes, & Smith, 2013; Rutledge, Cohen-Vogel, & Osborne-Lampkin, 2012). The shadowing data were collected in late spring, after the administration of state assessments, during the third and final visit to each school.

The Florida data were collected in Spring 2011 and the Texas data in Spring 2012. The data collection methods between the two districts were similar although since Texas was a year behind Florida, some changes were made to improve the research design. The differences between districts and the reasons for the change are noted in this section. In the Florida district, 24 10^{th} grade students were shadowed for an entire school day and in the Texas district 2, 37 10^{th} grade students were shadowed. The additional students shadowed in Texas were added to improve the representativeness of the student shadowing sample. In both districts, students were split evenly by gender and, when possible, students were shadowed by a researcher of the same gender. While Texas had slightly more students per school, within each district, students were evenly distributed among the four schools.

The sampling strategy was intended to sample based on the educational track students tend to take (i.e., half the students were to be selected among honors/advanced track students and half were to be selected among regular/remedial track students). This sampling strategy was chosen because earlier visits to the schools in Florida suggested that student experiences differ by student track trajectory and the research team wanted to understand the extent of these differences. However, given the nature of student course assignments, it was not always easy to identify a single track. In Florida, there were difficulties in two of the schools (one lower value-added and one higher value-added) in identifying the track of the students. In Florida, 12 students took primarily "high" track classes (i.e., honors, Advanced Placement), 8 students took primarily regular or "low" track classes, and 4 could not be identified by track. Due to the difficulties in identifying the track of each student in Florida, one difference in the shadowing design in Texas was the decision to identify the track of specific courses in which the student was enrolled, rather than just the overall track of the student. Thus although the sampling strategy still selected students based upon track (i.e., students who took mostly advanced courses and students who did not), we have track information at the level of each course rather than the student. In Texas, 23% of observational segments were in advanced track courses, 40% were in regular or remedial track courses, and 36% were in courses with other or unknown track placements. This third type of courses was often elective courses or those for a designated student population, such as ELL students.

All students were followed by a researcher for a full school day, shadowing the student from class to class and during transition periods and breaks. The goal was to understand how students experienced the school. Starting at the beginning of the school day, the researcher logged the student's activities every 5 minutes on log with predetermined categories. The log asked for several pieces of information: the time, the period of the day (i.e., first period, second period), where the student was located, what the student was doing, with whom the student was interacting, and whether the student was on-task or off-task. The log had specified categories for the location, activity, and with whom the student was interacting, although the researcher could also write in other activities or provide more details. By the end of the day,

the researcher has a picture of what the student experienced that day in 5 minute intervals. Horng, Klasik, and Loeb (2010) used a similar shadowing log to observe principals as they engaged in their regular daily activities.

The basic shadowing log was similar for both districts, but there were five main improvements made to the log in Texas (see Table 1 for a comparison of the two logs). First, as mentioned above, there was additional information about the track for each course in Texas, while in Florida there was only data on the track/level for the student overall. Second, the Texas log included space for the course title, allowing us to determine the subject area of each course, which was not possible in Florida. Third, describing what the student was doing at each 5-minute interval was captured slightly differently in each district. In Florida, the researcher noted what the student was actually doing and then indicated whether the student was on-task or not (i.e., what the student was doing matched what the teacher expected the student to be doing). In Texas, the research first noted what the teacher expected the student to be doing, whether the student was in fact engaged in that activity, and if not, what off-task behavior the student was doing. Finally, the Texas log used a more nuanced indicator of on-task/off-task by having the researcher note whether the student was actively or passively engaged in the task rather than just on-task. Directions to researchers indicated that active student engagement included asking questions, responding to questions, volunteering information, sharing ideas, or manipulating materials. Students who are actively engaged are on task and focused on their class-related goals. Passive engagement includes behaviors such as listening but not responding to questions, not asking questions, and being involved but appearing disinterested in the assigned task. Students are not engaged if they are unresponsive, disinterested, distracted, or involved in off-task behaviors. Finally, the Texas log included more space for the researcher to make comments if they had difficulty using the predetermined codes in the log. The changes to the Texas log were made to improve the shadowing process and resultant data, while also trying to preserve the ability to make comparisons between districts.

Table 2 provides basic information on the data. A total of 1,670 five-minute segments were observed in

Florida and 2794 five-minute segments in Texas. These segments are roughly equally distributed across schools, although with more segments in Texas due to the increased number of students shadowed. Some observational segments were excluded from analysis. One student in Texas missed three class periods due to a dentist appointment. Because a dentist appointment does not help us see what typical student experiences are, these observational segments were excluded. The intention was to observe students during lunch, homeroom classes, and/or tutorial periods. However, whether the researcher was able to observe during these periods varied within and among schools, particularly during lunch when students moved frequently and the researcher needed to take a daily break. Because these parts of the day were not uniformly observed, they are excluded from the analysis. Finally, observations that occurred during the transition between classes are also excluded as this analysis is focusing on what happened during class. The analytic sample includes 1,521 five-minute observational segments in Florida and 2,436 five-minute observational segments in Texas.

The teacher expectation of the student included many different options for the researcher to choose (and an "other" category). This analysis combines some seldom-used categories. The type of activity in which teachers expected students to engage include: whole class discussion, direct instruction, pair or group work, individual work, other academic activities (includes taking a test or quiz, watching or giving a presentation, watching a film or video, or academically-oriented talk with the teacher), non-academic or transitional activities (such as socially-oriented talk with the teacher, handing out report cards, saying the pledge, school announcements, transitioning between activities), and other or unknown activity (for non-core subjects, this includes doing ROTC drills and playing sports in gym class, as well as activities that are not easily classified into the above categories. When students had no activity in which they were expected to be engaged, they are categorized as "non-academic" activity. Within-class transitions, no expected activity, and non-academic activities are considered together because notes written by the researchers indicated they had difficulty distinguishing between these activities. For example, when a student was one of the first to complete an assignment given by the teacher, it was not clear if they should

be categorized as being expected to still be doing the assignment that most classmates were still working on or something else.

Student observational logs are not frequently used in education research, which tends to rely on student surveys and classroom observations to gather information on student perspectives. While these methods have their purposes, their ability to provide reliable data on student experiences is limited. Classroom observations often focus on the teacher or the classroom as a whole rather than any particular student. Shadowing students, on the other hand, puts the focus on how an individual student experiences classroom instruction. Shadowing students is a unique tool that allows a researcher to capture the quality and quantity of student and teacher interactions in a typical school day (Wilson & Corbett, 1999). There is little research on the systematic use of student shadowing logs, although the use of Experience Sampling Method (ESM) is a similar technique and has been used to understand activities that facilitate student engagement in high schools (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003). ESM has advantages over year-end surveys and other one-point-in-time methods because they reduce the error resulting from poor recall and social desirability bias (Hektner, Schmidt, & Csikszentmihalyi, 2007). ESM is also more likely to capture activities that appear inconsequential but occur relatively frequently for brief moments (Csikszentmihalyi & Larson, 1987). One limitation of ESM is that it places the burden on participants to record the logs, potentially introducing bias and interrupting the activity. Our approach reduces the need to interrupt the participant and trained researchers can reliably code shadowed activities (Horng, Klasik, & Loeb, 2009).

Chi-square tests were used to test for differences between the higher and lower value-added schools (combining the two schools of each type in each district). There are two main analyses that explore differences between schools. First, observation segments in all courses are considered. Second, differences between tracks within schools are explored. There are three track groupings used in this analysis: advanced (e.g., AP, pre-AP, honors), regular, and other (which includes missing track information and electives). Inclusion, sheltered language, and remedial classes are considered together

with “other” track classes because there were too few observations in to consider these tracks separately. Note that students in Texas may have been observed in both advanced and non-advanced classes. For example, a student may be enrolled in a pre-AP English class but regular level mathematics class.

Results

How Do Students Spend Their Time?

Table 2 presents the results for the shadowing analyses for both districts. Students in Texas were somewhat more likely to have an expected task to be doing, although the difference is small and there were no differences between higher and lower value-added schools in either district. Students were observed to have a task they should be doing between 90-95% of their time in class.

The data provide many different insights into how students spend their class time in high school. First, the data suggest that students spend about 14-18% of their class time off-task or not engaged in the expected activity (in Florida and Texas, respectively). While being on-task 86% of the time in Florida appears to indicate a relatively high level of student engagement, the more detailed engaged data in Texas indicates that students are predominantly passively engaged. Students in Texas were actively engaged only 29% of the class time observed and were passively engaged for 54% of the observational segments. There were differences in off-task and student engagement behavior by school value-added status, although the direction varied by state and appeared to be driven by one lower value-added school in each state. In Florida, students in lower value-added schools were more often observed to be off-task while in Texas they were more often observed to be actively engaged compared to their peers in higher value-added schools.

In general, the amount of time students were expected to spend in each type of activity during class was more evenly distributed across the activity categories in Texas than in Florida. Listening to direct instruction was the most frequent activity in which students in Florida were engaged, with about 35% and 37% of observational segments in higher and lower value-added schools, respectively, in this category.

Listening to direct instruction was the second most frequent activity in Texas, which was the expected activity in about 20% of observational segments in higher value-added schools and 16% of lower value-added schools (the difference was statistically significant). In Texas, the most frequent expected activity was individual work, which students were expected to work alone on a class assignment. This was the most frequent activity in both higher and lower value-added schools in Texas, although students in lower value-added schools were slightly more often observed in individual work than students in higher value-added schools (27% and 23%, respectively).

Although the data not provide evidence on the *quality* of the activities in which the students engaged, the *structure* of these activities suggests that recent efforts to make instruction more collaborative with active participation by students are not evident in the high schools that participated in this study. Combining the observational segments in which students were expected to listen to direct instruction or work individually, students were expected to do these types of activities between 43% to 50% of their class time. On the other hand, students were only asked to participate in class activity with a partner or group between 5% to 15% of the time, although more often in Texas than in Florida. Students were expected to participate in a whole class discussion between 7% and 13% of the observed segments, again slightly more often in Texas than in Florida.

Table 4 explores the relationship between the type of activity the teacher expects the student to do and the level of student engagement. Students are most likely to be observed as on-task in Florida when they are asked to listen to direct instruction. However, the more nuanced engagement data in Texas suggests that, while students may be on-task when the teacher expects them to listen to direct instruction, students are overwhelmingly passively engaged (i.e., in 73% of the observational segments in Texas). Students in Texas are most often observed to be actively engaged when the teacher expects them to be engaged in group or pair work or a whole class discussion. However, students are also more likely to be not engaged when asked to participate in group or pair work, suggesting that group or pair work leads to either complete disengagement or active engagement. The expected activity that was associated with the highest

levels of active engagement was the "other/unclear" category. Future analyses should take a closer look at the written comments to explore what was happening during these observational segments.

There were few differences by value-added status of the schools in Florida in the type of expected activity of students. As noted above, students in higher value-added schools in Florida were more often observed to be working individually than students in lower value-added schools in Florida. There were no other statistically significant differences between higher and lower value-added schools in expected activity in Florida. Note that while there were few differences by value-added status, there were differences among the individual schools, although the patterns are masked when combining schools by value-added status.

More differences were evident in Texas between higher and lower value-added schools. Students in higher value-added schools in Texas were more often observed listening to direct instruction and being expected to engage in activities classified as "other academic", which includes tests, quizzes, watching videos, and academically-oriented talk with a teacher (such as a conversation about getting into college) than their peers in lower value-added schools. They were less likely than their peers in lower value-added schools to be expected to engage in whole class discussion and individual work.

The enhanced shadowing log in Texas made it possible to examine the student experience during class time by the subject area of the course. The analysis was repeated using only courses in core academic subjects (i.e., English, math, science, and social studies). Findings (not shown) are consistent with the patterns for all courses.

Who Do Students Interact With During Class?

Table 3 also provides data on the types of individuals with whom students were interacting during their school day. In Florida, students were most likely to be interacting with the teacher alone, followed by interacting with no one, and then interacting with other students. This is not surprising given that the most common expected activity was listening to direct instruction. There were also differences by school value-added status in all of these categories. Students in higher value-added schools were more likely to

spend time interacting with the teacher alone and with no one and less time interacting with students.

The interaction patterns were different in Texas. Students in Texas were most often observed not interacting with anyone. About 47% of the observational segments were characterized as the student interacting with no one and this did not vary by school value-added status. Students interacted with other students about 25% of the time during class, which again did not vary by value-added status. Texas students were observed to be interacting with the teacher alone in about 14% of observational segments, although students in higher value-added schools interacted more often with the teacher than students in lower value-added schools.

How Do Student Experiences Vary by Their Educational Track?

To understand the variation in student experiences during class time across tracks, this analysis compares student experiences across tracks within the same schools (see Tables 5 and 6). The evidence in both states suggests that students in the same school but in different course levels or tracks had different experiences, although there were few consistent pattern across states or by school value-added status. Looking within states, there were no consistent patterns among higher and lower value-added in the nature of the track variation in Florida. That is, while differences by track were observed in both higher and lower value-added schools in Florida, the direction and type of difference was not the same in higher and lower value-added schools. In Texas, there were three areas in which consistent track differences were observed across both lower and higher value-added schools. Thus if school effectiveness in Texas is related to compressing variability between tracks, it is probably not coming from these areas. Compared to students in advanced courses in Texas, students in regular track courses in both higher and lower value-added schools were less likely to be observed to have an expected task, less likely to be observed interacting with the teacher, and more likely to be observed to have an “other or unclear” expected activity.

There were three instances in which the track differences were similar across states. Students in regular track classes in higher value-added schools in both Texas and Florida were less likely to be observed

listening to direct instruction and interacting with the teacher than their peers in advanced courses. Students in regular track classes in lower value-added schools in both Texas and Florida were less likely to be observed engaging in group or pair work than their peers in advanced courses.

One hypothesis for school effectiveness is that it is achieved by compressing variation between tracks. If this hypothesis were true, we would expect to see track differences in lower value-added schools but not in higher value-added schools. The findings that students in regular track classes in lower value-added schools in both states were less likely to be observed engaging in group or pair work than their peers in advanced courses is consistent with this hypothesis, although there is no other evidence of compressed variation in both states. There are a couple of other instances of this compressed variation in Texas, with students in regular track classes being more often observed to be not engaged and less often observed to be expected to participate in whole class discussion when in lower value-added schools, but no track differences exist in higher value-added schools.

Conclusion

Overall, there was considerable variation between individual schools in both activities in which students were engaged and with whom they interacted during class time. While these patterns point to few consistent differences between higher and lower VA schools, they do paint a picture of how high school students spend their day. The data indicate that high school students spend most of their day in relatively passive activities, although the specific findings indicate variation by state, school effectiveness, and course track. In Texas, for example, students were observed to be actively engaged in academic work only one-quarter to one-third of their time in class. Detailed data on active or passive engagement in Florida is not available, but students in Florida were observed to be spending over a third of their time listening to direct instruction, an activity in Texas that was most closely associated with passive engagement. Whole class discussion and group or pair work was associated with the highest levels of active engagement, yet in both states students spent less time in those activities than listening to direct instruction or working individually. Further, students in Texas were most often observed not interacting

with anyone, spending almost half their class time interacting with no one. The persistence of this relatively passive engagement of students calls into question whether the reforms of recent decades aimed at making high school more engaging through more rigorous and relevant instruction have been successful. Link to other research

The data in this paper does provide some insight into instructional practices that can facilitate active student engagement. Similar to what has been found in other research on how students experience school (Shernoff et al., 2003), students are more likely to be actively engaged when they are involved in group work or a discussion than when listening to direct instruction. However, students are also more likely to be not engaged when asked to participate in group or pair work, suggesting that group or pair work leads to either complete disengagement or active engagement. Link to other engagement research

It should be noted that these data only provide evidence on the structure of instructional activities, not their quality. This is noteworthy because this paper is part of a study that is trying to identify practices that differentiate higher and lower value-added schools, yet there are few differences between the higher and lower value-added schools in Florida and the differences that are evident in Texas are sometimes contradictory to hypotheses (e.g., that students in higher value-added schools were less often observed to be actively engaged). This finding suggests that it is not just the practices themselves that lead to effectiveness, but the implementation and intention behind those practices. Other research on effective instructional practices indicate that teachers need to appropriately scaffold instructional strategies such as group work and discussions to promote student motivation and engagement (Kelly & Turner, 2009).

The findings also suggest that efforts to reduce the differences in student experiences by course level have also faced difficulty. Although it was difficult to sample students based on track due to complex course patterns—suggesting some success in overcoming past patterns of rigid tracks with little movement across tracks—there was still quite different student experiences based upon the track of the student or course in which the observation took place. There were few consistent patterns across states or by school value-added status, yet observations in advanced courses had different results than those in regular level

courses. Link to research on tracking.

Links differences by state to the design challenge – personalization in Florida rather than Texas linked to fact that students didn't seem to interact with anyone in either HVA or LVA in Texas. More individual work in Texas linked to ownership.

References

Allensworth, E., Correa, M., & Ponisciak, S. (2008). From High School to the Future: ACT Preparation--Too Much, Too Late. Consortium on Chicago School Research: Chicago, IL.

Applebee, A. N., Langer, J., Nystrand, M., & Gamoran, A. (2003). Discussion-based approaches to developing understanding: Classroom instruction and student performance in middle and high school English. American Educational Research Journal 40, 685–730.

Birch, S. H., & Ladd, G. W. (1997). The teacher-child relationship and children's early school adjustment. *Journal of School Psychology*, *35*(1), 61–79. doi:10.1016/S0022-4405(96)00029-5

Boaler, J., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *The Teachers College Record*, *110*(3), 608–645.

Buhs, E. S., & Ladd, G. W. (n.d.). Peer rejection as an antecedent of young children's school adjustment: An examination of mediating process. *Developmental Psychology*, *37*, 550–560.

Cannata, M., Taylor Haynes, K., & Smith, T. (2013). *Reaching for Rigor: Identifying Practices of Effective High Schools*. Nashville, TN: National Center on Scaling Up Effective Schools, Vanderbilt University.

Carbonaro, W. & Gamoran, A. (2002). The Production of Achievement Inequality in High School English. American Educational Research Journal, 39, 4: 801-827.

Csikszentmihalyi, M., & Larson, R. W. (1987). Validity and reliability of the experience sampling method. Journal of Nervous and Mental Disease, 175, 526-536.

Csikzentmihalyi, M., & Csikzentmihalyi (Eds.). (n.d.). The flow experience and its significance for human psychology. In *Optimal experience* (pp. 15–35).

Darling-Hammond, L., Ancess, J., Susanna, W. O. (2002). Reinventing high school: Outcomes of the coalition campus schools project. American Educational Research Journal, 39(3), 639-673.

Elmore, R. F. (1996). Getting to scale with good educational practice. Harvard Educational Review,

66(1), 1-26.

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and Inattentive-Withdrawn Behavior and Achievement among Fourth Graders. *The Elementary School Journal*, *95*(5), 421–434.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research*, *74*(1), 59–109. doi:10.3102/00346543074001059

Gates Foundation, 2010. Learning about teaching: Initial findings from the Measures of Effective Teaching project. Bill and Melinda Gates Foundation.

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). Experience sampling method: Measuring the quality of everyday life. Thousand Oaks, CA: Sage.

Horng, Klasik, D., & Loeb. (2010). Principal's Time Use and School Effectiveness. *American Journal of Education*, *116*(4), 491–523. doi:10.1086/653625

Kelly, S. & Turner, J. (2009). Rethinking the effects of classroom activity structure on the engagement of low-achieving students. Teachers College Record, 111, 7: 1665-1692.

Marzano, R.J., Pickering, D. J., & Pollock, J.E. (2001). Classroom Instruction That Works: Research-based strategies for increasing student achievement. Alexandria, VA: Association for Curriculum Supervision and Development.

Mitra, D. L. (2004). The significance of students: Can increasing "student voice" in schools lead to gains in youth development? Teachers College Record 106(4): p 651-688.

Nystrand, M., & Gamoran, A. (1991). Instructional discourse, student engagement, and literature achievement. *Research in the Teaching of English*, 261–290.

Oakes, J. (2005). Keeping track: How schools structure inequality (2nded.). New Haven, CT: Yale University Press.

Powell, A. G., Farrar, E., & Cohen, D. K. (1985). *The Shopping Mall High School: Winners and Losers in the Educational Marketplace*. Houghton Mifflin (T).

Rutledge, S. A., Cohen-Vogel, L., & Osborne-Lampkin, L. (2012). *Identifying the Characteristics of Effective High Schools: Report from Year One of the National Center on Scaling Up Effective Schools*. Nashville, TN: National Center on Scaling Up Effective Schools, Vanderbilt University.

Sedlack, M., Wheeler, C., Pullin, D., & Cusick, P. (1986). *Selling Students Short: Classroom Bargains and Academic Reform in the American High School*. Teachers College Press.

Shernoff, D. J., Csikzentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly*, *18*(2), 158–176.

Sizer, T. R. (1984). *Horace's Compromise: The Dilemma of the American High School* (None.). Mariner Books.

Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, *85*(4), 571–581. doi:http://dx.doi.org/10.1037/0022-0663.85.4.571

Walker, C., & Greene, B. (2009). The relations between student motivational beliefs and cognitive engagement in high school. Journal of Educational Research, 102 (6), 463-472.

Wallace J. & Wildy, H. (2004). Old questions for new schools: What are the students doing? Teachers College Record, 106(4), 635-650.

Watanabe, M. (2008). Tracking in the era of high-stakes state accountability reform: Case studies of classroom instruction in North Carolina. Teachers College Record, 110, 489–533.

Wenglinsky, H. (2002) The link between teacher classroom practices and student academic performance. Educational Policy Analysis Archives, 10(2).

Wilson, B. & Corbett, D. (1999). Shadowing Students. Journal of Staff Development, 20(3), 47-48.

Yazzie-Mintz, E., & McCormick, K. (2012). Finding the Humanity in the Data: Understanding, Measuring, and Strengthening Student Engagement. In *Handbook of Research on Student Engagement* (pp. 743–761). Springer Science+Business Media, LLC.

Tables

Table 1. Number of Students and Five Minute Observational Segments for Student Shadowing

Log category	Florida Log	Texas Log
Identification codes	Date, Student ID, Researcher ID, School Code, Time, Track	Date, Student ID, Researcher ID, School Code, Time, Number of advanced courses
Course Codes	Class period, Location	Class period, Location, Course title, Course track/student population
Student activity codes	The activity in which the student is engaged, an on-task/off-task designation that indicates if the student is engaged in the activity the teacher expects him/her to do	The activity in which the teacher is expecting the student to engage, academic nature of that activity, indication of whether student is not engaged/passively engaged/actively engaged in that assigned activity, type of off-task behavior (if applicable)
Student interaction codes	The individual(s) with whom the student is currently interacting	The individual(s) with whom the student is currently interacting

Table 2. Number of Students and Five Minute Observational Segments for Student Shadowing

	Florida		Texas	
	LVA	HVA	LVA	HVA
Number of students	12	12	19	18
Total observation segments	830	840	1415	1379
Excluded segments				
Dentist	0	0	0	35
Transition	43	38	74	100
Homeroom/Tutorial/Lunch	33	35	36	113
Number observation segments in analysis	754	767	1305	1131

Table 3. Results from Student Shadowing Logs

	Florida			Texas		
	LVA	HVA	Ch-sq	LVA	HVA	Ch-sq
Time with expected task	91.1%	89.6%	1.04	94.9%	93.8%	1.47
Engagement						
Not engaged/Not on task	17.8	9.9	16.1***	16.9	18.9	11.2**
On-task	82.4	90.1		n/a	n/a	
Actively engaged	n/a	n/a		31.5	25.1	
Passively engaged	n/a	n/a		51.7	56.0	
Type of expected activity						
Whole class discussion	8.4	6.5	1.86	12.5	8.9	7.95**
Direct instruction	37.3	35.1	0.794	15.6	20.3	9.48**
Group/pair work	7.0	4.7	3.76	14.7	14.9	0.0254
Individual work	12.3	16.3	4.86*	26.8	23.1	4.52*
Test/Other academic	14.2	17.6	3.30	13.0	16.5	5.68*
Transition/Non-academic/No task	12.2	12.3	0.001	9.7	11.5	2.18
Other/unclear	8.6	7.6	0.573	7.7	4.8	8.94**
With whom interacting						
Teacher	40.1	47.9	9.38**	10.3	17.7	27.5***

Student(s)	23.1	13.8	21.7***	26	22.9	3.10
Teacher and student(s)	4.6	5.4	0.397	8.9	6.8	3.60
No one	20.8	25.6	4.78*	45.7	47.5	0.798
Other/unclear	11.4	7.4	7.05**	9.1	5.1	14.3***

*difference at $p<.05$

Note: this does not include homeroom, transition, tutorials, or a student who left for dentist appt

Table 4. Percentage of Observational Segments in Each Expected Activity by Student Engagement

Type of expected activity	Florida			Texas			
	Off-task	On-task	Ch-sq	Not engaged	Passively engaged	Actively engaged	Ch-sq
Whole class discussion	8.2	91.8	3.18	9.9	55.7	34.4	15.1**
Direct instruction	4.9	95.1	56.0***	13.9	72.6	13.5	84.0***
Group/pair work	5.8	94.3	5.07*	29.4	26.9	43.6	120.7***
Individual work	13.2	86.8	0.072	20.4	58.6	21.0	22.4***
Test/Other academic	16.6	83.4	1.91	16.3	66.0	17.7	30.5***
Transition/Non-academic/No task	78.7	21.3	281.97***	31.0	37.9	31.0	16.8***
Other/unclear	21.0	79.0	2.82	1.5	13.1	85.4	220.7***

*difference at $p<.05$

Table 5. Results from Student Shadowing Logs in Florida, by Track

	LVA				HVA			
	Advanced	Regular	Other	Ch-sq	Advanced	Regular	Other	Ch-sq
N	443	259	52		342	274	151	
Time with expected task	91.7	88.8	98.1	4.98	90.9	88.9	88.7	
Engagement								
Not engaged/Not on task	19.4	15.5	11.9	2.53	12.7	10.4	3.1	9.10*
On-task	80.6	84.1	88.1		87.3	89.6	96.9	
Type of expected activity								
Whole class discussion	7.5	11.6	0.0	8.74*	8.2	6.9	2.0	6.73*
Direct instruction	33.9	35.9	73.1	30.9***	31.6	26.6	58.3	46.1***
Group/pair work	8.6	5.8	0.0	6.16*	5.6	4.0	4.0	1.03
Individual work	13.3	10.4	13.5	1.33	13.7	20.8	13.9	6.73*
Test/Other academic	14.2	15.1	9.6	1.05	21.9	20.1	3.3	26.8***
Transition/Non-academic/No task	14.9	9.7	1.9	9.71**	13.5	11.7	10.6	0.925
Other/unclear	7.7	11.6	1.9	6.35*	5.6	9.9	8.0	4.06
With whom interacting								
Teacher	34.5	42.5	75.0	32.7***	48.3	41.2	58.9	12.3**

Student(s)	27.3	19.3	5.8	15.3***	16.1	14.6	7.3	7.02*
Teacher and student(s)	6.1	3.1	0.0	6.06*	1.5	12.4	1.3	42.0***
No one	19.9	23.2	17.3	1.50	26.3	26.3	22.5	0.912
Other/unclear	12.2	12.0	1.9	4.98	7.9	5.5	9.9	3.01

*difference at $p<.05$

Note: this does not include homeroom, transition, tutorials, or a student who left for dentist appt

Table 6. Results from Student Shadowing Logs in Texas, by Track

	LVA				HVA			
	Advanced	Regular	Other	Ch-sq	Advanced	Regular	Other	Ch-sq
N	312	523	470		406	465	260	
Time with expected task	98.1	93.5	94.5	8.87*	96.8	95.7	85.8	38.1***
Engagement								
Not engaged/Not on task	14.1	20.3	15.1	13.8**	15.9	21.6	18.7	8.63
On-task								
Actively engaged	27.6	29.6	36.2		22.6	26.0	27.9	
Passively engaged	58.2	50.1	48.7		61.4	52.4	53.4	
Type of expected activity								
Whole class discussion	25.0	9.9	7.0	60.6***	10.3	8.6	7.3	1.90
Direct instruction	12.5	16.1	17.0	3.088	31.8	16.8	8.9	57.6***
Group/pair work	21.8	14.2	10.6	18.8***	16.3	16.3	10.4	5.52
Individual work	24.4	24.5	31.1	6.74*	14.8	27.5	28.1	24.6
Test/Other academic	9.0	15.7	12.8	7.80*	19.5	13.8	16.5	5.12
Transition/Non-academic/No task	6.1	10.9	10.6	5.99*	7.4	9.7	21.2	32.1***
Other/unclear	1.3	8.8	10.9	25.4***	0.0	7.3	7.7	31.8***
With whom interacting								

Teacher	15.4	7.8	9.8	12.2**	23.9	16.1	10.8	20.1***
Student(s)	29.2	25.6	24.3	2.41	20.4	24.1	24.6	2.19
Teacher and student(s)	12.2	7.3	8.5	5.96	9.9	4.7	5.8	9.53**
No one	43.0	45.3	47.9	1.88	42.4	50.5	50.0	6.67*
Other/unclear	0.32	14.0	9.6	44.0***	3.5	4.5	8.9	10.1003 **

*difference at p<.05

Note: this does not include homeroom, transition, tutorials, or a student who left for dentist appt