Reader's Guide

The Condition of Education is available in two forms: this print volume for 2007 and a Web version on the National Center for Education Statistics (NCES) website (http://nces.ed.gov/programs/coe). The Web version includes the following: the 2007 Commissioner's statement, a user's guide, special analyses from 2000 through 2007, all indicators from this edition, and selected indicators from earlier editions of *The Condition of Education*. (See page xxvi for a list of all the indicators that appear on *The Condition of Education* website.)

The print volume of *The Condition of Education 2007* includes the 2007 special analysis and five sections of indicators. Each section begins with a summary of the general topic areas covered by the indicators in the section. Each indicator contains a discussion along with a graph or table on the main indicator page, and one or more supplemental tables found in *appendix 1*. The supplemental tables provide data tables of the estimates used in the indicator discussion as well as additional estimates related to the indicator. Tables of standard errors for applicable estimate tables are available on the Web (http://nces.ed.gov/programs/coe). Additional information on data sources, analyses conducted, and definitions of variables and measures can be found in the supplemental notes in *appendix 2*. Finally, a glossary of key terms, bibliography, and index are provided at the end of the volume.

The "eye" icon on the main indicator page is located to the side of the graph or table and provides references for supplemental notes, supplemental tables, or other source(s) for more information relating to the indicator.

Indicators use the most recent national data available from either NCES or other sources serving the purposes of the indicator. When the source is an NCES publication, such as the *Digest of Education Statistics, 2006* (NCES 2007-017), the publication can be viewed at the NCES website (http://nces.ed.gov/pubsearch).

Data Sources and Estimates

The data in this report were obtained from many different sources, including state education agencies, local schools, and colleges and universities using surveys and compilations of administrative records. Users of *The Condition of Education* should be cautious when comparing data from different sources. Differences in procedures, timing, question phrasing, interviewer training, and so forth can all affect the comparability of results.

Most indicators in *The Condition of Education* summarize data from surveys conducted by NCES or by the Census Bureau with support from NCES. Brief explanations of the major NCES surveys used in this edition of *The Condition of Education* can be found in *supplemental notes 3* and *4* of this volume. More detailed explanations can be obtained at the NCES website (http://nces.ed.gov) under "Surveys and Programs." Information about the Current Population Survey (CPS), another frequent source of survey data used in *The Condition of Education*, can be obtained in *supplemental note 2* and also at http://www.census.gov/cps/.

Data for indicators reported in this volume are primarily from two types of surveys: universe surveys and sample surveys. First, some

indicators report data from entire populations (universe surveys), such as *indicator 40* (Public Elementary and Secondary Expenditures by District Poverty). With this type of survey, information is collected from every member of the population. For example, data for *indicator 40* was obtained for each school district (approximately 17,000) in the United States. When data from an entire population are available, estimates of the total population or a subpopulation are made by simply summarizing the units in the population or subpopulation. A universe survey is usually expensive and time consuming, so researchers often collect data from a sample of the population of interest (sample survey). Other indicators report data from such sample surveys, such as *indicator 14* (Trends in the Achievement Gaps in Reading and Mathematics). *Indicator 14* reports information from the National Assessment of Educational Progress (NAEP), which assesses a representative sample of students each year, rather than the entire population of students. When a sample survey is used, the statistical uncertainty introduced from having data from only a portion of the entire population must be considered in reporting estimates and making comparisons.

Various types of estimates are reported in *The Condition of Education*. Many indicators report the size of a population or a subpopulation, and often the size of a subpopulation is expressed as a percentage of the total population. In addition, the average (or mean) values of some characteristic of the population or subpopulation may be reported. The average is obtained by summing the values for all members of the population and dividing the

sum by the size of the population. An example is the annual average salaries of professors at degree-granting institutions (*indicator 44*). Still another population measure sometimes used is the median. The median is the value of a population characteristic above which 50 percent of the population is estimated to fall. An example is the median annual earnings of young adults who are full-time, full-year wage and salary workers (*indicator 20*).

Estimates based on universe and sample survey data may be affected by a wide range of potential data collection errors, such as coverage errors, response errors, coding errors, and data entry errors. Estimates of the size of these errors are typically not available.

Estimates calculated from data based on a sample of the population requires consideration of several factors before the estimates become meaningful. However conscientious an organization may be in collecting data from a sample of a population, there will always be some margin of error in estimating the size of the actual total population or subpopulation because the data are available from only a portion of the total population. Consequently, data from samples can provide only an estimate of the true or actual value. The margin of error, or the range, of an estimate depends on several factors, such as the amount of variation in the responses, the size and representativeness of the sample, and the size of the subgroup for which the estimate is computed. The magnitude of this margin of error is measured by what statisticians call the "standard error" of an estimate.

Reader's Guide

Continued

STANDARD ERRORS

When data from samples are reported, as is the case with most of the indicators in *The Condition of Education*, the standard error is calculated for each estimate. The standard errors for all estimated totals, means, medians, or percentages reported in the supplemental tables of *The Condition of Education* can be viewed at the NCES website (http://nces.ed.gov/programs/coe).

The standard errors of the estimates for different subpopulations in an indicator can vary considerably. As an illustration, *indicator 18* reports on the adult literacy scores of adults age 16 or older in the United States in 2003. The average quantitative scores of adults who spoke only English and those who spoke English and a language other than Spanish were each 289 (see supplemental table 18-1). In contrast to the similarity of these scores, their standard errors were 1.2 and 4.1, respectively (see table S18-1 in http://nces.ed.gov/programs/coe/2007/section2/table.asp?tableID=757). The average score with the smaller standard error provides a more reliable estimate of the true value than does the average score with a higher standard error. Standard errors tend to diminish in size as the size of the sample (or subsample) increases. Consequently, for the same kinds of data, such as reading, mathematics, and science scores on the National Assessment of Educational Progress (*indicators 11, 12*, and *13*), standard errors will almost always be larger for Blacks and Hispanics than for Whites, who represent a larger proportion of the population.

For *indicator 20*, which reports median annual earnings, special procedures are followed for computing the standard errors for these medians. See *appendix G* of the source and accuracy statement for the Current Population Study (CPS) 2006 Annual Social and Economic supplement (ASEC) for information on how to calculate the standard errors (http://www.census.gov/apsd/techdoc/cps/cpsmar06.pdf).

DATA ANALYSIS AND INTERPRETATION

Due to standard errors, caution is warranted when drawing conclusions about the size of one population estimate in comparison to another or about whether a time series of population estimates is increasing, decreasing, or staying about the same. Although one estimate may be larger than another, a statistical test may find that there is no measurable difference between the two estimates because of a large standard error associated with one or both of the estimates. Whether differences in means or percentages are statistically significant can be determined using the standard errors of the estimates.

Readers who wish to compare two sample estimates to see if there is a statistical difference will need to estimate the precision of the difference between the two sample estimates. This would be necessary if one wanted to compare, for example, the mean proficiency scores between groups or years in the National Assessment of Educational Progress. To estimate the precision of the difference between two sample estimates, one must find the standard error of the difference between the two sample estimates (sample estimate A, or E_A, and sample estimate B, or E_B). Expressed mathematically, the difference between the two estimates E_A and E_B is $E_A - E_B$.

The standard error of the difference (or se_{A-B}) can be calculated by taking the square root of the sum of the two standard errors associated with each of the two sample estimates (se_A and se_B) after each has been squared. This can be expressed as

$$se_{A-B} = \sqrt{se_A^2 + se_B^2}$$

After finding the standard error of the difference, one divides the difference between the two sample estimates by this standard error

to determine the "*t* value," or "*t* statistic," of the difference between the two estimates. This *t* statistic measures the precision of the difference between two independent sample estimates. The formula for calculating this ratio is expressed mathematically as

$$t = \frac{E_A - E_B}{se_{A-B}}$$

The next step is to compare this *t* statistic to 1.96, which is a statistically determined criterion level for making a decision as to whether there is a difference between the two estimates. If the *t* statistic is greater than 1.96, then there is evidence that there is a difference between the two populations. Note that one cannot say for certain that the two estimates are different, only that there is evidence that the difference in estimates is not due to sampling error alone. If the *t* statistic is equal to or less than 1.96, then one is less certain that the observed difference is not due to sampling error alone. This level of certitude, or significance, is known as the ".05 level of (statistical) significance."

As an example of a comparison between two sample estimates to determine whether there is a statistically significant difference between the two, consider the data on the performance of 12th-grade students in the reading assessment of the 1992 and 2005 National Assessment of Educational Progress (see supplemental table 11-1). The average scale score in 1992 was 292, and the average scale score in 2005 was 286. Is the difference of 6 scale points between these two different samples statistically significant? The standard errors of these estimates are 0.6 and 0.6, respectively (see standard error table S11-1 on the NCES website). Using the formula above, the standard error of the difference is 0.85. The *t* statistic of the estimated difference of 6 scale points to the standard error of the difference is 7.07. This value is greater than

1.96—the critical value of the *t* distribution for a 5 percent level of significance with a large sample. Thus, one can conclude that there was a statistically significant difference in the performance of 12th-graders between 1992 and 2005 in reading and that the reading score for 12th-graders in 2005 was lower than the reading score for 12th-graders in 1992.

For all indicators reporting estimates based on samples in *The Condition of Education*, differences between estimates (including increases or decreases) are stated only when they are statistically significant. To determine whether differences reported are statistically significant, two-tailed *t* tests, at the 0.05 level, are typically used. The *t* test formula for determining statistical significance is adjusted when the samples being compared are dependent. When the difference between estimates is not statistically significant, tests of equivalence will often be run. An equivalence test determines the probability (generally at the .15 level) that the estimates are statistically equivalent; that is, within the margin of error that the two estimates are not substantively different. When the difference is found to be equivalent, language such as x and y "were similar" or "about the same" has been used; otherwise, the data will be described as having no measurable difference.

When the variables to be tested are postulated to form a trend, the relationship may be tested using linear regression, logistic regression, or ANOVA trend analysis instead of a series of *t* tests. These other methods of analysis test for specific relationships (e.g., linear, quadratic, or cubic) among variables.

A number of considerations influence the ultimate selection of data years to feature in *The Condition of Education*. To be as timely as possible, the latest year of data is shown if available during report production. The choice

Reader's Guide

of comparison years is based on the need to show the earliest survey year, as in the case of the National Assessment of Educational Progress and the international assessment surveys. In the case of surveys with long time frames, such as for enrollment, the decade's beginning year (e.g., 1980 or 1990) starts the trend line. Intervening years are selected in increments to show the general trend in the figures and tables. The narrative for the indicators typically compares the most current year's data with those from the initial year and then with those from the recent period. The narrative may also note the years in which the data begin to diverge from previous trends.

VARIATIONS IN POPULATIONS

In considering the estimates in the tables and figures shown in this volume and on the NCES website, it is important to keep in mind that there may be considerable variation among the members of a population in the characteristic or variable represented by the population estimate. For example, the estimated average mathematics reasoning score of 4th-graders in the United States in 2003 was 519 (see supplemental table 17-1). In reality, many students scored above 519 points, and many scored below 519 points. Likewise, not all faculty salaries, benefits, and total compensation at postsecondary institutions were the same at each type of institution in 2005–06 (*indicator 44*). Because of this variation, there may be considerable overlap among the members of two populations that are being compared. Although the difference in the estimated means of the two populations may be statistically significant, many members of the population with the lower estimated mean may be above the estimated mean of the other population, and vice versa. For example, some percentage of young adults with a high school diploma or equivalent have higher earnings than young adults with a bachelor's degree or higher (*indicator 20*). The extent of such overlap is not generally considered in the indicators in this volume. Estimates of the extent of variation in such population characteristics can be computed from the NCES survey datasets or are available in published reports. For example, estimates of the variation in students' assessment scores can be found using the NAEP Data Explorer at http://nces.ed.gov/nationsreportcard/nde/ or in the appendixes to most NAEP reports.

ROUNDING AND OTHER CONSIDERATIONS

Although values reported in the supplemental tables are generally rounded to one decimal place (e.g., 76.5 percent), values reported in each indicator are rounded to whole numbers (with any value of 0.50 or above rounded to the next highest whole number). Due to rounding, cumulative percentages may sometimes equal 99 or 101 percent, rather than 100 percent.

In accordance with the NCES Statistical Standards, many tables in this volume use a series of symbols to alert the reader to special statistical notes. These symbols, and their meaning, are as follows:

— Not available.
Data were not collected or not reported.

† Not applicable.
Category does not exist.

\# Rounds to zero.
The estimate rounds to zero.

! Interpret data with caution.
Estimates are unstable.

‡ Reporting standards not met.
Did not meet reporting standards.

* $p < .05$ Significance level.[1]

NOTES

[1] The chance that the difference found between two estimates when no real difference exists is less than 5 out of 100.