Introduction

In 2018, the National Science Foundation supported the sixth in a series of surveys through a grant to Horizon Research, Inc. (HRI). The first survey was conducted in 1977 as part of a major assessment of science and mathematics education and consisted of a comprehensive review of the literature; case studies of 11 districts throughout the United States; and a national survey of teachers, principals, and district and state personnel. A second survey of teachers and principals was conducted in 1985–86 to identify trends since 1977. A third survey was conducted in 1993, a fourth in 2000, and a fifth in 2012. This series of studies has been known as the National Survey of Science and Mathematics Education (NSSME+).

The 2018 iteration of the study included an emphasis on computer science, particularly at the high school level, which is increasingly prominent in discussions about K–12 STEM education and college and career readiness. The 2018 NSSME+ (the plus symbol reflecting the additional focus) was designed to provide up-to-date information and to identify trends in the areas of teacher background and experience, curriculum and instruction, and the availability and use of instructional resources. The research questions addressed by the study are:

1. To what extent do computer science, mathematics, and science instruction reflect what is known about effective teaching?
2. What are the characteristics of the computer science/mathematics/science teaching force in terms of race, gender, age, content background, beliefs about teaching and learning, and perceptions of preparedness?
3. What are the most commonly used textbooks/programs, and how are they used?
4. What influences teachers' decisions about content and pedagogy?
5. What formal and informal opportunities do computer science/mathematics/science teachers have for ongoing development of their knowledge and skills?
6. How are resources for computer science/mathematics/science education, including well-prepared teachers and course offerings, distributed among schools in different types of communities and different socioeconomic levels?

Data for the study come from six instruments:

School-level questionnaires

1. School Coordinator Questionnaire;
2. Mathematics Program Questionnaire;
3. Science Program Questionnaire;

Teacher-level questionnaires

4. High School Computer Science Teacher Questionnaire;[1]
5. Mathematics Teacher Questionnaire; and
6. Science Teacher Questionnaire.

The design and implementation of the 2018 NSSME+ involved developing a sampling strategy and selecting samples of schools and teachers, developing and piloting survey instruments, collecting data from sample members, and preparing data files and analyzing the data. These activities are described in the following sections. The final section of this chapter outlines the contents of the remainder of the report.

Sample Design and Sampling Error Considerations

The 2018 NSSME+ is based on a national probability sample of schools and science, mathematics, and computer science teachers in grades K–12 in the 50 states and the District of Columbia. The sample was designed to yield national estimates of course offerings and enrollment, teacher background preparation, textbook usage, instructional techniques, and availability and use of facilities and equipment. Every eligible school and teacher in the target population had a known, positive probability of being sampled.

The sample design involved clustering and stratification prior to sample selection. The first stage units consisted of elementary and secondary schools. Science, mathematics, and computer science teachers constituted the second stage units. The target sample sizes were designed to be large enough to allow sub-domain estimates, such as for particular regions or types of community.

The sampling frame for the school sample was constructed from the Common Core of Data and Private School Survey databases—programs of the U.S. Department of Education's National Center for Education Statistics—which include school name and address and information about the school needed for stratification and sample selection. The sampling frame for the teacher sample was constructed from lists provided by sample schools, identifying current teachers and the specific science, mathematics, and computer science subjects they were teaching.

Because biology is by far the most common science course at the high school level, selecting a random sample of science teachers would result in a much larger number of biology teachers than chemistry or physics teachers. Similarly, random selection of mathematics teachers might result in a smaller than desired sample of teachers of advanced mathematics courses. In order to ensure that the sample would include a sufficient number of advanced science and mathematics teachers for separate analysis, information on teaching assignments was used to create separate domains (e.g., for teachers of chemistry and physics), and sampling rates were adjusted by domain. In addition, because the number of computer science teachers in high schools is small compared to the number of science and mathematics teachers, all high school teachers who taught computer science were sampled for that subject.

[1] Based on the recommendation of the project's Advisory Board, high school computer science was defined for this study as courses that teach programming or have programming as a prerequisite.

The study design included obtaining in-depth information from each teacher about curriculum and instruction in a single, randomly selected class. Most elementary teachers were reported to teach in self-contained classrooms; i.e., they were responsible for teaching all academic subjects to a single group of students. Each such sampled teacher was randomly assigned to 1 of 2 groups—science or mathematics—and received a questionnaire specific to that subject. Most secondary teachers in the sample taught several classes of a single subject. Some secondary teachers taught multiple subjects addressed by the study. If such a teacher taught high school computer science, s/he was selected to respond to the computer science questionnaire; if s/he taught science and mathematics, s/he was randomly assigned to receive the science or mathematics teacher questionnaire. In addition, for all teachers responsible for more than one class in their designated subject area, one class was randomly selected.

Whenever a sample is anything other than a simple random sample of a population, the results must be weighted to take the sample design into account. In the 2018 NSSME+, the weight for each respondent was calculated as the inverse of the probability of selecting the individual into the sample multiplied by a non-response adjustment factor.[2] In the case of data about a randomly selected class, the teacher weight was adjusted to reflect the number of classes taught in that subject, and therefore, the probability of a particular class being selected. Detailed information about the sample design, weighting procedures, and non-response adjustments used in the 2018 NSSME+ is included in Appendix A.

The results of any survey based on a sample of a population (rather than on the entire population) are subject to sampling variability. The sampling error (or standard error) provides a measure of the range within which a sample estimate can be expected to fall a certain proportion of the time. For example, it may be estimated that 7 percent of all elementary mathematics lessons involve the use of computers. If it is determined that the sampling error for this estimate was 1 percent, then according to the Central Limit Theorem, 95 percent of all possible samples of that same size selected in the same way would yield computer usage estimates between 5 percent and 9 percent (that is, 7 percent ± 2 standard error units).

In survey research, the decision to obtain information from a sample rather than from the entire population is made in the interest of reducing costs, in terms of both money and the burden on the population to be surveyed. The particular sample design chosen is the one that is expected to yield the most accurate information for the least cost. It is important to realize that, other things being equal, estimates based on small sample sizes are subject to larger standard errors than those based on large samples. Also, for the same sample design and sample size, the closer a percentage is to zero or 100, the smaller the standard error. The standard errors for the estimates presented in this report are included in parentheses in the tables. The narrative sections of the report generally point out only those differences that are substantial as well as statistically

[2] The aim of non-response adjustments is to reduce possible bias by distributing the non-respondent weights among the respondents expected to be most similar to these non-respondents. In this study, adjustment was made by region, school metro status, grade level, type (public, catholic, other private), and student body race/ethnicity.

significant at the 0.05 level.[3] All population estimates presented in this report were computed using weighted data.

Instrument Development

Because one purpose of the 2018 NSSME+ was to identify trends in science and mathematics education, the process of developing survey instruments began with the questionnaires that were used in the 2012 NSSME. The project's Advisory Board, composed of experienced researchers in computer science, science, and mathematics education, reviewed the 2012 questionnaires and made recommendations about retaining or deleting particular items. Additional items that were needed to provide important information about the current status of computer science, science, and mathematics education were also considered.

Preliminary drafts of the questionnaires were sent to the professional organizations that endorsed the study for review (listed in Appendix B); these included the American Federation of Teachers, the Computer Science Teachers Association, the National Council of Teachers of Mathematics, the National Education Association, and the National Science Teachers Association.

The survey instruments were revised based on feedback from the various reviewers, field tested, and revised again. The instrument development process was lengthy, constantly compromising between information needs and data collection constraints. There were several iterations, including rounds of cognitive interviews with teachers and revisions to help ensure that individual items were clear and unambiguous and that the survey as a whole would provide the necessary information with the least possible burden on participants. Lastly, because of the large number of questions stakeholders (e.g., advisors, endorsers) wanted to include in the study, all teachers sampled for science or mathematics teacher responded to a core set of items plus 1 of 3 sets of items randomly assigned to respondents. The relatively small sample size of high school computer science teachers would not support random assignment of items, thus these teachers were presented only with core items. Copies of the questionnaires are included in Appendix C.

Data Collection

HRI secured permission for the study from education officials at various levels. First, notification letters were mailed to the Chief State School Officers. Similar letters were subsequently mailed to superintendents of districts including sampled public schools and diocesan offices of sampled Catholic schools, identifying the schools in the district that had been selected for the survey. (Information about this pre-survey mail-out is included in Appendix B.) Copies of the survey instruments and additional information about the study were provided when requested.

Principals received a mailing asking them to log on to the study website and designate a school contact person or "school coordinator." The school coordinator designation page was designed to confirm the principal's contact information, as well as to obtain the name, title, phone number, and email address of the coordinator. (The mailing also included a printed copy of the form and postage-paid return envelope.) Of the 2,000 target slots, 1,273 schools were successfully

[3] The False Discovery Rate was used to control the Type I error rate when comparing multiple groups on the same outcome. Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, B, *57*, 289–300.

recruited; 41 slots were ineligible (e.g., the school had closed, should have been excluded from the sampling frame, merged with another school already in the sample). Thus, 65 percent of eligible slots were filled.

An incentive system was developed to encourage school and teacher participation in the survey. School coordinators were offered an honorarium of up to $200 ($100 for completing a teacher list and school questionnaire, $15 for completing each program questionnaire (optional), and $10 for each completed teacher questionnaire). Teachers were offered a $25 honorarium for completing the teacher questionnaire.

Survey invitation letters were mailed to teachers beginning in February 2018. In addition to the incentives described, phone calls and emails to school coordinators were used to encourage non-respondents to complete the questionnaires. In May 2018, a final questionnaire invitation mailing was sent to teachers who had not yet completed their questionnaires. The teacher response rate was 78 percent. The response rate for the school-level questionnaires was 86 percent. A detailed description of the data collection procedures is included in Appendix B.

Outline of This Report

This report of the 2018 NSSME+ is organized into major topical areas. In most cases, results are presented for by grade level—elementary, middle, and high.[4,5] Factor analysis was used to create several composite variables related to key constructs measured on the questionnaires. Composite variables, which are more reliable than individual survey items, were computed to have a minimum possible value of 0 and a maximum possible value of 100. The definitions of these and other reporting variables used in this report are included in Appendix D.

Chapter Two focuses on teacher backgrounds and beliefs. Basic demographic data are presented along with information about course background, perceptions of preparedness, and pedagogical beliefs. Chapter Three examines data on the professional status of teachers, including their opportunities for continued professional development.

Chapter Four presents information about the time spent on science and mathematics instruction in the elementary grades and about course offerings at the secondary level. Chapter Five examines the instructional objectives and the activities used to achieve these objectives, followed by a discussion of the availability and use of various types of instructional resources in Chapter Six. Finally, Chapter Seven presents data about a number of factors that are likely to affect science, mathematics, and computer science instruction, including school-wide programs, practices, and problems.

In addition, each chapter contains a set of analyses that examine the distribution of key outcomes across schools and classes of different demographic characteristics. For these analyses, data from the school-level questionnaires are examined by four factors:

---

[4] The computer science teacher questionnaire was administered only to high school teachers; thus, results from this survey are shown only for high school. In addition, because it was not possible to matrix sample items on this questionnaire, some questions asked of science and mathematics teachers could not be asked of computer science teachers in order to keep response burden reasonable.

[5] Results by grade range for all applicable items can be found in Craven, L. M., Bruce, A. D., and Plumley, C. L. (2019). *The 2018 NSSME+ compendium of tables.* Chapel Hill, NC: Horizon Research, Inc.

1. Percentage of students in the school eligible for free/reduced-price lunch,
2. School size,
3. Community type, and
4. Region.

Data from the teacher questionnaires are examined by an additional two factors based on the randomly selected class:

1. Prior achievement level of students, and
2. Percentage of students in the class from race/ethnicity groups historically underrepresented in STEM fields.[6]

Additional information about these factors is included in Appendix D. Although the specific equity factors displayed in the body of the report vary by outcome, tables showing each examined outcome by all relevant equity factors are included in Appendix E.

[6] It is important to note that high school computer science classes tend to have many fewer students from these groups than science and mathematics classes. Consequently, the highest quartile of this variable for high school computer science is defined as the class having more than 39 percent of its students from a race/ethnicity group historically underrepresented in STEM compared to more than 76.9 and 76.2 percent in science and mathematics, respectively.