A BIBLIOMETRIC-BASED ASSESSMENT OF RESEARCH ACTIVITY IN THE PEOPLE'S REPUBLIC OF CHINA

H. Roberts Coward & Ronald R. Fresne
Science and Technology Policy Program,
SRI International, Arlington (USA)

Introduction

Publications from the People's Republic of China (PRC) appearing in the 1988-1990 "Research Front Databases" of the Institute for Scientific Infomation (ISI) were identified in order to profile research activity in the PRC and identify areas in which that activity merited a more in-depth assessment of Chinese capabilities. The Research Front Databases are produced by a clustering technique and consist of two sets of papers:

a) "current" papers from the current indexing year (e.g., 1989) that cite one or more papers in
b) a cluster of highly cited and frequently co-cited (i.e., appearing together in the reference lists of current papers) defining publications often referred to as the "core" papers of the specialty, or "Research Front," identified by the technique of co-citation analysis.

During the time period covered by available data, the PRC presence in the ISI databases in general and the Research Front Databases in particular (as well as an earlier database produced by the same technique of co-citation analysis) has shown a pattern of notable growth. Since the Research Front Databases represent focused problem areas within the mainstream development of international science, the PRC's pattern of growth, which appears primarily in fields important for their potential applications and impact on economic growth and competitiveness, implies substantial success on the part of PRC efforts to develop links between a strengthened R&D infrastructure and the country's prospective economic development. The data are only partial indicators, but the increasing links between PRC research and the international scientific community are illustrated by the figures in Table 1.

The growing presence of the PRC in the case of both current and core papers is significant: current papers represent a measure of activity, or productivity, in terms of mainstream international publications and not the long-term impact of these papers. The core papers represent a very selective subset of high-impact papers representing contributions that have significantly influenced the current work in each Research Front. Part of the growth in the number of PRC core papers may be attributable to self-citation

Table 1. PRC Publications in ISI's Research Front Databases

Current (citing) papers	1985 (1)	1988	1989	1990	% Incr. 1988-90
Total current papers	339,524	345,080	361,052	375,346	8.8 %
Current papers with a PRC address	1,571	2,746	3,304	3,842	39.9 %
PRC per cent current papers	0.46 %	0.80 %	0.92 %	1.02 %	27.5 %
Current PRC papers involving international collaboration	NA	729	884	1,077	47.7 %
Current US-PRC collaborative papers	NA	366	441	484	32.2 %
Core (cited) papers from PRC	NA	45	61	71	57.8 %
Current paper PRC addresses	NA	3,537	4,427	5,188	*
Unique unified PRC addresses 1988-1990			3,018		

(*) Not meaningful.

practices in the growing presence of the PRC among current papers, but the co-citation methodology usually requires more than national self-citation tendencies to include a paper in a core cluster.

It is particularly notable that the PRC does not show a pattern associated with some developing countries and the former Soviet Union, in which a combination of relative isolation and national self-citation tendencies has the effect of producing a number of clusters that are entirely, or almost entirely, from the country concerned. In no case does PRC participation among the current papers of a Research Front exceed 33.3 percent nor do PRC papers dominate any set of core papers: these are research areas of internationally heterogeneous interest – not parochial science. Also, it is relatively recent PRC work that is helping to define these Research Fronts: 36 of the 61 PRC core papers in 1989 were published between 1986 and 1989, including 7 in 1988 and 1 in 1989. One of the 1989 PRC core papers was published in 1989 and 5 of the 1990 PRC core papers in 1990. As discussed below, these represent contributions to research areas that appear to be developing rapidly.

ISI includes all addresses associated with a given paper that are published in the indexed journal in its database, which is better than most in its standardization of institutional names. Nevertheless, variants for the same institution are not uncommon, and these variants, multiple author affiliations, and collaborative papers all contribute to the figures in Table 1 showing fewer unique PRC current papers than PRC addresses. An extensive effort was undertaken to develop a standardized approach to PRC addresses in the interest of developing unified publication counts for institutional analyses. The impact of this effort is shown in the table: more than 13,000 address records were unified to a little over 3,000 (2).

An initial review of 1988 and 1989 data revealed the strong presence of an organization, the China Center of Advanced Science and Technology, most often appearing as the "CHINESE/CHINA CTR ADV SCI & TECHNOL", for which the abbreviation CCAST was adopted for unification of the data. With a total of 511 papers showing at least one CCAST affiliation in the combined two years of data, this represented 8.4 percent of the PRC papers. Subsequent investigation showed strong ties between CCAST and the Chinese Academy of Science, as well as a complex of international relationships, and CCAST's role in the PRC visibility in the data is discussed below.

Multiple PRC addresses on a single paper indicate internal collaboration among PRC institutions (or multiple affiliations on the part of one or more of the authors), while addresses from more than one country indicate international collaboration. A PRC author might also list both his home and a visiting affiliation in, for example, the United States. Close examination of listings for individual papers in the database suggests that this does occur with some regularity, but an earlier analysis of PRC data for the purpose of testing the hypothesis that much of the PRC activity in the Research Front databases was the result of international collaboration suggested that this was not the case. Of the 2746 PRC papers in the 1988 database only slightly more than one quarter (729) represented internationally collaborative papers, half of which (366) involved U.S. institutions (3). The data are updated for 1989 and 1990 in Table 1. As noted below, core papers from the PRC are more likely to be collaborative. In 1989 there were 61 core papers that included PRC addresses: 29 (48 percent) represented international collaboration, 17 of which were with the United States (59 percent of the collaborative papers).

PRC Core Papers

The increase in PRC core papers shown in Table 1 is an important indicator that PRC research is making significant contributions to mainstream international science. A high proportion of them are recent publications, and international collaboration seems to be important, although not necessary, to achieving core status. As noted above, core papers are of particular interest because they represent highly cited papers that have had an important impact on the research being reported in the currently indexed papers in the Research Front Databases. They do not represent *all* highly cited papers from a given year of the ISI database: only those papers whose frequent appearance with one or more other particular highly cited papers in the co-cited context means that there is a degree of consensus among publishing scientists that these papers are both important and associated with each other in contributing to the work being reported in the current publication. These are the contributions to the progress of the specialty that help to define the state of the art and, in co-citation analysis, define the cluster that represents the Research Front. The co-citation algorithm is deliberately designed to eliminate papers, such as methodologies or analytic procedures, that are often highly cited, but seldom so with any particular other paper. That some core papers are highly cited for negative reasons, as in the case of the "polywater" and "cold fusion" controversies, in no way detracts from their import in the course of international scientific activity.

Table 2 summarizes some of the characteristics of the PRC core papers. Among the

Table 2. 1988-1990 PRC Core Paper Characteristics

Field	1988	1989	1990
Clinical Medicine	7	13	19
Basic Bioscience	2	5	5
General Biology	0	1	2
Agriculture & Animal Sci.	4	2	0
Chemistry	4	7	5
Physics	6	7	18
Solid State Physics/Mat.Sci.	16	15	15
Earth & Space Sciences	4	5	3
Mathematics & Computer Science	2	4	4
Engineering	0	2	0
Other	0	0	0
TOTAL	45	61	71
Publication date			
1990	—	—	5
1989	—	1	12
1988	0	7	9
1987	6	17	14
1986	7	11	7
1985	10	4	9
1980-84	21	20	14
Before 1980	1	1	0
TOTAL	45	61	71
Collaborations			
International collaborations	22	29	44
With the United States	12	17	15
With the Fed.Rep.Germany	1	4	8
With the United Kingdom	1	3	3
With Canada	0	2	4
Multiple	8	6	12
Others	Only 1 each year		

scientific fields, it shows a steady increase in Clinical Medicine and a consistent steady base in Solid State Physics & Materials Science, as well as a surge in 1990 in Physics. (Five high energy physics papers involving CERN and very large numbers of authors and institutions from around the world appeared in the 1990 database as both current and core papers.) The numbers of papers in other fields are too small for any robust interpretations, but there were small surges in Basic Bioscience and Mathematics/ Computer Science, and fairly steady patterns in Earth & Space Sciences and Chemistry.

The year of publication of the PRC core papers shows an interesting pattern in the time-series data. The expected paradigm is most closely approximated by the 1989 data, in which the number of core papers published two-to-three years before the database year forms a peak in the distribution. This is because of the tendency of all papers to have peak annual citation rates about 2.5 years after publication, falling off thereafter. The PRC data show a skew toward older papers in 1988, an approximation of the normal pattern in 1989, and a skew toward more recent papers by 1990. This indicates that increasing attention is being paid to PRC research in areas that are developing quite rapidly and being influenced by recent findings. It is further evidence of the PRC's success in moving into important areas of mainstream international science. Many of the "hotter" core papers in 1990 are in the fields of superconductivity and microelectronics (see the discussion of PRC activity in rapidly developing areas, below).

Finally, the data on international collaboration show a rise consistent with the PRC increase in core papers between 1988 and 1989, with about half of the PRC core papers having involved collaboration with researchers from another country, then a jump to close to 60 percent in 1990. The papers produced with the United States are consistent with the pattern in the current papers of about half of all collaborative papers in 1988, but after a jump to nearly 60 percent in 1989, this falls off in 1990 - possibly an effect of the 1989 political developments. The numbers are small, but there is a steady rise in the number of core papers involving cooperation with Germany and Canada.

Not surprisingly, the institutional sources of PRC core papers is headed by various institutes of the Academy of Sciences, which often appeared in conjunction with one or more other institutions. (See infra, our analysis of the Institutional and Geographical Centers of Activity).

PRC Field Profile

Current papers in the clusters have been compared with groups of interactive journals with a common subject matter in order to classify the Research Fronts into a series of fifteen scientific fields (Clinical Medicine, Chemistry, Mathematics & Computer Science, Social Sciences, etc.) and a large number of subfields (e.g., organic chemistry, optics, probability and statistics). The subfield classifications are in an experimental stage of development and should be interpreted with caution. Since only a few Research Front clusters are represented in many of the subfields, and others are of only peripheral interest for this study or show little or no PRC activity, analysis proceeded by collapsing such small or peripheral subfields into the broader field data.

The PRC data were tabulated to show the scope of PRC activity (proportion of Research Front clusters in the field worldwide in which the PRC was involved) and a measure of the degree to which the data show Chinese emphasis on the field in comparison with the world average (an "Emphasis Ratio," or ER, that is greater than one if the proportion of PRC activity in a given field is higher than the average world distribution and less than one if lower). Since the denominator used in calculating the Emphasis Ratio is the proportion of world activity in each field and subfield as it is represented in a database dominated by the United States and other industrialized nations with a broader research base than smaller or developing countries, a high Emphasis Ratio in a given field for the PRC should be interpreted to represent international visibility as much as it does PRC investment and activity in the field. Thus, the PRC's generally low ERs in biological sciences is probably less a reflection of a lack of activity in these fields than it is that there is less international interest in PRC work in biology and a lack of access to the journals covered by ISI for that or some other reason (see the discussion of journal access in the context of CCAST, below).

This analysis focuses on the data for the median database year of 1989, in which the PRC's papers represented slightly less than one percent of the database (4). The 1988 and 1989 data are shown in Table 3 and, in graphic form, in Figure 1. In general, very small subfields (less than 25 Research Front clusters) and most of biomedical research and the social sciences are not detailed to the subfield level, except for the retention of some of special interest. For example, only 25 Research Fronts represent "biotechnology" in the SRI subfield system, and the PRC is active in only two. There is considerable interest in Chinese biotechnology, and so the PRC's apparent weakness in a subfield that is usually comprised of highly applied enzyme-based biotechnology processes, as well as Agriculture and Animal Science, where some work on plant protoplasts is likely to appear, is of some interest. On the other hand, genetics and biochemistry/ molecular biology are important to biotechnology and show growing emphasis in the PRC data.

Small subfields in which the PRC emphasis was quite strong included mechanics & thermodynamics, ceramics, applied mathematics, aerospace technology and engineering, nuclear engineering, and telecommunications. While the PRC Emphasis Ratio was well above 1 in all of these subfields, the data are too sparse to draw much in the way of conclusions.

All countries tend to have somewhat specialized research portfolios: only the United States is active in more than 90 percent of the research areas defined in the Research Front Databases. In the case of the PRC, the scope of its research activity ranges from ten percent or less in Psychology/Psychiatry and the various social science fields to a high of almost half (48 percent) of Table 3: Field distribution the Solid State Physics/ Materials Science field. Aside from Chemistry (29 percent), the PRC is active in at least one-third of Engineering and all of the physical sciences, but only 10-20 percent of the medical and biological fields. Among subfields of meaningful size, nuclear and particle physics, metallurgy, optics, and solid state physics stand out.

The Emphasis Ratios show the pattern expected from earlier analyses that included PRC data: the typical profile of a successful Newly Industrializing Country (NIC) that emphasizes Chemistry, Physics, Materials Science, Math, Computer Science, and Engineering (see Figure 1). The PRC data also show a strong emphasis on the Earth and Space Sciences. Absent was the typical developing country emphasis on biological science: overall, the PRC was below average in the biological sciences, although advanced biological subfields of clinical oncology and basic research on genetics had ERs of 1.3, and were followed by biochemistry/molecular biology and neuroscience with about 1.0. Within the physical sciences, the PRC "portfolio" shows more than twice the average emphasis on nuclear and particle physics, metallurgy, and optics. The emphasis ratio is between 1.5 and 2 in solid state physics, mathematics, computer science and cybernetics, materials science, as well as astronomy and geoscience. The confluence of strong emphasis on computer science & cybernetics and solid state

Table 3. →
Research Scope and Emphasis Ratio:
People's Republic of China (1988-1989)

	1988				1989			
Fields	**Clusters in Model**	**Clusters with PRC**	**Scope**	**Emphasis Ratio**	**Clusters in Model**	**Clusters with PRC**	**Scope Ratio**	**Emphasis**
Totals	8177	1721			8424	1999		
Clinical medicine	1941	241	12%	0.59	2047	305	15%	0.63
Dentistry & Odontology	30	4	13%	0.63	33	7	21%	0.89
Hematology	59	13	22%	1.05	65	12	18%	0.78
Oncology	152	32	21%	1.00	153	47	31%	1.29
Radiology, Imaging & Nuclear Med.	94	13	14%	0.66	87	19	22%	0.92
Rheumatology	26	4	15%	0.73	30	6	20%	0.84
Basic Biomedical Research	1210	217	18%	0.85	1180	234	20%	0.84
Biochemistry & Molecular Biology	472	88	19%	0.89	394	87	22%	0.93
Genetics & Heredity	73	17	23%	1.11	82	25	30%	1.28
Neuroscience	187	34	18%	0.86	190	42	22%	0.93
Virology	36	7	19%	0.92	25	8	32%	1.35
Other Life Sciences	570	87	15%	0.73	563	91	16%	0.68
Botany	158	32	20%	0.96	156	33	21%	0.89
Agriculture & Animal Sci.	215	14	7%	0.31	213	22	10%	0.44
Chemistry	977	238	24%	1.16	917	269	29%	1.24
Analytical Chemistry	99	34	34%	1.63	102	38	37%	1.57
Electrochemistry	35	7	20%	0.95	26	10	38%	1.62
Inorganic & Nuclear Chemistry	106	20	19%	0.90	93	23	25%	1.04
Organic Chemistry	183	33	18%	0.86	140	31	22%	0.93
Physical Chemistry	157	37	24%	1.12	150	43	29%	1.21
Polymer Science	78	25	32%	1.52	102	35	34%	1.45
Physics	637	222	35%	1.66	627	267	43%	1.79
Atomic, Molecular & Chem. Physics	106	22	21%	0.99	96	29	30%	1.27
Nuclear & Particle Physics	87	47	54%	2.57	82	51	62%	2.62
Optics	58	20	34%	1.64	54	27	50%	2.11

Fields	1988				1989			
	Clusters in Model	**Clusters with PRC**	**Scope**	**Emphasis Ratio**	**Clusters in Model**	**Clusters with PRC**	**Scope Ratio**	**Emphasis**
Totals	8177	1721			8424	1999		
Solid State/materials Sci.	699	315	45%	2.14	726	348	48%	2.02
Crystallography	38	15	39%	1.88	34	8	24%	0.99
Materials	25	7	28%	1.33	37	15	41%	1.71
Metallurgy	63	21	33%	1.58	57	30	53%	2.22
Solid State Physics	521	253	49%	2.31	541	261	48%	2.03
Earth & Space Sciences	354	111	31%	1.49	350	117	33%	1.41
Astronomy & Astrophysics	99	31	31%	1.49	83	36	43%	1.83
Geology & Mineralogy	55	25	45%	2.16	53	19	36%	1.51
Geosciences	82	29	35%	1.68	96	37	39%	1.62
Meteorology & Atmospheric Sciences	29	6	21%	0.98	35	7	20%	0.84
Oceanography & Limnology	31	5	16%	0.77	26	5	19%	0.81
Math/computer Science	470	171	36%	1.73	482	190	39%	1.66
Computer Science & Cyberneties	205	81	40%	1.88	238	95	40%	1.68
Mathematics	178	66	37%	1.76	151	69	46%	1.93
Probability & Statistics	59	12	20%	0.97	65	14	22%	0.91
Engineering	191	56	29%	1.39	209	69	33%	1.39
Biotechnology	18	5	28%	1.32	25	2	8%	0.34
Chemical Engineering	36	10	28%	1.32	35	11	31%	1.32
Nuclear Technology	21	8	38%	1.81	16	4	25%	1.05
Telecommunications	17	3	18%	0.84	16	6	38%	1.58
Psychology & Psychiatry	378	13	3%	0.16	349	34	10%	0.41
Economics/management Sci.	164	6	4%	0.17	181	14	8%	0.33
Social Sciences	244	11	5%	0.21	248	9	4%	0.15
Education	36	2	6%	0.26	46	0	0%	0.00
Inter/multidisciplinary	91	17	19%	0.89	104	30	29%	1.22
Instrumentation & Techniques	49	13	27%	1.26	71	24	34%	1.42

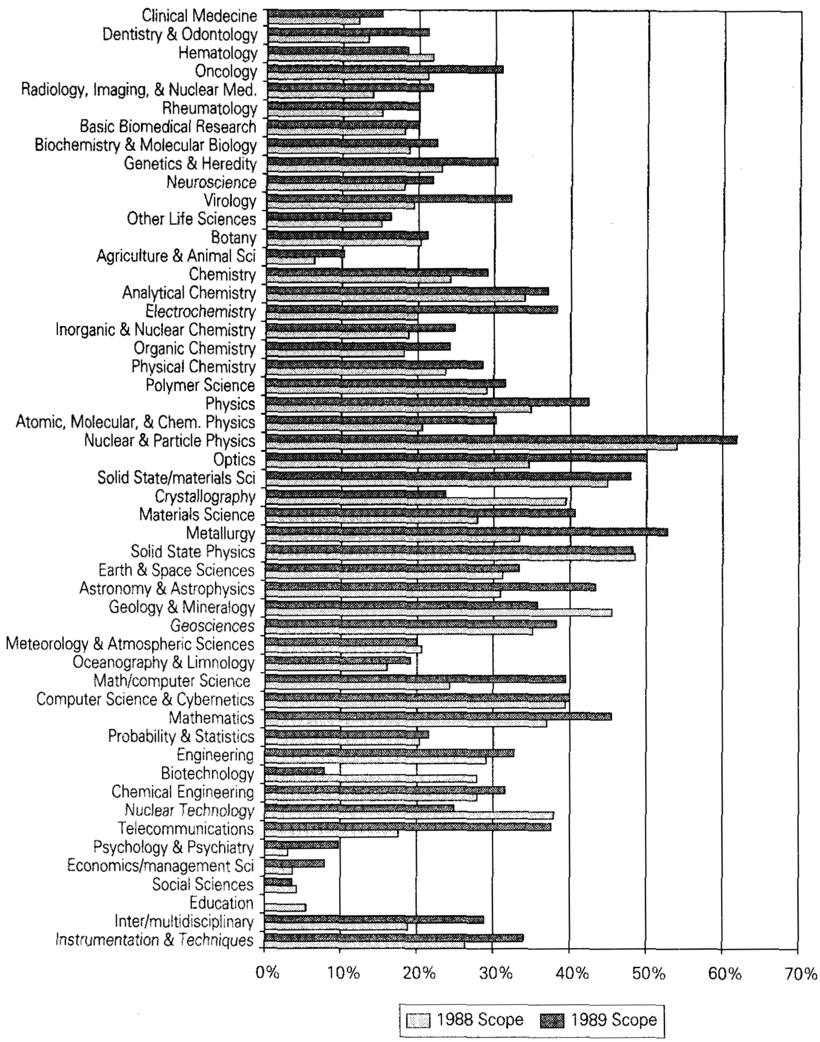


Figure 1a. Scope of Research Activity
People's Republic of China (1988-1989)

← **Table 3 continued.**

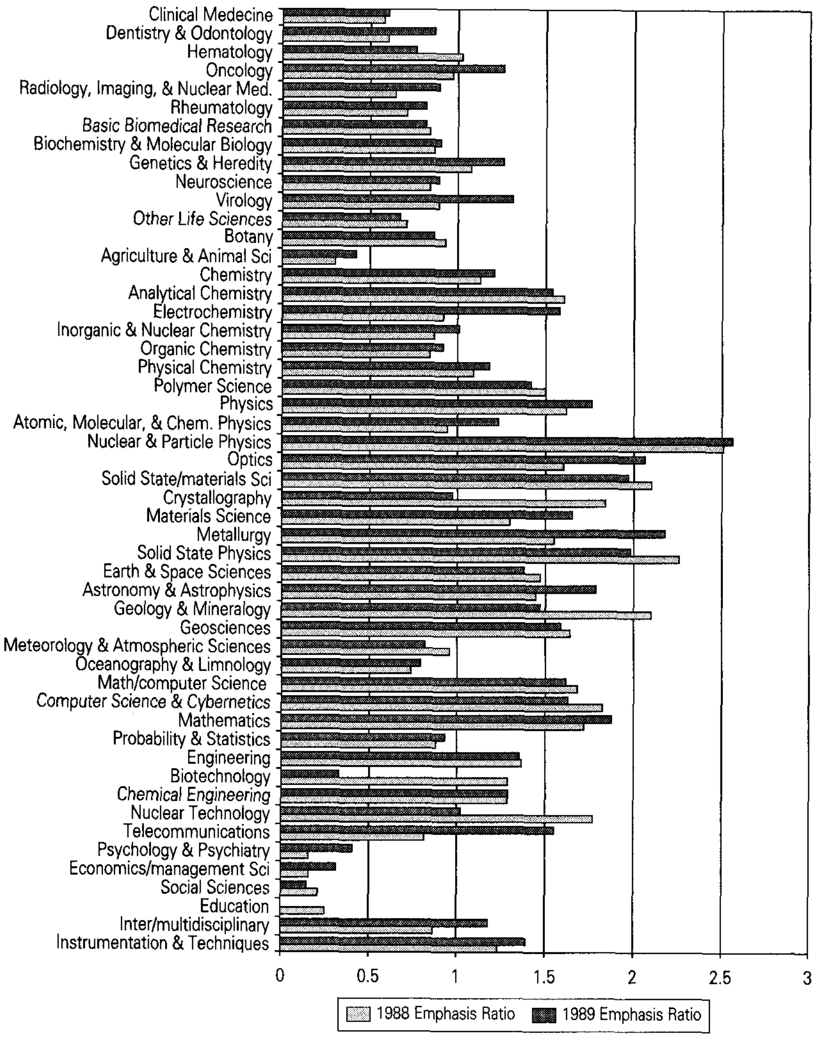


Figure 1b. Emphasis Ratio of Research Activity
People's Republic of China (1988-1989)

Table 4. PRC Field Distribution and Specialization

Field	1989 Clusters	1989 PRC Clusters	PRC > 0.92%	Percent >.92%
Clinical Medicine	2047	305	191	62.62%
Basic Biomedical Research	1180	234	119	50.85%
Other Life Sciences	563	91	60	65.93%
Agriculture & Animal Sci.	213	22	17	77.27%
Chemistry	917	269	216	80.30%
Physics	627	267	237	88.76%
Solid State Phys./Mat. Sci.	726	348	326	93.68%
Earth & Space Sciences	350	117	94	80.34%
Mathematics & Computer Sci.	482	190	177	93.16%
Engineering	209	69	64	92.75%
Psychology/Psychiatry	349	34	10	29.41%
Economics/Management Sci.	181	14	8	57.14%
Social Sciences	248	9	8	88.89%
Education	46	0	0	0.00%
Inter-/Multidisciplinary	104	30	20	66.67%
Totals or Average	8244	1999	1547	77.39%

physics (which is both the largest subfield in the current system and the one in which much of the work on semiconductors falls) suggests a highly active effort to bring PRC research to international standards in inter-related "high technology" fields concerned with computers and electronics. The solid state physics subfield is also the strongest for CCAST, primarily from its Center of Theoretical Physics, considerably outweighing its presence in nuclear & particle physics, the field with the highest PRC Emphasis Ratio.

Both scope and emphasis represent ways of looking at the uneven distribution of national participation across the range of research activity worldwide. Countries with a limited research infrastructure can, in principle, approach the problem of distributing their limited resources by either spreading these thinly across a broad range of research areas, or concentrate them in a limited number of fields and problem areas. Table 4 shows PRC data that suggests a strong pattern of specialization: while the scope of PRC activity in the 1989 Research Fronts in the various fields of science is limited, the PRC presence in those areas in which it is active is strongly polarized toward strength.

In more than three-quarters of the clusters in which the PRC is active, there is a higher proportion of current PRC papers than its 0.92 percent of the 1989 model overall. This degree of specialization is highest in the field of Solid State Physics/Materials Science, followed closely by Mathematics & Computer Science, and Engineering, then Physics and the Social Sciences. The numbers in the case of the Social Sciences are too small to be very meaningful, but it is a field in which problem areas often have a parochial, national focus. Only in the peripheral field of Psychology/Psychiatry is the PRC share below average in more than half of the Fronts in which it is active. The PRC is not spreading its efforts thin by this measure, which merits further exploration through the development of comparative international data.

Cluster Demographics Analysis

The overall field data suggested that PRC activity in various aspects of physics, materials science, and computer science represented a significant area of capabilities on which to focus, although there are important strengths appearing in the fields of Chemistry and the Earth and Space Sciences. In order to develop a more detailed view of PRC activity in these fields, data at the individual cluster level was generated for the fields of Physics, Solid State Physics/Materials Science, and Math/Computer Science. Each Research Front in the model represents a single primary-level cluster generated by ISI at what is referred to as the C_1 level. ISI generates a second, or C_2, level of clusters that group about three-fourths of the C_1 clusters into closely related research areas, and SRI has assigned "false" C_2 numbers for most of the remainder using data on links between clusters that fall below the threshold used by ISI to generate the C_2s. Data for each field were grouped so that all C_1 clusters containing 1989 PRC citing papers were listed within the context of their C_2 assignment, with the C_2s listed in descending order of PRC percent of papers in the larger C_2 grouping. Because of the close relationship between the fields and the evidence of PRC strength in both of them, the Physics and Solid State/Materials Science fields were combined for the purpose of inspection.

Examination of the Physics/Solid State/Materials Science fields showed further evidence of a tendency on the part of the PRC to be specialized in terms of patterns in the data:

1) PRC activity frequently appeared in only one of the primary clusters in the larger C_2 grouping. This means that the PRC papers are relative isolates in comparison with other current papers that cited into more than one C_1 member of the C_2 group, contributing to the links used to create the larger groupings. The PRC papers either relate to only one of the linked research areas, or their authors lack the awareness of the relevance to their work of core papers in the other C_1 clusters that would lead them to cite those papers as have other authors in the research area (5).
2) When PRC papers appeared in more than one of the primary clusters in a larger C_2 group, they were likely to be concentrated in only one or two of these, often the larger ones. For example, the C_2 group with the second highest percent of PRC papers dealt with magnetic materials, with slightly more than 9 percent of the papers including PRC addresses. Of the 59 PRC papers in the larger group, 48 were concentrated in one 300 paper primary cluster. Three other related primary clusters ranged from only 1 to 6 PRC papers (although these represented as much as 22 percent of these smaller clusters) (6).

While the field data show nuclear and particle physics as the most strongly emphasized subfield, it does not appear frequently among the top C_2s in the combined Physics/Solid State/Materials fields. Nearly all of the top 25 C_2s represent materials-oriented research. The largest number – seven – of C_2s in the top 25 deal with various specialized alloys or other metallurgy topics, while the largest number of primary clusters are in two C_2 groups dealing with High T_C Superconductivity. The largest group of related superconductivity clusters contains 424 PRC papers (including double counting) in 54 clusters representing most of superconductivity research in the 1989 model, with a PRC share of 3.29 percent, more than three times its overall database share. The number of PRC papers in individual superconductivity clusters ranges from 1 to 115, the latter being a very large cluster containing more than 1,000 papers, while PRC percentages range from slightly over 1 to 12 percent (i.e., above the PRC average of 0.92 percent in every case). Other sources point to superconductivity research as a focus of PRC work in materials science.

In addition to the strong presence in superconductivity research, the amount of PRC work in areas relating to the development of semiconductor and optoelectronic materials stands out, both in the top 25 C_2s and in the overall combined listing. A great deal of research appears to be going on in molecular beam epitaxy, sputtering, lasers and optical fibers, and semiconductor materials such as gallium arsenide.

The Mathematics/Computer Science situation does not appear to display such clear patterns, at least without seeking input from technical experts. The PRC's C_2 percentage peaks at 8.5 percent compared to a top figure of 13.6 percent in the Physics/Materials Science fields. Research Fronts classified by SRI as falling into the computer science subfield appear to be applications of computers in areas involving theoretical problems or computational problems, rather than areas dealing with computer software, hardware, architecture, or materials (most of the latter appear in the Solid State/Materials Science field in any event).

There is, however, a one large group involving 58 Research Front clusters (26th-ranked by PRC percentage C_2) that is homogeneously computer science – although including some applications. PRC papers represent 1.9 percent of the C_2 group, about twice its database share, and range from 1 to 10 percent. The linked research areas include parallel computer architecture, and robotic controls. A closely allied C_2 ranked as number 20 and included research on parallel pipelined computers and flexible manufacturing systems. Other computer science areas in which PRC activity was above average included artificial intelligence and expert systems.

In both the Physics/Materials Science and Math/Computer Science fields, technical experts are needed to determine the degree to which limited access to computer facilities and advanced laboratory instrumentation constrains PRC work in experimental and applications-oriented research. The highly theoretical character of much research in the former Soviet Union and many developing countries stems from such constraints. The 1989 core papers from the PRC in these fields, as well as others, appear to be more theoretical or based on techniques requiring limited facilities when the paper does not involve international collaboration. A scanning of 73 CCAST papers appearing in a number of leading U.S. physics journals indicated that all but three were essentially theoretical. On the other hand, a recent survey of three important condensed matter physics laboratories lists a variety of sophisticated instruments and equipment and claims substantial experimental capabilities (7).

Rapidly Developing Areas

When a Research Front cluster is defined by core papers that are of recent vintage, it suggests that the specialty is developing rapidly – that it is a "hot research topic" being driven by recent discoveries and ideas. It was pointed out above that the core papers from the PRC have become increasingly biased in their distribution toward recently published papers and represent contributions to areas that may be rapidly developing. An exploration of potentially "hot" topics was undertaken.

There are several possible criteria for identifying rapidly developing research areas based on the years of publication of the core papers. One is to take the average year of publication of the core papers and consider those clusters with a recent average with respect to the database year; another is to consider clusters with one or more recently published core paper even though the average is older on the theory that the impact of a recent piece of work may represent a reinvigoration and prospective take-off of the area. For the PRC, a combined approach was taken: 1990 Research Fronts with an average core publication date of 1987-1990 and ones with at least one 1990 core paper were identified and those that included current PRC papers listed, a total of 170 clusters. Table 5 lists the field distribution of these Research Front clusters.

The high number in Basic Bioscience (primarily biochemistry and molecular biology) and even higher in Solid State/Materials Science stand out. (Because the fields were classified by manual inspection the latter may include a few clusters that the automated procedure would assign to Chemistry or Physics.) Both are, in any event, fields in which rapid development is expected in general. Still, it is evidence of PRC activity in some cutting-edge areas of mainstream science. It is, however, highly concentrated in the field of superconducting materials: about two-thirds of the Research Fronts falling in the physics, materials, and computer science fields of particular interest dealt with HT_C superconductivity. Others dealt with epitaxy, laser and organometallic deposition of thin films (with some overlap with superconductivity topics), optics, semiconductor, and other materials.

Table 5. Field Distribution of Rapidly Developing Research Fronts with PRC Participation

Field	Number
Clinical Medicine	23
Basic Bioscience	34
General Biology	3
Agriculture & Animal Sci.	2
Chemistry	10
Physics	17
Solid State/Materials Science	64
Earth & Space Sciences	4
Mathematics & Computer Science	8
Engineering	0
Other	5
Total	170

Chinese Center of Advanced Science and Technology

The Chinese Center of Advanced Science and Technology (CCAST), representing 8.4 percent of the PRC papers in the combined 1988 and 1989 databases, stood out in initial examinations of the data, prompting a more in-depth investigation. Adding the 1990 papers dropped CCAST to slightly less than 8 percent. In the manner typical of ISI's corporate file entries, several variants were found, and the CCAST abbreviation was adopted for unification of the data. CCAST addresses often included "WORLD LAB" as part of the comma-delineated parts of ISI's address entry, and was also frequently associated with several research centers in specialized areas of physics. A number of the addresses included a street address or post office box that represented the Beijing headquarters of the Chinese Academy of Science (CAS). Of 788 papers with CCAST as the first element ("left-of-[first]comma") of one address, 31 percent also included the Chinese Academy of Science as the left-of-comma element of another address on the paper. Overall, the pattern in the combined three-year database was:

Total current PRC papers (1988-1990)	9,892
Total PRC addresses	12,313
Papers with CAS	3,130
Papers with CCAST	788
Papers with CAS and other institution	843
Papers with CAS and CCAST	247
Papers with CCAST and other institution	746
Papers with CAS, CCAST, and other institution	98

The World Laboratory represents a project founded in 1986 by an Italian physicist, Antonio Zichichi, with the goal of training scientists from developing countries and aiding them in the development of constructive opportunities in their home countries. It is currently headquartered in Lausanne, Switzerland, is largely funded by the Italian government, and has three programs other than the one involving the establishment of research centers in the PRC (8). More recent information indicates that the organization has opened additional offices in Moscow and Kiev in the former Soviet Union and in Brazil. Despite this, the visibility of "WORLD LAB" in the ISI database is quite limited beyond papers involving CCAST, and usually appears on high energy physics papers that include CERN as a source (some with Zichichi as one of the authors), or collaborations between Italy and other developing countries, especially India. On such papers, "World Lab" appears in conjunction with the Swiss headquarters as one collaborator and not with the developing country, as is the frequent case with the PRC.

The 1988-89 Research Front databases listed three research centers in conjunction with CCAST:

- Center of Theoretical Physics (388 papers)
- Center of Condensed Matter and Radiation (50 papers)
- Center of Astronomy and Astrophysics (14 papers) (9).

An English-language PRC *Report on the CCAST-WL Project* also listed three centers (referring to them as "Institutes"): the third was an Institute of High Energy Physics and Synchrotron Radiation, which does not appear in the ISI databases through 1990, while the Astronomy and Astrophysics center was not listed in the PRC report. The report indicates that CCAST was headquartered at the "Institute of Theoretical Physics" during the 1986-89 start-up phase. Since the ISI database frequently lists an "INST THEORET PHYS" in conjunction with the Academy, the "Center/Institute" inconsistency makes it unclear whether this means the Academy's Institute, CCAST's "Center," or, as seems likely, the colocation of the two. A new building housing computer and conference facilities was scheduled for completion in 1989. T.D. Lee, the Columbia Nobel physics laureate, has been deeply involved in the establishment of CCAST and an office is maintained for him in the headquarters building. A committee selects about 70 Regular Members, usually individuals recently returned from foreign post-doctoral training, and more senior Special Members for joint support of their research by World Lab and PRC funding. An additional category of Associate Members receive no funding. Substantial emphasis is apparently placed upon assisting Members in the preparation and placement of publications in international journals. The PRC *Report...* indicates an output of 450 publications in the first two years of operation (apparently 1986-87), and the ISI data attest to continuing success in an area in which developing country scientists face a number of hurdles, including the problem of hard-currency payment of page charges (10).

The World Laboratory and CCAST place emphasis on developing local opportunities for advanced research. As part of the CCAST program a number of conferences have been held on topics that are specifically chosen as having potential for development in China. The topics are therefore presumably ones in which there is a PRC self-perception of strength. Subject matter of 1987 and 1988 conferences included:

- Lattice gauge theory using parallel processors
- Charm physics
- Fractional quantum Hall effect
- High T_C superconductivity
- Two dimensional strongly correlated electronic systems
- Applications of synchrotron radiation

A search for PRC papers on related topics in the 1989 Research Front databases using words or phrases from the above topic descriptions found publications clearly related to each specific topic except the lattice gauge theory and quantum Hall effect. While there are ample PRC papers on both topics in generic terms, the context of parallel processors and "fractional" did not appear specifically, making their relevance hard to interpret in the absence of a technical advisor. High T_c superconductivity represented the greatest amount of PRC publishing activity, with two-dimensional electronic structures a distant second. About half of the identified papers included a CCAST address, but the only one dealing with synchrotron radiation originated from the Department of Physics of Nankai University.

CCAST is heavily concentrated in the Physics and Solid State Physics/Materials Science fields, as expected, but is also represented in a number of fields outside of physics. As noted above, the Center of Theoretical Physics is by far the most frequently appearing of CCAST's Centers and its appearance in Basic Biomedical Research (three 1988-89 papers), nuclear technology (3 papers), and some other engineering subfields, as well as some more closely allied fields in Chemistry and the Earth & Space Sciences, may reflect the headquartering of CCAST with this Center during start-up and the organization's role in aiding publication, more than the actual field of the papers concerned.

Other Institutional and Geographical Centers of Activity

The publication data for all three years were combined and tallied geographically and institutionally. Geographically, in all fields and virtually all of the SRI-derived subfields PRC research appears as being highly concentrated in Beijing (Table 6). This is exaggerated, however, by the tendency of many PRC authors to list multiple affiliations on their papers, especially by including the Academy of Science and/or CCAST in Beijing along with a provincial university or research institute. ISI attaches all of these addresses to individual papers. The combined 1988 and 1989 RF databases included 2754 papers with at least one Beijing address. Of these, nearly one-fourth (598) had at least one other PRC city listed and 475 more than one Beijing address; 74 of the papers had more than one Beijing *and* another PRC city listed.

Table 6. PRC Cities and Provinces Ranked by 1988-1990 Research Front Publications

City	Papers	City as % Province	Province	Papers
Beijing	4380	100 %	Beijing Shi	4380
Shanghai	1565	100 %	Shanghai Shi	1565
Nanjing	600	85 %	Jiangsu	708
Hefei	480	96 %	Anhui	500
Wuhan	396	96 %	Hubei	411
Chengdu	323	75 %	Sichuan	429
Xian	311	97 %	Shaanxi	320
Guangzhou	284	91 %	Guangdong	313
Tianjin	283	100 %	Tianjin A.R.	283
Shenyang	264	68 %	Liaoning	390
Hangzhou	262	95 %	Zhejiang	276
Changchun	250	97 %	Jilin	258
Lanzhou	241	100 %	Gansu	241
			Fujian	141
Ranked by City			Hainan	6

Even given the tendency of multiple affiliations to bias the data toward Beijing, the difference between Beijing and Shanghai is dramatic: a ratio of nearly three-to-one. Following Beijing and Shanghai, the distribution falls off fairly rapidly and shows a strong pattern of centralization within major cities. Only the provinces of Jiangsu, Sichuan, and Liaoning show significant activity outside their capital cities. The third centrally administered municipality, Tianjin, is far down the rankings as a center of research activity, but its proximity to Beijing probably affects the allocation of resources.

Guangzhou, hub of the three rapidly developing provinces in the southeast – Guangdong, Fujian, and Hainan – ranks only eighth, its province ninth, and the other two provinces are far down the list.

After Beijing and Shanghai, there is a substantial drop in paper counts to the next two major centers, Nanjing and Hefei. The University of Science and Technology of China at Hefei and Nanjing University represent the second and fourth, respectively, most productive independent institutions in the PRC, and appear representative of an infrastructure that is more focused on one primary institution at the provincial level than is the broader institutional base of Shanghai's output.

Table 7. High Ranking PRC Institutions Based on 1988-1990 Research Front Publications Counts

3130	Chinese Academy of Science
788	CCAST
484	Beijing University
420	University of Science & Technology of China (Hefei)
395	Fudan University (Shanghai)
379	Nanjing University (Jiangsu)
247	Qinghua University (Beijing)
177	Lanzhou University (Gansu)
173	Chinese Academy of Medical Science
171	Nankai University (Tianjin)

Institutionally speaking, Table 7 lists the most productive institutions in all fields combined.

Within the three fields that have been considered in some detail, various units of the Chinese Academy of Science are at the top of all three fields, although the CCAST Center of Theoretical Physics tops both the Academy's Institute of Theoretical Physics and Institute of High Energy Physics in the field of Physics, and comes in second to the Academy's Institute of Physics in Solid State Physics and Materials Science. In Mathematics and Computer Science, the Academy's Institute of Systems Science leads its Institute of Mathematics slightly if probable variants in the latter are taken into account, followed by its Institute of Computer Technology. Unification of these central institutions presents formidable barriers because the hierarchical rank of the units as entered in the database is often not clear, and some units, such as the Academy's Shanghai Institute of Optics and Fine Mechanics appears an equal number of times with Shanghai and with Beijing as the city element in the address. The research task of linking bibliometric data to detailed organization charts and accounting for local listing practices places the development of indicators for these organizations beyond what is immediately practical.

Aside from these central institutions, Table 8 shows the academic centers identified as ranking high in terms of output in each of these fields (the number in front of each

Table 8. 1988-89 Leading PRC Academic Centers in Three Scientific Fields Physics

	Physics
100	University of Science & Technology of China (Hefei)
96	Fudan University (Shanghai)
82	Beijing University
71	Nanjing University (Jiangsu)
56	Shanghai Jiaotong University
55	Northwest University (Xian, Shaanxi)
	Solid State Physics & Materials Science
231	Beijing University
212	Fudan University (Shanghai)
211	University of Science & Technology of China (Hefei)
143	Qinghua University (Beijing)
79	Shanghai Jiaotong University
71	Shandong University (Jinan)
51	University of Science & Technology Beijing
	Mathematics & Computer Science
60	Beijing University
44	Fudan University (Shanghai)
32	Qinghua University (Beijing)
31	Nankai University (Tianjin)
31	Zhejiang University (Hangzhou)
27	East China Normal University (Shanghai)
26	Xian Jiaotong University
24	Shanghai Jiaotong University
23	Wuhan University (Hubei)

represents the number of 1988 and 1989 Research Fronts in which papers appeared, double counting papers that appeared in more than one cluster).

The data are generally consistent with a PRC assessment and policy directed at encouraging the development of microelectronics in China, emphasizing the strengths of Shanghai, Beijing, and Jiangsu, with Sichuan and Shaanxi as secondary centers (11). However, Jiangsu is represented only by Nanjing University in Physics and Shaanxi by Northwest University in Physics and by Xian Jiaotong University, but in Math/Computer Science rather than in Solid State Physics/Materials Science as would be expected, while the strong position of the University of Science & Technology of China, located in Anhui province, does not figure in the paper referenced.

PRC-U.S. Collaboration

Because U.S. collaboration with PRC researchers was of intrinsic interest and as one way of identifying U.S. researchers familiar with PRC work in the three fields of Physics, Solid State/Materials Science, and Math/Computer Science, collaborative papers were identified as papers with addresses from both countries. Institutional and author counts were compiled in order to provide a list of prospective technical advisors to assist in focussing the assessment, reviewing the bibliometric data, and providing a technical perspective based upon review of PRC publications and their own experience.

In the combined three fields, the top institutional collaborators with the PRC are all universities, although all of the National Laboratories – Lawrence Berkeley, Los Alamos, Oak Ridge, Brookhaven, FermiLab, and Lawrence Livermore – are further down the list in the order indicated. Many U.S. laboratories, including those of NIST, had visiting programs that have been effectively terminated following the 1989 political problems in the PRC. The top twelve universities are shown in Table 9, along with the number of collaborative papers in 1989.

Table 9. Top 1989 U.S.-PRC Collaborative Institutions in Physics, Solid State Physics/Materials Science and Mathematics/Computer Science

Papers	Institution
18	University of Illinois
16	University of California, Davis
15	Virginia Polytechnic Institute & State University
15	Rutgers University
15	University of Rochester
15	University of Houston
14	Louisiana State University
14	Northwestern University
13	University of Minnesota
12	University of South Carolina
12	State University of New York, Stony Brook

Technical Assessment Results

This bibliometric analysis represented part of a broader effort to assess PRC capabilities, and was used to narrow the focus of the second stage, in which a panel of experts was asked to extend the perspective on Chinese work in their field on the basis of their technical knowledge, as well as experience in collaborating with PRC researchers, either in China or the United States. The bibliometric data highlighted a number of areas of PRC strength, including oncology in Clinical Medicine, genetics and virology among the Basic Biosciences, most of Chemistry, and the field of Earth and Space Sciences (primarily geology). However, they also appeared to indicate that PRC investment in several interrelated fields that provide a foundation for the development of computers and microelectronics is generating results. Strong emphasis and growth, as well as influence in terms of core papers appeared in the fields of Physics and Solid State Physics/Materials Science, and some in the field of Mathematics/Computer Science. There was also significant PRC activity in Research Fronts that appear to be developing rapidly in those three fields. Consequently, a focus on the fields of condensed matter physics and materials science was the basis for selecting the panel of four technical experts to review PRC capabilities in their research areas (12).

The panel technical experts were unanimous in their judgment that the bibliometric analysis and data had provided an excellent starting point for considering PRC science, providing systematic and in-depth information that transcended the particularist perspectives of individual experts. The panel's broad support of the bibliometric approach as a point of departure was conditioned by a sophisticated awareness of the limitations of both the databases used and the bibliometric analyses as a way of viewing PRC science. Quantity is typically more easily accessed by bibliometric data than quality, which was perceived as particularly important for the PRC, where competent human resources abound and a quantity of activity is more easily produced than work of genuine excellence. While the number of PRC papers indexed by ISI is growing, there was concern about overinterpreting relatively small numbers, particularly as the data were broken down into fields and subfields. It is also the case that the problem of a lag in the cycle from research to publication may be exaggerated for the PRC. This made the time-series of several years of data an important resource: the analysis revealed some trends that were not familiar to experts who had collaborated with Chinese scientists in the institutions involved.

From the panel's perspective, the PRC showed particularly strong capabilities in the materials science fields of high T_C superconductivity and, somewhat surprisingly, the recently developed field of carbon fullerenes. They are also ahead of the rest of the world in certain techniques of crystal growth, especially barium borate, that are of importance in optics and electronics, and all of these areas show strong experimental capabilities. In other areas, such as surface physics, quantum Hall effect research and semiconductor materials, the PRC work is of reasonable quality, but handicapped by the requirement of sophisticated and expensive instrumentation and other equipment. PRC strategy in these fields represents a concentration of resources in a limited number of centers and a pressure on researchers to engage in applications-oriented work of an experimental nature. This is an appropriate strategy for a developing country, and, while it has thus far only resulted in average work in the selected fields by PRC researchers, this is no small accomplishment in the context of mainstream international science in some highly competitive areas.

Aside from the specified topics and work emanating from the major well-equipped centers, PRC science in the selected fields remains second tier, but has the human resource availability to move up quickly if certain barriers are overcome. The barriers are cultural, social, and economic, and include the need for better scientific leadership, greater internal and external interaction and mobility, and the investment of financial resources to overcome inadequate equipment and provide attractive opportunities to offset the ongoing scientific brain drain. The existence of certain areas of Chinese leadership, such as that in crystal growth, and strong human resources were cited as reasons to expect that scientific cooperation between the United States and the PRC would provide mutual benefits for both parties.

NOTES

1) The 1985 data were drawn from a database produced by the Center for Research Planning using a variation on the co-citation clustering technique, but comparable in terms of the number of current papers to the databases produced by ISI. See H.R. Coward and R.R. Fresne, *Research Activity in the Pacific Rim Nations: Survey of a 1988 Bibliometric Database* (Arlington, VA: SRI International, March 1990).

2) Unification is probably incomplete: the two sources used to check ISI's abbreviated entries did not always agree: Office of the Academic Degrees Committee of the State Council, the People's Republic of China, *Directory of Chinese Institutions of Higher Learning and Research Institutes Authorized to confer Doctor's and Master's Degrees* (Beijing: Higher Education Press, 1988), and *The World of Learning 1991* (London: Europa Publications Limited, 1990). The University of Science and Technology Beijing did not appear in the official PRC listing, but may represent an effort to collect several geographically proximate technical institutions in Beijing into a new and more broadly based institution.

3) Coward and Fresne, *op.cit.*, p. 5.

4) Although data have been compiled on PRC papers for all three years, the analyses presented here vary in the number of years used. The 1989 data is the most recent for which the field classification has been fully implemented and time has not been available to explore ways of improving the method before applying it to the 1990 database. Some field-oriented analyses include both 1988 and 1989 data, and all three years have been used when field considerations are not an issue.

5) Awareness of the international literature may represent one possible explanation for this isolation, and may also be one reason that the PRC papers often seem the least focussed of those attached to a given Research Front. Current papers are assigned to each Front in which they cite at least one of the core documents, which may number from the minimum of two required by the definition of co-citation to a few macroclusters with several hundred. The more core documents cited, the more focussed on the problem area the citing paper is likely to be. Only a quarter to a third of PRC papers cite more than one paper, varying among fields. Only Physics and Solid State Physics/Materials Science exceed one-third. Data for international comparisons have not been compiled, so it is difficult to estimate the impact of this as a limitation on the interpretation of the data.

6) For details, see C2 Number 10 on page 1 of Annex II-B.

7) Norman J. Horing, "Aspects of Solid State/Semiconductor Physics Research in China", *Scientific Information Bulletin* 17 (2) 1992, pp. 47-55.

8) *Physics Today* (January 1990), p. 66.

9) The ISI translations of Chinese organizational names is fraught with potential confusion, muddling distinctions (or suggesting invalid ones) between universities, colleges, schools, departments, centers, and institutes. A number of Institutes (INST) are listed as part of the Chinese Academy of Science, but all CCAST addresses listed Centers (CTR) and not Institutes *unless* the Academy was included as part of the address listing, with the Institute appearing after the Academy's comma delineator. Examination of a number of CCAST articles in various (U.S.) physics journals showed that "Center" is the term being used by the journals and being entered by ISI.

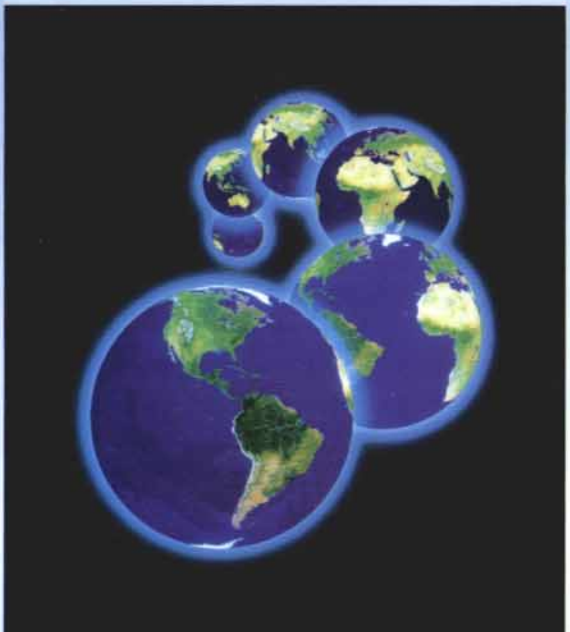
10) Catherine P. Ailes, *et al. New Directions for U.S.-Latin American Cooperation in Science and Technology* (Arlington, VA: SRI International, June 1988).

11) Wu Kangsheng *et al.*, "A Survey and Analysis on the Development of Microelectronics Technology in China," paper presented at *U.S.-PRC Joint Workshop on Measurement of the Impact of S&T on Development*, Sponsored by the U.S. National Science Foundation and the PRC State Science and Technology Committee, Arlington, VA, September 17, 1987.

12) The following paragraphs are based on the Executive Summary of the separate report on the panel's evaluation of PRC capabilities: H.R. Coward and R.R. Fresne, *A Mini-Review of Condensed Matter Physics and Materials Science Research in the People's Republic of China* (Arlington, VA: SRI International Science and Technology Policy Program, November 1992).

LES SCIENCES HORS D'OCCIDENT
AU XX^E SIÈCLE

SÉRIE SOUS LA DIRECTION
DE ROLAND WAAST



VOLUME 6

LES SCIENCES AU SUD
ÉTAT DES LIEUX

ROLAND WAAST
ÉDITEUR SCIENTIFIQUE

ORSTOM
éditions

LES SCIENCES HORS D'OCCIDENT
AU XXᴱ SIÈCLE

20TH CENTURY SCIENCES:
BEYOND THE METROPOLIS

SÉRIE SOUS LA DIRECTION
DE ROLAND WAAST

VOLUME 6

LES SCIENCES AU SUD
ÉTAT DES LIEUX

SCIENCES IN THE SOUTH
CURRENT ISSUES

ROLAND WAAST
ÉDITEUR SCIENTIFIQUE

ORSTOM Éditions
INSTITUT FRANÇAIS DE RECHERCHE SCIENTIFIQUE POUR LE DÉVELOPPEMENT EN COOPÉRATION
PARIS 1996