SOCIOLOGY

"It's hard to think of a more prescient moment to interrogate the role of science, knowledge, and higher education globally. An essential read."
—CYNTHIA MILLER-IDRISS, American University

"A tour-de-force combination of empirics and theorizing rendered with concision and wit."
—MITCHELL L. STEVENS, Stanford University

"*Global Mega-Science* is thoroughly detailed and remarkably expansive, blending empirics and theory in a comparative-historical study of science, universities, globalization, and world society."
—GILI S. DRORI, Hebrew University of Jerusalem

"This wonderful, timely book on the truly global character of science and higher education interweaves rich quantitative data with a rare sense for storytelling."
—GEORG KRÜCKEN, INCHER-Kassel

"*Global Mega-Science* chronicles the incredible pace of scientific discovery and collaboration, revealing a dynamic, globally interconnected scientific landscape. A must-read."
—KATHRIN ZIPPEL, Freie Universität Berlin

"Baker and Powell's grand tour through Germany, the US, China, and across the globe describes an explosion of science. Their brilliant analysis of university-based science brings startling clarity."
—EVAN SCHOFER, University of California, Irvine

NEVER HAS THE WORLD BEEN AS RICH IN SCIENTIFIC KNOWLEDGE as it is today. But what are its main sources? *Global Mega-Science* examines the origins of this unprecedented growth of knowledge. David Baker and Justin Powell argue that at the heart of this phenomenon is the unparalleled cultural success of universities and their connection to science: the university-science model. The age of global mega-science was built upon the symbiotic relationship between higher education and science, especially the worldwide research collaborations among networked university-based scientists. These relationships are key for scholars and citizens to understand the future sustainability of science.

DAVID P. BAKER is Professor of Sociology, Education, and Demography at Pennsylvania State University and the 2023–24 George Sarton Chair for the History of Science at Ghent University, Belgium.

JUSTIN J.W. POWELL is Professor of Sociology of Education at the University of Luxembourg.

Cover design: Martyn Schmoll
Cover art: Joel Filipe / Unsplash and iStock

Stanford University Press
www.sup.org

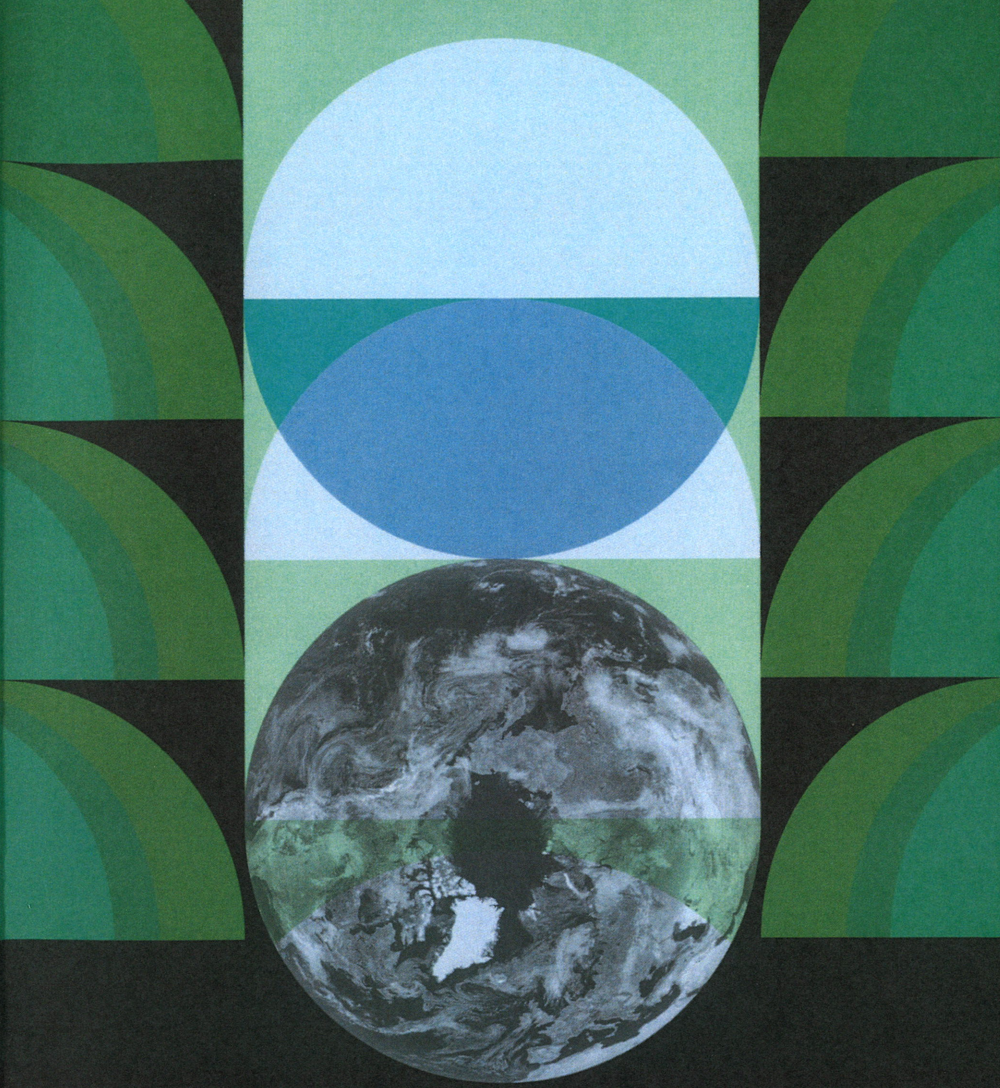
ISBN 978-1-5036-3789-4
90000
9 781503 637894

BAKER
POWELL

Global Mega-Science

Universities, Research Collaborations, and Knowledge Production

Stanford



Global Mega-Science

Universities, Research Collaborations, and Knowledge Production

David P. Baker and **Justin J.W. Powell**

GLOBAL MEGA-SCIENCE

UNIVERSITIES, RESEARCH COLLABORATIONS, AND KNOWLEDGE PRODUCTION

David P. Baker and Justin J.W. Powell

Stanford University Press
Stanford, California

Stanford University Press
Stanford, California

Printed in the United States of America on acid-free, archival-quality paper

Library of Congress Cataloging-in-Publication Data

Names: Baker, David, 1952 January 5– author. | Powell, Justin J.W., author.
Title: Global mega-science : universities, research collaborations, and knowledge production / David P. Baker and Justin J.W. Powell.
Description: Stanford, California : Stanford University Press, 2024. | Includes bibliographical references and index.
Identifiers: LCCN 2023038227 (print) | LCCN 2023038228 (ebook) | ISBN 9781503602052 (cloth) | ISBN 9781503637894 (paperback) | ISBN 9781503639102 (ebook)
Subjects: LCSH: Science—History. | Science—Social aspects—History. | Research—History. | Universities and colleges—History. | Education, Higher—Social aspects—History.
Classification: LCC Q125 .B26 2024 (print) | LCC Q125 (ebook) | DDC 509—dc23/eng/20231026
LC record available at https://lccn.loc.gov/2023038227
LC ebook record available at https://lccn.loc.gov/2023038228

Cover design: [designer]
Cover [image/photo/illustration]: [credit from database]

Contents

Preface. A Journey Across the Century of Science 1

1 Professor Price's Error: The Rise of Global Mega-Science 17

2 Talcott's Prediction: Why a Century of Science? 35

3 Göttingen and Beyond: The Ascendant German Research University 53

4 Modest Origins: The Expansive American University-Science Model 73

5 C. N. Yang's Children: Globalize or Fade 93

6 The Theologian's Institutes: A Culture of Scientific Genius as Counterfactual 107

7 It's Simple Engineering: Pursuing the World-Class University in East Asia 123

8 Mega-Science Goes Global: Investing in the Twenty-First-Century University 141

9 Chasing Neutrinos Through Networks of Science: University Collaborations and Scientization 155

10 Conclusion: Global Mega-Science, Universities, and Their Joint Future 173

Acknowledgments 199

Notes 203

Bibliography 223

Index 243

PREFACE

A Journey Across the Century of Science

This book was finished during the throes of the COVID-19 pandemic. As the world's media covered the terrifying emergence of a novel coronavirus and efforts to stop its global spread, another, originally less noticed, story has become part of the current zeitgeist. The spread of the pandemic was matched by the tremendous pace at which scientists generated new knowledge about the virus: its biology, contagiousness, therapeutic treatments, and the holy grail—effective vaccines.[1] Just six months into the pandemic, scientists had already published results from more than twenty-three thousand studies of the virus; this flow of new discovery had doubled every twenty days—certainly among the biggest explosions of research ever on a specialized scientific topic.[2] The devastation from the virus and mounting scientific knowledge about its origins and impacts are still ongoing.[3]

Obviously, the pandemic was highly motivating for the world's scientists, encompassing many disciplines and specialties, far beyond health. Yet this incredible scientific response did not come about by chance, and it could have been anticipated. Often missed in the coverage of this impressive response is the now massive and continuously expanding reach and interconnectedness of the world's scientific infrastructure, built over decades. This infrastructure serves as the backbone of an astonishing global capacity to undertake focused, frequently collaborative research at an unprecedented pace—now not only about a threatening disease but on literally thousands of topics.

Consequently, in 2020 alone, the world's scientists published over three million articles about their studies in more than nine thousand leading scientific and technical journals! These research articles, or "papers" in the everyday jargon of scientists, are where new discoveries, minute to monumental, are vetted and communicated. The sheer volume of papers reflects an extensive capacity for new discovery, increasingly done by globe-spanning networks of scientists carrying out research that continuously builds upon each other's findings. The pace of discovery and boundary-crossing worldwide collaboration are quintessential dimensions of what can be called *mega-science*. This exceptional change in the scope and dimensions of science—evolving for a long time, but not fully evident until recently—has inexorably transformed, for better or worse, the volume, breadth, and depth of knowledge production. The consequences of this vast growth and unparalleled collaboration for scientific discovery drive the centrality of science ever deeper into human societies in all regions of the world.

Yet mega-science was not supposed to happen, or at least not anywhere near to this level. As recently as the 1980s, prominent science-watchers predicted an end to the rising pace of science; it was, they thought then, completely unsustainable. Some pundits even went so far as to forecast a severe enough reduction in the world's capability for research to trigger a depression of the global economy. What occurred afterwards would have amazed them. Why their predictions were so wildly incorrect is because of a persistent misunderstanding of what facilitated the coming of global mega-science emerging at the turn of the twentieth century. Generalities about government policies and spending, geopolitical struggles, armaments, space races, economic demand, technological breakthroughs, and pressing societal crises are frequently trotted out as explanations.[4] And while these are certainly part of the story, none have the immediacy or consistent historical presence to be *the* foundation for what has transpired. They mostly play supporting roles, maybe necessary but not sufficient on their own to have set the far-reaching stage for global mega-science.

Instead, as argued and explored here, an often-ignored, yet major force behind science is the globally unfolding phenomenon known as the *education revolution* and its long-coming transformation of universities and their relationship to society, including science. This cultural process lies behind not only more people attending schooling for ever-longer phases of their lives than ever before but also, throughout the twentieth century and onwards, a steady inclusion of greater numbers of youth and young adults in universities and other postsecondary organizations across the world. In and of itself though, growing attendance alone could not have been the midwife to the mega-science we witness today. Also required was a concurrent influence of the culture of the education revolution on the very essence of what it means to be a university—and how this organizational form provides the crucial, well-resourced forum for the exchange of ideas and research for everyone devoted to scientific discovery.[5]

Over the same period, along with changing ideas about education in general, came the notion that universities could, and should, be places for generating new knowledge, science included. Of course, scholarship has always been part of the eight-hundred-plus-year history of the Western university, although mostly in a restrained fashion, with scientific research long considered beneath the lofty classical university of philosophy, theology, medicine, and law. At the middle of the nineteenth century something changed, first barely noticeable at a handful of universities, and thereafter, although contested and sporadic, spreading to universities across the world. Not only did universities become the organizational platform for teaching and research, but they also increasingly expressed their cultural power in recognizing and giving definition and boundaries to new areas of science, reinforced by discipline-based training and degrees in specific curricular areas. The cultural change born out of the education revolution and its new type of university was a melding together of the human energy and societal backing involved in expanding advanced education with a platform for the nurturing of scientific research. Of those approximately three million scientific journal articles in 2020,[6] the vast majority existed due to contributions of university-based scientists, frequently the sole source of even collaborative research efforts.[7] If earlier prognosticators of science had appreciated the maturing education revolution and its support of high-powered research occurring at most universities, they might not have missed the further explosion of mega-science that, even then, was brewing right under their noses.[8]

Usually, observations of a connection between education and science stop at noting that the former spreads scientific literacy and trains the talented few to become scientists. Required to sit through typical science courses in school, most people are more scientifically literate than were their grandparents or parents, and there are increasing opportunities to gain advanced scientific training. But this does not reflect the impact of the spread and growing intensity of science, trivializing what has occurred.

Alternatively, what we will refer to in shorthand as the *university-science model* formed a very potent and self-reinforcing connection between mass education and science, manifested in the symbiotic relationship between the university and science, without which it is unlikely that mega-science could ever have spread successfully to all continents and nearly every country.

Not long after the faulty forecasts of stagnant science, yet another tribe of science-watchers predicted that the university would lose its relevance to future science production. The notion was that many kinds of nonuniversity organizations, including science-based industries, would increasingly take over the knowledge-production enterprise, and thus universities would not keep pace in the science game. While the former turned out to be true, the latter did not. Even in an expanded environment with scientists working in many types of organizations, the university remains the powerful heart of mega-science, pumping out fresh ideas throughout society and providing the main platform upon which to research them. This is evidenced by the fact that, among a growing trend of studies from nonacademic organizations, such as businesses, most scientists in these organizations contribute to the research base by collaborating with university-based colleagues. Even the Internet itself was originally designed in universities to facilitate communication and collaboration among scientists across the world. And as will be shown, this university-based science has been cross-subsidized by the expanding role of advanced education within world society and the mass enrollments in postsecondary education in country after country that follow.

How Real Is Mega-Science?

Too often, reactions to the growth in volume of scientific papers dismiss it as merely reflecting hyperinflation in publication rather than real advances in discovery. Or some say that much of the rising volume of new discoveries could not have commensurate scientific value.[9] And there seems to be no shortage of new gloom-and-doom, end-of-science predictions, often based on scant empirical evidence. These knee-jerk reactions are untrue, especially considering that the incremental advance and development of fields over long periods of time is also essential to address current and future challenges, anticipated or not. Indeed, the present volume of papers directly corresponds to a concurrent explosion in the world's capacity to generate research. Employed scientists and researchers in all fields in wealthier countries tripled from 1980 to 2015 and grew worldwide by 50 percent over the past four decades.[10] While recently scientists have certainly been motivated to publish more, the fact that significantly more of them have been researching in more universities and other research-producing organizations over the century yields the steady increase in the rate of the world's output. Since the early twentieth century, for example, the number of graduate programs in science and technology and their volume of newly minted PhD scientists in the U.S. alone doubled many times. By 2015, about twenty-five thousand new PhDs in the sciences and related fields were graduating from the hundreds of research-active universities in the United States—every year. And in many countries, such as Denmark, Finland, and Israel, a once-miniscule research and development workforce now represents a substantial share of workers. With more opportunities as well as pressure for explicit and sustained collaboration among scientists, a multiplier effect of resources also leverages new scientific knowledge—and ensures it reaches wider audiences. Of course, with intensifying competitive pressure to publish comes some elevated incidents in errors, fraud, or misbehavior; for example, an estimated eighty-five hundred papers in 2005 among the over one million papers in STEM+ and social sciences journals were retracted, essentially a public notification that the paper is not of value, because of problems with the data, results, or other factors.[11] At the same time, reflective of the interconnectedness of mega-science is the growth in the sophistication in the identifying and tracking of such problems once found, and in replicating published results. Replication and synthesis of the huge volume of papers are perhaps even more important challenges to address; contemporary artificial intelligence (AI) transforms how we access vast quantities of existing knowledge, with dramatic technological and ethical implications.

As to quality, scientific discovery continually breeds novel areas for additional research, including, of course, high-value science arising from brand-new subtopics. This trend, part of the overall process of *scientization* also at the heart of the argument here,[12] is reflected in the creation since 1980 of some four thousand new major journals for papers on emerging subtopics of science, often at the intersection of disciplines, including parallel growth in journals publishing articles with the highest scientific impact. It is similarly reflected in the steady growth from nineteenth-century *fin-de-siècle* universities onwards in the creation of ever more hybrid academic departments and their graduate programs.[13] These continuously blend established topics of science into new research subfields to advance often multidisciplinary solutions to scientific and related social problems, and, of course, just to do more science. As with any robust institution within a specific cultural period, global mega-science becomes an intensified version of itself, an institution differentiated from the "inside out," or what can be referred to as the *scientization of science*.[14] This term reflects the greater institutionalization of a broadening scope of scientific inquiry through expanding domains and a deepening of the scientific activity within existing disciplines. In other words, although scientization is often taken to mean mostly the outward influence of scientific findings on nonscientific activities within society, the concept also reflects an emerging fundamental autonomy of internal organization and concentration of internal strategic actions by the thousands of contemporary scientists expanding this social institution globally as never before.

Such processes, farther down the line, also lead to robust applied uses that have revolutionized industrial production as well as human services. Imagine any area of life not affected by continuously advancing applications from basic research and its relationship to universities, as demonstrated by the smartphone or the Internet as a whole.[15] A concurrent spectacular growth has occurred in book publishing and patents, as the designs of useful, marketable things derived from the discoveries of basic research reported in innumerable papers. Universities and other research-producing organizations may capitalize on their discoveries, generating further scientization.[16] And all of these dimensions of mega-science are accompanied by vast—and increasing—global spending on research and development that grew at a rate faster than the world's economy over recent years, with larger investments, despite the pandemic's economic shock.[17]

Tracking a Cultural Model Through Time and Space

The readily apparent size of mega-science fascinates, but the why and how of it remains in many ways a hardly anticipated story, one that when missed fosters profound misunderstandings about the world's capacity for scientific research—and its sustainability in the future. Before now, the cultural side of the education revolution has not been thought of as a foundation for mega-science. Histories of the university, of course, have chronicled some of the key transformations behind the argument here, such as the rise in the conferment of scientific and professional degrees, competition for faculty and its output, and the differentiation of fields as disciplines develop.[18] The growth of research activities at universities in selected countries is well known. Though useful, such trendspotting and historical reconstruction nevertheless lacks an appreciation for what caused the historical events in the first place. That requires tracking the development of the cultural ideas underneath the surface of the everyday understandings and motivations of the people and the expansive—and interrelated—social institutions of education and science. Like individuals, universities as complex organizations are also influenced by prevailing cultural ideas about what they should be and how they should operate. Increasingly rationalized in organization, their enactments nevertheless result from cultural ideas about goals and purposes, which at any one time form a socially constructed and accepted model that imperceptibly dictates organizational design and behavior.

Essential to understanding social change, cultural models also dynamically evolve. Their long-term consequences can be difficult to detect and interpret at any single moment or in any particular context. So we undertake a journey through a history of ideas behind the university-science model specifically and the scientization of society broadly across the long century of science from the late nineteenth century until the present, when university-based research paved the way for unparalleled advances in all scientific fields, not least communication and information technologies facilitating unparalleled collaboration or the vaccines protecting us against both common and rare or novel diseases. In various contexts, this model became the main aspirational guide to how countries would grow their systems of higher education and reorganize their faculties for both teaching and scholarship, including increasingly specialized and pathbreaking scientific research.

Our journey starts in the drawing rooms of wealthy individuals and in the universities of Europe, particularly in the German-speaking region, moving west to the U.S. and then back again towards Europe and on to Asia and, ultimately, worldwide. We track a university-science model developed to such a degree over the twentieth century that most research-oriented universities are organized this way—everywhere from Berlin to Berkeley to Beijing.[19] Adopting some parts of the model at different points in time with various wrinkles from national history and culture, many thousands of universities and, recently, other postsecondary institutions joined the research game as well. Universities in countries that produced little to any globally accessible science before the 1980s, including those in Turkey, Brazil, Egypt, Qatar, Luxembourg, and Iran with their contrasting political systems and religious beliefs, now regularly contribute an appreciable flow of papers to the world's major science journals, usually in English. This widespread orientation towards producing research not only underpins the training of scientists but also makes universities the main place where research occurs. Faculty scientists are recruited not only nationally but increasingly internationally as universities become similarly guided by the universal research ethos. A distinctive shift has occurred that would have seemed odd even at the start of the nineteenth century, when most universities were primarily devoted to teaching and professional preparation, often centering on a traditional canon and serving state power. The global ubiquity of the contemporary university, with its similar structures and, most important, a similar culture based on the university-science model spread to all regions of the world, facilitates the intercultural exchanges and collaborations so essential to contemporary scientific advancement.

Following the pathways that the world's universities took collectively to reach mega-science, however, is not obvious. Assisting in the tracking of this expansive but underappreciated cultural model are guiding analyses of a unique and extensive set of information about the global flow of scientific papers from 1900 onwards. Analyzing the bibliometric information of the who, what, and where of 3.3 million published papers collected from the first year of each decade since 1900 offers a richly detailed map for the engrossing journey through time and space leading to what we call global mega-science. A more detailed note on data sources follows in the next section, including a brief description of the global collaborative process behind the project that well exemplifies the motivations, challenges, and benefits of transnational, intercultural, and multidisciplinary research.[20] In 1859, Darwin published his theory and empirical research in a book, *On the Origin of Species*; forty-one years later Einstein's first publication on relativity theory was in a paper among others in the journal *Annalen der Physik*; and since then the paper and the journal have been the main formal way scientists communicate about science.[21] At least since 1900, the volume of papers over time has arguably been the best direct metric of the flow of the newest scientific knowledge that is accessible worldwide. In their aggregate, the papers make up the sum of new findings on thousands of topics from basic science to advanced analyses of technology, engineering, mathematics, plus health, or what is now commonly termed STEM+. The 2015 wave of papers, around two million, was almost twice as large as the number published just five years before, and a volume that would have been unimaginable in 1900, when a grand total of only ninety-five hundred papers were published, largely by gentlemen scholars, in a small number of then-leading exclusive journals. Today, many universities and a few other large research organizations each produce more than that 1900 world volume of cutting-edge research every year. Since then, the annual volume of publication of scientific discoveries has increased by a staggering 21,000 percent, doubling its volume about every fifteen years.

Mega-science, though still not well understood, is very real, and it continues to spread everywhere and into further research fields. The flow of the millions of STEM+ papers analyzed is not a result of trivial change. Instead, it corresponds to a 120-year development in the ever-growing quantity of scientists across the globe, now collaborating in ways once unfathomable, thus expanding the awareness and often the quality of their increasingly jointly conceived and coauthored research. It has become routine to generate prodigious amounts of new discoveries and promising breakthroughs on a multiplying set of freshly scientized topics. And, as we will show, the ideas of the education revolution enacted through a changing university hatched and supported these trends with far-reaching implications for the nature of the world's science and its sustainability into the future.

Note on Data Sources and Collaboration

Funded by Qatar's equivalent of a National Science Foundation through a grant to Georgetown University at Education City in Doha, the assembled international team of the "Science Productivity, Higher Education, Research & Development, and the Knowledge Society" (SPHERE) project faced the formidable task of collecting enough valid and reliable information to reflect the entire development of mega-science, from 1900 into the twenty-first century, worldwide, and also information on the unfolding of the education revolution and university institutionalization. The team of sociologists of science and education, economists, and experts on higher education turned to several well-known scientific techniques, a full description of which can be found in this book's technical companion volume, *The Century of Science: The Global Triumph of the Research University*.[22] Briefly, first the team relied on the idea of a quantifiable indicator of scientific discovery, and others indicating advanced education development. Second, the team employed some sampling of discovery and years over the time period but endeavored to collect data for the broadest viewpoint possible. Last, some assumptions were made to make the task doable.

The team decided that scientific journal articles published in peer-reviewed journals is the most valid, historically consistent, and readily obtainable indicator of the volume of scientific inquiry at any one time, point, and place. Bibliometric analyses of science can include a range of additional published materials, although usually within a limited time and topical scope. Scientists do write books and government reports from time to time, and there are unpublished "grey" series of reports and correspondence that circulate online, but the globally recognized gold standard of declaring any and every discovery is to write an article on it and have it reviewed, vetted, accepted, and published in what is known as a scientific journal. The SPHERE project focused on published "papers," as they are routinely referred to by scientists, as a broad indicator capable of showing the transformation and global spread of the landscape of where science was produced.[23] There are other indicators—patents, R&D expenditures, and expert judgments—and of course no one indicator is perfect, since what each is ultimately asked to measure is very complex.[24] Especially in the contemporary era of increasingly coauthored papers, bibliometric counting is a conservative measure that reflects only a portion of myriad forms and results of research collaboration.[25] But papers offer a number of advantages over others, and important for the scope of our analysis, for the most part by 1900 published papers had become a highly recognized record of scientific discovery.

At the heart of the project, then, is an extensive dataset representing all *research* papers—omitting editorials, debates, conference reports, book reviews, and so forth—in STEM+ journals from selected years from 1900 to 2011. In fall 2012, the research team purchased the publication data of the Science Citation Index Expanded (SCIE) in an analyzable format from the bibliometric platform known as the Web of Science (WoS), maintained and marketed by Clarivate Analytics (formerly Thomson Reuters).[26] Data crucial for our analyses of the who, what, where of science—paper title, authors' institutional affiliations and addresses, journal citation impact factors (in later years), and subject area—every five years from 1900 to 1980 and every year from 1980 to 2012—were developed from the main dataset of papers. Since data files for 2012 were not finalized at the time of delivery, 2011 was the final year fully analyzed. Although the focus is on the time period of 1900 to 2011, we have recently augmented this with limited data up to 2020 and report these newer analyses when possible.[27] In addition to WoS, the Scopus database, maintained by Elsevier, has also evolved and expanded, along with the more inclusive (but less detailed and unedited) Google Scholar, and even "altmetrics" of nontraditional bibliometrics that complement traditional citation impact metrics.[28] At the time, WoS proved the team's best option and, in separate verification analyses completed since, very closely matches results from Scopus.

The purchased dataset sounded perfectly promising, but upon obtaining it the team quickly realized that the information prior to 1980 had limitations. Its completeness waned the further back in time the data went. Because the digitizing of journals proceeds in reverse chronological order, information from earlier years is less developed and often incomplete; many pre-1980 electronic files on papers did not include essential information such as affiliations or addresses of authors, and this usually was worse earlier in the century, when the scientific communities were small and providing organizational affiliation data was not standardized. For example, the proportion of papers without country of author information from 1900 to 1940 ranged from 56 percent to 90 percent. This is why contemporary analyses of scientific papers mostly focus on the decades since 1980 or even just the Internet era since the 1990s.[29] Another limitation was the breadth of STEM+ journals. Scientific journals come in all flavors—different languages, different subjects, different intended scope of readership, and different levels of selectivity of papers. Added to this is that no one knows for sure exactly how many journals there were worldwide at various points over the twentieth century, and while the ongoing digitizing and indexation by WoS, Scopus, and other platforms will no doubt become more inclusive in the future, historical analyses are particularly vulnerable to this limitation. Therefore, the team decided that it had to code the information by hand from the original hardcopy of the journal (or from microfiche) for pre-1980 journals to augment the WoS data, and because this represented millions of papers, it was only possible through a process of scientifically sampling a manageable amount and then weighting to estimate the total.

In many ways this is what WoS and Scopus have partially done anyway. WoS post-1980 is not technically a full census of all journals in the world. Instead, it includes a large volume of leading journals that have attracted more than 95 percent of the citations (cross-references) among scholarly articles. For example, in the most recent years our data is from all research papers in approximately ninety-two hundred such STEM+ journals, or the journals with many articles on discoveries frequently cited and thus built upon by other scientists (and with citations to articles in the same journal now a frequently used indicator of quality known as a "journal impact factor").[30] Some estimate that the WoS and other platforms are digitizing information about no more than a quarter of all academic journals on all topics, science and otherwise, but most agree that they are capturing the most influential, largest, most global, and most accessible (languagewise) journals across all STEM+ disciplines. Small journals pertaining to limited issues in one part of the world, maybe even just one nation, in a language other than one of the main ones of science publications tend to have miniscule readership and few citations, and hence are less likely to have had their papers indexed in these core bibliometric platforms. The main action of mega-science occurs mostly through the journals we sampled earlier in the century and that WoS had digitized post-1980. We try to remind readers of that in our subsequent analyses, but also note that the trends of mega-science are so clear that it is doubtful that inclusion of even more papers (in marginal journals; in many other languages) would much change the overall picture.

Connected to university libraries' journal archives around the world, the multinational team coded pre-1980 journal articles for the first and fifth year of each decade in a stratified, random sample of journals, and a random sample of research papers with each issue of a sampled journal. An initial sampling frame of all the journals for each year was used from the pre-1980 WoS data and stratified by four categories of science, technology, health, or all other. Also, in addition to English, for representation purposes the sample was stratified by German, French, and Russian journals, common languages for papers before English became the dominant publishing language of science.[31] The essential information was then coded by multilingual project staff, who if necessary used other archival sources to complete the crucial affiliation addresses of authors. This more comprehensive pre-1980 sample data, with weights, was merged with the WoS post-1980 data. A painstaking, multiyear process of countless staff hours yielded a uniquely comprehensive dataset of information on 3.3 million (weighted) of the world's scientific papers in major journals for a historical period that had rarely been represented by such data.

Obtaining indicators of the education revolution was easier. Gross-enrollment rates, the percentage of an age group attending a particular level of schooling, have been compiled for all countries across the full time period. More challenging though was that the actual number of universities has not been collected from each country over time (although we do estimate the numbers of those engaged in scientific research through author addresses).[32] To examine the interplay between not just higher education enrollment but also university creation, the team did exhaustive collections of foundings of universities in Germany, the U.S., South Korea, and Japan, and merged them with the historical papers dataset.

The team had to make some assumptions and rules for how to count papers and assign them to the university sector versus other organizational forms, and to countries and regions.[33] When papers have multiple authorships and cross-national collaborations—a huge trend driving mega-science particularly from 1980 onwards—this gives rise to significant not only technical issues in counting publications.[34] When counting total publications worldwide, the team used the number of unique research articles regardless of the organizational affiliation and address(es) of each article. That is, for global totals, single-authored and coauthored papers are counted once, regardless of the number of authors and countries involved. In other words, we do not double (or multiple) count collaborative publications for world totals. But when we compare across countries, papers are counted for each different country-based author on a "whole count" basis; for example, if a paper has two authors based in German universities and two based in South Korean ones, each country's total count would receive one paper. The same was done to estimate university involvement; papers with at least one university-based author were assigned to the university-sector count. Again, although some bibliometricians use fractional counting methods (in this example, each country receives one-half of a paper), we found that the large trends examined here are relatively insensitive to different approaches to paper counting.

Because of the significance of an author's country affiliation—not their actual citizenship status, but rather the host country of the research organization with which they are affiliated—for the analyses here the dissolution or unification of Germany, a major site of science and university development since 1900, required careful attention. Because of the lag time between research completion, article submission, and article publication date, a decision rule was adopted that allowed an article to be attributed to the former country up to three years after the date of transformative political regime change. For example, when occupied Germany was divided into the Federal Republic of Germany (West Germany) and the German Democratic Republic (East Germany), articles were thus coded.

During the period prior to 1949, all articles published by scientists in research organizations in the territories belonging to Germany were counted under "Germany." After reunification in 1990, articles from authors in both parts of the country were again attributed to "Germany," despite the different geographical borders. Similar rules were made for border changes among other countries, usually with contemporary borders applied retrospectively. Other coding rules related to historical and political changes in major science-producing countries, such as China's rapid rise to top producer of science after the end of the cultural revolution and the government's commitment to science and technology in 1978, can be found in the individual and comparative chapters of *The Century of Science: The Global Triumph of the Research University*.[35]

Last, over the later course of our project, more studies began to appear from a revival of the field known as "the science of science" using similar bibliometric data as we do here.[36] While to our knowledge none have examined the role of the university in global mega-science as we do, they report on recent trends such as growth, collaboration, impact, and so forth. Importantly, these findings corroborate the longer historical developments we found with the SPHERE project data.

Notes

PREFACE

1. The 2023 Nobel Prize in Physiology or Medicine was awarded to Katalin Karikó and Drew Weissman for their identification of a crucial chemical tweak to messenger RNA that enabled the development of effective COVID-19 vaccines that have saved millions of lives.

2. Brainard, "Scientists Are Drowning in COVID-19 Papers." Indeed, the extraordinary volume of research papers published now demands AI-based information retrieval and natural language processing should scientists hope to find the most relevant work related to their own research. The COVID-19 Open Research Dataset (CORD-19) was designed to connect diverse expertise in the "machine learning community with biomedical domain experts and policymakers in the race to identify effective treatments and management policies for COVID-19" (Wang et al., "Cord-19"). From March 13, 2020, to June 2, 2022, the date the final version was released, the weekly updated version of the dataset had grown to index over one million relevant papers (https://github.com/allenai/cord19).

3. Perhaps the most comprehensive global source of data on COVID-19 was provided by a university: the Johns Hopkins Coronavirus Resource Center (https://coronavirus.jhu.edu/), developed by the Center for Systems Science and Engineering in the Department of Civil and Systems Engineering, collected worldwide data on cases, deaths, vaccines, testing, and demographics from January 22, 2020, to March 10, 2023. The database registered 676,609,955 total cases; 6,881,955 deaths; and 13,338,833,198 total vaccine doses administered.

4. See, e.g., Diamond, *Guns, Germs, and Steel*.

5. See Schofer and Meyer, "The Worldwide Expansion of Higher Education in the Twentieth Century"; Frank and Meyer, *The University and the Global Knowledge Economy*.

6. National Science Board, "Publications Output: U.S. and International Comparisons."

7. Godin and Gingras, "The Place of Universities in the System of Knowledge Production"; Powell and Dusdal, "Science Production in Germany, France, Belgium, and Luxembourg"; Dusdal et al., "University vs. Research Institute?"

8. Peter Weingart ("Growth, Differentiation, Expansion and Change of Identity—The Future of Science") recounts that the growth of science has often inspired "visions of doom and destiny" while early analyses of scientific growth suffered from simplifications, false assumptions, and limited imagination to correctly foresee how science would develop over the long run.

9. Sparking debate, recent studies show that despite the growth in papers, the proportion of papers that are disruptive or considered breakthroughs—as opposed to incremental papers—has declined (see Park, Leahey, and Funk, "Papers and Patents Are Becoming Less Disruptive Over Time"; Kozlov, "'Disruptive' Science Has Declined").

10. Mike Zapp ("Revisiting the Global Knowledge Economy") shows that almost all countries had striking increases in their knowledge personnel, with OECD countries more than tripling their R&D personnel since the 1980s and non-OECD countries more than doubling theirs since the mid-1990s; see also Powell and Snellman, "The Knowledge Economy."

11. "Zombie Research Haunts Academic Literature Long After Its Supposed Demise."

On scientific fraud, see Franzen, Rödder, and Weingart, "Fraud"; see also Retraction Watch, "Retraction Watch."

12. Usually, observations of a connection between education and science stop at noting that the former spreads scientific literacy and trains the talented few to become scientists. Required to sit through typical science courses in school, most people are more scientifically literate than were their grandparents or parents, and there are increasing opportunities to gain advanced scientific training. But this leaves scientization—the greater institutionalization of a broadening scope of scientific inquiry through expanding domains and a deepening of the scientific activity within existing disciplines and society (see Murakami, "Scientization of Science" and "Transformation of Science"; Drori et al., *Science in the Modern World Polity*; Drori and Meyer, "Scientization" and "Global Scientization")—far too narrow, trivializing what has occurred. This includes the ways in which science has expanded and differentiated as well as how scientific knowledge has become a source of legitimacy across societal sectors (Gauchat, "The Cultural Authority of Science"). Alternatively, what we will refer to in shorthand as the *university-science model* formed a very potent and self-reinforcing connection between mass education and science, manifested in the symbiotic relationship between the university and science, without which it is unlikely that mega-science could ever have spread successfully to all continents and nearly every country.

13. David Frank and Jay Gabler (*Reconstructing the University*) uncover the shifting patterns of development in universities around the world of curricular offerings, disciplinary fields, and types of knowledge. Postsecondary curricula are becoming increasingly internationalized and isomorphic (Zapp and Lerch, "Imagining the World"), especially with the increasingly dominant *lingua franca* of English and scientific dominance of the United States since World War II (Stevens, Miller-Idriss, and Shami, *Seeing the World*).

14. Abrutyn, "Toward a General Theory of Institutional Autonomy"; see also Drori et al., *Science in the Modern World Polity*.

15. Jason Owen-Smith (*Research Universities and the Public Good*) emphasizes that research universities not only create new knowledge but also anchor (scientific) communities as well as serve as hubs for evolving diverse societal activities, which enables their continuous contributions to problem-solving and innovation; see also Stevens, Armstrong, and Arum, "Sieve, Incubator, Temple, Hub."

16. In the U.S., the *Bayh–Dole Act* of 1980 changed the game for university patenting: the university share of U.S. patents grew exponentially for two decades through 1998 (Leydesdorff, Etzkowitz, and Kushnir, "Globalization and Growth of US University Patenting (2009)."

17. UNESCO reports (2022) that R&D expenditures have continued to increase over the past decade, underscoring the fundamental importance of science, technology, and innovation (STI), especially during the COVID-19 pandemic. Countries strive to meet the Sustainable Development Goals (SDGs) Agenda 2030 of encouraging innovation, substantially increasing their research workforce and their public and private R&D spending (monitored in SDG 9.5). Data for 155 countries and territories shows continued growth in global R&D investment, despite the pandemic challenges: the average annual growth rate over the last decade (2010–2020) was 4.7 percent (3 percent adjusted for inflation). Over the decade, the proportion of global GDP invested in R&D went up significantly, from 1.61 percent to 1.93 percent.

18. Rudolf Stichweh (*Wissenschaft, Universität, Professionen: Soziologische Analysen*, first and second editions) charts the differentiation of disciplines and professions and further emphasizes that the university is simultaneously a local and global organi-

zation, with further regional and national relevance, due to embeddedness in networks and the building of universal knowledge systems (Stichweh, "The University as a World Organization").

19. In his tome based on extensive interviews and management experience, William Kirby (*Empires of Ideas*) charts the evolution of the modern university from Germany to the U.S. to China today, comparing the organizations, their resource bases, and the political conditions in which they operate.

20. Dusdal and Powell, "Benefits, Motivations, and Challenges of International Collaborative Research."

21. Murakami, "Transformation of Science"; Tenopir and King, "The Growth of Journals Publishing."

22. Powell, et al., "Introduction: The Worldwide Triumph of the Research University and Globalizing Science"; Dong et al. ("A Century of Science") also emphasize the shift from individual to teamwork amid the globalization of scientific development, analyzed on the basis of Microsoft Academic Graph (MAG), a dataset with 100 million publications from 1800 to 2016; Okamura ("A Half-Century of Global Collaboration in Science and the 'Shrinking World'") uses OpenAlex, another large-scale bibliometrics platform, launched in 2022 to replace MAG and now including 239 million works, to replicate change in the countries collaborating and contributing most to fifteen natural science disciplines over the past half-century.

23. See, e.g., Weingart and Winterhager, *Die Vermessung der Forschung*; Ball and Tunger, *Bibliometrische Analysen*; Gingras, *Bibliometrics and Research Evaluation*.

24. See, e.g., Gingras, *Bibliometrics and Research Evaluation*; Wang and Barabási, *The Science of Science*.

25. Laudel, "What Do We Measure by Co-Authorships?"

26. Officially launched in 1964, the Science Citation Index was created by Eugene Garfield at the Institute for Scientific Information (ISI); see https://clarivate.com/the-institute-for-scientific-information/history-of-isi/. The more selective Science Citation Index Expanded now includes over 9,200 of the world's most impactful journals across 178 scientific disciplines. From 1900 up to today, the index catalogs more than 53 million records and 1.18 billion cited references.

27. The follow-up project, "Relational Quality: Developing Quality Through Collaborative Networks and Collaboration Portfolios" (Q-KNOW) extended the analysis for Germany, focused especially on coauthorship patterns across disciplines and organizational forms, through 2020 (see Dusdal, Oberg, and Powell, "Das Verhältnis zwischen Hochschule und Wissenschaft in Deutschland").

28. The three main databases are sufficiently stable in their coverage for in-depth cross-disciplinary comparisons (Harzing and Alakangas, "Google Scholar, Scopus and the Web of Science"); on the evolution of altmetrics and their use for measuring societal impacts of research, see Tahamtan and Bornmann, "Altmetrics and Societal Impact Measurements"; and more generally, see Krüger ("Quantification 2.0?") on the implications of the massive expansion of reliable bibliometric infrastructures via digital publishing and automated data processing for data production and assessment.

29. See, e.g., Godin and Gingras, "The Place of Universities in the System of Knowledge Production"; Adams, "Is the U.S. Losing Its Preeminence in Higher Education?"; Bornmann and Mutz, "Growth Rates of Modern Science"; Bornmann, Wagner, and Leydesdorff, "BRICS Countries and Scientific Excellence."

30. Adams, "Global Research Report: United Kingdom," p. 6.

31. See Powell et al., "Introduction: The Worldwide Triumph of the Research University and Globalizing Science," for assumptions and procedures on stratifying by language, sample sizes, and weighting, adding to the sampling frame in earlier years, and other details.

32. But see the historical analysis and discussion of the proliferation of universities in Frank and Meyer, *The University and the Global Knowledge Economy*, chapter 2, which is based on *Minerva Jahrbuch der gelehrten Welt*.

33. See Dusdal, *Welche Organisationsformen produzieren Wissenschaft?*, for the organizational form matrix applied to the data; Dusdal et al., "University vs. Research Institute?", for comparison of German research universities and research institutes.

34. See, e.g., Gauffriau et al., "Publication, Cooperation and Productivity Measures in Scientific Research" and Gauffriau et al., "Comparisons of Results of Publication Counting Using Different Methods"; for a critical overview of various metrics and mechanisms and their implications, see Wu et al. "Metrics and Mechanisms."

35. See Powell, Baker, and Fernandez, *The Century of Science*.

36. Sugimoto and Larivière, *Measuring Research*; Wang and Barabási, *The Science of Science*.

CHAPTER 1

1. Shapin, *The Scientific Revolution*.

2. Price's *Little Science, Big Science* was his best-known book, published in 1963 (republished in 1986, with a foreword by Robert K. Merton and Eugene Garfield, other pioneering figures). Two years later, Price held the first Science of Science Foundation lecture, "The Scientific Foundations of Science Policy," at London's Royal Institution to promote evidence-based science policymaking. He published articles based on his lectures, including "Networks of Scientific Papers," and "The Science of Science."

3. Price, *Little Science, Big Science* and *Little Science, Big Science. . . And Beyond*.

4. The periodization little, big, and mega-science are used heuristically to makes sense of developments during the long "century of science," but these should not be taken as sharp demarcations.

5. Czaika and Orazbayev ("The Globalisation of Scientific Mobility, 1970–2014") show several supporting trends to this in their analysis of bibliometric data of global scientific mobility over the past four decades, including increased diversity of origin and destination countries; the continuous eastward shift of the center of gravity of scientific knowledge production; and increased average migration distances of scientists associated with enhanced integration of more peripheral knowledge-producing organizations into the global science system. As we will see in the case of neutrinos research (Chapter 9), the "nature of different types of research practices implies different spatial relations that in turn influence the motivations for and outcomes of academic mobility and collaboration" (Jöns, "Transnational Mobility and the Spaces of Knowledge Production," p. 111; see also Jöns, Meusburger, and Heffernan, *Mobilities of Knowledge*; Gui, Liu, and Du, "Globalization of Science and International Scientific Collaboration"). Across disciplines, international research collaboration leads to more influential, often-cited research, especially among scholars in different countries—a key argument for further globalizing the scientific enterprise and recognizing the brain circulation and intercultural teamwork that facilitates recognition and impact across scientific communities (see Sugimoto et al., "Scientists Have Most Impact When They're Free to Move").