- Academic background, including: overall undergraduate GPA; graduate school GPA; and verbal, analytical, and/or quantitative scores on the GRE.
- Participation in extracurricular programs during the undergraduate years, including participations in undergraduate research programs or in minority STEM undergraduate programs.
- Sources of graduate school financial support, including fellowships, scholarships, dissertation grants, teaching assistantships, research assistantships, traineeships, internships, loans, personal savings, family savings, and employer reimbursement.
- Duration and continuity of financial support.
- Amount of education and non-education debt at time of bachelor's degree completion.
- Amount of education and non-education debt at time of Ph.D. completion.
- Highest educational attainment of mother and/or father, including less than high school/secondary school, high school/secondary graduate, bachelor's degree, master's degree, professional degree, or doctoral degree.
- Time-to-Ph.D. degree.

Information about classification on colleges and universities can also be located at the following websites:

- Historically Black Colleges and Universities (HBCUs): http://www2.ed.gov/about/inits/list/whhbcu/edlite-list.html.
- Hispanic-Serving Institutions (HSIs): http://www.molis.org/hsis.asp.
- Tribal colleges: http://www.aihec.org/.
- Women's colleges: http://www.womenscolleges.org/.
- Colleges and universities with high enrollment of disabled students: Gallaudet University (http://www.gallaudet.edu/) or the National Technical Institute for the Deaf at Rochester Institute of Technology (http://www.ntid.rit.edu/).

The U.S. Department of Education also has a list of minority postsecondary institutions: http://www2.ed.gov/about/offices/list/ocr/edlite-minorityinst.html.

Comparison Groups

In order to make a strong case that programs, practices, or interventions are having a positive effect or resulting in positive outcomes for participants, it is important to identify an evaluation design that includes comparison groups. It is also important to ensure that sufficient data can be collected to permit appropriate and convincing comparisons.

In general, comparison groups are usually made up of members who are (a) similar to participants and/or (b) have not participated in the program or activity. For an educational study, depending on its objectives, it is generally important that the two groups be matched on characteristics that are correlated with:

- Students' educational achievement prior to graduate school entry;
- GRE scores;
- Type or rank of undergraduate institution where bachelor's degree was earned;
- Age when entering graduate school;
- Race/ethnicity, sex, disability, or citizenship;
- STEM department or undergraduate major;
- Number of years in a graduate program; or
- Type, duration, or timing of financial aid.

In evaluation studies about URM graduate students, there is a tendency to only ask the target population critical questions about the educational program, practices, or interventions. The problems with this evaluation design are as follows:

1. Members of the targeted group might be reluctant to share any negative information with evaluators out of concern for potential repercussions.
2. It is difficult to determine the extent to which data from such a small sample can be used to paint a representative picture of the larger context in which the graduate educational process occurs. Also, with small numbers, it might be difficult to keep information from being associated with individuals.
3. Conclusions cannot be used to indicate that problems are particular to race/ethnicity. In addition, it is difficult to separate out problems not associated with race/ethnicity but related to more general issues that affect all graduate students regardless of race/ethnicity, sex, disability, or citizenship.

The example on the next page illustrates how comparison groups were used for a large-scale group study.

A Large-Scale Comparison Group Study: Evaluation of the Louis Stokes Alliances for Minority Participation (LSAMP)

In 2005, a team of researchers at the Urban Institute in Washington, DC, completed an evaluation of the Louis Stokes Alliances for Minority Participation (LSAMP), an undergraduate-focused program that had been funded by NSF since 1991. The purpose of the program is to increase the representation of underrepresented minorities in the STEM disciplines. At the time the evaluation report was written, there were 34 alliances involving 450 different institutions across the country.

As a large-scale evaluation, many dimensions of the program required investigation. This example discusses only one of these dimensions: the number of LSAMP participants who completed their degrees compared to minority students not in the program. The example illustrates important points related to presenting data and the issue of comparison groups. The full report is available online at http://www.urban.org/UploadedPDF/411301_LSAMP_report_appen.pdf.

Comparison groups are critical in large-scale evaluations. Unlike a control group, however, comparison groups often differ in important ways from the subject "treatment" group. The term "control group" is used only when the people participating in a program and those in the comparison group are assigned to their respective groups via a random process.

To set up a control group for a program such as LSAMP, all interested students would be required to apply, followed by a random selection process for participants. This approach is not usually feasible with programs of this type. Rather, students are selected via non-random processes (e.g., selected due to high grades or as a result of a recommendation by a professor). This means that the participants in such programs, by definition, are a self-selected group.

To a large extent, self-selected groups generally differ from those who do not choose to participate in programs, making a classical experimental design impossible. Instead, quasi-experimental designs are more commonly used for evaluations. Quasi-experimental designs follow a similar framework as the "gold standard" – the randomized control trial – but deviate by lacking random assignment to treatment (i.e., the program) and control (i.e., not participating in the program). The goal, then, becomes to identify appropriate comparison groups. For the LSAMP evaluation, two natural comparison groups emerged: URM students who had not participated in the LSAMP program and U.S. students of white, non-Hispanic, and Asian descent who were not the original targets of the program.

By comparing the outcomes for LSAMP participants to URM non-participants, a general sense of the program impact can be gained – but with the important caveat that the participants are already known to be a self-selected group (and, hence, possibly already more personally motivated, with higher GPAs and stronger connections to professors in their universities). That said, however, the comparison to the white and Asian students does provide a sense of the extent to which the playing field for the program participants was effectively leveled. That is, how did the graduate school outcomes compare for the LSAMP participants to those of students who were not eligible to participate in the program, presumably those students who were already implicitly advantaged within the higher education system?

Figure 2-1, taken from the Urban Institute report, illustrates findings from LSAMP participants and those from the two comparison groups. The data are from two sources: (1) LSAMP participant data gathered from a survey of those who had graduated between 1992 and 1997 and (2) for the comparison groups, nationally representative data available from NSF via the National Survey of Recent College Graduates.[2]

The graph in Figure 2-1 shows the overall rates of transition to graduate school as well as the specific rates of transition into STEM fields, which were the ones of predominant interest to the LSAMP program. In this case, the two comparison groups are, indeed, random samples of people who were recent bachelor's and master's degree recipients and were contemporaries of the LSAMP participants.

FOR MORE INFORMATION, SEE HTTP://WWW.NSFAGEP.ORG/

This example shows the effective use of the comparison group method. Researchers from the Urban Institute took advantage of large-scale, national data, which are accessible to everyone through a public-use tool known as SESTAT (see Chapter 8 for more information about national sources of data). These data also can be used for more detailed analyses by those who have access to the restricted-use datasets. Chapter 8 also discusses these data.

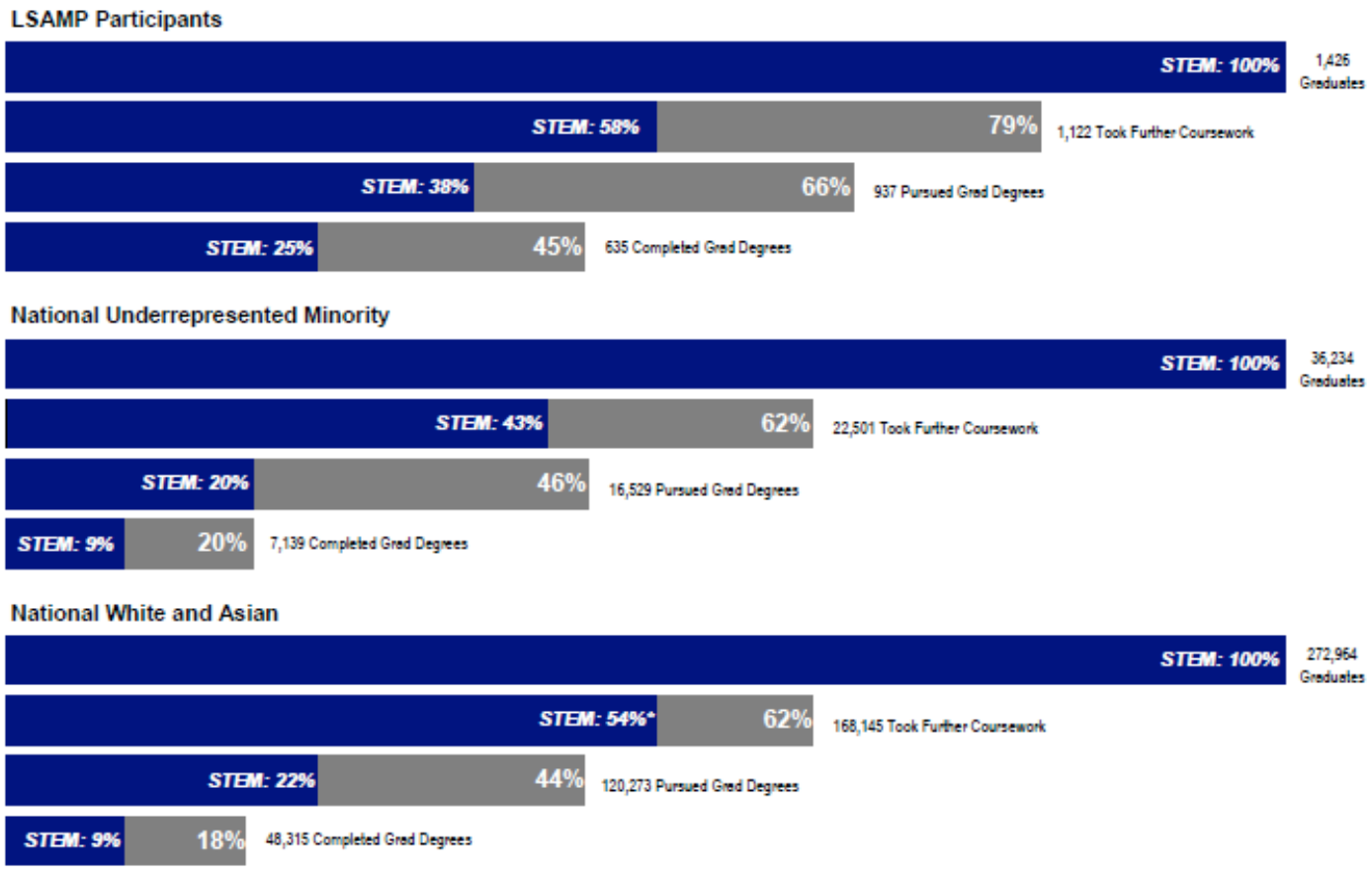
Because the researchers decided before starting the study that they were going to use the national data, they crafted the survey questions to allow for national-level comparisons. Even with the recognized flaws associated with the self-selected LSAMP participants, the use of the comparison data strengthens the case that the LSAMP program did have a positive impact. Without these data, the question of whether the LSAMP students were more or less successful than others in their graduation cohorts in making the transition to graduate school would persist, making it more difficult to prove the value of the program.

Source: Clewell, B. C., C. Cosentino de Cohen, L. Tsui, L. Forcier, E. Gao, N. Young, N. Deterding, & C. West. 2005. Final Report on the Evaluation of the National Science Foundation Louis Stokes Alliances for Minority Participation Program. pg. 47.Urban Institute, Program for Evaluation and Equity Research http://www.urban.org/UploadedPDF/411301_LSAMP_report_appen.pdf. Accessed August 8, 2010, p. 47.

[2]For more information about the National Survey of Recent College Graduates, including reports of findings, go to the National Science Foundation, Science Resource Statistics site at: http://www.nsf.gov/statistics/srvyrecentgrads/.

Figure 2-1

Graduate Coursework, Degrees Pursued and Degrees Completed: LSAMP Participants Compared to National Data



Sources: LSAMP Graduate Survey (UI) and NSRCG longitudinal file (NSF).
*National comparison group statistic is not significantly different from LSAMP.