MESSAGE FROM THE PRESIDENTS

A Data-Based Assessment of Research-Doctorate Programs in the United States (NRC, 2010) and its large collection of quantitative data will become, in our view an important and transparent instrument for strengthening doctoral education in the United States. The report follows in the tradition of assessments conducted by the National Research Council for almost 30 years, but with important changes. Beyond the traditional, printed document, the data that inform and grow out of this report are being made available electronically to promote widespread use and analysis of many characteristics of doctoral programs. The unparalleled data set covers twenty important variables for an enormous number of programs in 62 major fields. It enables university faculty, administrators, and funders to compare, evaluate and improve programs; it permits students to find those programs best suited to their needs; and it allows for updating important information on a regular basis to permit continuous improvement.

Much has been learned from this study, which turned out to be more challenging and to take longer than we originally expected. An enormous effort was contributed by universities to collect and recheck the data, demonstrating their desire to identify comparative strengths and weaknesses and to show accountability. The study committee had to refine and revise its methodology as it sought to provide tools for evaluating and comparing programs. Although the data are based on the 2005-2006 academic year, they permit many useful comparisons of programs across many dimensions. All those interested in graduate education can learn much from studying the data, comparing programs, and drawing lessons for how programs can be improved. The data for many variables can be updated and made current on a regular basis by universities.

In order to identify variables most valued by doctoral faculty as well as to avoid using exclusively reputational rankings as was done in earlier graduate doctorate assessments, the committee employed two alternative ranking methods. The first method asked faculty in each field to assign a weight to each of the quantitative variables in the institutional surveys, and the weighted variables could then be used to determine ratings and rankings of programs. The second method was to survey a subset of faculty to ask them to rank a sample of programs in their field, and then to use principal components and regression analyses to obtain the implied weights for the institutional variables that would most closely reproduce the results. The committee initially envisioned combining the results of these two methods into a unified set of rankings. The production of rankings from measures of quantitative data turned out to be more complicated and to have greater uncertainty than originally thought. The committee ultimately concluded that it should present the results of the two approaches separately as illustrations of how individuals can use the data to apply their own values to the quantitative measures to obtain rankings suitable for their own specific purposes. The illustrative rankings, which are provided with ranges to show some of the statistical uncertainties, should not be interpreted as definitive conclusions about the relative quality of doctoral programs. Doctoral programs

are valued for a variety of reasons, and their characteristics are valued in different ways by stakeholders; there is no single universal criterion or set of criteria.

The illustrative rankings and their ranges do provide important insights on how programs can be ranked according to different criteria and what variables are most important to faculty, which typically are variables that measure per capita scholarly output. Faculty generally do not assign great importance to program size when assigning weights directly -- but when they rank programs, program size appears to implicitly carry large weight. It is our view that strengthening graduate education will require paying attention to all of the variables in the dataset, not just those most important to faculty. Three additional metrics were presented in the report for each program; these focused separately on research activity, student support and outcomes, and diversity of the academic environment. A major value of the study is that this data set allows all stakeholders to assign weights which they believe to be important and then compare the programs on that basis.

If a process of continuous improvement is to result from this exercise, all of the stakeholders interested in graduate education will need to focus upon steps to improve performance across the board. A major commitment by universities will be needed to update the data set on a regular basis, so that programs can continue to be compared and evaluated. If this is done with the updated dataset as an important new tool, and we strive to improve what is already the world's strongest system of higher education, we believe that American doctoral education can continue to bring enormous benefits to our citizens and remain the envy of the world.

Ralph J. Cicerone, President, National Academy of Sciences
Charles M. Vest, President, National Academy of Engineering
Harvey V. Fineberg, President, Institute of Medicine