A Guide to GEAR UP Program Evaluation: Optimal Research Design, Methodology, and Data Elements

Commissioned Expert Paper
Developed by CoBro Consulting for
RTI, International

Final Report
October 2009

CoBro Consulting

Table of Contents

Table of Contents (continued)

List of Figures

List of Tables

A Guide to GEAR UP Program Evaluation:
Optimal Research Design, Methodology, and Data Elements

Overview

The present paper details effective research practices that may be employed at the local program level as well as those viable for national GEAR UP evaluation. Topics to be addressed include 1) identification of an optimal minimum set of data elements for meaningful GEAR UP evaluation, 2) development of an effective evaluation design and methodology, and 3) key analyses that will aid in determining effective and efficient program processes and help identify the effects of particular program services as well as overall program impact. Additionally, the paper will describe how study findings may be used to inform program processes and guide continuous program improvement.

This document is offered for practical application, rather than theoretical discourse. The aim is to provide a guide to serve as a useful tool for GEAR UP evaluators and practitioners in their GEAR UP program evaluation efforts. To this end, the report will focus on realistic methods that are transferrable and adaptable to a variety of GEAR UP programs. This includes an effective approach to conducting a systematic, multi-factor cost-benefit assessment of GEAR UP program services. At the national level, the guide will offer an evaluation approach that attempts to mitigate the challenges posed by the variety of GEAR UP program resources, goals, structure and composition.

This guide builds upon past successful efforts of GEAR UP evaluation experts and directors. Specifically, it incorporates findings from the 2006 NCCEP/AT&T Foundation National Evaluation Project. The NCCEP/AT&T Foundation project involved determination of an optimal set of data elements common to GEAR UP programs nationally. The resulting *Evaluation Framework* consisted of 73 variables that were organized into four GEAR UP student cognitive and personal *readiness stages*. The optimal data elements identified in this current guide are a modified set of the *Evaluation Framework* indicators.

The first two sections of this guide detail the study methodology adapted and the optimal data elements identified. Sections III and IV then address collection and use of survey and academic GEAR UP program data, respectively. The subsequent four sections (V through VIII) are organized in accordance with the four readiness stages noted above. These cognitive and personal developmental stages correspond to clusters of GEAR UP student grade levels as follows:

Stage 1: GEAR UP students in 6th through 8th grade
Stage 2: GEAR UP students in 9th and 10th grade
Stage 3: GEAR UP students in 11th and 12th grade
Stage 4: GEAR UP students in their first year of college

This structure reflects the fact that GEAR UP is a six-year transformational process whereby the emphasis in program goals shifts as students progress through school. Likewise, the variables required to assess progress will change accordingly. Therefore, within this guide, issues of data analysis and program evaluation that are particular to the specific developmental stages are addressed in these corresponding sections.

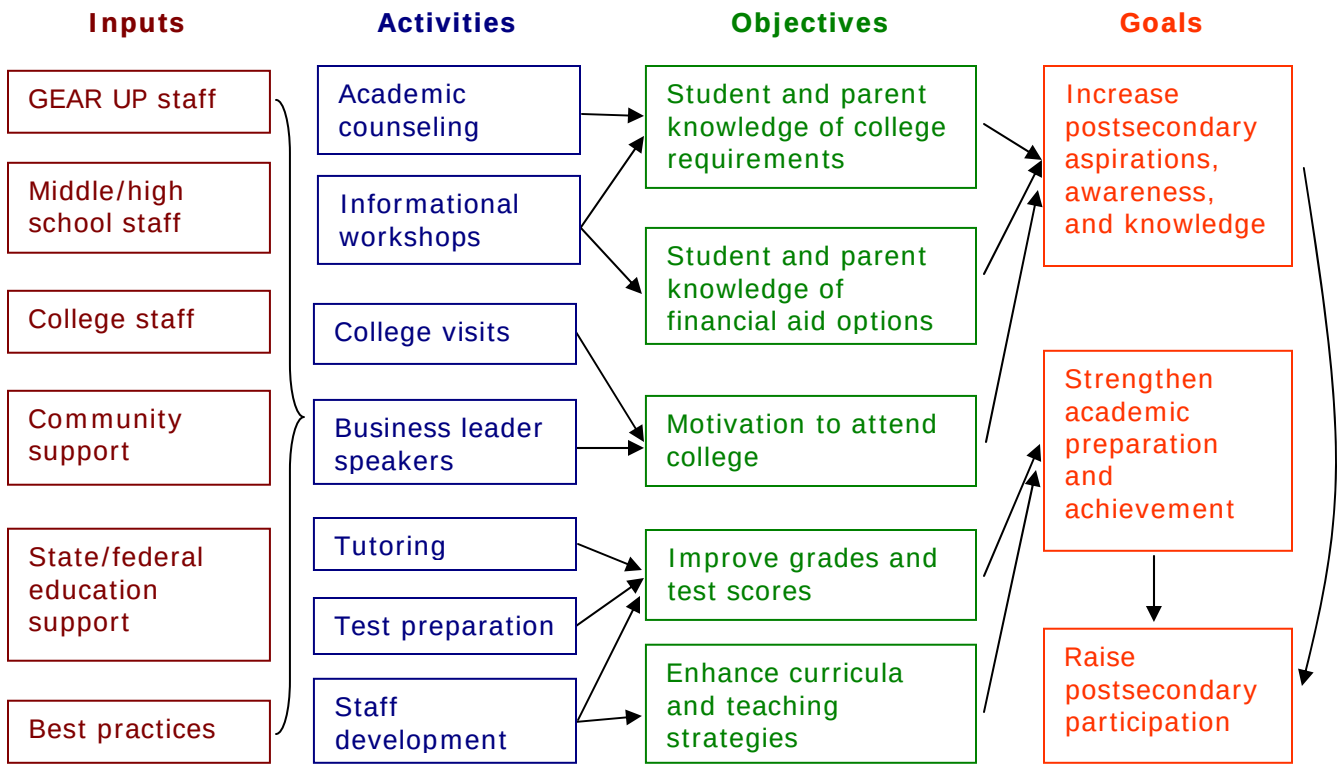
Within each section of this guide, local program evaluation is first addressed, followed by discussion of how such evaluation practices may be adapted across programs for a national evaluation effort.

The final section of this guide, Section IX, addresses interpretation and use of study findings to guide ongoing program improvement. This is in light of the fact that, even when rigorous evaluation

procedures result in useful findings, they have no value if the results are not then properly interpreted and leveraged to inform program processes. In this section, the guide offers effective methods to convey research findings to program staff, participants, and stakeholders. Examples are provided in the form of clear, simple, summary sheets graphically illustrating study findings. Such summaries convey results in a non-intimidating manner and have been extremely well-received by programs. These media reports also serve to engage and motivate program staff, participants, and community members. Thus, the expert paper will conclude with suggestions for development of user-friendly, graphically-oriented results reports that will help empower grant staff and administrators in their efforts toward continuous program improvement.

Note that, while GEAR UP programs share several essential components and overall goals and objectives, it is recognized that each program is unique and involves services and goals particular to their regional population needs. As such, it is acknowledged here that the evaluation methods presented in this guide may require modification for adaptation within individual GEAR UP programs. Moreover, the development and use of program-specific logic models in guiding evaluation processes for individual GEAR UP programs is strongly encouraged to ensure the evaluation effort is both comprehensive and targeted to particular programmatic needs. Figure 1, below, presents an example of such a GEAR UP program logic model.

Figure 1. Sample logic model



Section I:
Methodology

Sampling

With respect to evaluation at the program level, this guide advocates inclusion of data from the entire population of participants, rather than a sample. This is in consideration of the fact that often the tasks involved in identifying data records to track a sample of GEAR UP cohort students is more time consuming and labor intensive than obtaining and analyzing entire datasets of the cohort. Therefore, discussion of sampling techniques is not relevant for individual GEAR UP program evaluation.

Regarding nationwide studies of GEAR UP, it should be noted that currently there are over 200 GEAR UP state and partnership grants, serving over three quarters of a million low-income students. Given the magnitude of such a broad-scale evaluation project, the evaluation approach detailed in this guide emphasizes simple and practical methods at each stage of the process. For such studies across GEAR UP programs, this guide recommends inclusion of either the entire GEAR UP population, or, if resource limitations preclude this, a stratified proportionate sample of GEAR UP programs. Whether the evaluation model includes the entire population of programs or a well-constructed sample, since GEAR UP programs cannot be compelled to participate, it is expected that some programs would opt out of the study. There are two primary effects of such attrition that need consideration.

First, the number of GEAR UP programs included in the analyses must be large enough to ensure an adequate level of statistical power. Otherwise, meaningful program effects would likely go undetected. Even with adjustments for use of a small sample size, the reduction in statistical power would correspondingly decrease the sensitivity of analyses. Since it is expected that some attrition will occur due to programs opting out, an adequate number of GEAR UP programs must be identified initially to counteract the expected decrease in sample size. Thus, if sampling us used, it is recommended that no less than 35% of the GEAR UP programs are selected as participants, to help ensure the resulting sample is of adequate size to ensure statistical power and, in turn, sensitivity of analyses.

Additionally, regardless of whether the evaluation model includes the entire GEAR UP program population or a sample of the programs, the natural effect of voluntary attrition would likely result in systematic bias. Therefore, a concerted effort must be made to ensure the programs included in the study reflect the characteristics of the larger population. For example, if sampling is used to identify GEAR UP programs for inclusion, it is suggested that proportionate stratified sampling procedures are employed to promote a representative sample in terms of particular program characteristics (e.g., urban versus rural regions, state versus partnership structure, school district versus university grant holder). Moreover, as noted, preliminary analyses across programs would involve comparisons of characteristics of the resulting participating programs with those of the entire population of GEAR UP programs, based on general APR data indicators obtained for all programs. In this manner, the resulting group of participating programs will be examined to ensure it is representative of the total population of GEAR UP programs in terms of general characteristics.

Level of Analysis

It is possible for GEAR UP program data to be collected at varying levels of detail, from aggregated program-level data (the least specific) to course-level information for each participating student (the most detailed). The level of specificity of the data brings both drawbacks and benefits. Namely, although more detailed data often increases the sensitivity of the research design, such data are also typically more difficult for programs to obtain and format into consistent and meaningful values. This is a

challenge at the local program level, and becomes virtually prohibitive when collecting data on a national level, across GEAR UP programs.

However, there are also difficulties associated with collecting data in aggregate, at the cohort- or school-level. Because GEAR UP programs provide services primarily to an at-risk, lower socioeconomic population, students in the participating schools are more likely to be transient or migratory. In fact, many GEAR UP programs report annual turnover rates of 30% or higher. As a result, when data are collected in aggregate (e.g., at the school- or grade-level), such turnover works to attenuate program effects, since outcomes of students new to the school, and thus to the GEAR UP program, are included in the aggregate outcome (post-treatment) scores. Consequently, any program effects are diluted, artificially repressing program impact metrics and thus affecting study findings. Moreover, opting for aggregate data also prevents examination of dosage effects, which could reveal any associations (positive or negative) between service participation and academic achievement for students. Determining such dosage effects for particular services is a critical component of formative GEAR UP program evaluation. Further, the attenuation prevents aggregate examination of GEAR UP and comparison group students with particular characteristics expected to affect student outcomes, such as socioeconomic status, race/ethnicity, and baseline academic achievement.

For example, a common objective of GEAR UP programs is increased achievement on standardized tests. When examining aggregate data of average test scores for students in the program, it is unknown whether any changes in the average scores from year to year apply to students who received GEAR UP services at all. Average scores from one year to the next may be derived from a very different group of students, depending on the level of student turnover. Such turnover may result in very misleading results for such analyses. If lower performing students are those who leave the school (or drop out) while higher performing ones enter, average test scores will appear to increase. Thus, examination of only aggregate test score data would not reveal whether any meaningful change in test scores is occurring. Likewise, there would be no way to examine effects of individual GEAR UP services or any potential dosage effects, because aggregate data would only provide information on the program as a whole, though service participation levels often vary widely for individual students.

Use of student-level data will allow for linkage between participation in services and academic outcomes, in conjunction with examining change scores and employing a retrospective comparison group as discussed in more detail below. Tracking student test scores from year to year on a student-level dataset provides an understanding of whether students show improvement. At the student level of analysis, change score analyses automatically result in the exclusion of students without data during both timeframes examined, thus ensuring that test data included in the analyses are based on only those students who would have experienced GEAR UP, as opposed to those who left the GEAR UP schools or those are newly entering and so have little experience of the program. Additionally, test scores from individual students can be used to study any effects of particular services differing among students within a program, or level of participation within each service.

Given these issues, for GEAR UP evaluation at both the program-level as well as nationally (across programs), this guide adopts a research design that uses student-level data. However, because course outcome information is a key indicator of GEAR UP program success, the guide proposes a method to retain useful elements of course-level information as well. Namely, as an alternative to collecting the more specific course outcome data for each student (such as letter grade, pass/no pass status), it is suggested that dichotomous variables are created to flag students' enrollment in, and successful completion (i.e., receiving marks of A, B, C, or *Pass*) of core academic courses. Additionally, given that the large majority of turnover in program participants noted above occurs between school years, in contrast to the designated cycle for the U.S. Department of Education (DOE) GEAR UP Annual Performance Report (APR), the current study approach will use the academic year as the unit of time with respect to analyses.

With regard to nationwide evaluation of GEAR UP, across programs, every effort will be made to obtain viable student-level data; however, it is not expected that such data will be available for all programs. Therefore, although the evaluation approach aims to include the entire population of GEAR UP programs, from a practical standpoint, the actual goal will be to obtain student-level data from as many programs as possible. Meanwhile, for those programs that do not have such data available, aggregate data at the program level will be collected to allow for identification of any systematic biases inherent in the study due to characteristics of participating programs. In other words, the programs that have access to, and use, student-level data may differ from those programs that do not use data to the same extent. These variations in access to resources can result in differences in program effectiveness. For example, such programs may be more technically sophisticated and/or research-oriented, or simply have more programmatic resources available to them in comparison with GEAR UP programs that do not rely on data to the same degree. Thus, preliminary analyses will involve comparisons of GEAR UP APR aggregated data for participating and non-participating programs to detect and, if evident, control for this disparity. Because all GEAR UP programs are mandated to provide data for the APR, it is expected that all, or nearly all, programs will be included in this stage of the analysis.

Evaluation Model

As with most social program evaluation, it is not feasible or practical to randomly assign participants to treatment and control groups, which is required within a randomized experimental design. GEAR UP programs are structured as cohort models, priority models, or a combination of the two. In a cohort design, an entire grade of students in the schools receives services, generally beginning in the seventh grade and continuing in those same schools one grade level higher each year. The vast majority of GEAR UP programs uses this cohort program design; a small subset employ an *add-a-cohort* model in which consecutive seventh grade classes receive services. A priority model involves selection of students from GEAR UP schools based on particular characteristics, such as low test scores or socioeconomic status. Therefore, given that randomization is not inherently employed as a method of selection of students for GEAR UP participation, it cannot be utilized as part of a GEAR UP program study design.

Thus, a quasi-experimental approach is taken for both program-level and nationwide GEAR UP program evaluation. To maximize design rigor, a comparative change model is employed. This model incorporates both baseline and comparison group measures to control for alternative explanations for research findings such as maturational and/or selection effects. Note that this comparative change model applies to programs with cohort program structures, as it involves comparison of entire classes of students within the GEAR UP schools. Methods of incorporating these measures are detailed below.

Comparison Group Measures

To account for changes in outcomes due to historical or maturational effects, the design proposed in this guide includes comparison group measures. Often, when outcomes of students from socioeconomic and demographically matched schools are used as GEAR UP student comparisons, many differences remain due to issues such as school policies, resources and budget, leadership, and campus culture, which can influence student academic outcomes. Given this, a retrospective comparison group is incorporated in the present evaluation design. Specifically, the design includes a comparison group based on student-level academic outcomes data from students one year prior to GEAR UP implementation. For example, if the GEAR UP program began serving 7th graders in five schools in 2003/2004, the corresponding comparison group would be academic data from the 7th grade students in 2002/2003 from the same five schools. In this manner, changes in academic performance or other outcomes found for GEAR UP

participants over time are less likely to be dismissed as naturally-occurring changes due to student social developmental factors that would happen regardless of program participation.

It is acknowledged that use of a retrospective comparison group also brings potential drawbacks. Namely, changes revealed between groups of students from year to year may arise from extraneous temporal changes rather than actual treatment effects. For example, school district or state-wide policy changes affecting course enrollment requirements, standardized tests/proficiency levels, or academic rigor may result in spurious evaluation findings. To identify and control for such occurrences, this guide details further analyses and examination to be conducted if statistically significant differences are found between student groups from year to year.

As another precaution to this approach, GEAR UP programs in their second grant award are excluded from analyses involving retrospective comparison groups unless the prior grant program would not be expected to have affected the comparison group. However, due to the GEAR UP emphasis of sustainable change, it is expected that most GEAR UP programs in their second grant will have implemented services and policies that would impact the comparison group.

Baseline Measures

In addition to collecting student outcome data during the years of program administration, the evaluation design in this guide involves obtaining academic information for participating students from the year prior to grant implementation. If available, existing baseline survey information is also incorporated into the evaluation.

The baseline measures are used to derive change scores for students within both GEAR UP and comparison groups. Specifically, the difference from baseline (pre-treatment) outcome measures (such as course grades and standardized test scores) and outcomes obtained after participation in GEAR UP services (post-treatment) are calculated. These resulting change scores are then employed in the proposed analyses rather than relying on post-treatment scores alone. This method will help guard against selection effects, whereby differences are found that are actually due to pre-existing differences between groups of students prior to treatment.

For example, examination of test scores might indicate that GEAR UP students, after receiving one year of services, show average scores lower than those of the comparison group. Without baseline scores with which to compare these test scores, this result is disappointing. However, baseline scores might show that scores for the GEAR UP students increased, while those of the comparison group decreased, indicating progress made by the GEAR UP students. A finding such as this could lead to additional questions, such as reasons why the GEAR UP and comparison group students differed at baseline, which could in turn steer the evaluation towards analyses that account for statistical issues such as regression toward the mean or covariates inherent in the groups.

As noted, any analyses involving comparison with baseline measures would likely need to exclude those programs implementing a second multi-year grant. Because schools often implement activities intended to impact all students, baseline figures for these programs would be inflated compared with pure pre-treatment measures.

Data Collection

For evaluation at both the program level and nationwide, this guide advises using available archival data whenever possible to minimize the disruption to program participants. This involves working with

program personnel to determine and obtain all existing and relevant data, such as course grades, survey results, service and activity participation, and staff development attendance records. It is also stressed that any procedures involving transmission of electronic data must adhere to Family Education Rights and Privacy Act (FERPA) guidelines by applying appropriate security measures.

As noted above, the approach also includes use of available survey data. Both program-level and national GEAR UP evaluation involve the DOE-mandated student and parent survey items, which are included within the set of Optimal GEAR UP Program Evaluation Indicators.

Additionally, for individual program evaluation, the guide emphasizes the value of administering augmented student and parent surveys to obtain information beyond the mandated DOE items. This allows for more in-depth evaluation of program effectiveness, such as gap analysis to examine alignment (or lack thereof – i.e., disparities) between participant needs and service offerings. Results from these supplemental survey items can also offer insight into student and parent satisfaction and perceptions of effectiveness of GEAR UP services. Figures 3 and 4 (pages 14-17) present examples of such augmented survey forms.

At the national level, the approach includes conducting a survey of GEAR UP program top administrators (one per program) to obtain information about perceptions of program and service effectiveness and opportunities for improvement.

Additionally, the benefits of targeted focus groups of GEAR UP program staff and faculty are addressed. Such focus group sessions have been invaluable in examination of program processes such as service delivery, staffing, and organizational structure. Table 2 (page 21) provides suggested GEAR UP focus group questions.

Data Analysis

As noted above, the methodology recommended for individual GEAR UP program evaluation calls for nearly all resulting data to be maintained at the student level. However, some variables are included in aggregate form only, at the school- or program-level. This is primarily due to the nature of these variables, such as C*ost of each type of student service*, or *Number of GEAR UP schools.* Other data elements are collected in aggregate because they are typically unavailable for GEAR UP program use at the individual level (e.g., *Total hours of GEAR UP teacher service participation, by type).*

The nationwide approach will begin with inferential statistical analyses (such as t-tests) to identify any existing differences between those programs that have access to student-level data and those that do not. Beyond these preliminary analyses, the methodology for nationwide GEAR UP evaluation also relies upon use of student-level data.

Analyses included in the current approach involve both descriptive and inferential statistics that have been used to provide meaningful and useful information for GEAR UP programs. In evaluation of individual GEAR UP programs, occasionally data are available only in aggregate form. For example, the number, or percent, of students scoring within specified test proficiency levels may only be available in aggregate via a school or school district website. Analyses are employed that optimize their value, and include models using baseline and/or comparison measures. The most sensitive inferential statistical analyses appropriate to the type of data are recommended. Where necessary, such as when only categorical variables are available, analyses include the chi-square test for two independent samples, or the Fisher exact test. For interval data available, such as standardized test scores and grade point averages, more sensitive inferential statistics may be used, such as t-test, ANOVA, and correlation analyses (point biserial and Pearson r, as appropriate). The methodology also includes disaggregation of the analysis results by race, gender, and socioeconomic status in order to determine any disparate program effects related to student characteristics.

With respect to both nationwide and program-level evaluation, the benefits of studying participation in particular GEAR UP services are emphasized. These analyses allow identification of program elements that are effective, as well as those that are dispensable or require modifications, thus enabling continuous program improvement. Such formative program evaluation often involves examination of dosage effects, through correlating student participation in particular program activities and services (e.g., student participation in the after-school tutoring programs) with changes in outcomes such as academic achievement (e.g., GPA and course completion). These analyses can reveal whether greater student and parent participation in particular program services results in increases in student academic achievement, thus revealing program service dosage effects.

Data Organization

Due to common, practical challenges inherent in collection of GEAR UP participant data, this guide proposes establishing three separate participant-level data sets of GEAR UP indicators for local GEAR UP program evaluation, in addition to the school- and program-level data elements discussed in the prior section. The first set of indicators is comprised of student-level academic (course and standardized test) and program participation data. The student records in this dataset are linked via a unique identifier (ID) – namely a school or district-level ID assigned to each student. The additional two data sets contain student and parent survey data that is also maintained at the student/parent level. Although these elements may be used in multivariate analyses, they do not contain unique student identifiers and thus are not linked with quantitative indicators of achievement. Further rationale for these separate datasets is provided in Section 5 of this guide (i.e., *Stage 1: 6th to 8th Grade*, page 30-34).

With respect to nationwide GEAR UP evaluation, as noted above, after preliminary evaluation at the program level, data for participating programs will then be analyzed at the student level. Therefore, the same three types of datasets described above are included for national GEAR UP evaluation. However, the datasets containing student and parent survey elements will likely contain fewer variables, since only elements common across the majority of participating programs would be included.

Section II:

Determination of an Optimal Set of Data Elements

The optimal GEAR UP data elements identified in this guide are derived from past successful efforts of GEAR UP program evaluation. The majority of the variables, as well as the organizational features of the dataset, are adapted from results of a GEAR UP evaluation project led by the National Council for Community and Education Partnerships (NCCEP). The resulting set of *Evaluation Framework Indicators* was modified in development of the present Optimal GEAR UP Program Evaluation Indicators. Some variables were added, revised, or removed, based on extensive experience with evaluation of GEAR UP programs that took place in the three years following the 2006 NCCEP/AT&T project.

Development of the NCCEP/AT&T Foundation Evaluation Framework Indicators

The primary goal of the 2006 NCCEP/AT&T Foundation National Evaluation Project was to identify a set of transferable GEAR UP evaluation indicators that would empower local GEAR UP sites to better document their successes. To accomplish this, NCCEP established a National GEAR UP Evaluation Council comprised of GEAR UP practitioners, evaluators, and academic experts. This Council was then charged with creating a comprehensive *Evaluation Framework*, consisting of an optimal set of data elements common across programs that could be used for meaningful program evaluation.

It was critical that the resulting variables identified were transferrable across GEAR UP programs to ensure their utility on a national scale. To this end, the Council constructed and administered a nationwide survey of GEAR UP program directors to learn about data uses and availability at the local program level. The primary goal of the survey effort was to elicit information about current capacity and practice of GEAR UP grantees to collect and integrate their program data with other student data, such as academic outcome information. The survey included items to determine data collection methods, types and level of data collected (i.e. aggregate program-level or student-/course-level), and surveying frequency and response rates. It also inquired as to how programs use the resulting information and the types of analyses conducted.

The survey was administered to the Program Directors or Managers of the approximately 245 active GEAR UP programs in 2006. Responses were received from 85 GEAR UP programs (response rate = 35%). Among other findings, results revealed that most programs (88%) collect participation data at the student level and the large majority of programs (93%) have access to course-level academic data for each GEAR UP student. The Council considered the reported availability of particular data elements in determining the optimal set of GEAR UP evaluation variables.

This process resulted in the development of an Evaluation Framework, which consisted of 73 GEAR UP data elements. The Framework included 37 variables that GEAR UP programs are currently required to collect annually for DOE APR reporting purposes, and an additional 36 supplementary items.

In development of the optimal set of GEAR UP variables, the Evaluation Council also considered the fact that GEAR UP is a progressive, six-year program with evolutionary stages in terms of attainment of essential program goals. In acknowledgement of this evolution of GEAR UP goals, the Council developed the Evaluation Framework indicators within the structure of four *Readiness Stages.* The program emphasis at each Readiness Stage is as follows.

- Stage I (6th to 8th grade): Predisposition, awareness, and knowledge acquisition
- Stage II (9th and 10th grade): Motivation, expectations, and aspirations
- Stage III (11th to 12th grade): Academic enrollment, preparation, and achievement
- Stage IV (1st and 2nd postsecondary years): College course success and persistence

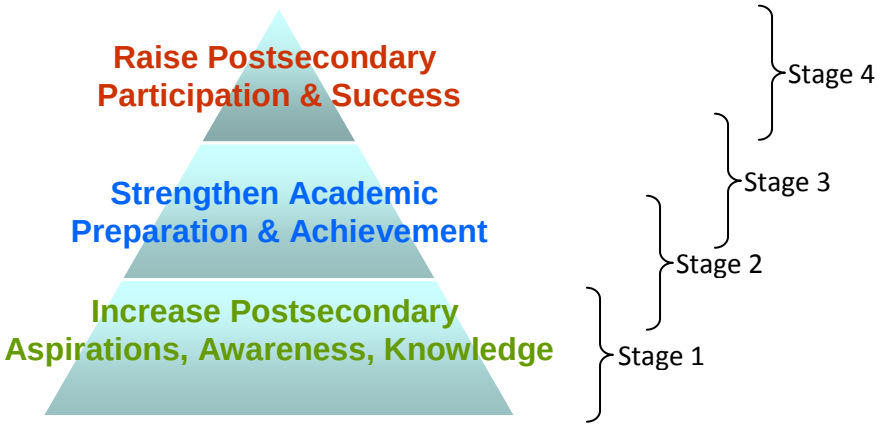
Appendix A presents the final NCCEP/AT&T Evaluation Framework Indicators.

The Optimal GEAR UP Evaluation Indicators

The optimal set of GEAR UP indicators identified in this present guide are based on the NCCEP/AT&T Evaluation Framework. However, as noted, these were modified in consideration of subsequent experience using such data elements in evaluation of a variety of GEAR UP programs nationwide. Likewise, although the general structure of the four Developmental Stages applied in this guide is adapted from the Readiness Stages described above, they have also been modified based on noted shifts occurring in GEAR UP goal emphasis.

The evolution of GEAR UP program goals applied within the present guide is depicted in Figure 2, below, which illustrates that each program goal builds upon previously attained goals.

Figure 2. Progression of GEAR UP Program Goals



The primary goals within each of the four GEAR UP Developmental Stages applied within in this guide are as follows:

- Stage 1 (6th to 8th grade): Aspirations, awareness, and knowledge acquisition
- Stage 2 (9th & 10th grade): Aspirations, knowledge acquisition, academic enrollment, & preparation
- Stage 3 (11th to 12th grade): Academic enrollment, preparation, and achievement
- Stage 4 (1st postsecondary year): College course enrollment, success, and persistence

Note that, while Stage 2 continues to emphasize increasing aspirations and knowledge attainment, academic course enrollment and preparation also become important during this stage. This reflects a noted nationwide trend toward increased course enrollment requirements to ensure students are *on track* to graduate high school.

Additionally, given the fact that GEAR UP grants span a maximum of seven years, Stage 4 was reduced from spanning the first two years of college, to just including the first year. This change increases alignment between the four stages and GEAR UP programs that begin with students in the 7th grade.

Within the NCCEP/AT&T Evaluation Framework, the variables were separated into those available from the APR and those that are not. The present set of GEAR UP indicators is not divided in this manner. This is because the strategy of the present guide involves collection of GEAR UP indicators at the student level (aside from preliminary analyses to examine differences between programs participating in the national study and those that are not).

For each of the four Developmental Stages, the present set of GEAR UP indicators are organized into three categories corresponding to the three data sets employed in this guide: 1) aspirations, knowledge, and awareness data, 2) academic & participation data, and 3) aggregate data used for preliminary program-level analyses and later cost-benefit assessment. The first two categories also include control and predictor variables found valuable in such analyses.

Table 1, on the following page, presents the optimal GEAR UP evaluation indicators suggested for each of the four Developmental Stages, categorized into the noted three data set groupings. Note that, in contrast with the NCCEP/AT&T Evaluation Framework, Table 1 begins with the aspirations, awareness, and knowledge indicators. This is in consideration of the fact that these indicators are emphasized more during the first GEAR UP program developmental stage.

Note that several of the variables were modified due to the shift from program-level variables (as conceptualized in development of the Framework Indicators) to student-level data. For example, *Number of students by grade* was modified to *Grade of student.* However, other variables, such as those relating to participating teachers, were retained but are to be collected at the school level.

Many of the Awareness Acquisition variables were also modified based on the practical utility of the resulting data. This group of indicators was also expanded to include, *Student Aspirations and Knowledge and Awareness Acquisition* data.

Likewise, many indicators within the Student Readiness Behaviors grouping were omitted or modified based on previous experience collecting such data. For example, obtaining viable student absence data from schools or districts is particularly difficult. Other indicators in this grouping (such as SAT score), did not align with the corresponding grade range (i.e., Stage 1: 6th – 8th Grade).

The data elements relating to cost of service were retained within the present set of optimal indicators. Such variables are vital in assessing the cost-benefit of the individual GEAR UP services and activities. Section IX of this guide describes the process of conducting a systematic, multi-factor cost-benefit assessment that incorporates these indicators.

The resulting set of Evaluation Indicators include both survey results and academic outcomes variables. Moreover, a variety of outcome indicators are included, thus allowing for triangulation of the data to increase the credibility of future research findings, in accordance with convergent validity theory.

For each of the four noted GEAR UP program stages, subsequent sections of this guide offer methods and strategies to collect, analyze, and interpret these indicators.

Table 1. Optimal GEAR UP Program Evaluation Indicators

Stages ↓	Student & Parent Survey Data		Student Academic & Participation Data		Aggregate Data
Readiness Types →	Aspirations, Knowledge & Awareness Acquisition	Control & Predictor Variables	Student Readiness Behaviors	Control & Predictor Variables	Control Variables
Stage 1: 6th – 8th grade	**Student** 1. Student education expectations 2. Student education aspirations 3. Student perceptions of college affordability 4. Student knowledge of college requirements 5. GU impact on post-HS intentions 6. Student knowledge of college costs **Parent** 7. Parent knowledge of college requirements 8. Parent perceptions of college affordability 9. Parent educ expectations of students 10. Parent educ aspirations of students 11. Parent knowledge of college costs	**Student & Parent** 12. School 13. Gender 14. Race/ethnicity 15. Time in GU prog 16. Yrs enrolled in GU school 17. Level of GU participation 18. Eligible for free/red lunch program 19. Parent highest education level attained	20. At/above grade level lang arts on state standardized test 21. At/above grade level math on state standardized test 22. At/above grade level science on state standardized test 23. EXPLORE English 24. EXPLORE math score 25. EXPLORE reading score 26. EXPLORE science score	27. Student ID 28. School 29. Grade 30. Gender 31. Race/ethnicity 32. LEP[1] Status 33. IEP[2] Status 34. Free/Reduced Lunch Status 35. # hrs GU student service by service type 36. # hrs GU parent services by service type	**School-level** Variables 20-26 & 29-36, in aggregate 37. # hrs GU teacher svcs by type 38. Cost of student svcs by type 39. Cost of parent svcs by type 40. Cost of teacher svcs by type **Program -level** 41. # GU schools 42. # of GU teachers by grade
Stage 2: 9th – 10th grade	Variables 1-11	Variables 12-19	Variables 20 – 26 43. Completed Algebra I 44. Completed Algebra II 45. Completed Calculus 46. Completed Chemistry 47. Completed Physics 48. Enrolled in AP math 49. Enrolled in AP English 50. Enrolled in AP Science 51. Completed AP math 52. Completed AP English 53. Completed AP Science 54. PLAN English score 55. PLAN math score 56. PLAN reading score 57. PLAN science score 58. PSAT reading score 59. PSAT math score 60. PSAT writing score	Variables 27-36	Variables 20-26 & 29-36, in aggregate Variables 37-42
Stage 3: 11th – 12th grade	Variables 1-11	Variables 12-19	Variables 20-22 & 43 – 57 61. ACT English 62. ACT math score 63. ACT reading score 64. ACT science score 65. SAT reading score 66. SAT math score 67. SAT writing score 68. Enrollment in College	Variables 27-36 69. First Name 70. Last Name[3] 71. Middle Initial[3] 72. Date of birth[3]	Variables 20-26 & 29-36, in aggregate Variables 37-42
Stage 4: 1st Year of College			68. Enrollment in College 73. Cont'd college enrollment	Variables 69-72	Variables 20-26 & 29-36, in aggregate Variables 37-42

[1] LEP: Limited English Proficiency.

[2] IEP: Individualized Education Plan.

[3] Required by National Student Clearinghouse to obtain student college enrollment information.

Section III:
Collection and Use of Data to Gauge College Aspirations, Knowledge, and Awareness

Survey Data

Survey data may be collected and analyzed at either the participant (student or parent) level or aggregated to grade, school, or program level. For evaluation at the local, program level, as noted, this guide advises that such data are collected and analyzed at the individual participant level. Specifically, survey data would be entered into a data set whereby each record in the data set pertains to information on an individual student and/or parent.

Although there are advantages to linking student survey data with other GEAR UP indicators such as service participation, the added complexity and resources required to incorporate unique student identifiers within such datasets are often prohibitive. Therefore, unless GEAR UP programs have found a practical and feasible method to obtain unique student identifiers as part of their surveying process, it is advised that survey data collected from individual students and parents are managed within separate datasets from the dataset containing academic and participation data (which would include unique school/district ID numbers for students).

The primary rationale for including student school/district IDs among survey information obtained is to later merge the resulting survey data with other sources of information such as student transcript data and hours of GEAR UP program participation. This allows for analysis of predictor and outcomes variables. Possible analyses would include correlations between *hours of program participation* and reported levels of student knowledge, awareness, and college aspirations. An additional reason for inclusion of student IDs is to track results for students over time, by merging the survey information with data obtained from subsequent survey administrations.

Although the above benefits are acknowledged, including IDs on student surveys typically brings a host of challenges as well. Very often, GEAR UP programs must cope with low survey response rates. When students know their survey responses are identifiable, they are often even less apt to submit completed surveys and, those that do are often less candid with their responses. Moreover, tracking student survey forms by unique identifiers is often cumbersome or, at minimum, problematic. For either paper-based or online surveys, researchers are limited to two general approaches to obtaining unique student identifiers. First, the survey form would require students to write, type, or *bubble-in* their ID numbers. The primary drawbacks to this method are 1) students may not know their ID number and leave it blank, 2) if students write in their number, some responses will remain illegible, and 3) if a bubble-in approach is used, the bubble-in box of numbers requires substantial survey form space. As an alternative, the survey forms may be *preslugged* with student names and IDs, or linked to student IDs via barcodes. The drawback to this approach is that the students must each be issued the form containing their preprinted name or linked barcode, requiring staff to shuffle through numerous forms and, inevitably, a percentage of students in other locations or classrooms during administration will not have their proper form. Such methods increase the level of complexity and negatively impact the feasibility of surveying students.

As an alternative, online or paper-based surveys may contain some key outcomes-oriented and control items for use in analysis without the challenges of collecting unique identifiers. Figure 3 presents a GEAR UP student survey that contains a sample of such items. Note that item 26 inquires about GEAR UP program participation the prior year. Likewise, item 13 asks about the frequency of participation. Correspondingly, Figure 4 presents a sample GEAR UP parent survey containing similar items. Relationships between these variables and student levels of college aspirations, knowledge, and awareness may be examined via nonparametric statistics such as chi-square analysis. Note that, while this alternative approach still does not allow for student-level analysis across survey administrations, longitudinal analyses at the aggregate, cohort-level are possible, and useful to examine descriptive indicators for trends across survey administrations.

Figure 3. Sample GEAR UP Student Survey

City College GEAR UP

Student Survey: Fall 2008

The purpose of this survey is to help us learn whether the City College GEAR UP Program is doing a good job. Please answer the following questions about your school experiences, future plans, and opinions about GEAR UP. The information you give will be confidential. Your individual responses will not be shared with your teachers and will not affect your grades. Also, survey responses will be combined before they are presented -- Individual responses will not be reported.

Note if you have already completed and submitted this survey online, please do NOT submit a second survey.

Please **completely fill in the bubble** for your response as shown: CORRECT: ● INCORRECT: ⊘ ⊗ ⊘

Please indicate your school: ○ School A ○ School B ○ School C ○ School D ○ School E ○ School F

ACADEMIC INVOLVEMENT	High school or less	Some college, no two- or four-year college degree	Two-year college degree	Four-year college degree or higher
1. What is the highest level of education that you would like to obtain?	○	○	○	○
2. What is the highest level of education that you expect to obtain?	○	○	○	○

3. How important to your future is getting an education beyond high school?
   ○ Very important ○ Important ○ Somewhat important ○ Not important

4. If you do not continue your education after high school, what would be the main reason? (Mark only one)
   ○ I will continue
   ○ It costs too much or cannot afford it
   ○ I need to work
   ○ I want to work
   ○ My grades are not good enough
   ○ I want to join the military service
   ○ Family issues
   ○ Other: (Please write in)_____

KNOWLEDGE ABOUT COLLEGE	Yes	No
6. Do you know what you need to do to get accepted into the college/university you plan to attend?	○	○
7. Has anyone from your school or GEAR UP ever spoken with you about the availability of financial aid to help you pay for college?	○	○
8. Do you have enough information about financial aid to help you pay for college?	○	○
9. Has anyone from your school or GEAR UP ever spoken with you about college entrance requirements?	○	○
10. Do you have enough information regarding college entrance requirements?	○	○

11. Do you think that you could afford to go to a four-year college using financial aid, scholarships, and your family's resources?
    ○ Definitely ○ Probably ○ Not sure ○ Probably not ○ Definitely not

12. How much do you think it costs (tuition and fees only) to attend a four-year public college in your state for one year?
    ○ $1-$1,900 ○ $1,901-$4,500 ○ $4,501-$7,400 ○ $7,401-$10,000 ○ $10,001-$18,000 ○ More than $18,000

13. How often did you participate in GEAR UP activities last year?
    ○ Never ○ Rarely (1-3 times/year) ○ Sometimes (~Once a month) ○ Frequently (~Every week)

14. How much has being in GEAR UP helped you to be more successful in school?
    ○ Very much ○ Somewhat ○ A little ○ Not at all ○ N/A – I haven't participated in GEAR UP

Page 1 of 2

Figure 3. Sample GEAR UP Student Survey (continued)

15. Has being in GEAR UP helped you prepare to successfully graduate from high school when the time comes?

○ Yes ○ No, I was already on track to graduate. ○ No, I still don't know if I will graduate.

16. Has being in GEAR UP helped you to decide to go to college after high school graduation?

○ Yes ○ No, I was already planning on going to college. ○ No, I still don't plan to go to college.

17. Have any of the following GEAR UP resources helped you with your decision to go to college?

Yes	No		Yes	No	
○	○	Tutoring or help with school work	○	○	College preparation workshops
○	○	Visiting a college campus	○	○	Information about the benefits of going to college
○	○	Mentoring from GEAR UP Staff	○	○	Information about financial aid and how much college costs
○	○	Other: (Please write in)_____			

18. If you participated in any of the following GEAR UP or school sponsored activities/services during this school year, how effective was it in helping you prepare for college?

	Very Effective	Effective	Somewhat Effective	Not Effective	N/A - I did not participate
a. Tutoring in math	○	○	○	○	○
b. Tutoring in English	○	○	○	○	○
c. Academic counseling/advising	○	○	○	○	○
d. GEAR UP workshop	○	○	○	○	○
e. Visit to a high school or college	○	○	○	○	○
f. Standardized test prep (e.g., state test, ACT, SAT, PLAN)	○	○	○	○	○
g. Work packets/materials to help you catch up or move ahead in school	○	○	○	○	○
h. Interest/career exploration activities	○	○	○	○	○

19. Overall, how satisfied have you been with the GEAR UP program this year?

○ Very satisfied ○ Satisfied ○ Neutral ○ Dissatisfied ○ Very dissatisfied

BACKGROUND

20. What is your current grade level? ○ Grade 7 ○ Grade 8 ○ Grade 9 ○ Grade 10 ○ Grade 11 ○ Grade 12

21. What is your gender? ○ Male ○ Female

22. What is your race/ethnicity? (Mark all that apply)

○ Asian ○ Hispanic or Latino/a ○ Pacific Islander
○ African-American/Black ○ Middle Eastern ○ White/Caucasian
○ Filipino ○ Native American ○ Other: (Please write in)_____

23. What is your primary language? (Mark all that apply)

○ English
○ Spanish
○ Other: (Please write in)_____

24. Do you qualify for the free/reduced lunch program? ○ Yes ○ No

25. Were you enrolled in this same school during the last school year (2007/2008)? ○ Yes ○ No

26. Did you participate in a GEAR UP program during the last school year (2007/2008)? ○ Yes ○ No

Thank you for your participation!

Page 2 of 2

Figure 4. Sample GEAR UP Parent Survey

City College GEAR UP

Survey of Parent/Guardians of GEAR UP Students: Fall 2008

Please answer the following questions about your child participating in GEAR UP. If you have more than one child in GEAR UP, please turn in a survey for each child. These questions are about his/her experiences in school and your expectations for your child's future.

The purpose of this survey is to help us learn whether the GEAR UP Program in which your child is participating is doing a good job. The information you provide will remain confidential. Also, responses from those returning the survey will be combined before they are presented -- Individual responses will not be reported.

Note if you have already completed and submitted this survey online, please do NOT submit a second survey

Please **completely fill in the bubble** for your response as shown: CORRECT: ● INCORRECT: ⊘ ⊗ ⊘

Please indicate your child's school: ○ School A ○ School B ○ School C ○ School D ○ School E ○ School F

ACADEMIC INVOLVEMENT	High school or less	Some college, no two- or four-year college degree	Two-year college degree	Four-year college degree or higher
1. What is the highest level of education that you would like your child to achieve?	○	○	○	○
2. What is the highest level of education that you think your child will achieve?	○	○	○	○

	Yes	No
3. Have you talked with your child about attending college?	○	○
4. Has anyone from your child's school or GEAR UP ever spoken with you about college entrance requirements?	○	○
5. Do you have enough information about college entrance requirements?	○	○
6. Do you know what your child needs to do to get accepted into college?	○	○
7. Has anyone from your child's school or GEAR UP ever spoken with you about the availability of financial aid to help you pay for college?	○	○
8. Do you have enough information about financial aid to help you pay for college?	○	○

9. How much do you think it costs (tuition and fees only) to attend a four-year public college in your state for one year?

○ $1-$1,900 ○ $1,901-$4,500 ○ $4,501-$7,400 ○ $7,401-$10,000 ○ $10,001-$18,000 ○ More than $18,000

10. Do you think that your child could afford to attend a public four-year college using financial aid, scholarships, and your family's resources?

○ Definitely ○ Probably ○ Not sure ○ Probably not ○ Definitely not

11. If your child participated in any of the following school-sponsored activities/services last year, how effective was it in helping your child with school work or otherwise preparing your child for college?	Very Effective	Effective	Somewhat Effective	Not Effective	N/A – S/he did not participate
a. Tutoring in math	○	○	○	○	○
b. Tutoring in English	○	○	○	○	○
c. Academic counseling/advising	○	○	○	○	○
d. GEAR UP workshop	○	○	○	○	○
e. Visit to a high school or college	○	○	○	○	○
f. Standardized test prep (e.g., state test, ACT, SAT, PLAN)	○	○	○	○	○
g. Work packets/materials to help your child catch up or move ahead in school	○	○	○	○	○
h. Interest/career exploration activities	○	○	○	○	○

Page 1 of 2

Figure 4. Sample GEAR UP Parent Survey (continued)

12. Do you feel GEAR UP activities are helping your child prepare to go to college after high school?

○ Yes ○ No ○ Don't know ○ Child is not going to college ○ Child does not participate in GEAR UP

13. If you participated in any of the following GEAR UP activities/ services last year, how effective was it in helping you assist your child with school work or prepare him/her for college?

	Very Effective	Effective	Somewhat Effective	Not Effective	N/A I did not participate
a. Meeting(s) with GEAR UP staff	○	○	○	○	○
b. GEAR UP parent workshop(s)/informational meeting(s)	○	○	○	○	○
c. Financial assessment of your qualifications for grants, scholarships, and other opportunities	○	○	○	○	○
d. Home visit(s) by GEAR UP staff	○	○	○	○	○
e. Meetings with an academic counselor	○	○	○	○	○

14. Overall, how satisfied have you been with the GEAR UP program this year?

○ Very satisfied ○ Satisfied ○ Neutral ○ Dissatisfied ○ Very dissatisfied

15. What is your highest level of education? (Mark only the highest level achieved.)

○ Less than high school diploma
○ High school diploma or equivalent
○ Certificate program (less than 2-year degree)
○ Associate's degree (2-year degree)
○ Bachelor's degree (4-year degree)
○ Graduate or professional degree (MA, MS, Ph.D., law, MD)

16. What is your child's current grade level?

○ Grade 7 ○ Grade 8 ○ Grade 9 ○ Grade 10 ○ Grade 11 ○ Grade 12

	Male	Female		Male	Female
17a. What is your gender?	○	○	17b. What is your child's gender?	○	○

18. What is your race/ethnicity? (Mark all that apply)

○ Asian ○ African-American/Black ○ Filipino ○ Hispanic or Latino/a ○ Middle Eastern ○ Native American ○ Pacific Islander ○ White/Caucasian ○ Other: _____ (Please write in)

19. What is your primary language? (Mark all that apply)

○ English
○ Spanish
○ Other: _____ (Please write in)

20. Does your child qualify for the free/reduced lunch program? ○ Yes ○ No

21. Was your child enrolled in this same school during the last school year (2007/2008)? ○ Yes ○ No

22. Did your child participate in a GEAR UP program during the last school year (2007/2008)?

○ Yes ○ No ○ Don't Know

Thank you for your participation!

Page 2 of 2

Survey Administration Methods

The primary means of administering surveys are paper-based, online, or via telephone. Each of these methods brings advantages and drawbacks, which are described below.

Depending on the method of administration, paper-based surveys are often the most effective in obtaining higher response rates for GEAR UP student and parent surveys. With respect to surveying students, highest response rates have been obtained when survey forms are administered and collected from groups of GEAR UP students while they are in the classroom. It has been found that students are more apt to complete a survey when it is a whole-class effort with class time devoted to the task. The method also allows all students within a classroom to receive consistent responses when unanticipated questions are raised during survey administration. In contrast, if survey forms are distributed to students (via postal mail or while on school grounds) with instructions to complete and return the form on their own time, rates of returned surveys drop dramatically.

Although paper-based surveys typically result in the highest response rates for parents, many GEAR UP programs remain challenged with low parent survey response rates. However, several strategies have been helpful to increase these rates. For example, sending home the paper-based survey forms to parents within the fall student registration packets and requesting they return their completed forms with those packets of mandatory registration materials dramatically increases response rates of GEAR UP parents. Offering particular incentives also helps to increase response rates. For example, increased response rates have been found for GEAR UP programs offering field trips or Friday pizza parties for students in classes reaching a target percentage of returned completed parent surveys. Providing raffle tickets for a chance to win a prize, such as an iPod, in exchange for completed parent surveys is also an effective approach.

Online survey administration is gaining popularity among GEAR UP programs. Advantages of this form are ease in administration, minimal data entry processes, and increased accuracy in recording open-ended responses (since they are typed in directly by the respondent and do not require hand-entry). However, the response rate of this method is directly dependent on computer internet access. Certainly, if schools have computer labs available to them, the above approach to classroom administration of paper-based surveys may also be adopted for online surveys. However, although access to such technology has increased dramatically over the past few years, many GEAR UP participants live in remote, rural locations and virtually all are within lower socioeconomic regions of the country. Given this, many segments of the GEAR UP population (particularly parents) still do not have regular internet access. Therefore, while this method is expected to gain in effectiveness as access to technology increases, it is currently best used in conjunction with administration of paper-based forms, as an alternate means for respondents to complete the survey.

Telephone surveys offer a third method of obtaining information from GEAR UP participants, typically parents of GEAR UP students. Advantages of such surveys are the immediacy of responses, the ability to ask respondents follow-up questions or get clarification of comments, the ease of respondents to provide information, and the level of detail obtained for responses. However, this method typically requires substantial staff resources in time and effort. Moreover, even when a structured questionnaire is used for the call, often responses deviate from the topic points and information obtained is not consistent across respondents. Additionally, respondents also tend to provide answers that are more socially desirable/acceptable when the survey is conducted via telephone.

Survey Administration Cycle

Although the U.S. DOE only requires administration of a survey to students and parents every two years, to increase the sensitivity of the overall research design, it is advised here that survey data are collected twice annually: at the beginning and end of each school year. This is because GEAR UP programs are typically located in regions with high student mobility rates, resulting in annual turnover rates for GEAR UP cohorts that often exceed 30%. While it is acknowledged that student migration is a constant year-round challenge, the majority of such turnover occurs between school years, during the summer months. Therefore, administration of surveys at the beginning and end of each school year increases the consistency within each survey sample by ensuring that student turnover during the summer does not impact findings. Maximizing consistency in this way is vital for pre- and post-treatment comparative analyses.

Administering a survey at the beginning and end of each academic year also allows programs to adapt the survey focus accordingly. Specifically, surveys administered at the beginning of the school year serve primarily for needs assessment purposes. As such, along with baseline informational items, the form should include questions about services or resources students feel would be beneficial to them. Such information is very useful in determination of program service offerings. Likewise, surveys administered at the end of the school year should include items to capture student perceptions of effectiveness of individual program services (See Figure 3, item 18).

If feasible, for comparison purposes, it is advised that surveys are also administered to the student cohort one grade above the first GEAR UP student cohort, which would thus serve as a retrospective comparison group for same-grade GEAR UP cohorts. As noted on page 4, a retrospective comparison group allows for comparison of students within the same schools during the year prior to program implementation, ensuring the most similar comparison group possible for the program structure. Many GEAR UP programs implement a planning year in conjunction with the schools included within the grant, and this year may provide them with access to survey students within the retrospective comparison group prior to GEAR UP service provision. If such pre-program access is not available, comparison group survey data might still be accessible beginning with the next grade level, such that, for example, comparison group eighth graders are surveyed while seventh grade students are served through GEAR UP, providing comparison data for the following year when GEAR UP students are in eighth grade.

Use of Survey Data for Nationwide GEAR UP Evaluation

Beyond the five mandated survey questions the U.S. DOE requires all GEAR UP programs include on student and parent surveys, additional information may be collected at the discretion of the individual programs. Therefore, aggregation of archival survey data for examination across programs nationwide would be restricted to only those items commonly collected across all programs (i.e., the five APR-mandated items). These items include the following:

1. What is your current grade level?
2. Has anyone from your school or GEAR UP ever spoken with you about college entrance requirements? (Yes/No)
3. Has anyone from your school or GEAR UP ever spoken with you about the availability of financial aid to help you pay for college? (Yes/No)
4. What is the highest level of education that you expect to obtain?
   (H.S. or less/Some college but less than four-year degree/four-year degree or higher)
5. Do you think that you could afford to attend a public four-year college using financial aid, scholarships, and your family's resources?
   (Definitely/Probably/Not sure/Probably not/Definitely not)

Analysis may be conducted across programs using items #2 and #3 as predictors and examining #4 and #5 as outcome variables reflecting student academic expectations, and examining results by grade level (item #1). However, resulting GEAR UP survey data from items about whether staff had spoken with students about college (i.e., items #2 and #3 above) often do not reflect actual student knowledge acquisition. This has been found by comparing results of items #2 and #3 with responses to questions asking if students feel they have enough information about college and financial aid. Typically, disparities exist in results such that substantially fewer respondents report having enough information. This suggests that, while staff may deliver information about college to groups of students in classroom settings, students are attending to, and/or retaining the information to a much lesser extent or they feel a disjoint between information provided and information needed. This finding illuminates the limitations of using such items as predictors of GEAR UP program success.

For this reason, to augment the existing five items, it is advised that a common student survey, such as that presented in Figure 3, is administered to collect data on aspirations, knowledge, and awareness from programs participating in the nationwide study. Although baseline data would not be included, results would likely reveal program effects such as student perceptions of effectiveness of various GEAR UP program services in helping them decide to go to college. Additionally, associations can be identified between student GEAR UP participation levels (based on survey self-report data) and student knowledge about college costs and entrance requirements. See page 24, *Analyses to Examine Program Impact* for further detail on use of GEAR UP survey items in statistical impact analyses.

To ensure a comprehensive nationwide study, it is suggested that supplemental data are collected via online survey from GEAR UP directors and managers. The survey would be administered to the population of GEAR UP director/managers. Topics important for inclusion in this survey include administrator perceptions of effectiveness of the primary individual types of GEAR UP services offered. Other topics include most challenging issues, and the most evident program effects.

Note that the approach adapted in the present guide does not include surveying participating school personnel such as teachers and counselors. Survey administration generally requires extensive resources in terms of GEAR UP program staff time and effort. Surveying students and their parents is often the optimal means of obtaining information about student and parent college awareness and knowledge acquisition and aspirations, given the large numbers of students and parents generally involved in any given GEAR UP program. However, the number of participating teachers and counselors in any given GEAR UP program is substantially fewer than the number of participating students and parents. For this reason, it is often more efficient and effective to conduct a series of structured focus groups of to elicit information from participating school personnel. See Supplementary Qualitative Data: Focus Groups, for further explanation of this process.

Supplementary Qualitative Data: Focus Groups

To ensure a comprehensive program evaluation, in addition to data obtained via surveys, focus groups can provide a wealth of information about the effectiveness of program processes. Such process evaluation includes examination of program functions and operations, including program staffing and organizational structure, clarity of staff roles and objectives, culture and climate, administrative leadership, and service delivery and coverage. Often the value of such program process evaluation is overlooked in the quest to identify program effects through impact analyses. However, if such program processes are not effective, this can directly hinder program impact. Thus, examination of these program features is a vital part of a comprehensive evaluation. Moreover, results often reveal opportunities for formative program improvement.

Effective focus group sessions have included groups of GEAR UP program staff, administrators and coordinators, and school personnel such as teachers and counselors at participating schools. The

sessions typically work best when they contain four to six people at a time who have similar roles in the GEAR UP program. The optimal duration of the focus groups is generally 45 minutes to an hour, depending on the number of participants. On the following page, Table 2 presents sample questions that have been useful in conducting such sessions. Note that separate lists of questions are developed for teachers. Additionally, some questions only pertain to GEAR UP program staff, counselors, or the administrative team.

Table 2. Sample Questions for Structured Focus Group Sessions

Focus Group Questions for Administrative Team, Program Staff, and Counselors
1. What are your roles within GEAR UP?
2. How does your role contribute to the overall goals of the GEAR UP program?
3. What challenges have you faced in implementation?
4. What improvements would you suggest for the future?
5. If you could go back with all you know now, what would you do differently in implementing the grant?
6. What has been the most beneficial aspect of your services?
7. What outcomes have you seen among the students, their parents, the teachers, the school?
8. What would you like to learn from this present evaluation?
Additional Questions for Administrative Team
• Describe your involvement with the counselors and school staff providing services to students as part of the grant.
• Describe your involvement with the participating schools/districts and community-based organizations.
Additional Question for Counselors
• Describe your involvement with the administrators and school staff providing services to students as part of the grant.
• Describe your involvement with the students as part of the grant.
Additional Questions for Program Staff
• Are you involved with other grants at the school as well? How are these coordinated?
Focus Group Questions for Teachers
1. Who do you teach (student characteristics)? What subject areas?
2. How long have you been a teacher? (this school and in general)
3. In which GEAR UP-sponsored professional development programs did you participate?
4. How are the training sessions conducted? (Multiple days? On/off site?)
5. What was the most important benefit of the training to you?
6. What aspect of the training did you feel was the least useful?
7. What impact has the training had on your teaching?
8. What is your level of involvement with the GEAR UP grant?
9. What outcomes, if any, have you seen among your students that you feel result from the training?

Section IV:

Collection and Use of Quantitative Data Relating to Student Readiness Behaviors

The majority of student outcome data is available through schools or school districts in the form of student transcript information (e.g., course enrollments and grades earned) and standardized test data. It is strongly recommended that, whenever possible, GEAR UP programs obtain student transcript data in electronic format. Although some programs continue to develop datasets of student academic information via manual data entry, this approach is both time consuming and costly in terms of staff resources required. Further, manual data entry increases opportunity for errors and thus reduces overall program data integrity. Additionally, as a rule, student transcript records are maintained within school and school district information technology (IT) offices, making hand entry of transcript data redundant. Therefore, it is advised to work with school IT staff to identify and transfer student academic data for evaluation purposes.

Ensuring Data Access and Security

Sometimes school district IT personnel or administrators are reluctant to share such data with GEAR UP evaluators, or they are under the impression they can only do so after obtaining parental consent. This is typically due to a concern that it may be in violation of the Family Educational Rights and Privacy Act (FERPA). However, according to U.S. DOE regulations, FERPA allows schools to disclose those records, without consent, to the following parties (among others):

- Specified officials for audit or evaluation purposes;
- Organizations conducting certain studies for or on behalf of the school;

(Source: http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html)

It is important to note that the evaluator remains responsible for ensuring that the security and confidentiality of such data are maintained. This includes 1) password-protecting files containing school or district student IDs or social security numbers, 2) ensuring files reside on computers requiring staff user IDs and passwords, 3) using a secure ftp format to transfer files containing student records, and 4) encrypting student IDs when emailing such files.

School district or educational institution policies may also result in restricted access to student academic data. In fact, some college or universities require clearance through an Institutional Review Board (IRB) prior to providing access to such data. It is advised here that GEAR UP programs confronted by such policies investigate whether allowances are made for program evaluation purposes.

Coordinating Data Transmission

When working with school or district IT personnel to obtain electronic student transcript data, it is best to provide the IT staff with a data record layout specifying the parameters of the indicators requested. The layout should also identify the specific cohort for which the data are needed (e.g., end-of-term data for 2007/2008 8th graders at Clover High School). Note that, if a retrospective comparison group is required for the evaluation, that cohort must also be identified within the layout. Moreover, since the data files are typically needed on a routine basis for ongoing evaluation, a data transmission calendar is also helpful to ensure IT staff are aware of future data needs. Table 3 presents an example of such a data record layout containing the transcript variables within the set of Optimal GEAR UP Indicators. Note that, often the optimal schedule for GEAR UP student data submissions includes 1) fall enrollment and demographic data for tracking students in the new academic year, and 2) January/February

submission of current course enrollments and the prior school year end-of-term course outcomes (to fulfill APR data reporting requirements).

It is advisable that the school or district IT staff are alerted as soon as possible about the data requirements and are given a reasonable amount of time to meet the request for data. Often school IT staff are very busy fulfilling an array of technical requests and appreciate as much advance notice as possible.

Table 3. Sample School Database Record Layout for 8th Grade GEAR UP Student Cohort

Field Name	Definition	Field Properties
SchID	School Identifier	Alpha or Numeric
Grade	Student Grade Level	Numeric
FName	Student First Name	Alpha
LName	Student Last Name	Alpha
Addss	Student Address	Alpha
City	Student City	Alpha
State	Student State	Alpha
Zip	Student Zip Code	Numeric
Phone	Student Home Phone # (include area code)	Numeric
StID	Student ID	Numeric
Gender	Student Gender	Alpha or Numeric
Eth	Student Ethnicity	Alpha or Numeric
DOB	Student Date of Birth	Numeric
LEP	Limited English Proficiency (LEP) flag	Numeric
IEP	Individualized Education Plan (IEP) flag	Numeric
S-DATE	Current school year start date	Numeric (MMDDYYYY)
TSQ	Trimester/Semester/Quarter (T1 or S1 or Q1)	Alpha Numeric
Class	Class Description	Alpha
ClassN	Class Number	Alpha or Numeric
Classlvl	Class Level (Advanced, AP, IB, College, etc.)	Alpha or Numeric
TID	Instructor ID	Alpha or Numeric
TFName	Teacher First Name	Alpha
TLName	Teacher Last Name	Alpha
Units	Units Earned	Numeric
GradeAch	Grade Achieved	Alpha
Abs_ex	Excused Absence*	Numeric
Abs_unex	Unexcused Absence*	Numeric
ST-M	Standardized Test Score - Math	Numeric
ST-L	Standardized Test Score – Language Arts	Numeric
ST-S	Standardized Test Score - Science	Numeric
EXPL-E	EXPLORE English Test Score	Numeric
EXPL-M	EXPLORE Math Test Score	Numeric
EXPL-R	EXPLORE Reading Test Score	Numeric
EXPL-S	EXPLORE Science Test Score	Numeric
Term GPA	Term GPA (Semester/Trimester/Quarter GPA)	Numeric
GradeLY	Student Grade Level Last School Year	Numeric

* Please provide per-term.

Notes.

1. Records within the provided datasets must be at the course-level for each student (e.g., students enrolled in five courses in a term will have five records in the dataset for that term. This will result in some repeated variables (such as demographics) in each record of individual students.
2. A minimum of two transmissions are required, annually, according to the following transmission schedule:

   1st transmission: Approximately two weeks after the beginning of the new school year (~ 2nd week of September).

   2nd transmission: At completion of the school term ending just prior to March 31st (In preparation of the reporting needs required by the Annual Performance Report, due April 15th.

Obtaining Educational Planning and Assessment System Data from ACT, Inc.

Sometimes it is necessary for GEAR UP practitioners to obtain Educational Planning and Assessment System (EPAS) data from ACT, Inc. directly. This is generally initiated by the GEAR UP program director or manager, who notifies the ACT representative of the need for the data. Typically the ACT, Inc. contact will transfer the GEAR UP student outcomes records data file to a CD and sent via postal mail to the GEAR UP program directly.

Analysis Across Programs

As noted, this current guide advises that a nationwide evaluation of GEAR UP programs begin with preliminary analyses to identify any systematic differences existing between those programs participating in the study and those who do not participate. This may be done using aggregate school-level and/or program data obtained from as many GEAR UP programs as possible. The primary concern is that, GEAR UP programs with the capacity to obtain student-level data for evaluation and/or the interest in participating in a nationwide evaluation may be more apt to have more resources available to that program and/or more technically sophisticated staffing. These inherent advantages seen for participating programs may influence the overall program findings. Specifically, while GEAR UP students participating in the study may show higher academic performance in comparison to non-GEAR UP students, the performance of students within GEAR UP programs opting out of the study may not differ as notably in comparison with non-GEAR UP students.

It is recommended that two approaches are employed to collect aggregate data from non-participating GEAR UP programs. First, a request may be issued to GEAR UP programs nationwide, requesting submission of the most recent APR reports for inclusion in the national evaluation. Such data received from non-participating GEAR UP programs may be compared against those programs participating in the study. Data examined for differences would include program-level indicators such as service coverage, by service type (located in APR Section IV) and student academic proficiency levels (within APR Section V). Note that if the DOE consents to providing GEAR UP program APR data directly to the evaluator, this would be preferred, as it is least intrusive and time-consuming with respect to individual GEAR UP program personnel.

As an additional evaluation for differences between participating and non-participating GEAR UP programs, school-level test proficiency rates can be collected from school district websites which typically post ongoing achievement rates for state standardized tests. Differences in proficiency rates between participating and non-participating programs may then be identified and further examined.

Analyses to Examine Program Impact

The Comparative Change Model

When examining differences between GEAR UP and non-GEAR UP students, as noted in Section I, a comparative change model of analysis is recommended. Such analyses may be used with continuous or categorical variables. This method involves subtracting baseline predictor scores from post-treatment scores to derive *change*, or *difference* scores. For example, an analysis may examine differences between GEAR UP and non-GEAR UP student achievement scores on a standardized test administered at the end of 8th grade. If 8th grade test scores are used on their own as the outcome measure, statistically significant differences may be found between the average scores for the two groups; however, these differences may simply be due to preexisting differences between the two cohorts of students. However, using an outcome metric that is based on the difference between 7th grade scores and 8th

grade scores accounts for any pre-existing disparities in average scores for the two groups. In this manner, any systematic differences existing between GEAR UP and comparison group scores prior to exposure to GEAR UP services will not interfere with examination of change. However, it should be noted that any extreme pre-existing disparities may make it difficult to judge the meaning of such change. Therefore, in addition to comparison of change scores as discussed, an Analysis of Covariance (ANCOVA) can be implemented to determine whether there is a significant difference between the comparison group and the GEAR UP students, once the baseline scores are accounted for as a covariate.

Note that, in examination of GEAR UP impact, it is also possible to employ linear regression, or hierarchical logistic regression models. However, such analyses are best used when multiple predictors are involved. Because the goal of GEAR UP evaluation is to determine differences, or associations due to participation or non-participation in GEAR UP, inferential analyses offering the most sensitivity and power include t-test and chi-square analyses (to detect differences between GEAR UP and non-GEAR UP outcomes) or correlation analyses (to reveal associations between GEAR UP participation and participant outcomes). Moreover, in determining the appropriate analysis, one must also consider that it is typically more difficult to convey the results of more complicated analyses to GEAR UP program stakeholders; the more simple inferential analyses, such as correlation and chi-square, are generally more intuitive for those who do not have a solid background in statistical analysis.

Some outcome variables are continuous in nature, such as the above example of standardized test scores. In these situations, a t-test may be used to statistically examine for differences between GEAR UP and non-GEAR UP student cohorts. The analysis would answer the question: *Is there a statistically significant difference in average change scores (e.g., derived from differences in 7th and 8th grade scores) for GEAR UP and non-GEAR UP students?* Table 4 below presents the statistics that would result from such an analysis. This includes the average change in test scores for each group, the standard deviation for each group, the t value, and the significance level attained.

Table 4. T-Test Results of GEAR UP and Comparison Group Change in GPA

	# of Students	Ave. Chg in GPA	Std. Dev.	t	Sig.
GEAR UP	112	0.49	0.05	-4.398	$\leq$ 0.0001
Comparison Group	360	0.44	0.03		
TOTAL	**507**				

Note. Sample data are presented for presentation purposes only.

Categorical or ordinal outcome variables may also be incorporated into a comparative change analysis. In these cases, non-parametric analyses, such as chi-square, may be used to statistically examine differences between GEAR UP and non-GEAR UP cohorts. For example, 9th and 10th grade student proficiency levels (e.g., below basic, basic, proficient, advanced) may be examined to determine those students moving up a level in 10th grade, those performing worse in 10th grade and those remaining at the same level both years. In this manner, a three-tier ordinal variable would result and the change in proportions of students in each category may be examined statistically.

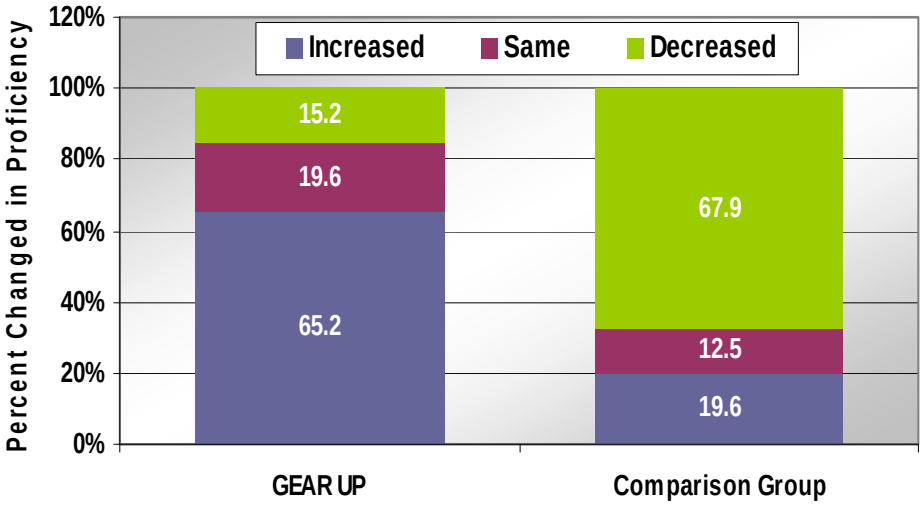
Table 5 presents an example of results of a chi-square analysis used to examine differences between GEAR UP and non-GEAR UP students with respect to change in proficiency level. Namely, comparisons were made between the percentages of students increasing or decreasing from one level to another, and those remaining at the same level.

Table 5. Chi-Square Results of GEAR UP and Comparison Group Change in Test Proficiency Level

	Decrease in Proficiency		No Change in Proficiency		Increase in Proficiency		Total		Chi Sq.	Sig.
	#	%	#	%	#	%	#	%		
GEAR UP	17	15.2	22	19.6	73	65.2	112	100.0	106.9	< 0.0001
Comparison Group	266	67.9	49	12.5	77	19.6	392	100.0		
TOTAL	**283**		**71**		**150**		**504**			

Note. Sample data are presented for presentation purposes only.

Figure 5. GEAR UP and Comparison Group Change in Test Proficiency Level



Note. Sample data are presented for presentation purposes only.

Of course, not all GEAR UP indicators are appropriate for incorporation into the comparative change model of analysis. For example, to examine program impact, chi-square analysis such as that described above may be employed to compare the percent of GEAR UP and non-GEAR UP 10th grade students successfully passing Algebra I. However, for such cases, given the nature of the variable, change scores cannot be derived for students, and a dichotomous pass/no pass variable must be used as the outcome variable. Since baseline measures of course competency would not be included in the analyses, it is not known if the students differed in performance prior to receiving GEAR UP program services. However, with that caveat, such analyses are still useful in providing insight into program impact on student course performance.

Additionally, as discussed earlier in this guide, covariates may be found when examining baseline data, indicating existing student-level differences. Likewise, program-level effects of self-selection into the study may occur. It is useful when such findings occur to incorporate them into a multi-level model of GEAR UP program outcomes to help provide an understanding of program-related effects.

Determining Associations between Variables

It is often necessary to examine associations between two variables within a group of GEAR UP students. In these cases, correlation analyses are conducted. Generally speaking, it is possible to examine differences in correlations between treatment and comparison groups. However, typically the associations examined for GEAR UP students use program participation indicators in such analyses, and, since non-GEAR UP students do not participate in such services, it is not possible to include comparison

groups in these cases. Therefore, within GEAR UP evaluation, correlation analyses are typically conducted to examine relationships between variables within groups of GEAR UP students.

Correlation analysis can be used with continuous or categorical variables. A Pearson product moment correlation coefficient may be used to examine relations between two continuous variables, such as hours in GEAR UP tutoring services and ACT PLAN test outcomes. To determine associations between a continuous and a dichotomous variable, such as months in GEAR UP and student indication of having enough knowledge of college requirements, a point biserial correlation coefficient may be used.

Dosage Effects

When positive relationships are found between GEAR UP service participation and academic achievement or improved attitude towards and/or knowledge of college, it is important to examine these relationships in more detail. Most types of effective GEAR UP services do not reflect a continuous linear relationship between the amount of service hours provided and the program impact on participants. For example, two campus visits may increase student aspirations to a moderate degree, but twenty would not increase aspirations another tenfold. Thus, it is helpful to determine the optimal number of GEAR UP service hours required to make a meaningful, positive impact on student academic achievement. Such analyses are conducted to determine the effects of various levels of service *dosage.* Examination of dosage is often done by using both correlation and chi-square techniques. Correlation analyses may be run first to examine any existing relationship between particular GEAR UP services and expected program outcomes. Next, student participation hours may be organized into groupings of high, moderate, low, and no participation levels. Associations between this new categorical variable and the outcomes of interest can be analyzed via chi-square analysis to examine for differences among the groupings.

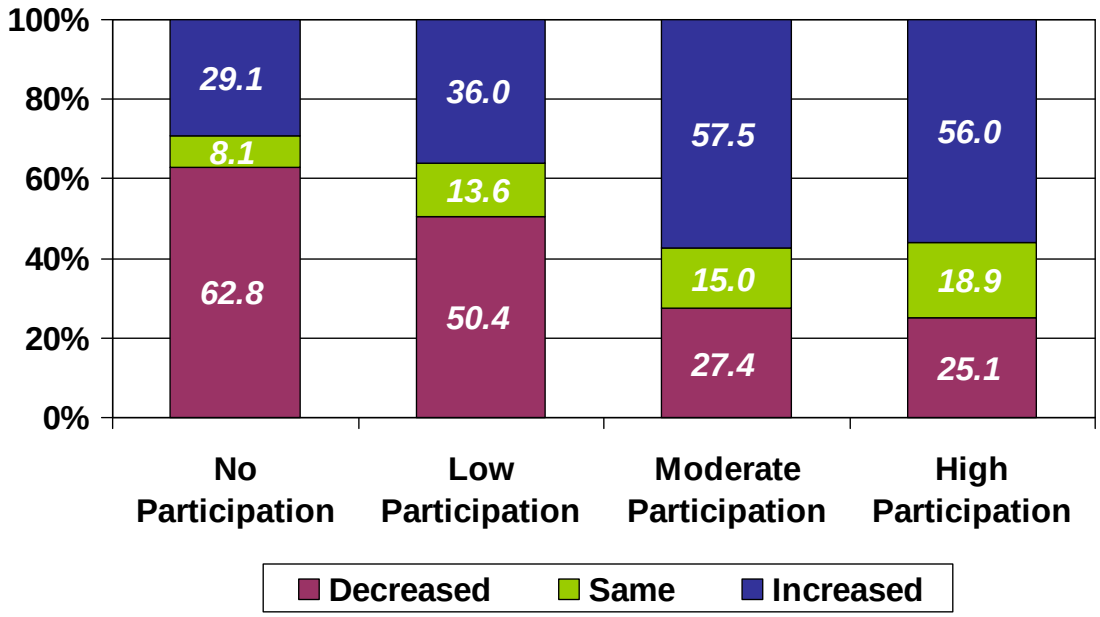
Table 6 below presents results for a preliminary examination for dosage effects. Note that the correlation analysis reveals a statistically significant positive relationship between number of hours in math tutoring and change in student standardized math test scores. Next, follow-up chi-square analyses would be conducted to examine differences in the math test score differences, based on student level of participation in math tutoring services (i.e., high, moderate, low, and no participation). In this example, follow-up analyses reveal that moderate and high participation in math tutoring produced similar increases in test scores.

Table 6. Correlation Results for Number of Hours of Math Tutoring and Change in Math Test Score

# of GU Students	Ave. # of Hrs of Math Tutoring	Ave. Change in Math Score	Correlation	Sig.
112	7.8	32	0.738	< 0.0001

Note. Sample data are presented for presentation purposes only.

Figure 6. Math Tutoring Participation and Difference in Math Test Scores

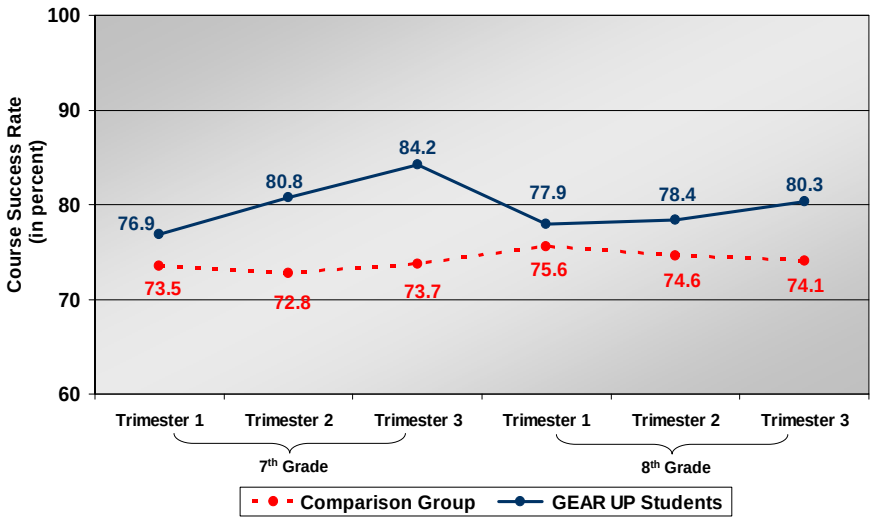


Note. Sample data are presented for presentation purposes only.

Evaluation of Longitudinal Data

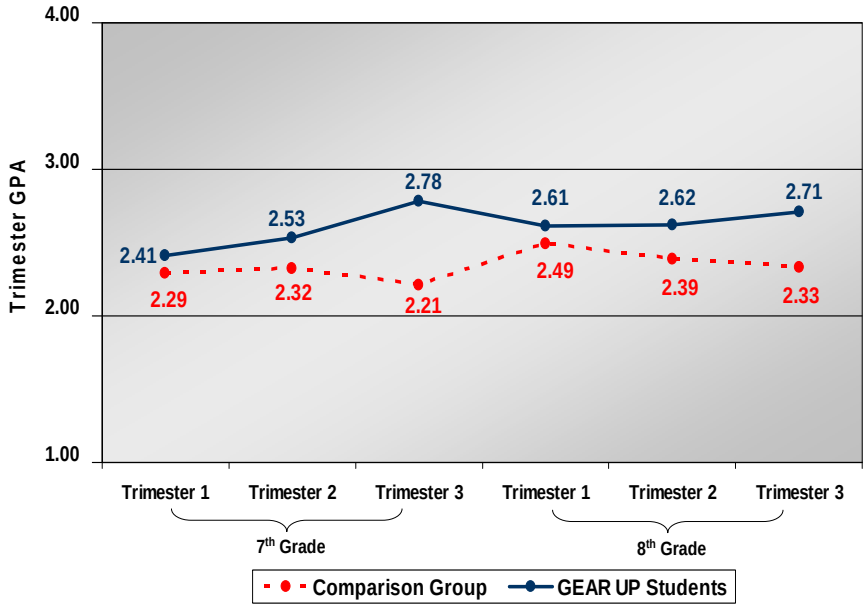
As GEAR UP programs progress, data across points in time may be evaluated to identify trends or differences in academic achievement between students in GEAR UP and comparison groups. Figures 7 and 8 illustrate how such data may be presented. The first figure compares average success rates for GEAR UP and comparison groups across six trimesters. The second figure presents the average trimester GPA for these same groups of students. Presenting academic outcome data in this manner is helpful in revealing trends or identifying opportunities for program improvement. Follow-up statistics, such as t-test analysis, may be employed to determine if differences are statistically significant.

Figure 7. GEAR UP and Comparison Group Average Trimester Success Rates



Note. Sample data are presented for presentation purposes only.

Figure 8. GEAR UP and Comparison Group Average Trimester Grade Point Averages

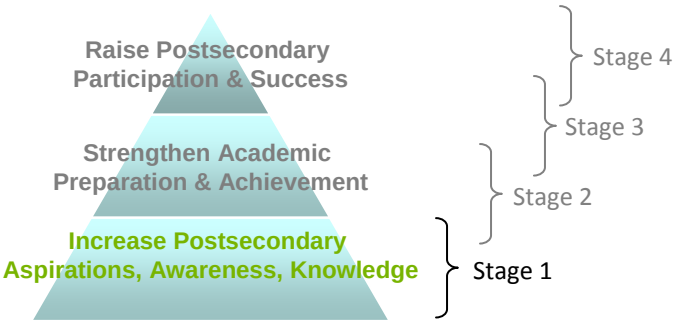


Note. Sample data are presented for presentation purposes only.

Section V:

Stage 1: Sixth through Eighth Grade

Aspirations, Awareness, and Knowledge Acquisition



The majority of GEAR UP programs begin by serving students in the 7th grade, although some start with students in lower grades. During this beginning phase of the program, it is most critical to inspire and educate students about the benefits of obtaining a college degree. At this early stage, it is also important to inform parents of GEAR UP students of the value of college and available options for funding college for their children. By devoting resources to increasing academic expectations, and motivating students to strive for a college degree, a college-going culture can be established within participating GEAR UP schools, thus laying the groundwork necessary to strengthen academic achievement and college preparation.

Table 7, below, presents the 42 Optimal GEAR UP Indicators for this first stage.

Table 7. Optimal GEAR UP Program Evaluation Indicators: Stage 1 (6th through 8th Grade)

Student & Parent Survey Data		**Student Academic & Participation Data**		**Aggregate Data**
Aspirations, Knowledge & Awareness Acquisition	Control & Predictor Variables	Student Readiness Behaviors	Control & Predictor Variables	Control Variables
Student 1. Student education expectations 2. Student education aspirations 3. Student perceptions of college affordability 4. Student knowledge of college requirements 5. GU impact on post-HS intentions 6. Student knowledge of college costs **Parent** 7. Parent knowledge of college requirements 8. Parent perceptions of college affordability 9. Parent educ expectations of students 10. Parent educ aspirations of students 11. Parent knowledge of college costs	**Student & Parent** 12. School 13. Gender 14. Race/ethn 15. Time in GU prog 16. Yrs enrolled in GU school 17. Level of GU participation 18. Eligible for free/red lunch program 19. Parent highest education level attained	20. At/above grade level lang arts on state standardized test 21. At/above grade level math on state standardized test 22. At/above grade level science on state standardized test 23. EXPLORE English 24. EXPLORE math score 25. EXPLORE reading score 26. EXPLORE science score	27. Student ID 28. School 29. Grade 30. Gender 31. Race/ethnicity 32. LEP[1] Status 33. IEP[2] Status 34. Free/Reduced Lunch Status 35. # hrs GU student service by service type 36. # hrs GU parent services by service type	**School-level** Variables 20-26 & 29-36, in aggregate 37. # hrs GU teacher svcs by type 38. Cost of student svcs by type 39. Cost of parent svcs by type 40. Cost of teacher svcs by type **Program -level** 41. # GU schools 42. # of GU teachers by grade

[1] LEP: Limited English Proficiency.
[2] IEP: Individualized Education Plan.

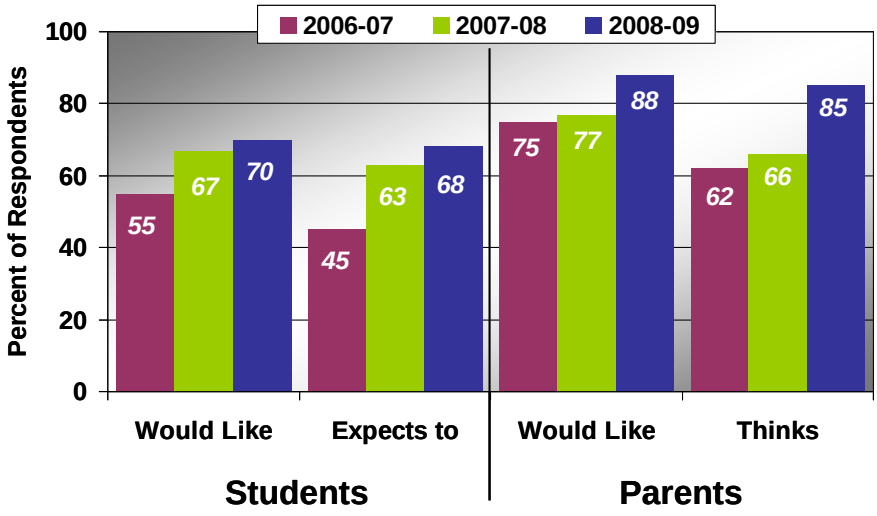
As discussed in Section I, during this first stage of the GEAR UP program, it is often premature to expect notable increases in academic achievement. While meaningful increases in academic achievement is a core long-term objective of all GEAR UP programs, often meaningful improvement in academic performance requires delivery of targeted program services over an extended timeframe. In light of this, program progress during Stage 1 is usually best measured by the more intermediate program objectives of college aspirations, knowledge, and awareness acquisition.

GEAR UP services and activities aimed at increasing aspirations, awareness, and knowledge include informational meetings, parent workshops, campus tours, motivational speakers, and academic counseling. To determine the effectiveness of such services, participant opinions are often collected in the form of survey and/or focus group data. The following sections detail methods and strategies found most effective to collect such data.

As noted, Figures 3 and 4 on pages 14-17 present sample student and parent survey forms that may be used to collect the above indicators. This includes variables #1 through #11, and control and predictor data elements #12 through #19.

Note that survey item 2 regarding student expectations is the mandated APR item, however the prior item addresses student *aspirations*. Examining disparities between these two indicators is often helpful in assessing student progress in aligning their desire to attend college with expectations that they can achieve this goal. Figure 9, below, present an example of how resulting data may be used in program gap analysis.

Figure 9. Expectations and Aspirations for a Four-Year Degree or Higher



Note. Sample data are presented for presentation purposes only.

Likewise, survey items 7 and 9 are the APR-mandated item to assess levels of service coverage (how many students report that school or program staff has spoken with them about college entrance requirements and financial aid). Additionally, examining this item against items 8 and 10, which ask if the respondent has enough information about college entrance requirements and financial aid, illuminates gaps between what the program is providing and how effective those services are at educating students about college costs and requirements. Further, item 12, which asks respondents to identify the correct range of college costs is effective in determining how many GEAR UP participants actually know how much college costs – not just their perceptions of what they know. This item is often

surprisingly low – underscoring the need to attend to what participants are learning and not just what staff are presenting.

Items 16 and 17 have also been valuable in eliciting information addressing indicator 5 from the set of Stage 1 Optimal Indicators regarding the effectiveness of GEAR UP services in student decisions to attend college.

Because a primary focus during Stage 1 is increased college aspirations, knowledge, and awareness, it is useful to examine crosstabulated results of survey items such as 1, 2, 10, and 12, which capture this information, by item 13 (frequency of program participation). Tables 8 and 9 below provide examples of such crosstabs.

Table 8. Crosstab: Student College Aspirations, by Frequency of Program Participation

Student College Aspirations	Frequency of Participation in GEAR UP								Total	
	Never		Rarely (define)		Sometimes (define)		Frequently (define)			
	#	%	#	%	#	%	#	%	#	%
HS or less	89	*73.0*	11	*9.0*	18	*14.8*	4	*3.3*	122	*100.0*
Some college	43	*30.7*	19	*13.6*	42	*30.0*	36	*25.7*	140	*100.0*
2-year degree	2	*0.5*	9	*2.3*	75	*19.1*	306	*78.1*	392	*100.0*
4-year degree	67	*4.4*	97	*6.4*	321	*21.3*	1025	*67.9*	1510	*100.0*
Total	**201**	***9.3***	**136**	***6.3***	**456**	***21.1***	**1371**	***63.4***	**2164**	***100.0***

Note. Sample data are presented for presentation purposes only.

Table 9. Crosstab: GU Assisted with Decision to Attend College, by Frequency of Program Participation

GU Assisted with Decision to Attend College	Frequency of Participation in GEAR UP								Total	
	Never		Rarely (define)		Sometimes (define)		Frequently (define)			
	#	%	#	%	#	%	#	%	#	%
Yes	68	*8.0*	72	*8.5*	325	*38.1*	387	*45.4*	852	*100.0*
No, already planning on attending	83	*7.1*	50	*4.3*	101	*8.7*	927	*79.8*	1161	*100.0*
No, still don't want to attend	42	*38.9*	8	*7.4*	23	*21.3*	35	*32.4*	108	*100.0*
Total	**193**	***9.1***	**130**	***6.1***	**449**	***21.2***	**1349**	***63.6***	**2121**	***100.0***

Note. Sample data are presented for presentation purposes only.

Additionally, examining *level of participation in GEAR UP activities* can provide insight into dosage effects when analyzed via chi square analyses in conjunction with perceptions of program effects.

The sample survey also contains demographic variables such as *race/ethnicity*, *gender, parent level of education,* and *free/reduced lunch program eligibility.* These correspond to the Optimal GEAR UP Indicator Control and Predictor Variables #13 - #18 for students and #13 - #19 for parents. (Note that, since free/reduced lunch program eligibility directly corresponds to parent reported household income, this indicator is useful as a proxy measure to identify economically disadvantaged students.) Such information allows for disaggregation of data to examine for disparities based on demographic or economic factors.

Stage 1 Evaluation of Student Readiness Behaviors

Although statistically significant increases in academic achievement often do not occur during the initial years of GEAR UP, it is vital that such data are collected while students are in Stage 1, both for use as baseline measures as well as evaluation for evidence of academic improvement. Moreover, because student IDs are included within the dataset of Student Academic and Participation Data, more sensitive, student-level analyses may be conducted to examine for effects of student and parent service participation upon academic indicators. Additionally, the inclusion of student IDs allows for longitudinal analysis of trends across school terms and academic years. The specific academic achievement indicators useful for Stage 1 evaluation are discussed below.

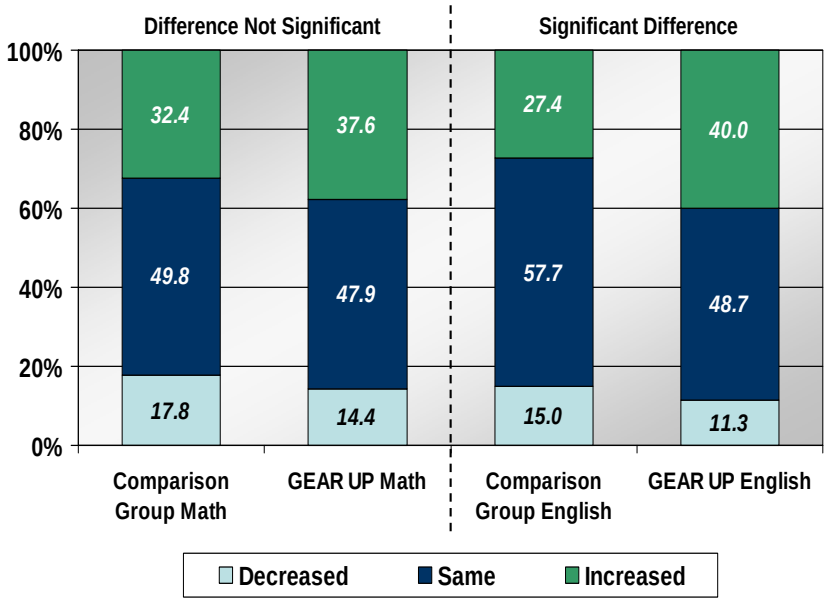
State Standardized Test Scores

The first three student readiness behavior indicators pertain to statewide standardized test outcomes. Note that, although three elements are listed corresponding to language arts, math, and science, the number and content areas of particular state standardized subtests may vary.

Maintaining the standardized scores at the subtest level allows for examination of individual GEAR UP services corresponding to the particular academic content area. For example, the number of hours of student participation in math tutoring can be correlated with changes in statewide standardized math subtest scores. As noted earlier, this analysis would provide preliminary information about existence of dosage effects. If, in fact, a statistically significant correlation is found, further analyses may be used to examine optimal dosage by categorizing the participation data into groups of high, moderate, low, and no participation.

Typically, school and district IT datasets include test outcomes as continuous variables. This allows for analyses such as t-tests and correlations to identify statistically significant differences between GEAR UP and non-GEAR UP students. However, often GEAR UP program goals (as well as school administrator objectives) focus on increasing the percent of students qualifying as *proficient* on statewide tests. Thus, additional, follow-up analyses are recommended to examine for changes in proficiency status (e.g., far below basic, below basic, basic, proficient, and advanced), which may be a more meaningful measure of success for program administrators. Moreover, some schools and districts only maintain statewide test outcomes for students in terms of their proficiency level; consequently, such analyses may be the only means of examining changes in standardized test outcomes. Figure 10 presents an example of how such analyses may be presented. Note that the following analysis employed change scores, based on whether students increased, decreased, or remained at the same proficiency level from 7th grad to 8th grade.

Figure 10. Changes in State Standardized Test Scores: 7th Grade to 8th Grade



Note. Sample data are presented for presentation purposes only.

EXPLORE Scores

The EXPLORE assessment is comprised of four multiple-choice tests: English, math, reading, and science. The tests measure student curriculum-related knowledge and cognitive skills that are deemed important for future education and careers. Scores for each test range from 1 to 25. These indicators (listed as Optimal Indicators 23 through 26) have proven to be very useful GEAR UP outcome measures. As with the standardized test scores noted above, the original scores can serve as continuous outcome variables in correlation and t-test analyses. Additionally, as with all ACT, Inc. EPAS assessments, the EXPLORE test also includes *College Readiness Benchmarks*. The two scores identified for each test reflect: 1) a 50% chance of obtaining a *B*, and 2) a 75% chance of obtaining a *C* in corresponding credit-bearing college courses. These benchmark scores are currently as follows: English: 13/14, math: 17/18, reading: 15/16, and science: 20/20). Note, these scores may change over time and it is advised that actual benchmarks be verified with ACT, Inc. contact personnel.

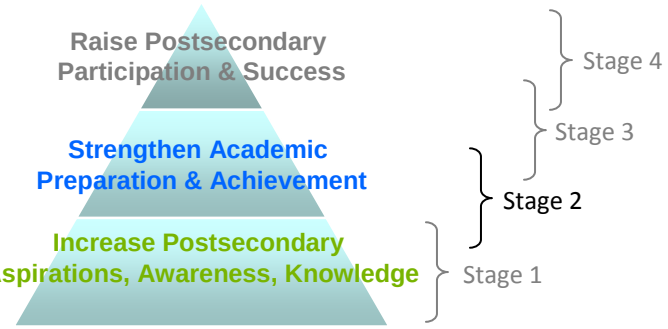
Control and Predictor Variables

The primary objective of GEAR UP is to increase academic achievement and college-going rates for at-risk student populations. As such, it is critical that analyses include disaggregation of data by demographic and socioeconomic indicators to examine for disparate impact. Optimal Indicator variables 30 through 34 include student gender, race/ethnicity, limited English Proficiency (LEP), Individualized Education Program (IEP), and Free/Reduced Lunch Status. In addition to demographic and economic indicators, flags for students identified as IEP and LEP within their school systems are included because the DOE considers these students as particular targeted groups for GEAR UP services. Therefore, it is important to examine outcomes data by these indicators to ensure academic progress is achieved proportionately for LEP and IEP student populations.

Section VI:

Stage 2: Ninth and Tenth Grade

Aspirations, Knowledge Acquisition, Academic Enrollment, and Preparation



As GEAR UP programs mature beyond the first few years, the focus shifts to a greater extent on strengthening academic preparation and achievement. However, increasing aspirations, awareness, and knowledge about college options continues to be important at this stage.

The following table presents optimal data indicators for evaluation of GEAR UP during this period.

Table 10. Optimal GEAR UP Program Evaluation Indicators: Stage 2 (9th through 10th Grade)

Student & Parent Survey Data		**Student Academic & Participation Data**		**Aggregate Data**
Aspirations, Knowledge & Awareness Acquisition	Control & Predictor Variables	Student Readiness Behaviors	Control & Predictor Variables	Control Variables
Student 1. Student education expectations 2. Student education aspirations 3. Student perceptions of college affordability 4. Student knowledge of college requirements 5. GU impact on post-HS intentions 6. Student knowledge of college costs **Parent** 7. Parent knowledge of college requirements 8. Parent perceptions of college affordability 9. Parent educ expectations of students 10. Parent educ aspirations of students 11. Parent knowledge of college costs	**Student & Parent** 12. School 13. Gender 14. Race/ethn 15. Time in GU prog 16. Yrs enrolled in GU school 17. Level of GU participation 18. Parent highest education level attained 19. Eligible for free/red lunch program	20. At/above grade level lang arts on state standardized test 21. At/above grade level math on state standardized test 22. At/above grade level science on state standardized test 23. EXPLORE English 24. EXPLORE math score 25. EXPLORE reading score 26. EXPLORE science score 43. Completed Algebra I 44. Completed Algebra II 45. Completed Calculus 46. Completed Chemistry 47. Completed Physics 48. Enrolled in AP math 49. Enrolled in AP English 50. Enrolled in AP Science 51. Completed AP math 52. Completed AP English 53. Completed AP Science 54. PLAN English score 55. PLAN math score 56. PLAN reading score 57. PLAN science score 58. PSAT reading score 59. PSAT math score 60. PSAT writing score	27. Student ID 28. School 29. Grade 30. Gender 31. Race/ethnicity 32. LEP[1] Status 33. IEP[2] Status 34. Free/Reduced Lunch Status 35. # hrs GU student service by service type 36. # hrs GU parent services by service type	**School-level** Variables 20-26 & 29-36, in aggregate 37. # hrs GU teacher svcs by type 38. Cost of student svcs by type 39. Cost of parent svcs by type 40. Cost of teacher svcs by type **Program -level** 41. # GU schools 42. # of GU teachers by grade

[1] LEP: Limited English Proficiency.
[2] IEP: Individualized Education Plan.

This stage represents a particularly vulnerable period for students. Ninth graders typically are coping with the major transition from middle school to high school. Nationwide, this is seen as a volatile period during which academic declines are evidenced. Subsequently, as students progress to 10th grade, most become of age for legal employment. This results in conflicting demands for many GEAR UP students, who understand the value of a college education but are also confronted with family economic needs compelling them to seek employment. Given these circumstances, students at this stage are most vulnerable to decreased academic performance and worse, failure to complete high school.

Stage 2 Evaluation of Aspirations, Knowledge, and Awareness Acquisition

While academic indicators take on more significance during this stage, metrics of college aspirations, knowledge, and awareness remain important to ensure students and parents are acquiring a solid understanding of the value of college and the means to gain entrance and success in college.

Survey results during Stage 2 often show genuine increases in student and parent knowledge and awareness about college costs and entrance requirements. Likewise, the GEAR UP impact on post high school intentions is a valuable metric to assess at this time. Additionally, it is important to continue monitoring disparities between student aspirations and expectations, which should begin to reduce during this stage as student expectations rise to align with their academic goals.

Stage 2 Evaluation of Student Readiness Behaviors

Although the importance of statewide and national test assessments increase during this time, enrollment in, and successful completion of particular high school courses also become prominent indicators of academic achievement. These metrics are addressed in further detail below.

Course Enrollment and Success

Optimal GEAR UP Indicators 43 through 53 pertain to student course enrollment and successful completion. (Note that successful completion is defined as students earning marks of A, B, C, or Pass for a given course.) As noted in Section I, actual course grade information will not be maintained for students. Rather, to increase data integrity and ensure consistency across schools, districts, and GEAR UP programs, a dichotomous variable will be retained for each of these variables, indicating whether or not the student had enrolled in and successfully passed the given course.

In addition to offering insight into program impact, the course enrollment and outcome indicators are useful in formative GEAR UP program assessment. For example, comparison of the percent of students enrolling in and successfully completing a particular AP course may reveal that, while a growing percentage of students are enrolling in the AP course, the percent of students successfully completing the course has reduced over time. This may be due to AP course enrollment by GEAR UP students who are not as prepared for the course, and thus pass at a lower rate, in comparison with students who would have taken an AP course even without GEAR UP. Such higher enrollment in AP courses by GEAR UP students may be related to GEAR UP services or information provided, indicating that GEAR UP staff may want to supplement an emphasis on enrollment in such challenging courses with additional support for the students. Alternatively the lower pass rate may be caused by recent changes to the AP course curricula, or student academic preparation services. Thus, further examination of this finding should include consideration of such factors.

The PLAN Assessment

As with the EXPLORE assessment, the four PLAN tests also measure student academic development. However, the PLAN assessment is focused on academic skills students are expected to attain later in their educational experience (i.e., 10th grade). Each PLAN test ranges in scores from 1 to 32.

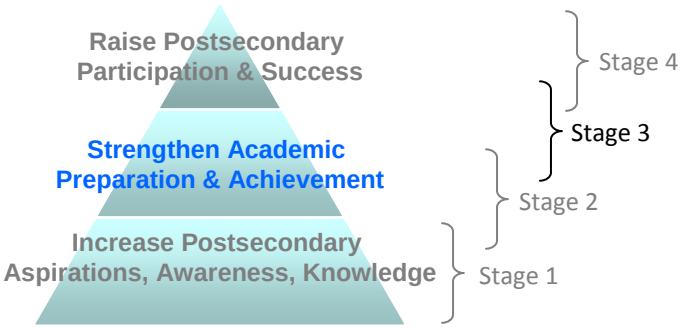
Preliminary Scholastic Aptitude Test (PSAT)

In preparation for the College Board Scholastic Aptitude Test (SAT), students in some regions of the country have an opportunity to take the Preliminary SAT (PSAT), which is typically offered in 10th and 11th grade. The PSAT is comprised of three sections that correspond to reading, math, and writing content areas. Scores for each section range from 20 to 80, with a total score of 240.

As with the EPAS assessments, for GEAR UP evaluation purposes, both at the program level as well as nationwide, the PSAT offers a standardized outcome measure that serves to gauge program progress with respect to core academic areas.

Stage 3: Eleventh and Twelfth Grade

Academic Enrollment, Preparation, and Achievement



During the later years of GEAR UP (i.e., years 5 and 6), an increasing emphasis is placed on strengthening academic preparation and achievement. Heightened student aspirations, awareness, and knowledge about college options provide the foundation on which such student readiness behaviors are developed and strengthened. Table 11, below, presents the optimal evaluation indicators corresponding to this GEAR UP program stage.

Table 11. Optimal GEAR UP Program Evaluation Indicators: Stage 3 (11th through 12th Grade)

Student & Parent Survey Data		Student Academic & Participation Data		Aggregate Data
Aspirations, Knowledge & Awareness Acquisition	Control & Predictor Variables	Student Readiness Behaviors	Control & Predictor Variables	Control Variables
Student 1. Student education expectations 2. Student education aspirations 3. Student perceptions of college affordability 4. Student knowledge of college requirements 5. GU impact on post-HS intentions 6. Student knowledge of college costs **Parent** 7. Parent knowledge of college requirements 8. Parent perceptions of college affordability 9. Parent educ expectations of students 10. Parent educ aspirations of students 11. Parent knowledge of college costs	**Student & Parent** 12. School 13. Gender 14. Race/ethn 15. Time in GU prog 16. Yrs enrolled in GU school 17. Level of GU participation 18. Parent highest education level attained 19. Eligible for free/red lunch program	20. At/above grade level lang arts on state standardized test 21. At/above grade level math on state standardized test 22. At/above grade level science on state standardized test 43. Completed Algebra I 44. Completed Algebra II 45. Completed Calculus 46. Completed Chemistry 47. Completed Physics 48. Enrolled in AP math 49. Enrolled in AP English 50. Enrolled in AP Science 51. Completed AP math 52. Completed AP English 53. Completed AP Science 58. PSAT reading score 59. PSAT math score 60. PSAT writing score 61. ACT English score 62. ACT math score 63. ACT reading score 64. ACT science 65. SAT reading score 66. SAT math score 67. SAT writing score 68. Enrollment in College	27. Student ID 28. School 29. Grade 30. Gender 31. Race/ethnicity 32. LEP[1] Status 33. IEP[2] Status 34. Free/Reduced Lunch Status 35. # hrs GU student service by service type 36. # hrs GU parent services by service type 69. First Name[3] 70. Last Name[3] 71. Middle Initial[3] 72. Date of Birth[3]	**School-level** Variables 20-25 & 28-35, in aggregate 37. # hrs GU teacher svcs by type 38. Cost of student svcs by type 39. Cost of parent svcs by type 40. Cost of teacher svcs by type **Program -level** 41. # GU schools 42. # of GU teachers by grade

[1] LEP: Limited English Proficiency.
[2] IEP: Individualized Education Plan.
[3] Required by National Student Clearinghouse to obtain student college enrollment information.

Stage 3 Evaluation of Aspirations, Knowledge, and Awareness Acquisition

Because students must also actively apply to colleges during this stage, maintaining academic aspirations at this time becomes critical. Thus, although academic achievement is emphasized during these later GEAR UP program years, sustaining high academic aspirations remains important. Ongoing survey results should show alignment of student aspirations and expectations during these years. Additionally, indicators of college entrance requirements and costs ought to rise to acceptable levels in these final high school years.

As supplemental program information, it is recommended that 12th grade GEAR UP student and parent surveys also include additional items inquiring as to whether students have applied to college, been accepted, and enrolled. While there are other methods of obtaining such information directly from colleges (after students have enrolled), it is valuable to collect the information for concurrent validation purposes.

Stage 3 Evaluation of Student Readiness Behaviors

Successful Course Completion

As noted above, during this stage, students are in their last few years of high school and beginning to apply to college. As such, there is an increasing emphasis on academic preparation, which involves successful completion of the right courses to graduate high school and gain admission to college. As noted in Section VI, the Optimal GEAR UP Indicators 43 through 53 relate to student course enrollment and successful completion (course marks of A, B, C, or Pass). These are formatted and maintained in the dataset as dichotomous variables indicating enrollment/no enrollment and success/no success.

The ACT Assessment

During this stage, most students will have the opportunity to take the American College Testing (ACT) assessment, which colleges commonly use as part of their admissions criteria. The ACT assessment contains four tests assessing English, math, reading, and science. Scores for each of the four tests range from 1 to 36. As with EXPLORE and PLAN assessments, the test scores may be used as continuous variables within analyses, or dichotomous scores may be developed based on the *College Readiness* benchmark scores provided by ACT, Inc. In this manner, the percent of students demonstrating College Readiness may be compared for GEAR UP and comparison group students.

The SAT Assessment

Many students take the College Board SAT test as an alternative, or in addition to the ACT. Students typically take the SAT while in 11th or 12th grade. The test is comprised of three sections: critical reading, math, and writing. Scores for each section range from 200 to 800, with a total possible score of 2400 on the SAT test. As with the PSAT, the SAT section scores can serve as objective, continuous metrics of GEAR UP student academic achievement.

National Student Clearinghouse College Enrollment Data

Over the years, National Student Clearinghouse (NSC) has served as a vital resource for GEAR UP programs in terms of identifying students attending college after graduating high school. Currently, approximately 92% of colleges nationwide subscribe to NSC. As a result, institutions or grant programs

working with NSC can obtain data identifying any of their students enrolling in public and private, two- and four-year college institutions throughout the country. Information provided by NSC on students includes institution(s) attending, date of first enrollment, enrollment in subsequent terms, declared college major, and degrees/certificates awarded.

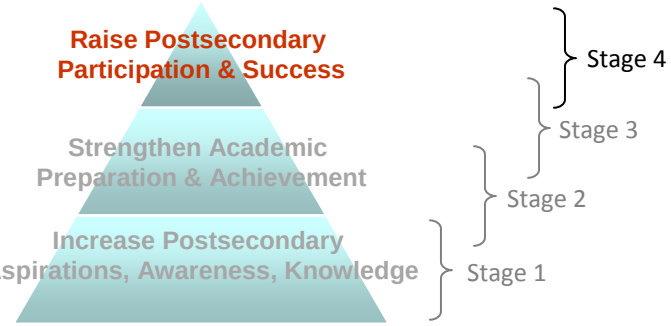
To match student records, NSC typically uses an algorithm based on student first and last name, middle initial, and date of birth. Thus, indicators 69 through 72 are included among the Stage 3 optimal indicators.

Because students must be currently enrolled in a college to be included in the NSC dataset, NSC is not able to provide such data on the GEAR UP cohort until after the student's 12 grade year. However, during this Stage 3, it is recommended that NSC college enrollment data are acquired for students one or two grades above the GEAR UP cohort, to serve as a comparison group against which to compare subsequent GEAR UP student college enrollment figures.

Section VIII:

Stage 4: First Year of College

College Course Enrollment, Success, and Persistence



During this last GEAR UP program stage, the emphasis is on college enrollment and continued persistence with postsecondary participation. This is illustrated within Table 12, below.

Table 12. Optimal GEAR UP Program Evaluation Indicators: Stage 4 (1st Year of College)

Student & Parent Survey Data		**Student Academic & Participation Data**		**Aggregate Data**
Aspirations, Knowledge & Awareness Acquisition	Control & Predictor Variables	Student Readiness Behaviors	Control & Predictor Variables	Control Variables
		68. Enrollment in College 73. Cont'd college enrollment	27. Student ID 28. School 29. Grade 30. Gender 31. Race/ethnicity 32. LEP[1] Status 33. IEP[2] Status 34. Free/Reduced Lunch Status 35. # hrs of GU student service by service type 36. # hrs of GU parent services by service type 66. First Name[3] 67. Last Name[3] 68. Middle Initial[3] 69. Date of Birth[3]	**School-level** Variables 20-25 & 27-36, in aggregate 36. # hrs GU teacher svcs by type 37. Cost of student svcs by type 38. Cost of parent svcs by type 39. Cost of teacher svcs by type **Program -level** 40. # GU schools 41. # of GU teachers by grade

[1] LEP: Limited English Proficiency.
[2] IEP: Individualized Education Plan.
[3] Required by National Student Clearinghouse to obtain student college enrollment information.

Stage 4 Evaluation of Student Readiness Behaviors

As a rule, student descriptive information is not available once the student leaves high school. Therefore, data elements 28 through 36 in the above table pertain to descriptive data from the students' prior (12th grade) year. Because these elements (e.g., gender, race/ethnicity, Free/Reduced Lunch Status) are constant, they continue to be useful in disaggregating outcomes data.

National Student Clearinghouse College Enrollment Data

At this time GEAR UP cohort college enrollment information may be acquired from NSC. It is during this stage that data are first available to examine the overarching GEAR UP goals of increasing student enrollment and success in college. Note that, while NSC provides student-level information, the available indicators do not include GPA data or course outcome information. Therefore, it is not possible to obtain a direct metric of academic achievement, short of degree attainment, which typically takes several years for students to achieve. However, NSC data does include each college term attended, which reflects student persistence. This persistence indicator serves as an appropriate and useful proxy measure of student college success.

Section IX:

Using Research Results to Inform Program Processes

This section addresses interpretation and use of study findings to guide ongoing program improvement. This is in light of the fact that, even when rigorous evaluation procedures result in useful findings, they have no value if such study results are not then properly interpreted and leveraged to inform program processes.

At times, GEAR UP top administrators appear wary of program evaluation out of a concern the findings will reflect poorly on their program. However, typically, assurances that results (both negative and positive) can be leveraged to improve program processes work to alleviate this anxiety. Given this, the final section of this guide addresses interpretation of findings, cost-benefit assessment methodology, and the value of using results to encourage data driven decision-making.

Interpretation of Findings

Program evaluation results have the potential to reveal best practices and areas of opportunity. When conducted properly, the findings highlight the strengths and limitations of the program as a whole. In other words, evaluation outcomes can be used to determine program elements that are effective and those that are dispensable or need modification. Such formative program assessment at the local level empowers program administrators to effectively guide continuous program improvement.

Key to such practices is appropriate interpretation of evaluation outcomes. In examination of analysis results, it is important to consider alternative explanations for findings, due to confounding variables external to GEAR UP program interventions. As noted in Section I, when possible, including baseline data and comparison groups into the research design helps guard against selection and maturational effects. For example, often student course performance declines during 8th and 9th grade, due to social/developmental factors. Therefore, when comparison group data are not included, it may appear that GEAR UP students are declining academically two years into the program. If such findings are presented in isolation, they would no doubt be disconcerting for GEAR UP administrators. However, when a comparison group is included, it is often found that the non-GEAR UP students show a much greater rate of decline in 8th and 9th grade course performance, thus suggesting that GEAR UP services help to counteract the external factors negatively influencing student 8th and 9th grade academic performance. The sample GEAR UP program evaluation summary sheet presented in Figure 12 on page 49-50 illustrates how such findings may be articulated.

Other times, declines in student performance may be due to GEAR UP program practices that need modification. For example, in the past, preliminary analysis results for some GEAR UP programs have revealed that non-GEAR UP students achieved higher standardized math test scores in comparison with GEAR UP students. In isolation, this can be troubling to program administrators. However, upon closer examination, it was found that GEAR UP counselors or teachers were strongly encouraging their GEAR UP students to take a higher-level math test based on more challenging coursework implemented as part of the GEAR UP program. Thus, the results were not a reflection of GEAR UP student academic decline, but rather due to a procedural change based on higher expectations of GEAR UP students and their courses. When such procedural changes are identified, GEAR UP administrators may then examine and modify policy related to such matters. This may include additional counselor training to ensure alignment between student competency and test level, or modification of course curricula to better

reflect the standards of the higher-level math tests. The sample academic outcomes summary report presented in Figure 13 on page 51-54 shows how such findings may be reported.

Survey results may also be misleading if the findings are interpreted in isolation. For example, many GEAR UP programs include items asking if students know the cost of college. When interpreting responses to such items, it is important to consider that results reflect student *perceptions* of their knowledge of college costs. Responses to such items suggest that GEAR UP students and parents often overestimate their level of knowledge regarding college costs. This is evident when responses are compared with survey items directly asking students to identify the correct cost estimate of attending a four-year college in their state (see survey items 12 and 9 within Figures 3 and 4, respectively).

Likewise, caution must be taken when interpreting survey results inquiring if anybody from the school or GEAR UP has spoken with them about financial aid availability, or about college entrance requirements. Certainly these two DOE-mandated survey questions can offer insight into the extent of service coverage. However, often GEAR UP administrators use these items to make inferences about student and parent actual knowledge attainment. As noted earlier, when compared with items asking students if they have enough information about financial aid and college entrance requirements, a notable gap is revealed, such that substantially more students and parents report school or GEAR UP staff have talked with them in comparison to those reporting to have enough information on the particular topic.

Use of Nationwide GEAR UP Evaluation Results

At the national level, results from analyses across GEAR UP programs can also reveal best practices, thus identifying exemplary individual programs, or services that are found effective across an array of GEAR UP programs. In addition, results may highlight services that are effective for particular regions, or types of GEAR UP programs (e.g., urban versus rural). In this way, findings from such a national GEAR UP evaluation can serve to inform program practices on a broad scale.

Of course, it is anticipated that results from a nationwide study would be useful for summative evaluation of GEAR UP. Specifically, the study findings would help determine the level to which GEAR UP is attaining its fundamental goal of increasing college enrollment and success for disadvantaged students. Moreover, the results of the national-level evaluation could reveal other indirect program impacts. For example, examining the findings from focus groups and surveys across programs may suggest that participation in GEAR UP programs also serves to reduce juvenile delinquency rates and/or substance abuse.

Once the expected and unanticipated impacts of GEAR UP are determined at the national level, it is vital that such information is disseminated to stakeholders. The study findings may then help inform policy relating to future support of the national GEAR UP initiative.

Cost-Benefit Assessment

There are several factors one must consider when determining the efficacy of particular GEAR UP services. Results of statistical analyses are a vital step in this process, as they can reveal associations between a given service and established program goals. However, decisions to eliminate, modify, or retain services must also consider other aspects of service benefits such as relevance, feasibility, participant satisfaction and attendance levels. These factors must then be weighed against the dollar cost of the particular service.

Given this multitude of factors, it is often helpful to establish and apply a systematic method to assess the cost-benefit of GEAR UP services. This involves identifying the primary factors under consideration, and then defining the range of possible levels for each factor. Table 13, below, presents an example of a cost-benefit assessment grid that may result from this process. Note that five factors are identified as important when considering the benefits of an individual GEAR UP service. These factors are then considered against the *dollar cost per participant* for the particular service.

Table 13. Cost-Benefit Assessment Grid

Value of Service/Activity					**Cost**
Effectiveness*	Relevance to Goal*	Level of Participation	Participant Satisfaction	Feasibility	Dollar Cost per Participant
Very effective	Very relevant	Very high participation	Very satisfied	Very feasible	Inexpensive
Effective	Relevant	High participation	Satisfied	Feasible	Reasonable
Somewhat effective	Somewhat relevant	Moderate participation	Somewhat satisfied	Somewhat feasible	Somewhat costly
Not effective	Not relevant	Low participation	Not satisfied	Not feasible	Very costly

*Consider for each outcome examined (e.g., math tutoring could have outcomes related to math course grades, math test scores, attendance, attitude towards attending college, etc.).

Issues for Consideration

For each factor noted in the table, there are issues that must be taken into account when assessing individual GEAR UP services. Below, each factor is addressed in terms of questions to consider and issues to take into account when applying this process to the individual services.

Effectiveness: Is there evidence demonstrating the service is related to the outcome? How strong is the association? Effectiveness should be demonstrated through data analysis techniques as discussed in this guide.

Relevance to Goal: Does the outcome help further the program's objectives? Also consider whether the service results in disparate impact with respect to student subgroups. For example, an increase in test scores may be revealed, but when data are disaggregated, it may be found that the benefits are only found for students of higher socioeconomic status. Because the focus of GEAR UP is primarily to assist students of low socioeconomic status, the outcome would not be relevant to the goal. Such disparate effects may be revealed through examination of the research results, or identified in focus group discussions with program staff.

Level of Participation: What percentage of participants (i.e., students, parents, or staff) attended the activity or received the service? A service may be shown to be effective and relevant, but low participation may be cause to consider whether to continue the service or to alter implementation. This information may be obtained through examination of participation data for individual services.

Participant Satisfaction: Are participants satisfied with the service? A service may be effective, relevant, and have high participation rates, but low satisfaction may indicate a need for modification. This information may be obtained via surveys or focus groups. Discussion with staff implementing services or informal discussion with participants may also reveal levels of satisfaction with particular services.

Feasibility: What level of effort is involved to implement the service? A lack of resources or staff expertise may make some services very difficult, time-consuming, or logistically complicated to provide. Such issues should be taken into account when evaluating services. This may be determined through discussion with program staff and administrators.

Dollar Cost per Participant: How much does it cost, per participant, to provide this service? This metric should take into account materials, food, transportation costs, and staff resources. This includes any personnel and non-personnel resources donated as matching, in-kind funds.

Example of a Cost-Benefit Assessment

As a practical example of such a cost-benefit assessment, the following describes application of this process to determine the efficacy of a College Campus Tour activity.

Effectiveness: A chi-square analysis reveals a statistically significant ($p < .05$) relationship between participation in campus tours and student reported college aspirations. The service is deemed *Very effective*.

Relevance to Goal: Although, as noted above, aspirations increased significantly, no effects were evident for male students. The statistically significant correlation was found to be due to substantial increases in aspirations for female students only. The service is determined to be *Somewhat relevant.*

Level of Participation: A high level of participation was revealed for the service (> 80% of students attended at least two campus tours), resulting in an assessment of *Very high participation.*

Participant Satisfaction: Survey results revealed that the majority (60%) of students and parents reported to be *Satisfied* or *Very Satisfied* with the campus tours. This results in an assessment of *Satisfied.*

Feasibility: Focus groups reveal that staff can handle the logistics to carry out the tours with minimal challenges/concerns. The trips appear to go smoothly and an effective process is in place for conducting the tours. The service is deemed to be *Very feasible*.

Consideration of the five factors above results in an overall assessment that the service is beneficial as it is currently being delivered. At this point, if the cost per student is reasonable, the service will be retained, with modifications to encourage greater effectiveness among male students. Note that, if the service was rated low on multiple factors above, a decision may be made to eliminate the service, regardless of subsequent findings regarding service cost.

Dollar Cost per Participant: Each campus tour includes the following costs, per trip, for one bus holding 20 students: Bus rental ($225), Gas ($175), Driver ($150), Staff ($250). This results in $800 per bus, or $40.00 per student. Students are also each provided with a $5.00 boxed lunch and $2.00 snack. This results in a total cost, per student, of $47.00 for the service. This is determined to be *Reasonable.*

As noted, the service was deemed to be beneficial with respect to the five factors reflecting the Service Value. Subsequently, the service was also found to be reasonable in cost. Therefore, a decision is made to further examine causes of the gender disparity in service effectiveness, and make modifications as necessary. However, the overall assessment of the service is that it is both valuable and cost-effective.

Dissemination of Research Results

There is a variety of methods to communicate research findings to program staff, participants, and stakeholders. Results are best conveyed in a clear and concise manner. Figures 11 through 13 on the following pages present some examples of GEAR UP evaluation results reports.

Note that, as a rule, the usefulness and relevance of survey information declines over time. Thus, it is important that survey results are provided to program administrators and staff as early as possible after the survey administration. This allows GEAR UP program personnel to use the information to make informed decisions based on the results. Moreover, it is recommended that, when possible, relevant survey findings are also disseminated to participating students and their parents. This helps engage program participants and validates their efforts in completing the survey forms. Figure 11 presents some examples of how results for particular survey items may be presented to program stakeholders. Note these results correspond to select items within the student sample survey shown in Figure 3.

In presenting results of GEAR UP program evaluation, it is often helpful to convey findings using graphs, tables, and charts when possible. Figures 12 and 13 on the following pages offer examples of data summary reports graphically illustrating study findings. Such summaries present findings in a non-intimidating manner and have been very well-received by program administrators. The reports also serve to engage and motivate program staff, participants, and community members. Last, these documents serve to inform program processes and thus empower grant staff and administrators in their efforts toward continuous program improvement.

Figure 11. Sample survey results

City College GEAR UP Student Survey Results: Spring 2009

ACADEMIC INVOLVEMENT

1-2[1]. What is the highest level of education you…?

	...would like to obtain?		...expect to obtain?	
	#	%		
High school or less	122	5.6	106	4.9
Some college, no 2- or 4-year college degree	140	6.5	243	11.2
Two-year college degree	392	18.1	664	30.7
Four-year college degree or higher	1510	69.8	1151	53.2
Total	**2164**	***100.0***	**2164**	***100.0***
No Response	22		24	
Total	2186		2186	

KNOWLEDGE ABOUT COLLEGE

	Yes		No		Total		NR[1]	Total
	#	%	#	%	#	%	#	#
6. Do you know what you need to do to get accepted into the college/university you plan to attend?	1208	57.0	908	42.9	**2116**	***100.0***	70	2186
*7. Has anyone from your school or GEAR UP ever spoken with you about the availability of financial aid to help you pay for college?	1277	60.3	841	39.7	**2118**	***100.0***	68	2186
8. Do you have enough information about financial aid to help you pay for college?	696	33.2	1401	66.8	**2097**	***100.0***	89	2186
*9. Has anyone from your school or GEAR UP ever spoken with you about college entrance requirements?	1023	49.2	1058	50.8	**2081**	***100.0***	105	2186
10. Do you have enough information about college entrance requirements?	1253	59.6	851	40.4	**2104**	***100.0***	82	2186

[1]NR = No Response

12. How much do you think it costs (tuition and fees only) to attend a four-year public college in your state for one year?

	#	%
$1-$1,900	37	1.7
$1,901-$4,500	246	11.5
$4,501-$7,400	490	22.9
$7,401-$10,000	429	20.1
$10,001-$18,000	484	22.6
More than $18,000	452	21.1
Total	**2138**	***100.0***
No Response	48	
Total	2186	

Note. The category of $1,901-$4,500 represents actual 2008/2009 costs to attend a four-year public college in this state for one y

17. Have any of the following GEAR UP resources helped you with your decision to go to college?

	Yes		No		Total		NR[1]	Total
	#	%	#	%	#	%	#	#
Tutoring or help with school work	1101	55.4	886	44.6	**1987**	***100.0***	199	2186
Visiting a college campus	852	44.6	1060	55.4	**1912**	***100.0***	274	2186
Mentoring from GEAR UP Staff	812	43.6	1051	56.4	**1863**	***100.0***	323	2186
College Preparation workshops	675	36.6	1170	63.4	**1845**	***100.0***	341	2186
Information about the benefits of going to college	1222	63.7	696	36.3	**1918**	***100.0***	268	2186
Information about financial aid and how much college costs	1053	55.6	840	44.4	**1893**	***100.0***	293	2186
Other	110	25.8	316	74.2	**426**	***100.0***	1760	2186

*Items mandated by the DOE for inclusion within the Annual Performance Report.

COBRO CONSULTING, LLC

Page 1 of 2

Note. Sample data are presented for presentation purposes only.

Figure 12. Sample GEAR UP Program Evaluation Summary Sheet

City College GEAR UP Program
Summary of Progress

This summary report presents information about the City College GEAR UP (GU) Program at the midpoint of the program. Progress toward the three overarching objectives of the program is presented using data from student and parent surveys, academic course and test outcomes, and service participation records.

INTRODUCTION

Since its beginning in September 2005, the City College GEAR UP Program has provided services to cohort and priority students in 40 school districts and 45 schools in urban and rural areas throughout the state. Core services include test preparation, tutoring, workshops and summer programs about college admissions and financial aid, and college visits. Table 1 presents the number of cohort and priority students served during the past three years of the program.

Table 1. Number of students served

	2006-07	2007-08	2008-09
Cohort	**7th**	**8th**	**9th**
#	550	675	725
Priority	**#**	**#**	**#**
7th	155	140	0
8th	60	390	205
9th	100	400	502
10th	50	325	470
11th	15	210	350
12th	10	25	300
Total	940	2165	2552

OBJECTIVE 1: COMPETENCY

To increase basic skills competency and preparation for college.

Table 2. Student test outcomes: percent passing and proficient

Passed State Test	Math	Writing	Reading	All
State's GEAR UP Program 2007-08				
10th	40	50	38	35
11th	70	80	72	60
12th	81	90	83	71
State's GEAR UP Program 2008-09				
10th	65	70	88	45
11th	80	83	92	70
12th	78	83	90	75
Statewide Economically Disadvantaged 2008-09				
10th	55	60	65	50
11th	65	70	75	55
12th	67	72	80	60
Proficient on EPAS	**English**	**Math**	**Science**	
State's GEAR UP Program 2008-09				
EXPLORE	20	15	17	
PLAN	19	17	18	
ACT	19	18	19	
State Economically Disadvantaged 2008-09				
EXPLORE	18	16	16	
PLAN	15	13	13	
ACT	14	13	14	

As Table 2 presents, the percentage of 10th and 11th grade GU students passing the State Math, Writing, and Reading tests in 2008-09 increased notably from the 2007-08 rates. Although 12th grade students showed lower pass rates for 2008-09 State Math and Writing tests, they achieved higher rates in Reading. In addition, the State Test pass rates for GU students far exceeded those of statewide economically disadvantaged students.

Table 2 also shows GU student proficiency rates for the 2008-09 EPAS. Given the nature of the cohort/priority model structure, the number of students in each grade level of the program varies greatly each year. As a result, year-to-year comparisons of similar groupings of GU students are not feasible. Therefore, an external comparison group (i.e., statewide economically disadvantaged students) within the same time frame was used. As Table 2 illustrates, GU students achieved higher EPAS proficiency rates than did the comparison group in the majority of the subjects. It was only in Math Explore that GU student proficiency rates were lower than the comparison group.

Figure 1. GEAR UP student course success rates, by gender

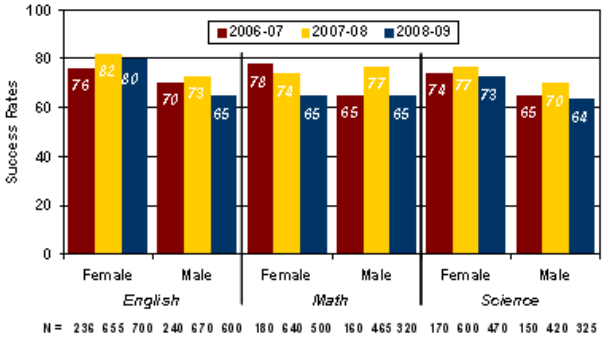


Figure 1 illustrates average success rates of GU students in English, Math, and Science courses, by gender. There appears to be a general increase in success rates from 2006-07 to 2007-08, followed by a decrease in 2008-09. This decline is likely due to student maturational factors. Namely, in 2007-08, the majority of GU students were in 8th and 9th grade. In general, declining course performance is typical for 8th and 9th grade students. Likewise, the consistent decline in math course success for females from 7th to 9th grade is also common. These results indicate a need for possible increases or modifications to program resources and services aimed to increase academic course performance.

Note. Sample data are presented for presentation purposes only.

Figure 12. Sample GEAR UP Program Evaluation Summary Sheet (continued)

OBJECTIVE 2: PREPAREDNESS

To increase high school graduation rates and college enrollment.

Table 3 displays the percentages of 2008-09 GU students reporting that particular activities or services were helpful in preparing for, or deciding to attend college. Figure 2 presents parent and student GU participation data, showing annual increases in the number of GU students participating in mentoring, counseling/advising, and workshops. Parents of GU students also showed increased participation in counseling/advising, workshops, and family events over the three years examined. Figure 3 presents student aspirations and expectations for obtaining a four-year college degree. While both of these indicators show consistent increases across the three years, a notable gap between student aspirations and expectations remains.

Table 3. Percent of students reporting in 2008-09 that services/activities were helpful

Helped in deciding to attend college	%
Info about benefits of going to college	70
Info about fin. aid & cost of college	65
Mentoring from GEAR UP Staff	60
Tutoring or help with school work	55
College preparation workshops	50
Visiting a college campus	45
Helped in preparing for college	**%**
Visit to a high school/college	75
Acad. Assist. Activities/classes/packets	72
Interest/career exploration activities	70
Math tutoring	70
English tutoring	66
Tutoring - another subject	65
GEAR UP workshop	65
Academic counseling/advising	63
Standardized test prep	60
Presentation by coll. staff/bus. leaders	59

Figure 2. Graph of student & parent participation

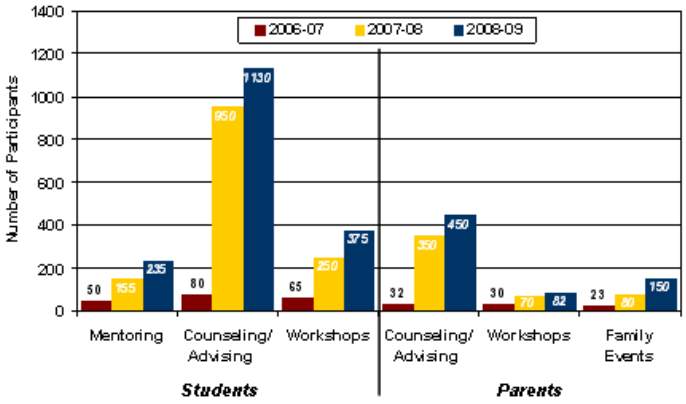
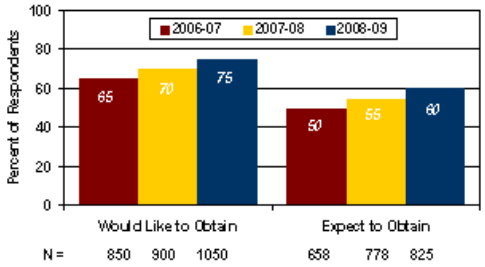


Figure 3. Student aspirations and expectations for transferring to a 4-year college



OBJECTIVE 3: SCHOLARSHIPS

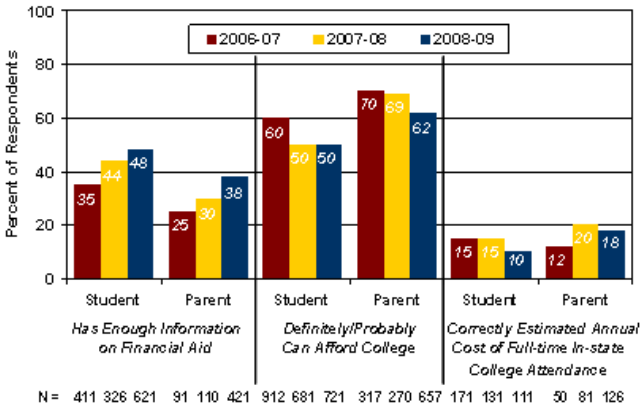
To increase GEAR UP student access to postsecondary education options and financing.

In 2008-09, 262 of the 300 seniors in the City College GEAR UP Program (87%) provided information on their postsecondary status. Of these, 150 (57%) reported to have enrolled in some kind of postsecondary education or training program.

With regards to financing their education, in 2008-09, 110 of the 262 GU seniors applied for financial aid. In addition, 95 seniors applied for non-GU scholarships, of which 75 received awards, and 67 seniors applied for GU scholarships, of which 55 received awards.

Figure 4 presents information about student and parent knowledge of financial aid and college costs. Note that less than half of the students and parents report having enough financial aid information, although the figures are increasing. Moreover, while most students and parents report they can afford college, these figures have remained stable or decreased from 2007-08 to 2008-09. Last, the chart reveals that, during all three years, just 10% to 20% of students and parents were able to identify the correct cost range of attending a state college. This indicates a need to increase current GU efforts to inform students and parents about actual college costs.

Figure 4. Financial aid and college costs



Note. Sample data are presented for presentation purposes only.

Figure 13. Sample GEAR UP Program Academic Outcomes Summary Report

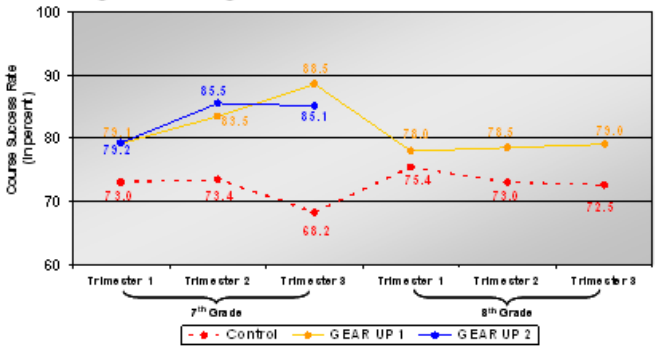
City College GEAR UP
Year Two Academic Outcomes

The following graphs present the results of analyses examining the academic impact, to date, of the City College GEAR UP program. These graphs show changes in course success rates, student GPAs, and State Standards Test (SST) scores for students in the GEAR UP 1 and GEAR UP 2 cohorts (7th grade in 2005-2006, and 7th grade in 2006-2007, respectively), as well as a control group (7th grade in 2004-2005) for comparison purposes.

Course Success Rates

Course success is defined as the percent of course enrollments resulting in a grade of A, B, C, or Credit. As shown in Figure 1, trimester success rates show differing trends for GEAR UP students when compared with the control group. While students in the comparison group tended to perform worse by the end of each academic year, GEAR UP students from both classes showed improvement. T-test analyses revealed these increases in GEAR UP student course success rates within each academic year to be significant ($p < .05$).

Figure 1. Average Trimester Success Rates



The following two graphs present success rates for the control and GEAR UP 1 7th and 8th grade courses. These data are disaggregated by gender and ethnicity.

As shown in Figures 2 and 3, similar patterns as described above are evident for both males and females, as well as for those racial/ethnic groups where there are adequate sample sizes (i.e., White, Hispanic, and African-American students). Though the results are not always consistent for 8th grade males, generally, the data suggest that GEAR UP students show greater progress than the control group, within each specific racial/ethnic and gender group examined.

Figure 2. Average Trimester Success Rates by Gender

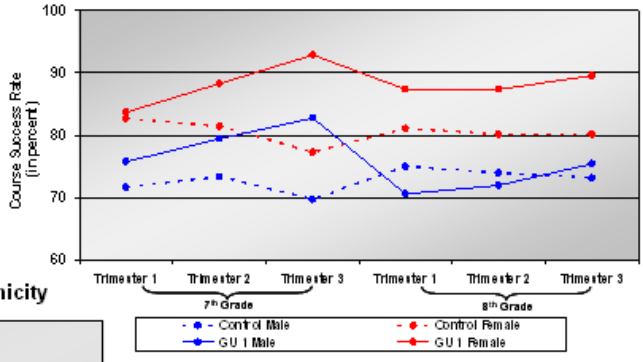
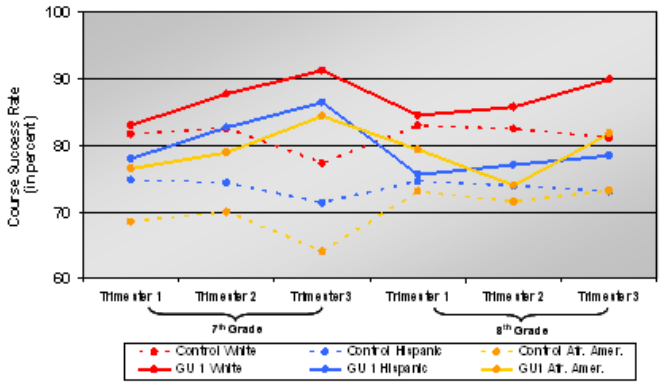


Figure 3. Average Trimester Success Rates by Ethnicity



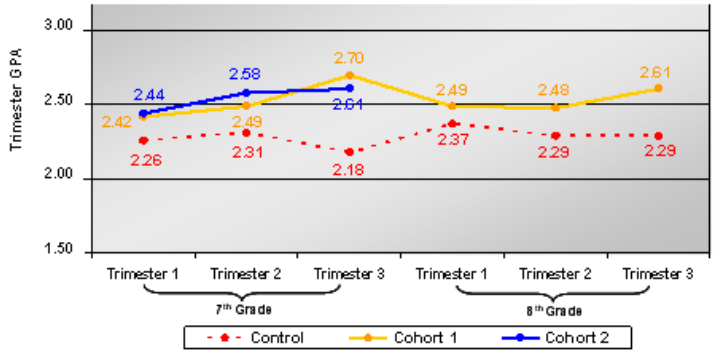
Note that for all demographic groups in GEAR UP 1, the improvement in 8th grade by the end of the academic year is generally less dramatic than their 7th grade improvement. Additionally, their 8th grade Trimester 1 success rates were lower than the prior 7th grade Trimester 3 success rates.

Note. Sample data are presented for presentation purposes only.

Figure 13. Sample GEAR UP Program Academic Outcomes Summary Report (continued)

Trimester GPA

Figure 4. Average Trimester GPA



As shown in Figure 4, an examination of GEAR UP and control group grade point averages (GPAs) yields similar results across the two years as those presented for success rates. The GPAs of both GEAR UP 1 and GEAR UP 2 students increased from Trimester 1 to Trimester 3, while the control group did not show a corresponding increase. T-test analyses revealed these differences to be significant for both GEAR UP 1 and GEAR UP 2 cohorts ($p < .05$).

Again, similar patterns emerge when the data are disaggregated for males and females (Figure 5), and by race/ethnicity (Figure 6). For all demographic groups studied, GEAR UP students consistently earned higher average GPAs than did students in the comparison group. This indicates that the effects of GEAR UP benefit all demographic groups.

Figure 5. Average Trimester GPA by Gender

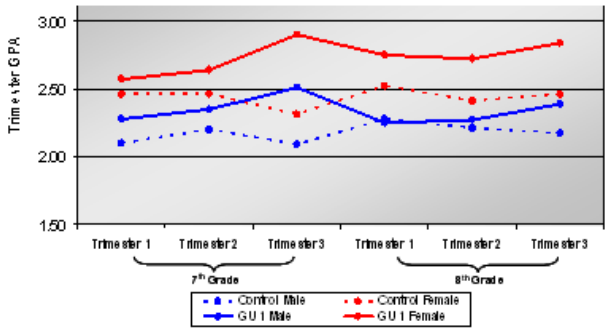
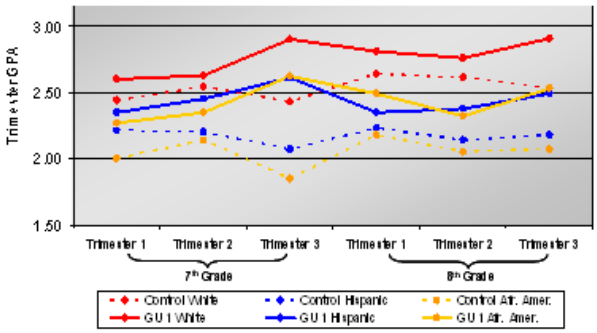


Figure 6. Average Trimester GPA by Ethnicity



Further Analyses of Course Success and GPA

The decrease in course performance from the end of 7th grade to the beginning of 8th grade was thought to potentially be related to the increased turnover (approximately 25%) in students that occurs before each school year begins. To examine this possibility, academic outcomes of students who had been part of the original first trimester cohort were examined separately. Their course success rates and trimester GPAs showed the same decrease between academic years, indicating that the decrease is not due to new students entering the cohort.

Tracking the original GEAR UP students yielded another observation. African-American students in this cohort began the program with 60 students, and at the end of two years, 34 students remained, representing approximately 57% of the original group. All other demographic groups studied showed a two-year retention rate of 70% to 80%. This finding has been brought to the attention of the GEAR UP program staff members, who are examining potential causes.

Summary: Course Success and GPA

Students in GEAR UP classes showed significantly greater increases in course grades in comparison with control group students, providing evidence that these increases in achievement are due to reasons beyond student maturation. Disaggregating the data based on gender and race/ethnicity showed similar results.

Note. Sample data are presented for presentation purposes only.

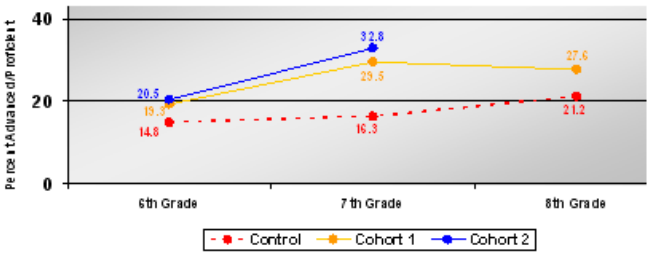
Figure 13. Sample GEAR UP Program Academic Outcomes Summary Report (continued)

State Standards Tests

Percentages on all graphs represent students who achieved a score of *Advanced* or *Proficient* on the test indicated. Figure 7 presents the percent of students in each GEAR UP group and the control group who earned a score of *Advanced* or *Proficient* on their English/Language Arts SST test.

Students in both GEAR UP groups showed improvements in English over the control group, and over their own baseline (6th grade) scores. Chi square analyses revealed that a significantly higher percentage of GEAR UP students obtained English scores of *Advanced/Proficient* during 7th grade than did students in the control group ($p < .05$). However, differences found for GEAR UP and control group 8th grade test scores were not significant.

Figure 7. Percent *Advanced/Proficient* in English



The following figures present SST English rates of *Advanced/Proficient* scores separated by gender and race/ethnicity.

As shown in Figure 8, separating the GEAR UP 1 and control group SST scores by gender reveals a similar pattern; GEAR UP 1 males and females performed better than those in the control group. On average, females in GEAR UP 1 performed better than GEAR UP 1 males; the same holds true for the control group. During 8th grade, GEAR UP 1 males achieved only slightly higher rates of *Advanced/Proficient* scores than did control group females. Both control group and GEAR UP 1 males showed improvement each year, as did control group females. GEAR UP 1 females showed a marked increase from 6th to 7th grade, but a decline in their 8th grade scores.

Figure 8. Percent *Advanced/Proficient* in English by Gender

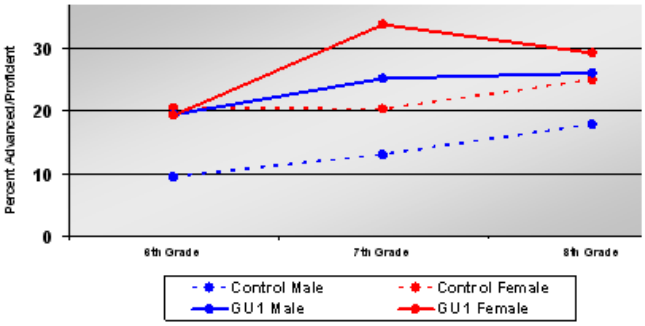


Figure 9. Percent *Advanced/Proficient* in English by Ethnicity

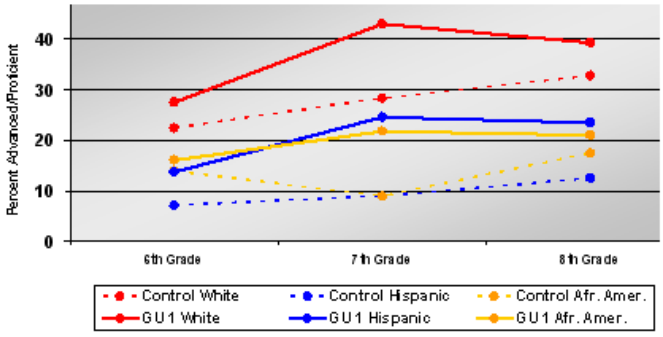


Figure 9 presents the rates of *Advanced/Proficient* scores for racial/ethnic groups where there are adequate sample sizes (i.e., White, Hispanic, and African-American students). All groups performed better than their respective control groups. Each GEAR UP 1 group showed increases from 6th to 7th grade, but no increase (and, in the case of White students, a decline of 4%) from 7th to 8th grade. White students in both the control and GEAR UP 1 groups performed better than Hispanic and African-American students, who showed similar rates of achievement in GEAR UP 1.

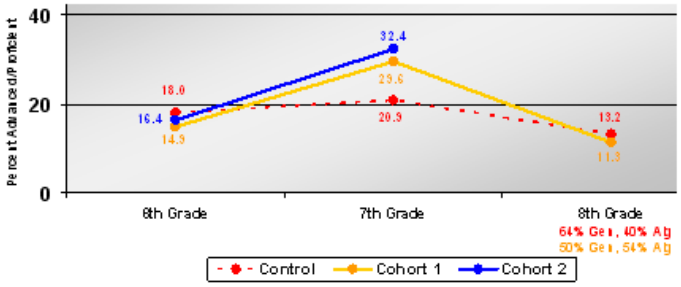
Results of the SST Math test showed different outcomes than those on the SST English/Language Arts test.

As with the English test, GEAR UP students showed increases over the control group on their 7th grade math SST scores, as seen in Figure 10, on the following page. Chi square analyses revealed these differences to be significant ($p < .05$). However, the differences in GEAR UP and control group 8th grade test scores were not significant. GEAR UP 1 students showed lower rates of *Advanced* and *Proficient* scores in 8th grade than both the control group 8th grade scores, and their own baseline 6th grade scores. This may be due to a change in school testing policies implemented at that time.

Note. Sample data are presented for presentation purposes only.

Figure 13. Sample GEAR UP Program Academic Outcomes Summary Report (continued)

Figure 10. Percent *Advanced/Proficient* in Math



Students in 8th grade take either the SST General Mathematics (Grades 6 & 7 Standards) test or the SST Algebra I test, as indicated by the color-coded percentages below the graph. According to Middle School staff, GEAR UP 1 8th grade students were strongly encouraged to take the Algebra I test, resulting in 54% taking this higher level test compared with 40% of the control group. This likely contributed to the lower scores, as students may not have been adequately prepared for the higher level test.

Examination of GEAR UP 1 and control group SST Math scores separated by gender yields interesting results, as shown in Figure 11. Females in 7th grade showed a striking increase in achievement over the 7th grade control group, as well as over their own 6th grade baseline rates. Males also showed an increase in their 7th grade achievement. However, both genders showed large decreases in achievement during 8th grade, dropping below the control group and, in the case of males, well below their own baseline scores as well.

Figure 11. Percent *Advanced/Proficient* in Math by Gender

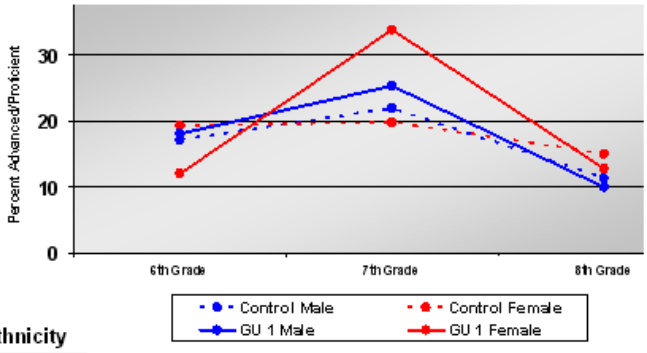
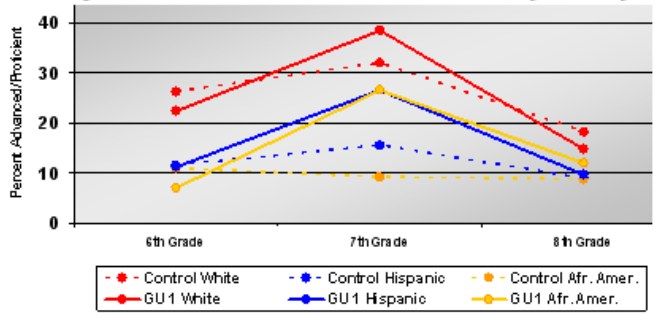


Figure 12. Percent *Advanced/Proficient* in Math by Ethnicity



As presented in Figure 12, all GEAR UP 1 racial/ethnic groups examined performed better than their respective comparison groups on the 7th grade SST Math test. Their scores declined on the 8th grade tests. White students' scores fell below those of their comparison group on the 8th grade SST Math test, and well below their own baseline 6th grade achievement rates. During 8th grade, African-American students showed a decline in achievement from their 7th grade scores but maintained a higher rate of achievement than their comparison group, as well as their own baseline 6th grade scores.

SST Math scores of students who had been part of the original 1st trimester 7th grade GEAR UP 1 cohort were examined separately. These results were similar as for the entire GEAR UP 1 group examined above. Analyses by gender and ethnicity also showed similar results for this original cohort as those presented above.

Summary: State Standards Tests

GEAR UP students show marked improvement in SST English scores over a comparison group of students. SST Math scores yielded less consistent results: GEAR UP students improved over the control group during their 7th grade tests, but on average performed worse than the control group during the 8th grade SST Math tests. This may relate to the fact that more GEAR UP students took the higher level SST Math test than did students in the comparison group.

Conclusion

The GEAR UP program's fundamental goal is to improve the academic success of low income and minority students. The information presented in this report confirms that the City College GEAR UP program is succeeding in this effort.

City College GEAR UP *Page 4 of 4*

Note. Sample data are presented for presentation purposes only.

APPENDIX A

NCCEP/AT&T Foundation GEAR UP Evaluation Framework Indicators

NCCEP/AT&T Foundation GEAR UP National Evaluation Project

Minimum Suggested Evaluation Variables

Variables [data collection tools – note: [1] = APR student survey; [2] = APR parent survey; [3] = school data**]**

STAGES ↓	Data Available from APR			Additional Minimum Data to be Collected			Unique Program Data
READINESS TYPES →	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	
Stage I: 6th – 8th grade	1. student educational expectations 2. parental educational expectations of students 3. student perceptions of college affordability 4. parental perceptions of college affordability	5. # ≥ grade level English/language arts 6. # ≥ grade level math	21. PR # 22. state 23. # of GU students by grade 24. # GU student by gender [3] 25. # GU students with LEP [1] 26. # GU students by race/ethnicity [3] 27. # students eligible for FRL in school 28. # of schools 29. # of GU students receiving service by service type 30. # of hours of GU services attended by students by service type 31. cost of student services by type 32. # of GU parents receiving service by service type 33. # of hours of GU services attended by parents by service type 34. cost of parental services by type 35. # of teachers receiving GU service 36. # of hours of GU services attended by teachers 37. cost of teacher services	38. student awareness of postsecondary educational options [1] 39. student awareness of ways to finance college [1] 40. parental awareness of ways to finance college [2] 41. parental savings for college [2] 42. post-HS intentions [1] 43. GU impact on post-HS intentions [1]	44. total # of GU student full-day absences [3] 54. # of hours worked/week [1] 55. # who completed PSAT [3 or SAT] 56. # who completed PLAN [ACT] 57. average PSAT score [3 or SAT] 58. average PLAN score [ACT]	64. # of all students in school 65. parental highest education level attained [2] 66. # all students by race/ethnicity [CCD] 67. # counselors in school [3] 68. # teachers certified in field [3] 69. # of teachers 70. HS grad rate [3] 71. HS PSE enrollment rate [NCS, state ED]	

Minimum Suggested Evaluation Variables

Variables [data collection tools – note: [1] = APR student survey; [2] = APR parent survey; [3] = school data**]**

STAGES ↓	Data Available from APR			Additional Minimum Data to be Collected			Unique Program Data
READINESS TYPES →	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	
Stage II: 9th – 10th grade	1. student educational expectations 2. parental educational expectations of students 3. student perceptions of college affordability 4. parental perceptions of college affordability	5. # ≥ grade level English/language arts 6. # ≥ grade level math 7. # who completed Algebra I 8. # who completed Algebra II 9. # who completed Calculus 10. # who completed Chemistry 11. # who completed Physics 12. # who enrolled in AP math 13. # who enrolled in AP English 14. # who enrolled in AP science 15. # who completed AP math 16. # who completed AP English 17. # who completed AP science	21. PR # 22. state 23. # of GU students by grade 24. # GU student by gender [3] 25. # GU students with LEP [1] 26. # GU students by race/ethnicity [3] 27. # students eligible for FRL in school 28. # of schools 29. # of GU students receiving service by service type 30. # of hours of GU services attended by students by service type 31. cost of student services by type 32. # of GU parents receiving service by service type 33. # of hours of GU services attended by parents by service type 34. cost of parental services by type 35. # of teachers receiving GU service 36. # of hours of GU services attended by teachers 37. cost of teacher services	38. student awareness of postsecondary educational options [1] 39. student awareness of ways to finance college [1] 40. parental awareness of ways to finance college [2] 41. parental savings for college [2] 42. post-HS intentions [1] 43. GU impact on post-HS intentions [1]	44. total # of GU student full-day absences [3] 45. # ≥ state minimum test score standards [3] 46. # who completed ≥ 1 computer science (or similar) course [3] 47. # who completed ≥ 2 English courses [3] 48. # who completed ≥ 2 social studies courses [3] 49. # who completed ≥ 1 years of non-English language courses [3] 50. # who completed ≥ 2 years of math courses [3] 51. average HS GPA [3] 52. # who took AP test by subject [3] 53. # who scored ≥ 3 (Qualified) on AP test by subject [3] 54. # of hours worked/week [1] 55. # who completed PSAT [3 or SAT] 56. # who completed PLAN [ACT] 57. average PSAT score [3 or SAT] 58. average PLAN score [ACT]	64. # of all students in school 65. parental highest education level attained [2] 66. # all students by race/ethnicity [CCD] 67. # counselors in school [3] 68. # teachers certified in field [3] 69. # of teachers 70. HS grad rate [3] 71. HS PSE enrollment rate [NCS, state ED]	

Minimum Suggested Evaluation Variables

Variables [data collection tools – note: [1] = APR student survey; [2] = APR parent survey; [3] = school data**]**

STAGES ↓	Data Available from APR			Additional Minimum Data to be Collected			Unique Program Data
READINESS TYPES →	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	Awareness Acquisition	Student Readiness Behaviors	Control & GU Program Variables	
Stage III: 11th – 12th grade	1. student educational expectations 2. parental educational expectations of students 3. student perceptions of college affordability 4. parental perceptions of college affordability	5. # ≥ grade level English/language arts 6. # ≥ grade level math 7. # who completed Algebra I 8. # who completed Algebra II 9. # who completed Calculus 10. # who completed Chemistry 11. # who completed Physics 12. # who enrolled in AP math 13. # who enrolled in AP English 14. # who enrolled in AP science 15. # who completed AP math 16. # who completed AP English 17. # who completed AP science 18. # who completed SAT 19. # who completed ACT 20. # who graduated from HS	21. PR # 22. state 23. # of GU students by grade 24. # GU student by gender 25. # GU students with LEP [1] 26. # GU students by race/ethnicity 27. # students eligible for FRL in school 28. # of schools 29. # of GU students receiving service by service type 30. # of hours of GU services attended by students by service type 31. cost of student services by type 32. # of GU parents receiving service by service type 33. # of hours of GU services attended by parents by service type 34. cost of parental services by type 35. # of teachers receiving GU service 36. # of hours of GU services attended by teachers 37. cost of teacher services	38. student awareness of postsecondary educational options [1] 39. student awareness of ways to finance college [1] 40. parental awareness of ways to finance college [2] 41. parental savings for college [2] 42. post-HS intentions [1] 43. GU impact on post-HS intentions [1]	44. total # of GU student full-day absences [3] 45. # ≥ state minimum test score standards [3] 46. # who completed ≥ 1 computer science (or similar) course [3] 47. # who completed ≥ 4 English courses [3] 48. # who completed ≥ 3.5 social studies courses [3] 49. # who completed ≥ 2 years of non-English language courses [3] 50. # who completed ≥ 4 years of math courses [3} 51. average HS GPA [3] 52. # who took AP test by subject [3] 53. # who scored ≥ 3 (Qualified) on AP test by subject [3] 54. # of hours worked/week [1] 55. # who completed PSAT 56. # who completed PLAN 57. average PSAT score [3 or SAT] 58. average PLAN score [ACT] 59. average SAT score [3 or SAT] 60. average ACT score [ACT] 61. # of GU students who applied to college [1] 62. # of GU students with ≥ 1 college acceptance [1] 63. # of GU students who completed FAFSA [1]	64. # of all students in school [3] 65. parental highest education level attained [2] 66. # all students by race/ethnicity [CCD] 67. # counselors in school [3] 68. # teachers certified in field [3] 69. # of teachers 70. HS grad rate [3] 71. HS PSE enrollment rate [NCS, state ED]	
Stage IV: 13th – 14th grade				72. # of GU students by postsecondary status type (4 yr public/private PSI, 2 yr public/private, technical, military, employed, unemployed, other) [NSC, state ED] 73. # GU students returning for second year [NSC, state dept of ed]			