

# How Accurate Are Self-Reported High School Courses, Course Grades, and Grade Point Average?

Edgar Sanchez  
Richard Buddin

WP-2015-03  
April 2015

## ACT Working Paper Series



ACT working papers document preliminary research. The papers are intended to promote discussion and feedback before formal publication. The research does not necessarily reflect the views of ACT.



# **How Accurate Are Self-Reported High School Courses, Course Grades, and Grade Point Average?**

**Edgar Sanchez and Richard Buddin**

**April 2015**

## **Abstract**

Education researchers and college admissions officials often rely on self-reported coursework and grade information provided by high school students. This study examines the accuracy of self-reported course and grade information from the ACT registration with high school transcript data from a Midwestern state. The evidence shows that students' self-reported information is an accurate representation of their high school experience. About 94% of students accurately reported their coursework. The correlation between self-reported and transcript course grades was 0.66 with 96% of self-reported grades within a single letter grade of their transcript grade. High school grade point average computed from self-reported course grades was highly correlated with transcript grade point average ( $r=0.83$ ). The accuracy of coursework and grades differed little by gender, race/ethnicity, and low-income status. The results indicate that self-reported coursework and grade are reasonably valid measures for education researchers and for preliminary screening of students by college admissions officials.

Keywords: self-reported student data, high school grades, high school coursework

## Introduction

Education researchers often study the relationship between high school coursework, grades, and admissions test scores on college enrollment and persistence, grade point average (GPA), and on-time degree completion (Adelman, 2006; Schmitt et al. 2009, Radunzel & Nobel, 2012; Sanchez, 2013; Mattern, Patterson, & Wyatt, 2013). Researchers seldom, however, have access to high school transcript data and typically rely on student self-reported course and grade information. One of the premises of this post-secondary research is that the self-reported information adequately reflects student background and high school preparation.

Student self-reported information is also used by college recruiters and postsecondary administrators to identify and encourage potential applicants with suitable preparation to apply to their institution. Ultimately, colleges obtain high school transcripts for applicants, but self-reported information provides an efficient way to identify students adequately prepared for college. While transcripts are the official record, course titles vary considerably from district to district, so postsecondary administrators may have difficulty assessing and categorizing high school coursework. Having a short list of discrete core courses may help postsecondary administrators efficiently identify qualified applicants.

Given the widespread use of self-reported coursework and grade information, this study focuses on whether self-reported information is an accurate indication of a student's high school experience. We obtained high school transcript data from a Midwest state and compared this data with course and grade information that students self-reported when registering for the ACT test. The registration form asks for course enrollment and grade information for 30 high school courses in English, math, science, social studies, foreign languages, and the arts.

## Literature Review

Valiga (1987) examined the accuracy of self-reported course-taking and grades for about 1,000 students from 26 Kentucky and Illinois high schools that registered to take the ACT in April or June of 1983. In this study the correlations between self-reported and transcript grades ranged from 0.75 for Computer Science to 0.92 for English 11 with a median correlation of 0.86. Overall, about 80% of students accurately reported their grades.

Sawyer, Liang, and Houston (1988) assessed the accuracy of self-reported coursework and grades using data from about 1,000 ACT-tested students in the 1985-1986 academic year from 53 high schools. Across 30 courses, they found that the rate of accurately reporting coursework taken ranged from 65% for American Government to 97% for English 9 and 10. The median accuracy rate across subjects was 87%. They also found that about 97% of students reported course grades that were within one letter grade of their transcript grade (e.g., A and B). Moreover, the correlations between self-reported and transcript reported grades ranged from 0.53 for Drama to 0.89 for Trigonometry with a median correlation of 0.80. The authors showed that these findings held across race/ethnicity, gender, and ACT Composite (ACTC) score. The exceptions to this trend included greater accuracy of self-reported grades for females and higher ACTC scoring students.

Schiel and Noble (1991) also investigated this topic by looking at almost 2,000 students from 55 high schools. This study used student data taken from the 1990 October and November administrations of the P-ACT+ in a southern state.<sup>1</sup> The percentage of students accurately reporting having taken a course varied between 83% for General Mathematics and 100% for courses such as other mathematics, astronomy, German, and other languages with a median

---

<sup>1</sup> The P-ACT+ was the early form of the ACT PLAN test.

percentage across courses of 96%. These percentages tended to be consistent across gender, minority, and P-ACT+ score ranges.<sup>2</sup> The correlations between transcript and self-reported grades ranged from 0.58 for Art to 0.85 for Spanish with a median correlation of 0.64. The percentage of students that accurately reported their high school grades ranged from 51% for General Mathematics to 83% for music. The percentage of students who reported their HSGPA within one letter grade of their transcript grade ranged from 89% for General Mathematics to 98% for Chemistry. While the correlations for males and females were similar across courses, they found that the self-reported grades of African American students were less accurate than those of white students across the P-ACT+ scale, with accuracy increasing as achievement increased.

Kuncel, Credé, and Thomas (2005) conducted a meta-analysis of 37 studies that examined self-reported high school grade point average (HSGPA), class rank, and test scores. They found that the correlations between self-reported and transcript HSGPA in English, math, science, and social studies ranged from 0.80 to 0.85. They also found that the correlation between overall self-reported and actual HSGPA was 0.82. This study found that about 82% of students tended to accurately report their overall HSGPA while about 12% and 4% over- and under-reported their overall HSGPA respectively. While the accuracy of self-reported grades was similar for males and females, white students were found to be more accurate than nonwhite students. In addition, they found that lower achieving students self-reported HSGPA less accurately than students with a higher HSGPA.

---

<sup>2</sup> This study only compared the difference between African American and white students due low numbers of other race/ethnicities.

Shaw and Mattern (2009) compared self-reported and actual HSGPA and SAT test score for over 40,000 students at 32 colleges. They found that the average difference between students' self-reported HSGPA and school-reported HSGPA was small (-0.04).<sup>3</sup> They also found a high correlation between the two HSGPA measures ( $r = 0.73$ ). Interestingly, they found a slightly lower correlation between the two measures for lower SAT achieving students than for students that scored higher on the SAT. This study found that 52% of students accurately reported their HSGPA while the two measures of HSGPA for an additional 37% of students were within one grade difference (e.g., A and A+). These percentages were fairly consistent across gender, race/ethnicity, parental education, parental income, and SAT score. They found higher correlations between self-reported and HSGPA for white versus African American students, for high- versus low-income students, and for students with high- versus low-SAT test scores. They further found that students were more accurate in their self-reported grades at the upper end of the HSGPA scale.

Allen (2013) investigated the accuracy of self-reported HSGPA for almost 2,000 students in 11 school districts from the graduating classes of 2010 and 2011. That study found that 83% of students accurately reported their HSGPA within 0.50 units, and 58% accurately reported their HSGPA within 0.25 units. Furthermore, they found that the correlation between the HSGPA reported by the student's high school and the students self-reported HSGPA was 0.84.<sup>4</sup> They also found that overall, students tended to over report their HSGPA by 0.07 points. Additionally this study found that lower achieving students tended to over report their HSGPA to a greater extent than higher achieving students.

---

<sup>3</sup> Most studies have relied on letter grades without pluses and minuses (e.g., A, B, C, D, F), but the data in the Shaw and Mattern (2009) study included plus and minus grades (e.g., A-, B+).

<sup>4</sup> These two HSGPA calculations may not, however, have included the same courses or weighting of each course.

These studies consistently show high correlations between self-reported and transcript data, but some of the studies are over 20 years old and included small numbers of students. More recent studies have focused on high school GPA (HSGPA) and not examined the accuracy of self-reported coursework. This study updates the prior research on course taking and grade accuracy, while providing further information on HSGPA accuracy. The study will provide insights into whether student self-reported information is sufficiently reliable for postsecondary research as well as for college administrators as a means to identify potential applicants. We do so by focusing on three questions:

1. Do students accurately self-report high school course and grade information?
2. Does accuracy differ by student gender, race/ethnicity, or family income level?
3. Has student accuracy changed over time?

### **Data**

This study relies on high school transcript and ACT test record data for over 15,000 students from a Midwestern state that were in the 2009 high school graduation class. The dataset was limited to students who took the ACT as juniors in 2007-2008 and who had transcript data for at least four academic years from 2005-2006 through 2008-2009. About 70 percent of high school students in this state take the ACT and 72 percent of these students enroll in college, so the sample is somewhat broader than college bound students and broader still than students that are accepted to, or enroll in, college. We compared transcript data with student-reported information on high school course and grades in the ACT registration file.

Student gender, race/ethnicity, and family income provided at the time of registration for the ACT were used in this study. Information on student gender and race/ethnicity had virtually

no missing information. Family income was not reported by about 15 percent of students. Low-income status (family income less than \$36,000) was imputed for students with missing family income using multiple chain methods (Little and Rubin, 2002). The imputation was based on gender, race/ethnicity, coursework taken, course grades, GPA, and ACT subject test scores.<sup>5</sup>

## Methods

We focus on two separate analyses in this report. First we examine the accuracy of students' self-reported coursework taking. This is followed by an examination of the accuracy of self-reported grades on high school coursework. The transcript data was provided by the Department of Education for a Midwestern state. This transcript data provided details for a total of 288 high school courses. By examining the course names with state sources, these 288 courses were mapped to the 30 courses listed in the Course/Grade Information Section (CGIS) of the ACT registration file. While not all transcript courses mapped to a CGIS course, there were many possible high school courses that were mapped to a given CGIS course (see Table 1). For example, English 9, English 9 Honors, and a state approved English 9 course were mapped to the English 9 CGIS course. Another example would be the nine possible courses that were mapped for the Trigonometry CGIS course. These courses included Trigonometry, International Baccalaureate Trigonometry, and Pre-Calculus/Trigonometry courses as well as the honors courses.

---

<sup>5</sup> Missing values were imputed for all variables using multiple chain techniques, but the analysis was conducted on the subset of observations with complete data on the dependent variables in the statistical models. Several studies have argued that the dependent variables should be included as part of the imputation procedure but then excluded from the analysis of particular dependent variables (Little and Rubin, 2002; Allison, 2001; Von Hippel, 2007; and White et al., 2011).



**Table 1: Number of Transcript grades that mapped to each CGIS course.**

CGIS Course	Number of transcript courses
<b>English</b>	
English 9	3
English 10	3
English 11	3
English 12	3
Other English	15
<b>Math</b>	
Algebra 1	8
Geometry	9
Algebra 2	7
Trigonometry	8
Calculus	5
Computer Math	10
Other Math	9
<b>Science</b>	
General Science	11
Biology	5
Chemistry	5
Physics	9
<b>Social Studies</b>	
Geography	3
Government	8
US History	6
World History	4
Economics	3
Psychology	3
Other Social Science	3
<b>Art</b>	
Art	17
Drama	4
Music	17
<b>Foreign Languages</b>	
French	9
German	7
Spanish	14
Other Foreign Language	33

It is not possible to accurately know which transcript course a student is reporting on the CGIS.<sup>6</sup> As such, if a student took any of the possible transcript courses that mapped to a CGIS course the student was considered to have taken the course. When considering grades, a similar situation arises. In the case of grades for the group of transcript courses, the highest grade

<sup>6</sup> It is possible that students would receive a grade for a given course per semester. For courses that span multiple semesters it is not possible to know if a student is reporting their more recent semester grades or not.

reported from any acceptable course is retained as the grade for that course. Given that we do not know which course grade is being reported in the CGIS, we give the student the benefit of their highest possible grade. While the transcript course grade was recorded on a continuous scale, the CGIS uses a discrete scale to represent grades. Both the transcript and CGIS grades were converted to their respective letter grades (i.e. A, B, C, D, or F). These letter grades were used for calculating agreement rates and correlations. These letter grades were then converted to a 4 point scale to calculate an estimated GPA.<sup>7</sup> The calculated HSGPA was based on 23 possible courses from English, math, science, and social studies.

Most students take the ACT as a junior or senior in high school, so most students have not completed their high school coursework. The survey asked students for courses they “have taken or am taking” and those they “have not taken but will take” before graduation. We measure the accuracy of these self-reports by comparing courses taken or planned on the survey with courses recorded on their high school transcript.<sup>8,9</sup>

## **Results**

### *Accuracy of self-reported coursework*

Student self-reported coursework was very consistent with courses that were recorded on their official transcripts. Table 2 shows that self-reported course information was generally

---

<sup>7</sup> Students who had less than 10 grades reported on their transcript were not included in the HSGPA calculation.

<sup>8</sup> Some students self-reported English 11 and English 12, but their transcript showed a state approved advanced course that was a substitute for these courses. In these cases, the analysis counted these courses as being consistent with the self-reported information.

<sup>9</sup> Students do not have grades for courses that they plan to take, so the number of students used in the grading analysis is typically smaller than that of the course taking analysis (this is especially true for coursework typically taken in the senior year). For example, an individual student may plan to take or is currently taking Beginning Calculus as a senior, but they will not have a grade to self-report until the course is completed.

accurate more than 90 percent of the time. The median accuracy across all subjects was 94 percent.

Student reports were especially accurate for fundamental coursework (courses taken by over 90 percent of students), but they were less accurate for more advanced coursework that was taken by a smaller percentage of students. For example, Physics was taken by 22 percent of students, and the percent of students correctly self-reporting taking the course was 82 percent. Similarly, only 5 percent of students took Geology, and self-reported course taking was accurate 53 percent of the time.

Several reasons may explain the reduced accuracy for advanced coursework. First, students may be embellishing their accomplishments. Students are much more likely to over report taking a course (the student lists the course in CGIS, but it is not on their transcript) than to under report taking a course (the student does not list the course in CGIS, but it is on their transcript). Second, the course names for advanced classes are sometimes more complicated than for fundamental coursework, so students may have difficulty mapping their coursework into CGIS categories. For example, students may not understand whether the CGIS category Other Advance Math should include courses like Mathematical Applications and Algorithms or Linear Systems and Statistics. Finally, some coursework may overlap between classes, so students are uncertain how the work should be classified in CGIS. For example, a Trigonometry course may include some calculus, so the student might self-report both Trigonometry and Beginning Calculus in the CGIS. Unfortunately, we have no basis to sort between the alternative conjectures for why self-reporting is less accurate for advanced coursework.

**Table 2: Accuracy of self-reported high school course taking**

Variable	Percentage Correct	Over Report	Under Report	Percentage Taking Course
<b>English</b>				
English 9	95	4	1	95
English 10	98	2	0	98
English 11	99	1	0	99
English 12	93	6	1	93
<b>Math</b>				
Algebra I	96	3	1	96
Geometry	98	2	0	98
Algebra II	95	5	0	91
Trigonometry	80	12	8	40
Beginning Calculus	82	18	0	11
Other Advanced Math	46	52	2	17
<b>Science</b>				
General Science	95	4	1	95
Biology	99	1	0	99
Chemistry	93	6	1	76
Physics	81	18	1	22
<b>Social Studies</b>				
US History	99	1	0	99
World History	97	2	1	98
Other History	70	23	7	17
American Government	93	3	4	97
Economics	82	17	1	4
Geography	53	46	1	5
Psychology	75	23	2	16
<b>Foreign Language</b>				
Spanish	96	2	2	81
French	96	3	1	11
German	98	2	0	2
Other Language	95	3	2	2
<b>Arts</b>				
Art	84	7	9	71
Music	85	12	3	42
Drama	86	11	3	23
<b>Median Values</b>	<b>94</b>	<b>5</b>	<b>1</b>	<b>74</b>

Note: The entries are based on a comparison of individual course taken or planned with high school transcript data.

Self-reported course taking was similarly accurate across various student groups. We estimated the accuracy of course taking in each subject as a function of student gender, race/ethnicity, and whether they came from a low-income family (see Appendix Table A-1). Accuracy varied little by gender. In more than half the classes, females were not significantly more or less likely to accurately report their course taking than males. In other cases, females

were sometimes more accurate than males and sometimes less accurate, but the differences were seldom more than 2.5 percentage points. African American students were less accurate than White non-Hispanics in most subjects, but the gap was only about 2 percentage points. Hispanic students also misreported more often than White non-Hispanics, but the gap was significant in fewer subjects and only by about 1 percentage point. Asian students were also less accurate than White non-Hispanics, but the gaps were statistically insignificant in most classes. Holding gender and race/ethnicity constant, students from low-income families were sometimes more and sometimes less accurate than students from more wealthy families, but the typical magnitude of the difference was generally less than 2 percentage points.

These small differences in the accuracy of self-report across groups are probably of little practical importance given the high overall accuracy in most subjects. For example, a 95 percent accuracy versus a 93 percent accuracy may be statistically significant, but this small inaccuracy is unlikely to have a substantive effect for postsecondary researchers or college administrators.

#### *Accuracy of self-reported grades*

The correlations of self-reported and transcript grades for 28 courses are reported in Table 3. Correlations for the courses ranged from 0.36 for Music to 0.78 for Chemistry. We can also see that the correlations for coursework that students would typically have taken by the time they take the ACT in their junior year tend to be higher than for the more advanced courses in English, math, and science. Moreover, while mathematics courses tended to have the highest correlations between self-reported and transcript grades, art courses had the lowest correlations.

**Table 3: Correlation of self-reported and transcript grades.**

Course	Correlation	N
<b>English</b>		
English 9	0.67	13,419
English 10	0.71	13,370
English 11	0.70	9,620
English 12	0.60	4,489
<b>Math</b>		
Algebra 1	0.69	13,025
Geometry	0.76	13,013
Algebra 2	0.77	11,505
Trigonometry	0.76	3,255
Calculus	0.63	768
Other Advanced Math	0.58	1,283
<b>Science</b>		
General Science	0.58	12,748
Biology	0.72	13,370
Chemistry	0.78	9,274
Physics	0.66	1,880
<b>Social Studies</b>		
US History	0.63	12,718
World History	0.67	12,284
Other History	0.60	966
American Government	0.66	11,555
Economics	0.60	257
Geography	0.40	421
Psychology	0.63	1,190
<b>Foreign Language</b>		
Spanish	0.70	10,783
French	0.67	1,462
German	0.69	290
Other Language	0.73	88
<b>Art</b>		
Art	0.51	7,559
Music	0.36	5,192
Drama	0.52	2,366

Note: All correlations are significant ( $p < 0.0001$ ).

In Table 4 we can see the percentage of students who self-reported the same letter grade as their transcript grade, reported a grade within one letter grade of their transcript grade, underreported their transcript grade, and over reported their transcript grade. The percentages of students reported the same letter grade as their transcript tended to be similar across courses. The exact grade agreement rate ranged from 58% for English 12 to 93% for music with the median agreement rate being 67%. The exact/adjacent agreement rates were very high for all courses

examined. This agreement rate ranged from 91% for economics to 100% for other foreign languages. In English, math, science, and social studies more students tended to underreport their grades than over report. For these courses, the percentage of students who underreported their transcript grades ranged from 15% for Trigonometry to 30% for economics with a median under reporting rate of 20%. Few students over reported grades for Arts coursework, but a fair number (about 1/5) of students over reported their foreign language grades.

**Table 4: Transcript and self-reported coursework grade agreement rates**

Course	Exact Agreement	Exact/Adjacent Agreement	Percent Under Reporting	Percent Over Reporting
<b>English</b>				
English 9	66	96	18	16
English 10	68	97	18	14
English 11	66	97	17	17
English 12	58	95	25	17
<b>Math</b>				
Algebra 1	64	95	19	17
Geometry	68	97	18	14
Algebra 2	68	96	18	14
Trigonometry	73	97	15	11
Calculus	69	97	21	11
Other Advanced Math	59	94	23	17
<b>Science</b>				
General Science	63	94	21	16
Biology	67	97	19	14
Chemistry	70	97	17	13
Physics	67	96	19	14
<b>Social Studies</b>				
US History	65	96	21	14
World History	65	96	21	14
Other History	70	95	22	8
American Government	68	96	20	12
Economics	65	91	30	6
Geography	60	93	24	15
Psychology	65	94	28	7
<b>Foreign Language</b>				
Spanish	69	96	9	21
French	67	97	10	23
German	69	97	9	22
Other Language	75	100	14	11
<b>Arts</b>				
Art	76	96	16	8
Music	93	99	3	4
Drama	75	96	18	7

### *Accuracy of Self-Reported HSGPA*

The average HSGPA calculated from transcript grades (3.08) and self-reported grades (3.16) were similar, and both measures had a standard deviation of 0.59. There was a high correlation between HSGPA calculated from self-reported grades and transcript grades ( $r = 0.84$ ). Figure 1 shows the regression of transcript grades on self-reported grades.<sup>10</sup> This figure also shows the 95% prediction interval for self-reported grades. We can see that there tends to be under reporting at lower self-reported HSGPA levels and slight over reporting at higher levels. Self-reports are only slightly greater than the transcript average at the 25<sup>th</sup> percentile of self-reported grades (2.75 versus 2.72). At the 75<sup>th</sup> percentile of self-reported grades, over reporting is larger (3.65 for self-reported versus 3.51 for transcript grades).

---

<sup>10</sup> Fewer than 3 percent of self-reported grades were less than 2.0, so the graph focused on students that reported at least a 2.0 grade point.



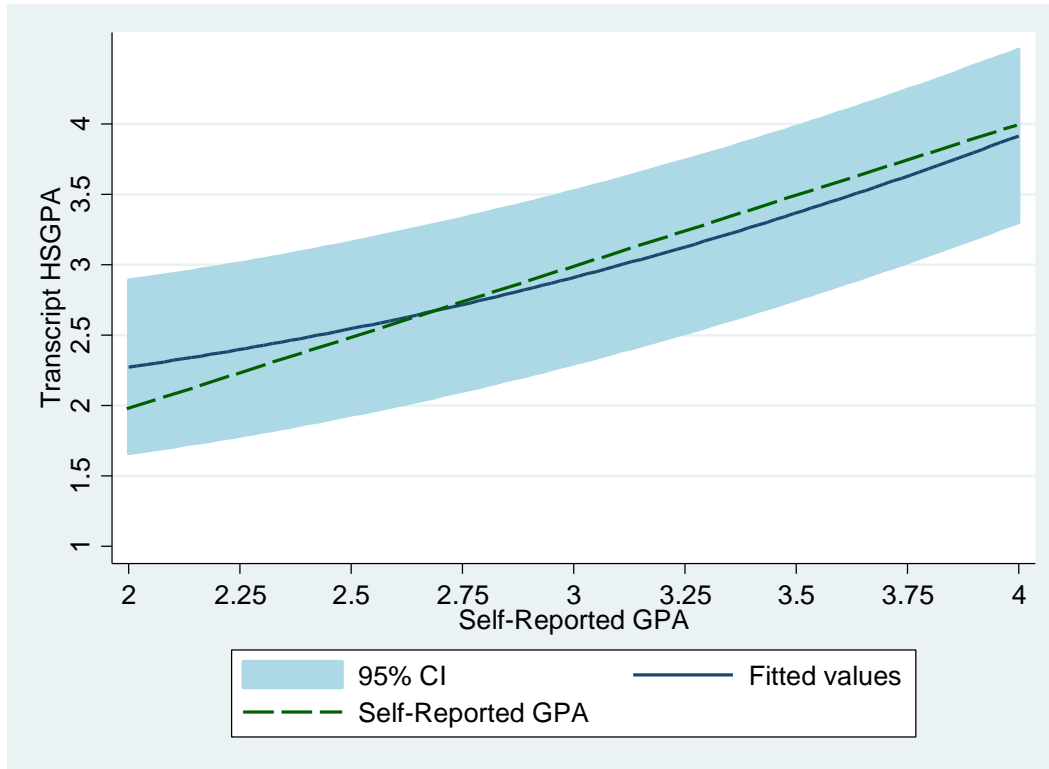


Figure 1: Regression of transcript grades with self-reported grades and 95% prediction limit.

We also examined whether the accuracy of self-reported HSGPA differed by gender, race/ethnicity, and low-income status. We estimated the difference between transcript HSGPA and self-reported HSGPA as a function of indicator variables for gender, race/ethnicity, and low-income groups (see Appendix Table A-2). The results were as follows:

$$\text{Transcript HSGPA} - \text{Self-Reported GPA} = -0.086 + 0.057 * \text{Female} - 0.013 * \text{African American} + 0.007 * \text{Hispanic} + 0.051 * \text{Asian} - 0.001 * \text{Low Income}.$$
<sup>11</sup>

Only the coefficients on the female and Asian variables were statistically significant at the 0.05 level. The average difference was -0.09 (3.08 for transcript versus 3.17 for self-reported grades), but the gap for females was about 0.06 larger than for males—females overstate their grades

<sup>11</sup> R<sup>2</sup>=0.008; N=14,204

slightly more than males. White students slightly overstate their grades (by about 0.06 grade points), but the self-reported HSGPA of Asian students match almost perfectly with their transcript grades (self-reported grades are 0.004 points higher than transcript grades).

On average, the differences in over/under reporting transcript GPA for these groups is less than one tenth of a grade point. These group-level differences in the accuracy of HSGPA are probably of little practical importance to educational researchers or postsecondary administrators.

## **Discussion**

*Do students accurately self-report high school course and grade information?*

Overall, students' self-reported coursework and grades are accurate representations of actual coursework taken and grades. We found that in over 30 courses in the areas of English, math, science, social studies, foreign language, and the arts, the median accuracy across all subject was 94%. We also found greater accuracy in reporting for more fundamental coursework such as English 11, Algebra II, Chemistry, and World History.

While the correlation between self-reported course grades and transcript grades varied considerably across courses, the median correlation for all courses was relatively high (0.66). This suggests fairly good accuracy across courses. Among courses, the median percentage of students who self-reported the exact letter grade as their transcripts was also relatively high. If we included self-reporting within a letter grade that median percentage was 96%. We also saw that students tended to under report their grades more than they over reported their grades. When a HSGPA was calculated based on all self-reported grades, the correlation with transcript HSGPA was high,  $r = 0.83$ . This suggests that a HSGPA based on self-reported course-level

grades is a good analog for a HSGPA based on transcript-reported course-level grades. We also found that students on the lower end of the HSGPA scale, below about 2.5, tend to under report their true HSGPA while students on the upper end of the HSGPA scale tend to slightly over report their true HSGPA.

*Does accuracy differ by student gender, race/ethnicity, or family income level?*

We found that the accuracy of self-reported coursework tended to be fairly similar for the student subgroups examined. There was little to no difference in the rates of accuracy of self-reported coursework between males and females. While Hispanic students tended to slightly under report their coursework taking for a handful of the courses examined relative to white non-Hispanic students, African American students tended to slightly under report their coursework taking for most courses.

Larger differences in accuracy were observed in the rates of self-reported grades exactly matching transcript grades for student subgroups. While female students were more likely to report their exact course grade, minority (African American & Hispanic) and low-income students were less likely than white non-Hispanic and other income students to report their exact course grades, respectively.

*Has student accuracy changed over time?*

Compared to prior studies that have examined the accuracy of self-reported coursework taking, the current study found very similar results. While there are certainly variations between courses examined, this study confirms the findings of prior studies that there is a high degree of accuracy in the self-reported coursework taking. The percentage of students that self-reported their exact grade in the current study was similar to that found by previous studies. That said, the

correlations between self-reported and transcript grades found in the current study, however, tended to be lower than those found for similar prior research when course level grades were examined. When overall HSGPA was examined, the correlation was consistent with prior research with the exception of Shaw and Mattern (2009) which found a lower correlation than other studies.<sup>12</sup>

### **Conclusion**

Based on our findings, coursework taken and overall HSGPA provided by students can be a good measure of academic preparation and achievement for researchers and postsecondary administrators. We find that using self-reported coursework taken is a reliable way to estimate course specific student level experiences. This may be particularly useful for advising purposes. A measure of caution, however, is appropriate for the use of course specific student-level grades. Given the correlations found in the current and prior studies, self-reported grades earned should be followed up with transcript grades for high-stakes student-level decisions. For overall HSGPA for an individual, however, we have a stronger basis for trusting self-reported values. Regardless of these student level cautions, self-reported coursework taking and grades earned appear to be reasonably valid measures for aggregate level decisions and research.

---

<sup>12</sup> The lower correlation in Shaw and Mattern (2009) may reflect the finer HSGPA gradations (plus and minus grades) than in the other studies.

## Appendix

**Table A-1. Regression Estimates for Accuracy of Course Taking by Subject**

Dependent Variable	Female	African American	Hispanic	Asian	Low-Income	Constant	Adj. R-Square	Students
English 9	-0.006	-0.021*	-0.008	-0.009	0.010*	0.957*	0.00129	14,165
English 10	0.002	-0.031*	-0.023*	-0.025*	0.004	0.983*	0.00657	14,164
English 11	0.006*	-0.005*	-0.019*	0.006	-0.002	0.992*	0.00331	14,168
English 12	-0.013*	0.028*	0.002	0.019	0.016*	0.925*	0.00360	14,118
Algebra I	0.003	-0.019*	-0.015	-0.017	0.019*	0.959*	0.00262	14,126
Geometry	0.005	-0.021*	-0.010	-0.025*	0.006*	0.981*	0.00341	14,136
Algebra II	0.015*	-0.056*	-0.007	-0.001	-0.016*	0.955*	0.0125	14,077
Trigonometry	-0.001	-0.030*	-0.064*	-0.007	0.014	0.806*	0.00118	13,381
Beginning Calculus	0.018*	-0.021*	-0.017	-0.035	0.024*	0.809*	0.00133	13,259
Other Advanced Math	-0.037*	0.027*	0.038	-0.021	0.036*	0.457*	0.00314	13,567
General Science	-0.005	-0.037*	-0.029*	-0.036*	0.022*	0.957*	0.00539	14,133
Biology	-0.000	-0.017*	-0.007	-0.003	0.006*	0.988*	0.00275	14,163
Chemistry	0.025*	-0.048*	0.011	-0.012	-0.009	0.933*	0.00835	13,857
Physics	0.028*	-0.051*	0.013	0.010	0.009	0.799*	0.00312	13,400
United States History	-0.001	0.010*	0.004	0.013	-0.007*	0.987*	0.00114	14,153
World History	-0.001	-0.011*	0.004	-0.013	0.010*	0.973*	0.00102	14,133
Other History	0.047*	0.037*	-0.006	0.020	0.031*	0.654*	0.00526	13,208
American Government	-0.003	-0.039*	-0.005	-0.010	0.011*	0.940*	0.00297	14,022
Economics	0.020*	-0.018	0.004	-0.034	0.008	0.808*	0.000733	13,175
Geography	0.002	-0.012	0.065*	0.069*	-0.009	0.530*	0.000846	13,491
Psychology	-0.022*	-0.026*	-0.016	0.008	0.024*	0.759*	0.00107	13,328
Spanish	0.014*	-0.017*	-0.024*	-0.025*	-0.017*	0.968*	0.00592	13,971
French	-0.001	-0.043*	-0.030*	-0.015	-0.007	0.978*	0.00953	13,265
German	0.005*	-0.019*	-0.006	-0.016	-0.010*	0.981*	0.00396	13,155
Other Language	0.002	-0.016*	-0.079*	-0.054*	-0.007	0.963*	0.00728	13,082
Art	-0.017*	-0.020*	0.036*	0.011	0.011	0.851*	0.00119	13,828
Music	-0.003	-0.023*	0.010	0.036	-0.002	0.852*	0.000564	13,490
Drama	0.000	-0.054*	-0.002	0.045*	-0.007	0.869*	0.00381	13,300
Median Value	0.001	-0.020	-0.007	-0.010	0.007			

Note: Entries are regression coefficients for a linear probability model. Asterisk indicates statistical significance at the 0.05 level. Omitted reference categories are male, white non-Hispanic, and not from a low-income family. “Median Value” refers to the median coefficient of column for all listed classes.

Source: Matched High School Transcripts and ACT Test Records.

**Table A-2. Regression Estimates for Exact Accuracy of Grades by Subject**

Dependent Variable	Female	African American	Hispanic	Asian	Low-Income	Constant	Adj. R-Square	Students
English 9	0.056*	-0.092*	-0.051*	0.056	-0.060*	0.675*	0.0148	12,585
English 10	0.042*	-0.096*	-0.040	-0.010	-0.047*	0.699*	0.0120	12,591
English 11	0.030*	-0.050*	0.012	-0.041	-0.042*	0.677*	0.00501	9,060
English 12	0.041*	-0.065*	-0.059	0.031	-0.035*	0.592*	0.00607	4,229
Algebra I	0.035*	-0.089*	-0.042	0.044	-0.051*	0.658*	0.0107	12,233
Geometry	0.034*	-0.081*	0.017	0.051	-0.049*	0.698*	0.00982	12,209
Algebra II	0.039*	-0.057*	-0.036	0.065*	-0.058*	0.687*	0.00860	108,19
Trigonometry	0.059*	-0.050	0.042	-0.034	-0.060*	0.720*	0.00827	3,049
Beginning Calculus	0.115*	-0.014	-0.106	0.044	-0.077	0.654*	0.0145	704
Other Advanced Math	0.021	-0.101*	-0.259*	-0.083	-0.054	0.632*	0.0177	1,192
General Science	0.044*	-0.088*	-0.044	0.043	-0.090*	0.658*	0.0181	11,938
Biology	0.042*	-0.072*	-0.016	0.073*	-0.066*	0.690*	0.0121	12,511
Chemistry	0.055*	-0.053*	-0.045	0.078*	-0.054*	0.694*	0.00925	8,691
Physics	0.029	-0.120*	-0.097	-0.008	-0.052*	0.697*	0.0121	1,749
United States History	0.044*	-0.097*	-0.039	0.086*	-0.040*	0.665*	0.0117	11,838
World History	0.032*	-0.130*	-0.034	0.070*	-0.055*	0.683*	0.0186	11,523
Other History	0.040	-0.020	-0.057	-0.012	-0.066	0.692*	0.00141	876
American Government	0.051*	-0.104*	-0.033	0.075*	-0.052*	0.690*	0.0151	10,830
Economics	0.063	0.022	-0.046	-0.211	0.012	0.608*	-0.00950	247
Geography	0.077	-0.202*	0.121	0.126	-0.083	0.608*	0.0245	392
Psychology	0.030	-0.062	-0.091	0.145	-0.022	0.645*	0.00358	1,122
Spanish	0.072*	-0.088*	0.052*	0.052	-0.047*	0.683*	0.0144	10,056
French	0.073*	-0.047	-0.084*	0.033	-0.026	0.657*	0.00848	1,359
German	0.079	0.080	0.103	0.125	-0.080	0.677*	-0.00276	265
Art	0.121*	-0.158*	-0.054*	0.090*	-0.044*	0.740*	0.0451	7,048
Music	0.032*	-0.092*	-0.052*	0.019	-0.035*	0.944*	0.0320	4,800
Drama	0.054*	-0.136*	0.024	-0.083	-0.027	0.758*	0.0198	2,202
Median Value	0.044	-0.088	-0.044	0.044	-0.052			

Note: Entries are regression coefficients for a linear probability model. Asterisk indicates statistical significance at the 0.05 level. Omitted reference categories are male, white non-Hispanic, and not from a low-income family. “Median Value” refers to the median coefficient of column for all listed classes.

Source: Matched High School Transcripts and ACT Test Records.

## References

- Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college*. Washington, DC: U.S. Department of Education.
- Allen, J. (2013). *Accuracy of Student-Reported High School GPA*, (ACT Information Brief 2013-12). Iowa City, IA: ACT.
- Allison, P.D. (2001). *Missing Data*. Thousand Oaks, CA: Sage.
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2005). *The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature*. Review of educational research, 75(1), 63-82.
- Little, R. J. A., and Rubin, D. B. (2002). *Statistical analysis with missing data*. 2nd ed. Hoboken, NJ: Wiley.
- Mattern, K., Patterson, B., & Wyatt, J. (2013). *How Useful Are Traditional Admission Measures in Predicting Graduation Within Four Years?* (College Board Report 2013-1). New York: The College Board.
- Radunzel, J. & Nobel, J. (2012). *Predicting Long-Term College Success through Degree Completion Using ACT Composite Score, ACT Benchmarks, and High School Grade Point Averages*, (ACT Research Report 2012-5). Iowa City, IA: ACT.
- Sanchez, E. I., *Differential Effects of Using ACT College Readiness Assessment Scores and High School GPA to Predict First-Year College GPA among Racial/Ethnic, Gender, and Income Groups*, (ACT Research Report 2013-4), Iowa City, IA: ACT.
- Sawyer, R., Laingm J., Houston, M. (1988). *Accuracy of self-reported high school courses and grades of college-bound students*. (ACT Research Report 88-1). Iowa City, IA: ACT.
- Schiel, J., & Noble, J. (1991). *Accuracy of Self-Reported Course Work and Grade Information of High School Sophomores*. (ACT Research Report 91-6). Iowa City, IA: ACT.
- Schmitt, N., Keeney, J, Oswald, F.L., Pleskac, T.J., Billington, A.Q., Sinha, R., & Zorzie, M. (2009). *Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students*. *Journal of Applied Psychology*, 94(6), 1479-1497.
- Shaw, E. J., & Mattern, K. D. (2009). *Examining the accuracy of self-reported high school grade point average*. (College Board Report No. 2009-5). New York: The College Board.
- Valiga, M. J. (1987). *The accuracy of self-reported high school course and grade information*. (ACT Research Report 88-1). Iowa City, IA: ACT.

Von Hippel, P. (2007). Regression with missing Y's: An improved strategy for analyzing multiply imputed data. *Sociological Methodology* 37(1), 83–117.

White, I., Royston, P., and Wood, A. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine* 30, 377–399.