International Collaboration in the Social and Behavioral Sciences

David Lightfoot

Assistant Director, NSF. Directorate for Social, Behavioral, and Economic Sciences

International collaboration has been a constant at the National Science Foundation, US scientists do a lot of it and increasingly. International activities are built into very many grant proposals and they are quite simply part of normal science. It is sometimes said that international collaboration is expensive and difficult but many kinds of collaboration are routine and involve no special mechanisms. Sometimes special mechanisms and efforts are indeed needed. Over recent years, after the September 11th attacks, there have been difficulties stemming from the issuance of visas but those difficulties have now evaporated in large part, as new procedures have begun to run smoothly.

In addition, NSF has conducted special workshops and run joint solicitations with funding agencies from other countries. A few years ago, to take an example of something of general interest for this meeting, we worked with the UK Economic and Social Science Research Council (ESRC) to fund the Social Science Research Council (SSRC) to make recommendations about how we might internationalize work in the social sciences more effectively. Recommendations were made and four have been acted upon: establishing the International Forum for Funding Agencies (IFFA), an International Data Forum (IDF), annual conferences on international cyberinfrastructure or e-social science, and collaborative work with international development agencies.

Alongside international collaboration, we should recognize that there is international competition and countries, including the US, are concerned about how well they compete in the science and engineering enterprise. This may work against collaboration in some instances.

In any case, international collaboration needs to be viewed alongside interdisciplinary and interinstitutional work and the move toward big science, tackling big questions with interdisciplinary teams of scientists, often with social and behavioral scientists working with scientists from other domains. This reflects the fact that the world has changed and made collaboration easier, mostly through easier travel and more effective telecommunications. Also science has changed and there is greater emphasis on interdisciplinary work as

we tackle bigger problems and cultivate more innovative science, which often stems from new connections. In this paper I will concentrate on new initiatives that encourage science tackling big questions that involve large-scale operations; such initiatives offer the richest possibilities for new international collaborations in the human sciences.

One of the major tools of this big science is what in the US we call cyberinfrastructure and what Europeans call e-science. Cyberinfrastructure has become the common tool and this represents a remarkable development in the history of science. Different sciences have always had their own tools: Astronomers have had their telescopes, physicists their colliders, biologists their sequencers, sociologists their surveys. But now for the first time in the history of science all scientists use a common tool, cyberinfrastructure; new computational facilities have become part of the infrastructure of all the sciences. To be sure, astronomers and sociologists use cyberinfrastructure differently but there are common issues, mostly relating to the establishment and curation of massive data archives, along with the accompanying issues of privacy and confidentiality, and virtual organizations are now re-shaping the way that science is conducted in almost all areas.

The development of cyberinfrastructure is of great interest to the human sciences, which study human behavior in many domains, including the behavior of scientists, and have much to contribute to developing the infrastructure associated with new computational capacities. Certainly cyberinfrastructure has enabled the advancement of social and behavioral sciences; the fields that I work in, language variation, acquisition and change, have been transformed over the last generation and papers look very different from what was published a generation ago, mostly in terms of reliance on massive data archives and the associated analytical techniques. In addition, the social, behavioral and economic (SBE) sciences inform the design of effective cyberinfrastructure and they assess the impacts of cyberinfrastructure and ways for it to benefit society; for these sciences, cyberinfrastructure is itself an object of study.

The contributions of the social and behavioral sciences are significant and many, are covered in detail in 2005 report from joint CISE-SBE Airlie House conference, and have been incorporated into NSF's *CI Vision for 21st Century Discovery* (NSF, 2007). Here I will use cyberinfrastructure as a lens on emerging areas of interdisciplinary, potentially transformative science, where there are opportunities for scientists from different disciplines, institutions and countries to collaborate. I will discuss the social and behavioral science contributions to cyberinfrastructure with respect to five key NSF investment areas:

- Adaptive Systems Technology
- Environment
- Science of Science & Innovation Policy
- Complex Systems
- Cyberinfrastructure

Adaptive Systems Technology

Much has been learned about how the brain functions in the past generation through various imaging tools. These tools and technologies include:

- More powerful computationally based imaging devices
- Tools for gathering coordinated, simultaneous data from different monitoring devices (SBE/CISE/OCI Next-Generation Cybertools award to U. Chicago)
- High performance computers capable of storing and analyzing massive data sets

However, brain science is still in its infancy and present machines measure what they measure and have severe limitations – new tools and technologies are necessary to help us better understand the anatomy, development, and physiology of the brain. New tools will come from collaborations between people working on brain function, on the one hand, and chemists, physicists and nanotechnologists, on the other. AST aims to promote the collaborations that could develop new technologies mimicking the brain processes we understand and new technologies that will enable new understanding of how human brains work. These collaborations cross disciplines and international borders.

Environment

NSF established its first ever inter-directorate standing program two years ago. The Dynamics of Coupled Natural & Human Systems is supported by the Biological Sciences, GEO and SBE and focuses on human factors in environmental change, both human causes and consequences for humans.

This is viewed as a vehicle for future investments in environmental matters, including work on climate change. Any modern work on environmental matters is based on new possibilities provided by Geographical Information Systems (GIS), which combine geospatial data with data gathered by social and behavioral scientists, allowing sophisticated research on environmental change, resource inequality, business networks, criminal justice, health and disease. There is a focus on disasters and that, too, integrates climate, environment and social science data to enable better prevention, preparation and mitigation. Environmental work also involves simulations of societies; current simulations are too simplistic to capture social processes in even small groups, so substantial high-speed computing resources are required. And new kinds of observatories enable fine-grained multidimensional recording of natural and human-built assets over time.

Science of Science & Innovation Policy

The new program in the Science of Science & Innovation Policy (SciSIP) aims to develop the data, tools and knowledge needed to establish a new social science of science and innovation policy. It promotes the study of the returns that public and private sectors receive from R&D investments in science and engineering; gathering data to improve and expand science metrics, indicators of research investments and returns on such investments; production of usable knowledge from data to understand innovation and returns to investments in the US and overseas; and developing new kinds of data extraction

and virtual collaboratories bringing social and behavioral scientists together with scientists from specific domain to examine and understand what it means to make progress in that domain. This work needs international collaboration for the purposes of comparison.

This work requires developing new indicators, both redesigns and new surveys. For example,

Redesign of the Industrial R&D Survey

Costs associated with a firm's R&D – U.S. and abroad

R&D workforce in industry

Conduct of R&D by industry

Intellectual Property and licensing – innovation related data

S&E workforce – Postdocs

Need statistical information on all postdocs in all sectors

Address data void on foreign degreed postdocs

S&E Workforce – National Survey of College Graduates

Decennial census – long form eliminated: problems and opportunities

New approach to include S&E immigrants and will provide more timely, comprehensive, & current data

These new indicators are critical to meet the challenge to develop a new social science of science policy, for a better understanding of the competitiveness element of the American Competitiveness Initiative, to fulfill NSF's legislative mandate to be a clearinghouse on the S&E enterprise, to improve the biennial Science & Engineering Indicators, and for the development of the SciSIP program. Improved workforce indicators are also critical for NSF's Congressional mandate to report on women, minorities, and persons with disabilities in science and for the agency's efforts to broaden participation in the S&E workforce.

This is an area ripe for international collaboration. Every government that runs any kind of science program is interested in this kind of work. Also different countries have different models of science funding and there is much to be learned from international comparative work.

Complex Systems

Complex systems have many interacting elements (neurons, individuals, societies), often with stochastic and intricate interactions. Behavior of the systems cannot be understood by examining the individual constituents only. They have "emergent" properties: complex, system-level behavior emerges from simple rules governing interacting elements. Systems require analysis from many viewpoints and at many levels and that kind of work is facilitated now by new computational capacities. A vehicle for supporting this kind of work at NSF is the new Cyber-Enabled Discovery & Innovation (CDI) initiative (http://www.nsf.gov/crssprgm/cdi). One theme of the CDI solicitation is Understanding Complexity in Natural, Built and Social Systems. CDI supports multidisciplinary, transformative research and not business as usual and seeks to promote innovation in, or innovative use of, computational thinking. It is a substantial five-year initiative with $52M in FY08, projected to rise to $250M in FY12.

A particular focus of work on complexity is on how complex systems change, sometimes dramatically,

comprehensively and quickly; such "catastrophic" changes constitute the "tipping points" of Malcolm Gladwell and others and such phenomena pervade the human sciences in the emergence of new properties at individual, group or system levels; examples are the collapse of economic markets, stampeding behavior on crowds, and catastrophic language change in developing children and across generations. Dynamic interrelationships between social, behavioral, biological and physical factors often lead to or reflect dramatic change in the environment. Complex networks of interrelationships across multiple scales raise intriguing questions in their dynamic properties and that has been a focus of our Human Social Dynamics (HSD) initiative, which we will now be embedding in our core programs.

Cyberinfrastructure

All of the initiatives I have discussed have cyberinfrastructure at their core and that has been a major focus at NSF over the last few years, with the establishment of the Office of Cyberinfrastructure within the Office of the Director three years ago. New cyberinfrastructure has brought progress within the human sciences as in the other sciences and engineering. But the human sciences are also concerned with cyberinfrastructure as an object of study. Social and behavioral scientists are interested in human behavior, including human behavior within the context of science.

One feature of this general interest is a concern for virtual organizations, how they work and how they might work better. Virtual organizations, drawing scientists together outside the bounds of geography, constitute a major tool for new international collaborations and we need to understand their possibilities as well as possible. A new solicitation invites work on this topic, dealing with social and technical issues. Virtual Organizations as Sociotechnical Systems (VOSS) supports scientific research directed at advancing the understanding of what constitutes effective virtual organizations and under what conditions virtual organizations can enable and enhance scientific, engineering, and education production and innovation.

VOSS funded research must be grounded in theory and rooted in empirical methods. It must produce broadly applicable and transferable results.

Basic science and parallel funding

It is worth pointing out that NSF is a basic science agency. We support work on basic science and generally not on policy. We are acutely aware of the needs of policy makers and seek to support science of interest to policy makers. The emphasis on basic science makes NSF different from some funding agencies for the social sciences in other countries where policy concerns have more direct consequences for the scientific agenda. International collaborators need to bear in mind NSF's goals in supporting the Social, Behavioral & Economic Sciences:

1. Increase fundamental understanding of human behavior and society by supporting basic research, infrastructure, and education in the SBE sciences.
2. Provide societal relevance by providing information on

5

critical national problems such as terrorism, business failures, global workforce, America's educational system, and the implication of large-scale trans-formational changes for ethnic and cultural diversity and equality.

Within those limits, we are always open for international collaboration on the basis of what our Director sometimes calls the "byob principle:" bring your own budget! This means parallel review and parallel funding:

The U.S. collaborator submits a description of the work and a budget for the U.S. activities to the NSF, while

The partner submits a parallel or even identical proposal to his/her funding agency along with a budget for the collaborative activities.

The NSF proposal undergoes the usual review process as does the non-NSF proposal.

Our experience is that joint proposals of this type review well, because each funder is getting more net research for their partial support of the overall project. A notice of this approach to joint funding is on the SBE website.

The Directorate is interested in supporting not only work that falls within the standard disciplines but also interdisciplinary work. We have disciplinary programs:

- Linguistics
- Physical Anthropology
- Cultural Anthropology
- Archaeology
- Social Psychology
- Economics

- Sociology
- Political Science
- And interdisciplinary programs:
- Cognitive Neuroscience
- Developmental & Learning Sciences
- Documenting Endangered Languages
- Perception, Action & Cognition
- HOMINID
- Geography & Regional Science
- Environmental, Social & Behavioral Science
- Decision, Risk & Management Sciences
- Science of Science & Innovation Policy
- Innovation & Organizational Sciences
- Methodology, Measurement & Statistics
- Science & Society
- Law & Social Sciences

Interdisciplinary programs reflect mature interdisciplinarity, where there is a community of reviewers and panel members. But we also do a lot of co-funding on an ad hoc basis, where a proposal submitted to one program may be of interest to other programs within the Directorate and in other directorates or even other agencies. This is worth bearing in mind particularly for topics that are themselves inherently of international interest, such as Migration, Comparative Science Policy, and the Science of Poverty Alleviation.

Infrastructure

When social and behavioral scientists think about international collaboration, there are aspects that are different from the other sciences and I

believe that we need to be thinking more, much more, about an international infrastructure. There are no Japanese or Italian data in physics or chemistry. Cannonballs fall from the Tower of Pisa as they fall from similar towers in Japan. But social and behavioral scientists are interested in the different social and political contexts of different countries and need data on education, health, security, etc in different countries. In our fields there are Japanese and Italian data that are different. Comparative work requires comparable data from different countries so that we are comparing apples with apples and not with oranges. This is of central concern to our sciences and much, much more needs to be done.

Our Human & Social Dynamics initiative has developed interesting international infrastructure that can be a model for other enterprises. The Integrated Public Use Microdata Series (IPUMS) has been developed by Principal Investigator Steven Ruggles of the University of Minnesota. The goal is to preserve and harmonize census microdata from around the world and make them freely accessible to researchers.

In much of the world, census microdata are either unavailable or difficult to obtain. In the few countries where census microdata are readily available to researchers, they have become an indispensable part of social science infrastructure. In the journal Demography, the leading U.S. journal of population, census microdata are used three times as often as any other source for studies of the U.S. or Canada. No alternate source offers comparable sample sizes, chronological depth, or widespread availability across countries.

We began funding IPUMS in 1994 and the scope has extended progressively.

1999-2004: International Integrated Microdata Access System (NSF Social Science Infrastructure Project), $3.5 million; 28 censuses from 7 countries

2004-2009: International Integrated Public Use Microdata Series (NSF HSD Project), $5 million; 128 censuses from 37 countries

1999-2009 Additional funding from 3 NIH grants to assist with Latin American and European components of the project

IPUMS-International is now the largest public-use population database in the world, with extensive microdata from 26 countries, 80 censuses and data on 202 million individuals. But this is just the beginning: there are now agreements for 200 censuses from 70 countries, over 3 billion person records, and negotiations are in progress for another 100 censuses.

The work of IPUMS represents a considerable technical achievement. Investigators have gathered data from the last 50 years, compiled without modern computers and facilities, and stored under conditions that sometimes leave much to be desired. These data have now been translated into English and stored in a form where they can be used by current machines and have the right interoperability properties to be usable in the foreseeable future. The data are already the subject matter of thousands of articles and books and this will explode as the scale of the data is

increased significantly over the next few years. The impact is transformational and affects many disciplines. The data are good for:

- The study of large-scale trans-formational changes such as economic development, urbanization, fertility transition, large-scale migration, population aging, mass education.
- Relationships of social and economic change to variations in climate, geography, and environment.
- Human consequences of social, economic, and demographic trans-formations in such diverse areas as family structure, economic inequality, cultural diversity, and assimilation.

Conclusion

I have looked at international collaboration in the human sciences, focusing on scientific opportunities, dealing with big scientific questions and large-scale projects, where international collaboration would be particularly beneficial to our sciences. I have not focused on mechanisms and procedures and they should follow scientific needs. We have done much collaborative work with the Deutsche Forschungsgemeinschaft (DFG), the UK Research Councils for the Arts and Humanities (AHRC) and the Economic and Social Sciences (ESRC), the Agence Nationale de Recherches (ANR) in Paris, the National Natural Science Foundation of China (NNSF-C), and others. We have participated in several EUROCORES programs over the last few years. Let me be clear that NSF is interested quite generally in cultivating international collaboration at the level of principal investigators from different countries who want to undertake common research and at a broader institutional level where agencies from different countries can develop the international infrastructure that is so needed by our sciences.

References

1. Thomson JA, Itskovitz-Eldor J, Shapiro SS, et al. Embryonic stem cell lines derived from human blastocysts. Science 1998; 282:1145-1147.
2. Shamblott MJ, Axelman J, Wang SP, et al. Derivation of pluripotent stem cells from cultured human primordial germ cells (vol 95, pg 13726, 1998). Proceedings of the National Academy of Sciences of the United States of America 1998; 96:1162-1162.
3. Cowan CA, Atienza J, Melton DA, Eggan K. Nuclear reprogramming of somatic cells after.