

*National Longitudinal Study of
Adolescent Health*

*Wave III
Education Data
Academic Courses Component
Physical Education*

Chandra Muller, Jennifer Pearson, Catherine Riegle-Crumb,
Jennifer Harris Requejo, Kenneth A. Frank, Kathryn S. Schiller,
R. Kelly Raley, Amy G. Langenkamp, Sarah Crissey,
Anna Strassmann Mueller, Rebecca Callahan,
Lindsey Wilkinson, and Samuel Field



Carolina Population Center
University of North Carolina at Chapel Hill

January 2007

This research was funded by a grant from the National Institute of Child Health and Human Development under grant R01 HD40428-02 to the Population Research Center, University of Texas at Austin; Chandra Muller (PI) and the National Science Foundation grant number REC-0126167, Chandra Muller (PI). Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (addhealth@unc.edu).

WAVE III EDUCATION DATA
ACADEMIC COURSES COMPONENT
PHYSICAL EDUCATION

**Not all variables discussed in this codebook are in the accompanying data file.
Some will be released at a later date.**

I. INTRODUCTION

The official AHAA website (<http://www.prc.utexas.edu/ahaa/>) provides full documentation of the theoretical underpinnings of the Academic Courses component. This website also lists all publications, and other current research-related activities based on the Academic Courses data. Full documentation of the Academic Courses indicators included in the first release is also available in the Wave III education data codebooks on the Add Health website (<http://www.cpc.unc.edu/projects/addhealth>).

The following introductory paragraphs present a general overview of the Academic Courses indicators, and detail the organizational structure of the Academic Courses users' guide.

A. *Summary of Academic Courses Indicators*

The Academic Courses component of AHAA contains academic status and/or performance indicators for math, science, foreign language, English, history/social sciences, physical education (PE), and a combined category (overall) which includes all coursework taken including non-core and non-academic classes. Up to six types of Academic Course indicators were constructed for each of these subjects, and for the overall category. These include course sequence variables, course type indicators, course grade measures, course failures, semesters attempted, and credits earned variables. An overview of each indicator type is presented immediately below.

Course sequence variables measure the trajectories or strands of courses taken by students in highly differentiated and sequentially ordered subjects. Because math, science, and foreign language courses are typically organized into hierarchical, linear sequences – meaning successive courses are recognized as being more advanced and generally requiring more prerequisites – variables that capture students' location within these subjects' course hierarchies in each year and by the end of high school were created. 'B' versions of the math and science course sequence variables were also constructed where students were placed at a given level in the math and science course structures only if they received credit for courses taken. These two types of course sequence variables allow analysts the option of exploring whether students take certain levels of math or science courses, and/or if they successfully complete these courses.

Although English and history/social science courses arguably become more complex and rigorous over time (i.e., a 4th year English 'survey' course will generally

contain more challenging material than a standard 1st year ‘survey’ course), enrollment in later English and history/social science coursework is not necessarily contingent upon prior mastery of a specific set of concepts or acquisition of a particular knowledge base. The greater variety and incomparability of English and history/social science courses (because they tend not to be differentiated by topic) offered across and within AHAA schools also complicates attempts to organize them into a coherent hierarchical trajectory. Given these issues, no course sequence variables were constructed for English and history/social sciences. Course sequence variables were also not created for PE.

The second category of Academic Course indicators are course type variables which facilitate identification of the degree of difficulty and/or type of coursework taken by students in certain subjects, shedding insight into the grouping of these students according to ability. Specifically, analysts can explore these variables to determine if students took honors, Advanced Placement (AP), International Baccalaureate (IB), or remedial courses in English and history/social science in each year of high school course-taking. The English and history/social science course type variables are significantly detailed – representing particularly important measures of students’ academic position/positional advantage in these two core academic subjects. The course type indicators developed for PE, while not informative of ability level, enable identification of the specific kinds of PE courses students enrolled in during each year of high school.

Six course type variables were also developed for foreign language. Analysts can use foreign language course type variables to identify which students ever enrolled in AP/IB level foreign language during their high school careers, which students took multiple foreign languages, which students took English as a Second Language (ESL) classes, which specific types of ESL classes students enrolled in (either language-based or content area), and to discern which specific foreign languages were collapsed into the category ‘other’ foreign language.

The third set of constructed Academic Course measures includes grade point average (GPA) variables that capture students’ academic performance in the key curricular subjects (math, science, foreign language, English, history/social science, and PE) as well as across all subjects including non-core and non-academic courses (overall). Measures of students’ subject-specific and overall GPAs were produced for each year of high school course-taking and cumulatively across all years of high school.

The fourth set of Academic Course indicators are course failure or failure index variables. These were constructed to facilitate examination of the extreme low end of academic achievement. Separate failure index measures were produced for math, science, foreign language, English, and history/social science. Overall measures of the proportion of courses that students failed in each year of course-taking and cumulatively across all years of high school are also available. Course failure measures were not produced for PE.

The final two sets of constructed Academic Courses indicators consist of semesters attempted and credits earned variables which convey information about the

intensity of students' exposure to course content. Specifically, these variables inform the analyst of the number of semester-length courses students attempted in each of the core academic subjects per year of high school, and the amount of credit, if any, students received for them. Overall semesters attempted and credits earned variables were also produced to enable analytical assessment of the total amount of coursework students took and received credit for in each year and cumulatively. This information is critical to understanding Add Health students' academic experiences because learning opportunities and the benefits of education are largely determined by the *quantity* of course work taken. Importantly, these variables are not restricted to graded courses as in the case of the GPA and failure index variables, but capture information about all types of high school coursework taken by students including pass/fail and non-graded courses. Thus, they enable a comprehensive evaluation of students' high school course-taking patterns and academic achievement in the key curricular subjects and overall. It is important to note that only credits earned indicators were produced for PE, although all PE coursework attempted during high school is included in the year-specific and cumulative overall semesters attempted indicators.

All constructed Academic Course variables are named to indicate the students' course-taking year to which they correspond (i.e., years **0-6**, year **0** referring to high school courses recorded on students' transcripts taken prior to actual high school enrollment, years **1-4** relating to the standard four years of high school course-taking, and years **5-6** incorporating high school courses taken by those students who remained in high school beyond the typical 4 years). Similarly, the names assigned to cumulative and highest-level attained Academic Course measures (ending with a **C** or **H**, respectively) signify that they capture summarized information about students' high school course-taking experiences in the key curricular subjects. This standardized year (and cumulative) specification enables cross and within cohort comparisons. (See the analytic advice webpage on the AHAA website for more information about the significance of the standardized year specification particularly as it facilitates analyses using both AHAA and Add Health data). Complete information about the conventions used to name Academic Course variables is provided at the end of this document.

B. Special Considerations Concerning Academic Courses Indicators

1. Year 0

It is important to note that no year **0** variables were constructed for history/social science because fewer than 50 student transcripts list history/social science coursework taken prior to actual high school enrollment. Similarly, no year **0** indicators were produced for PE.

Since only successfully completed coursework taken before entrance into high school appears on student transcripts (transcript-based year **0** data is limited to coursework students earned credit for), values for year **0** semesters attempted variables would equal those assigned for year **0** credits earned variables. Therefore, in an effort to reduce the production of redundant information and because available year **0** information

is not necessarily reflective of all year 0 high-school level coursework students actually enrolled in, year 0 semesters attempted variables were not constructed.

2. Year 5-6 Variables

Year 5-6 variables were also not constructed for PE course-type indicators because of low student enrollment rates in those years.

C. Organization of Document

The documentation of the constructed Academic Course indicators is composed of three interrelated sections. The first contains extensive descriptions of these measures, including discussions of how they were constructed, and how they should be used and interpreted for analytical purposes. The last sections outline the standard conventions employed for naming Academic Courses indicators, and the missing codes applied to them. An inventory listing the data files constructed for the Academic Courses component is also presented. Note that all students with course-level information are included in each of these data files. (The student-level disposition file, edustdent, in the primary component of AHAA provides information about which students have course-level data.)

II. CONSTRUCTED ACADEMIC COURSES INDICATORS

A. Section Overview

This section presents detailed information about the six major types of indicators included in the Academic Courses domain of AHAA: course sequence variables, course type indicators, course grade measures, course failure (failure index) measures, semesters attempted, and credit earned variables. The descriptions of all these measures are divided into two complementary parts, general and subject-specific considerations. Displayed under general issues are an explanation of how the variables were constructed, and a discussion of each variable's analytical significance. Topics covered in subject-specific considerations are those unique to each key curricular subject (math, science, foreign language, English, history/social science, PE, and overall) which have critical analytical implications. This includes cautionary statements informing analysts of pertinent issues they must consider when using and interpreting particular subject-specific Academic Courses variables.

B. Description of Indicators

1. Course Sequence Indicators

Course sequence measures apply only to math, science, and foreign language courses, and are not discussed in this document.

2. Course Type Indicators

i. General Issues

In general, course type indicators are measures which capture the type and/or degree of difficulty of coursework taken by students. These indicators essentially function as the primary analytical means of identifying how students were differentiated in terms of ability (AP/IB, Honors, remedial) in the non-sequentially ordered subjects of English and history/social science. Course type measures enabling identification of the kinds of PE courses students took were also produced. In addition, variables measuring foreign language course type were constructed to supplement the information captured by the foreign language course sequence measures. Because course sequence variables provide substantial information about the ways students are ranked within the hierarchically organized subjects of math and science by the end of high school and in each year of course-taking, math and science course type variables were not generated.

Course type indicators for the subjects of English and history/social science are multi-tiered and provide information about the type and level of students' English and history/social science coursework taken in each year of high school. Dichotomous course type indicators were produced for three English categories – a general category encompassing all English courses ('any'), and two sub-classifications, 'survey' and 'other' English – which is represented in the variable name by variations of the fourth character (**A**, **S**, or **O**). English course type indicators also enable a comprehensive analysis of how Add Health students were grouped according to ability by further specifying at what level the course was taken (noted by variations in the 5th-6th characters of the variable names). For example, analysts interested in discerning which students took honors or AP 'survey' English courses in the second year of high school can examine the variables **EAESHN2** (indicating whether an honors-level 'survey' English course was taken in year **2** of high school) and **EAESAP2** (indicating whether an AP-level 'survey' course was taken in year **2** of high school).

In contrast to the English variables (which merged category and level identification into the same sets of type indicators), four separate course *category* indicators were developed to facilitate identification of the particular category of history/social science coursework students took in each year of high school. These category variable names feature **CAT** as the 5th-7th characters, but are differentiated through a fourth character variation (**A**, **B**, **C**, or **D**). The values assigned to these indicators correspond to a hierarchical taxonomy of history/social science course categories, with categories ranked according to frequency of student enrollment. A detailed explanation of the way the diverse range of history/social sciences courses recorded on student transcripts was ordered into this hierarchical taxonomy, a description of the taxonomy itself, and how analysts should interpret the **A-D** course category and type indicators is provided in the history/social sciences subject-specific section. Once the **A-D** designations are established for a particular student, there are sets of corresponding (through the fourth character denotation) binary course type indicators that identify the levels of these courses. These course level sub-classifications (denoted by variations in the 5th-6th characters of the variable names) facilitate an examination of how

students who took history/social science courses were stratified within these courses on the basis of ability-level.

Students who never enrolled in or who did not take English and/or history/social science courses in a given year were assigned ‘missing values’ (described later in more detail) for appropriate English and/or history/social science course type indicators. Importantly, students who took English or history/social science courses in a given year in one particular sub-category (i.e. ‘survey’ English or category **A** in history/social science) were assigned missing values for the other sub-categories in that subject in that year (i.e. ‘other’ English or categories C-D in history/social science). For example, students whose only English course was an honors ‘other’ English course during the third year of high school course-taking were assigned values of *1* for **EAEOHN3** (dichotomous variable indicating whether an honors-level ‘other’ English course was taken in the 3rd year) and **EAEAHN3** (indicating whether ‘any’ honors-level English course was taken in the 3rd year), values of *0* for **EAEOAP3** (indicating whether an AP/IB-level ‘other’ English course was taken in the 3rd year) and **EAEORM3** (indicating whether a remedial level ‘other’ English course was taken in the 3rd year), and missing values for **EAESHN3**, **EAESAP3**, and **EAESRM3** (‘survey’ type indicators for honors, AP/IB, and remedial-level ‘survey’ English courses, respectively). The missing value assigned depends upon whether or not students *ever* enrolled in ‘survey’ English courses during high school (*9993* if they never enrolled, and *9994* if they did take ‘survey’ coursework at some point in their high school careers). See the section on missing codes for more information about the application of missing values to multi-tiered English and history/social science variables.

Analysts can assume that students receiving values of *0* across all English and/or history/social sciences course-type variables measuring course level (AP/IB, Honors, remedial) in a given year were enrolled in regular-level coursework. (Note that these variables are not mutually exclusive, and students who took AP/IB, Honors, or remedial level courses in a particular year may also have taken regular-level English courses.)

The course-type indicators constructed for physical education are dichotomous measures enabling identification of the category (general physical education, competitive sports, marching band, or dance/pep squad/cheerleading/drill team) of physical education coursework students took in each year (**1-4**) of actual high school enrollment.

The foreign language course type indicator, **EAFAAP**, informs analysts if students *ever* took an AP/IB level foreign language course, enabling analysts to determine whether or not students assigned to level 5 for foreign language course sequence variables were enrolled in an AP/IB level course. Five additional foreign language course type measures, **EAFOSP**, **EAFAML**, **EAFESLA**, **EAFESLL**, and **EAFESLS** convey descriptive information about foreign language coursework taken by AHAA participants. The first of these measures presents a breakdown of the languages included in the broad category ‘other’ foreign language. The second is a dichotomous indicator enabling analysts to identify which students took multiple foreign languages. The third provides information about which students took ESL coursework while

enrolled in high school. The fourth and fifth measures indicate if the AHAA participant took language-based ESL coursework and/or content area Sheltered or SDAIE (Specially Designed Academic Instruction in English) coursework respectively.

ii. Issues Specific to Physical Education

Course Category Indicators

EAPG1 – EAPG4 (General Physical Education)

EAPS1 – EAPS4 (Competitive Sports)

EAPB1 – EAPB4 (Marching Band)

EAPDP1 – EAPDP4 (Dance/Pep Squad/Cheerleading/Drill Team)

These indicators are dichotomous measures indicating the category or ‘type’ of PE coursework students took in each year of high school. Most students participated in more than one PE course during high school, and some took more than one ‘type’ of PE coursework in a single year. For example, students enrolled in ‘competitive sports’ (e.g. basketball, gymnastics, track, etc.) often also took a general physical education course in the same academic year. The PE course-type indicators enable analysts to determine whether, when, and in what type of PE course(s) students enrolled in during high school (this applies to years **1-4** only due to low enrollment rates in years **0** and **5-6**).

All PE course type indicators are named to indicate a particular category or ‘type’ of PE coursework. The character **G** is used to designate the category ‘general physical education’ and includes the following kinds of classes: health-related activities, physical and health education, physical education, physical education leadership training, life-saving, leisure and recreational activities, and conditioning and athletics. The letter **S** denotes ‘competitive sports’ and includes: individual sports, team sports, gymnastics, track and field, and aquatics. The character **B** refers to ‘Marching Band’. The character **DP** refers to ‘dance, pep squad, cheerleading, or drill team’ and encompasses modern dance, dance, performing dance, ballet and jazz, ethnic dance, square dance, aerobic dance classes and student involvement in pep squad, cheerleading, or drill team activities.

There are five sets of PE course type indicators – each representative of one of the 5 categories of PE coursework described above. Students were assigned a value of *1* or *0* for each of these indicators depending upon whether or not they enrolled in the corresponding ‘type’ of PE coursework in a given year of high school. For example, a student who took dance and no other PE course during the first year of high school was assigned a value of *1* for **EAPDP1** and a value of *0* for **EAPG1**, **EAPS1**, and **EAPB1**. Similarly, a student enrolled in both competitive sports and marching band in the second year of high school was assigned the value *1* for **EAPS2** and **EAPB2**, and the value *0* for **EAPG2** and **EAPDP2**.

Students who did not take a PE course in a given year, but did take PE coursework in other years were assigned the missing value 9994 for all five sets/categories of corresponding course-type variables. (As detailed in the examples presented above, students who took a PE course in a given year, but not in a particular category of PE coursework are assigned a value of 0 for the corresponding course type indicator.) Students who *never* enrolled in a PE course during high school were assigned the missing value 9993 for all PE course-type variables. See the section on missing codes for more information about specific missing values.

3. Course Grades Indicators

i. General issues

These variables capture students' school performance for each year of their high school course-taking, as well as cumulatively across all years of high school. Grade point average (GPA) indicators were created separately for each of the key curricular subjects (math, science, foreign language, English, history/social sciences, and PE). GPA indicators were also created to measure students' performance in physical education, and in courses taken across all subjects (overall category).

The majority of students in AHAA took courses on a semester basis, such that schools recorded two separate entries for a year-long course on the transcript, each designated with a grade. The GPA variables are calculated as the average grade across semester-length courses in a given year (for the yearly indicators), or across all years of students' course-taking (for the cumulative indicators). Less than one percent of all courses taken by the entire sample of AHAA students occurred on a trimester basis. For the purposes of the construction of academic indicators, trimesters are considered equivalent to semesters. Students who took courses designated as year-long (and with only one grade recorded) are treated as having received the same grade for two semester-length courses. Fs are coded as 0, Ds are coded as 1, Cs are coded as 2, Bs are coded as 3, and As are coded as 4. Course in which students received a 'P' for pass, an 'NG' for not-graded, a 'W' for withdrew, a "WF" for withdrew failing, a 'WP' for withdrew passing, or an 'I' for incomplete, were not included in the calculation of GPA. Students who were not assigned a grade of A-F for a course that they took in a given year would have a missing value of 9995 (no grade received for a subject-specific course taken in a given year) on the corresponding GPA variable (for math, science, foreign language, English, history/social science courses, PE, or overall courses).

All GPA variables are named to indicate the students' course-taking year (**0-6**) to which they correspond (i.e., year **0** referring to high school courses recorded on student transcripts but completed prior to actual enrollment in high school, years **1-4** relating to the standard four years of high school course-taking, and years **5-6** incorporating high school courses taken by those students who remained in high school beyond the typical 4 years). For example, **EAEAGPA2** is the student's English GPA for the second year of high school course-taking. Cumulative measures are also specified to represent grade point average for all years of course-taking (designated with a **C** as the last character of the variable, such as **EAEAGPAC**). Extending the English example, **EAEAGPAC** is

the students' cumulative English grade point average. Lastly, overall GPA includes all graded courses that appear on the students' transcripts, including non-core and non-academic courses (**EAOGPA0-EAOGPA6, EAOGPAC**).

Additionally, analysts should note that the cumulative indicators represent the average across all years for which the student was taking courses. If a student has only two years of course-taking data, for example, his or her value on **EAOGPAC** (overall cumulative GPA) would be calculated based on only two years of data, in contrast to the typical student with four years of course-taking data.

Students not taking a course in a specified subject – math, science, foreign language, English, history/social sciences, or PE – in a given year are assigned a missing value for the corresponding GPA variable. The missing value assigned will be (1) *9993* if a student never enrolled in a specified subject throughout the duration of their high school career, (2) *9994* if a student did not take courses in a specified subject in a given year but did take courses in the subject in other years, or (3) *9995* if a student did not take a *graded* course in a specified subject in that year. Refer to the missing codes section for further clarification of these issues.

The course grades measures provide analysts with yearly indicators of students' academic performance in the core academic subjects of math, science, foreign language, English, and history/social sciences as well as in physical education and across all subjects taken. In contrast to self-reported data, these are official indicators of performance as recorded on the students' high school transcripts. As such, they provide analysts with parallel measures of academic performance for students from all the various cohorts included in Add Health/AHAA. For example, by using **EAOGPA1**, the analyst has a base measure of academic performance at the beginning of high school for all students, regardless of what calendar or school year they began high school. (Note that in the example mentioned, a careful analyst might choose to restrict the analysis only to students' whose transcript-indicated grade level in **ELYEAR1** corresponded to ninth grade by using **ELYININE**. See the Linking component for an in-depth discussion of these school-year and grade level indicators).

Finally, analysts interested in examining students' grades in conjunction with survey data from Add Health are advised to consider issues of temporal order of their variables. (See the analytical advice page on the AHAA website.)

ii. Considerations Specific to Physical Education

EAPAGPA1 – EAPAGPA6 and EAPAGPAC

These variables indicate the grade point average for all PE coursework taken in each year (**1-6**) and across all years of high school (**C**).

The PE GPA measures are not disaggregated by type or category of PE coursework. They represent averages across all PE courses taken in a given year and cumulatively.

4. Course Failures – Failure Index Measures

i. General Issues

While GPA captures the range of student performance, the failure index variables convey information about the extreme end of low academic performance. These variables measure the proportion of semester-length courses failed (in each year **1-6**), calculated as the number of semester-length courses failed divided by the number of semester-length courses attempted (in each year or cumulatively). Failure index variables are restricted to years **1-6** because only successfully completed high school courses taken prior to enrollment in high school are recorded on high-school transcripts. Therefore, by definition, failure indexes cannot be computed for year **0**.

As with grade point average indicators, separate failure index measures for math, science, foreign language, English, and history/social science are presented alongside an overall failure index which encompasses failures across all subjects including non-core or non-academic courses (course failure/failure index measures were not produced for PE). Failures are defined as they are for the grade point average variables (refer to the description of the grade point average measures above), relying only on the grade received, and not on whether the student's transcript indicate that he/she received credit for a course. (However, in the vast majority of cases, students who received a grade of 'F' in a given course did not receive any credit.) Specifically, only courses for which students' received standard grades (A, B, C, D, or F) are included in this measure.

A value of *0* on a failure index variable indicates no failures, while a value of *1* indicates that the student failed all courses in a given year or cumulatively (in math, science, foreign language, English, history/social science, or across all subjects). Therefore, students whose values fall between *0* and *1* failed some proportion of the courses they attempted for a grade. Although the failure index variables are continuous, the analyst should take care when using them in this format given that the majority of students have a value of *0* on each of the variables. Instead, analysts might choose to create a dichotomous indicator (any failure vs. none), or choose certain threshold values based upon the particular analysis or research question being addressed.

Additionally, unless a student fails all of his courses in a given subject or across subjects, the student will have a corresponding grade point average measure that is greater than *0*. For example, if a student took two semesters of math in **ELYEAR2** (see the discussion of school-year variables in the Linking component for information about this variable), and failed the first semester but received a C the second semester, he/she would have a value of .5 for **EAMFIX2** (the ratio of math courses student failed in year 2) and a value of *1* for **EAMGPA2** (grade point average for math courses in year 2).

Students not taking a course in a specified subject (math, science, foreign language, English, or history/social sciences) in a given year are assigned a missing value for the corresponding failure index variable. The missing value assigned will be 9993 if a student never enrolled in a specified subject throughout the duration of high school or 9994 if a student did not take a course in a specified subject in a given year but did take

courses in the subject in other years. Refer to the missing codes section for information about specific missing values.

ii. Considerations Specific to Physical Education

Course failure measures were not constructed for PE.

5. Semesters Attempted Variables

i. General Issues

Semesters attempted variables provide analysts with important measures of students' exposure to core course material during high school. Specifically, these variables document the total number of semester units of each key curricular subject (math, science, foreign language, English, and history/social sciences) taken by Add Health students in each year (**1-6**) (i.e, years **1-4** relating to the standard four years of high school course-taking, and years **5-6** incorporating all high school courses taken by those students who remained in high school beyond the typical 4 years) and cumulatively across all years of high school. (Semesters attempted variables were not constructed for PE.) Over 85% of the courses taken by AHAA students were of semester-length. However, some students took year-long courses which were viewed as consisting of two semester-length courses. Students who took these classes were, therefore, assigned a value of 2 for corresponding semesters attempted variables. For consistency purposes (i.e., to maintain the designation of each semester as having a value of 1), the very small number of trimester courses (less than 1% of all courses) were treated as semesters. Although the majority of students took between 0-6 semesters of each core academic subject per year of high school (and were, thus, assigned values of 0-6 for corresponding semesters attempted variables), a small number of students took 7 or more semesters of a single subject in a given year. The value 12 was selected as a logical cut-off point because it indicates a student took 6 year-long courses – the equivalent of the average number of courses taken by high school students in a single year of course-taking – in one subject in a given academic year. Therefore, the very small percentage of students (less than .5%) whose transcripts record more than 12 semesters were taken in one subject in a given year were all assigned the value 12 for relevant semesters attempted variables. It is important to note that overall semesters attempted variable values were not truncated because they reflect the amount of coursework taken by students in all subjects including non-core or non-academic courses in a given year of high school.

Cumulative measures for each subject represent the sum of the values recorded for truncated year-specific semesters attempted variables. Therefore, the highest assignable number for cumulative semesters attempted variables is 72 (if a student took 12 or more semesters of a given subject over 6 years of high school course-taking). Cumulative overall semesters attempted variables, however, indicate the total, non-truncated sum of coursework students enrolled in during high school.

Values assigned for semesters attempted variables were calculated independent of the credits students earned for coursework taken. For example, students who enrolled in

2 semesters of math in the third year of high school have a value of 2 for **EAMSA3** whether or not they received credit for these courses (i.e., students who passed as well as students who failed one or both courses were both assigned a value of 2 for **EAMSA3**).

Unlike GPA and failure index measures, which are limited to graded coursework, semesters attempted measures are inclusive of all high school courses (i.e., graded, pass/fail, non-credit), making them comprehensive indicators of the quantity of coursework taken by students while in high school. Referring to the example described above, this means that students who took 2 semesters of math in the third year of high school course-taking were assigned values of 2 for **EAMSA3** regardless of the type of math course taken (i.e., graded, pass/fail, or non-credit). (Thus, the missing value 9995, which denotes that student did not receive grade in course, is not relevant to semesters attempted variables.)

The missing value 9994 which indicates no subject-specific course was taken in a given year is inapplicable to semesters attempted variables. Instead, students who did not take subject-specific courses in a given year but whose transcripts indicate enrollment in the subject in other years were assigned a value of 0 for corresponding semesters attempted variables. For example, a student who did not take English in year 2 but took English in year 1 was assigned a value of 0 for **EAEASA2** (quantity of English courses student took in year 2 regardless of whether credit was earned). Only students who *never* enrolled in a particular subject throughout the duration of high school are assigned the missing value 9993 for appropriate semesters attempted variables.

Because only coursework for which students earned credit prior to actual high school enrollment are recorded on student transcripts, available year 0 information reflects the amount of credits students were awarded in specified subjects and not necessarily all high-school level coursework students attempted before starting high school. Therefore, year 0 semesters attempted variables were not constructed. (Year 0 credits earned variables were produced and will be discussed later.)

ii. Subject-Specific Considerations

None for Physical Education.

6. Credits Earned Indicators

i. General Issues

The credits earned indicators refer to the number of Carnegie units earned in each core curricular subject (math, science, foreign language, English, history/social sciences, and PE) in each year (0-6) (i.e., year 0 referring to high school courses recorded on student transcripts but taken prior to actual high school enrollment, years 1-4 referring to the standard 4 years of high school course-taking, and years 5-6 encompassing all high school courses taken by those students who remained in high school beyond the typical 4 years) and across all years of high school. A Carnegie unit is a standardized measure equating one unit to the completion of a course that meets one period per day for one

year. Therefore, a value of .5 is generally awarded for each successfully completed semester course. For example, if students took 2 semester-length science classes in the third year of course-taking and failed one, they would have a value of .5 for **EASCR3** (amount of science credits earned during year 3).

Students were assigned values ranging from 0-6 for credits earned indicators for all key academic subjects (math, science, foreign language, English, and history/social science) except PE. Six credits was selected as a logical cut-off point because it represents the average total number of credits high school students typically earn in a given year of course-taking (encompassing coursework taken in all core academic subjects and electives). The small percentage of students earning 6 or more credits in a single subject in a specified year of course-taking were all assigned the value of 6 for corresponding credits earned indicators.

In the case of PE, students were assigned values ranging from 0 to 3 for credits earned indicators. Three credits was selected as the cut-off point for PE because (1) this number represents 6 semester-length PE courses taken in a single academic year, and (2) a very small number of students received more than 3 Carnegie units of PE credit in a given year. These students were assigned the value of 3 for all corresponding credits earned indicators.

It is important to note that the values assigned to students' overall credit measures, both year-specific and cumulative, were not truncated because these measures incorporate credits earned in all subjects including non-core and non-academic courses.

Students who took subject-specific courses in a given year but did not receive credit for these courses have a value of 0 for corresponding credits earned measures (i.e., students who did not earn credit for any science classes taken during year three received a value of 0 for **EASCR3**).

Because only credits earned while students were actually enrolled in high school are considered for the cumulative measures, high school credits attained during year 0 are not included in their calculation. It is important to note that subject-specific cumulative measures reflect the sum of truncated year-specific credits earned variables. Therefore, the highest assignable number for cumulative credits earned variables in all subjects except PE is 36 (indicating a student earned 6 credits in one subject every year for 6 years of high school course-taking). (The highest assignable number for PE cumulative credits earned variables is 18.) (As already stated, the values assignable for overall cumulative measures were not truncated.)

Credits earned measures are not restricted to graded coursework but encompass all courses taken during high school (i.e., graded, pass/fail, credit/non-credit). This feature makes credit earned measures quality indicators of students' success in meeting curricular requirements in each of the core academic subjects during high school.

Because a value of 0 for credits earned variables indicates students either failed courses attempted or enrolled in non-credit courses in a given year, students who did not take subject-specific courses in a particular year, or never took courses in a specified subject during high school, were assigned missing values for corresponding credits earned variables. The missing value 9995 which indicates students did not receive a grade in a specified course, is not applicable to credits earned variables because students can receive credit for non-graded courses. See the section on missing codes for detailed information about specific missing values.

ii. Issues Specific to Physical Education

EAPACR1 – EAPACR6, and EAPACRC

These indicators refer to the number of Carnegie units students earned in PE coursework in each year (1-6) and across all years of high school. Similar to the PE GPA variables, the PE credits earned variables are not disaggregated by type or category of PE course (General Physical Education, Competitive Sports, Marching Band, and Dance/Pep Squad/Cheerleading/Drill Team). Instead, these measures reflect the amount of credit students earned in *any* PE coursework taken per year and cumulatively during high school. For example, students that received credit for a semester course of general physical education and for enrollment in marching band in year 3 were assigned the value 1 for **EAPACR3**. Likewise, students who received credit for two semesters of general physical education in year 3 (and who took no other PE coursework) were assigned the value 1 for **EAPACR3**.

Assignment of a value of 0 for these variables indicates a student either failed PE coursework attempted, or, as in most cases, enrolled in non-credit PE courses in a given year. Students who did not take any PE course in a specified year, but took PE coursework in other years were assigned the missing value 9994 for corresponding PE credits earned variables. Students who never enrolled in PE classes while in high school were assigned the missing value 9993 for all PE credit earned indicators. The missing value 9995 (student did not receive a grade in a specified course) is not applicable to physical education credits earned variables because students can receive credit for non-graded courses.

Students were assigned values ranging from 0-3 for year-specific credits earned measures. The above “General Issues” section within the “Credits Earned Indicators” section (II.B.6.i.) explains why the value 3 was chosen as the cut-off point for PE credits earned variables (and why 18 is the highest value assignable for cumulative PE credits earned indicators).

III. NAMING CONVENTIONS

A. Variable Names and Labels

Academic Courses variables were named and labeled according to a standard format:

1. Each variable consists of up to but not exceeding 8 upper-case characters.
2. The initial character of all variables is **E**, serving as a reference to the educational component of Add Health, AHAA.
3. The second character of all variables is **A**, specifying that each of these variables belongs to the Academic Courses component of the broader AHAA data set.
4. The next two characters relate to subject headings and sub-headings. The first of these two characters designates a particular subject (**M** = math, **S** = science **F** = foreign language, **E** = English, **H** = history/social sciences, **P** = physical education, **O** = overall), while the second character is used, if necessary, to denote more content specific sub-headings (**A**, **F**, **L**, **S**, and **O** represent 'any,' French, Latin, Spanish and 'other' within the foreign language indicators; **A**, **S**, or **O** denote 'any,' 'survey' or 'other' within the English variables; **A** – **D** represent the various categories history/social science courses were grouped into; and **A**, **G**, **S**, **B**, and **DP** classify physical education courses according to the categories of any, general, competitive sports, marching band, dance/pep squad/cheerleading/drill team, respectively).
5. The next three characters within these variable names indicate variable type, signifying if the variable contains information about sequences (**SQ** and **SQB**), course type (**HN**, **AP** and **RM**), grade point average (**GPA**), failure rates (**FIX**), semesters attempted (**SA**), or credits earned (**CR**).
6. Some variables have an extra character on the end which denotes that the variable data corresponds to a certain year (i.e., year of high school course taking **0-6**), or that the data reflects cumulative information (**C**), or the highest level of a subject attained by the end of high school course-taking (**H**).
7. While the variable names consist of eight or fewer characters, the variable labels assigned to each Academic Courses variable can contain up to 40 characters including spaces.

B. Examples

1. Example of a Course Grade Measure

EAMGPA1 – *grade point average of math courses taken in year 1.*

This variable adheres to the conventions outlined above. The first character is **E**, a general reference to Add Health/AHAA. The second character, **A**, demonstrates that this variable contains information pertaining to the substantive Academic Courses component of Add Health/AHAA. The third character, letter **M**, identifies this variable as relating to the subject math. Notice that a second character is not needed because the label **M** for math sufficiently describes the subject covered under the variable. The next three characters, GPA, describe the specific type of information conveyed. In this case, the variable refers to grade point average. The last character listed for this variable is **1**, indicating that it refers to information about math course-work taken in the first year of high school.

2. Example of a Course Sequence Indicator

EASSQBH – *‘B’ version, highest level science course for which the student received credit by the end of high school.*

Like the previously described variable, this variable begins with **E** and **A**, identifying it as (1) part of Add Health/AHAA and (2) belonging to the Academic Courses component of the Add Health/AHAA data set. The third character, letter **S**, shows that the variable contains information about the subject science. As in the first example, a second character is not needed to further describe the subject at hand because the letter **S** for science provides adequate detail. The letters **SQB** indicate that these are the variables for the ‘**B**’ version of science course sequences. More specifically, they indicate that the variable captures information about the highest level of science each student received credit for in a specified time interval. The ending letter **H** tells the analyst that the variable contains data about the highest science level attained by students at the end of high school course-taking.

C. Data File Names

The Academic Courses component data files were named in accordance with the following conventions:

- 1.** Data file name begins with the prefix **edu.**
- 2.** The remaining characters describe the academic subject(s) and corresponding variables included in the data file.

IV. MISSING CODES

A. Description of Missing Codes

1. 9991: No transcript data exists for a given student
2. 9992: Add Health/AHAA respondent's transcript reflects no course data in a given year
3. 9993: Respondent never took a given subject throughout her/his high school career
4. 9994: Transcript indicates that no subject-specific course was taken in a given year but shows that courses in the specific subject were taken in other years
5. 9995: Transcript does not have a typical grade (A, B, C, D, F) for a subject-specific course in a given year

B. Variable-Specific Considerations

1. Year 0 Variables

Students whose transcripts did not list any course work taken prior to actual high school enrollment are assigned the missing value 9992 for all year 0 variables.

2. Semesters Attempted Variables

Students cannot receive a missing value of 9994 (indicates no subject-specific course was taken in a given year) for semesters attempted variables. Instead, students who did not take subject-specific courses in a given year, but took courses in this subject in other years were assigned a value of 0 for corresponding semesters attempted variables.

3. Credit Earned Variables

Because a value of 0 for credits earned variables indicates students either failed courses attempted or took non-credit courses in a given year, students who did not enroll in subject-specific courses in a particular year but who took courses in the specified subject in other years are assigned the missing value 9994. The missing value 9995 (student did not receive a *grade* in a specified course) is not applicable for credit earned variables, because students can receive credit for non-graded courses.

C. Subject-Specific Considerations

None for Physical Education

V. DATA FILE INVENTORY

All students with course-level information are included in each of these Academic Courses data files listed below.

A. edumsov – this data file contains indicators that capture information about students’ math, science, and overall academic experiences and achievement. Indicator order: Course sequence (regular and ‘**B**’ versions), course grades (GPA), course failures (failure index measures), semesters attempted, and credits earned.

B. edueng – this data set consists of English Academic Courses indicators. Indicator order: Course type, course grades (GPA), course failures (failure index measures), semesters attempted, and credits earned.

C. eduflng – this data set contains foreign language Academic Courses indicators. Indicator order: Course sequence (no ‘**B**’ version), course type, course grades (GPA), course failures (failure index measures), semesters attempted, and credits earned.

D. eduhis – this data file consists of history/social sciences Academic Courses indicators. Indicator order: Course type, course grades (GPA), course failures (failure index measures), semesters attempted, and credits earned.

E. edupe – this data file contains the physical education Academic Courses indicators. Indicator order: Course type, course grades (GPA), and credits earned.

VI. QUICK REFERENCE GUIDE

Physical Education Course Indicators

1. Course Type Indicators

i. General Physical Education

EAPG(1-4) Binary variable indicating student participation in a general physical education class per year (1-4) of high school.

ii. Competitive Sports

EAPS(1-4) Binary variable indicating student participation in a competitive sports class per year (1-4) of high school.

iii. Marching Band

EAPB(1-4) Binary variable indicating student participation in a marching band class per year (1-4) of high school.

iv. Dance/ Pep Squad/Cheerleading/Drill Team

EAPDP(1-4) Binary variable indicating student participation in a dance/pep squad/cheerleading/drill team class per year (1-4) of high school.

2. Course Grades Indicators

EAPAGPA(1-6) Grade point average for physical education courses in each year of high school course taking (1-6).

EAPAGPAC Grade point average for physical education courses cumulatively across all years of high school.

3. Credits Earned Indicators

EAPACR(1-6) Students' physical education course completion rates, or the amount of physical education credits earned, during each year of high-school (1-6).

EAPACRC Students' physical education course completion rates cumulatively across all years of high school.

Education Data - Physical Education Courses

Education - Physical Education Courses

aid		Str 8	RESPONDENT IDENTIFIER NOTE: Smallest 5 and largest 5 values are displayed.
Frequency	Percent	Value	Label
1	0.0%	10316654	
1	0.0%	10316952	
1	0.0%	10506342	
1	0.0%	10606128	
1	0.0%	11316754	
12227	99.9%	11316958-99719976	NOTE: Range of values omitted from display
1	0.0%	99886991	
1	0.0%	99886994	
1	0.0%	99886995	
1	0.0%	99886996	
1	0.0%	99886999	

eapg1		Num 2	Physical Education: General Taken Year 1?
Frequency	Percent	Value	Label
485	4.0%	0	no
9889	80.8%	1	yes
601	4.9%	9993	No PE course on transcript in any year
1262	10.3%	9994	No PE course in specific year

eapg2		Num 2	Physical Education: General Taken Year 2?
Frequency	Percent	Value	Label
1026	8.4%	0	no
7581	62.0%	1	yes
301	2.5%	9992	No course-taking data in specific year
492	4.0%	9993	No PE course on transcript in any year
2837	23.2%	9994	No PE course in specific year

eapg3		Num 2	Physical Education: General Taken Year 3?
Frequency	Percent	Value	Label
1264	10.3%	0	no
3897	31.8%	1	yes
724	5.9%	9992	No course-taking data in specific year
430	3.5%	9993	No PE course on transcript in any year
5922	48.4%	9994	No PE course in specific year

eapg4		Num 2	Physical Education: General Taken Year 4?
Frequency	Percent	Value	Label
1089	8.9%	0	no
3205	26.2%	1	yes
1306	10.7%	9992	No course-taking data in specific year
370	3.0%	9993	No PE course on transcript in any year
6267	51.2%	9994	No PE course in specific year

eaps1		Num 2	Physical Education: Competitive Sports Taken Year 1?
Frequency	Percent	Value	Label
9489	77.5%	0	no
885	7.2%	1	yes
601	4.9%	9993	No PE course on transcript in any year
1262	10.3%	9994	No PE course in specific year

eaps2		Num 2	Physical Education: Competitive Sports Taken Year 2?
Frequency	Percent	Value	Label
7417	60.6%	0	no
1190	9.7%	1	yes
301	2.5%	9992	No course-taking data in specific year
492	4.0%	9993	No PE course on transcript in any year
2837	23.2%	9994	No PE course in specific year

eaps3		Num 2	Physical Education: Competitive Sports Taken Year 3?
Frequency	Percent	Value	Label
4151	33.9%	0	no
1010	8.3%	1	yes
724	5.9%	9992	No course-taking data in specific year
430	3.5%	9993	No PE course on transcript in any year
5922	48.4%	9994	No PE course in specific year

eaps4		Num 2	Physical Education: Competitive Sports Taken Year 4?
Frequency	Percent	Value	Label
3379	27.6%	0	no
915	7.5%	1	yes
1306	10.7%	9992	No course-taking data in specific year
370	3.0%	9993	No PE course on transcript in any year
6267	51.2%	9994	No PE course in specific year

eapb1		Num 2	Physical Education: Marching Band Taken Year 1?
Frequency	Percent	Value	Label
10024	81.9%	0	no
350	2.9%	1	yes
601	4.9%	9993	No PE course on transcript in any year
1262	10.3%	9994	No PE course in specific year

eapb2		Num 2	Physical Education: Marching Band Taken Year 2?
Frequency	Percent	Value	Label
8205	67.1%	0	no
402	3.3%	1	yes
301	2.5%	9992	No course-taking data in specific year
492	4.0%	9993	No PE course on transcript in any year
2837	23.2%	9994	No PE course in specific year

eapb3		Num 2	Physical Education: Marching Band Taken Year 3?
Frequency	Percent	Value	Label
4820	39.4%	0	no
341	2.8%	1	yes
724	5.9%	9992	No course-taking data in specific year
430	3.5%	9993	No PE course on transcript in any year
5922	48.4%	9994	No PE course in specific year

eapb4		Num 2	Physical Education: Marching Band Taken Year 4?
Frequency	Percent	Value	Label
4046	33.1%	0	no
248	2.0%	1	yes
1306	10.7%	9992	No course-taking data in specific year
370	3.0%	9993	No PE course on transcript in any year
6267	51.2%	9994	No PE course in specific year

eapdp1		Num 2	Physical Education: Dance/Pep/Cheer/Drill Taken Year 1?
Frequency	Percent	Value	Label
10180	83.2%	0	no
194	1.6%	1	yes
601	4.9%	9993	No PE course on transcript in any year
1262	10.3%	9994	No PE course in specific year

eapdp2		Num 2	Physical Education: Dance/Pep/Cheer/Drill Taken Year 2?
Frequency	Percent	Value	Label
8315	67.9%	0	no
292	2.4%	1	yes
301	2.5%	9992	No course-taking data in specific year
492	4.0%	9993	No PE course on transcript in any year
2837	23.2%	9994	No PE course in specific year

eapdp3		Num 2	Physical Education: Dance/Pep/Cheer/Drill Taken Year 3?
Frequency	Percent	Value	Label
4851	39.6%	0	no
310	2.5%	1	yes
724	5.9%	9992	No course-taking data in specific year
430	3.5%	9993	No PE course on transcript in any year
5922	48.4%	9994	No PE course in specific year

eapdp4		Num 2	Physical Education: Dance/Pep/Cheer/Drill Taken Year 4?
Frequency	Percent	Value	Label
4033	33.0%	0	no
261	2.1%	1	yes
1306	10.7%	9992	No course-taking data in specific year
370	3.0%	9993	No PE course on transcript in any year
6267	51.2%	9994	No PE course in specific year