Jointly published by Akadémiai Kiadó, Budapest and Kluwer Academic Publishers, Dordrecht

Scientometrics, Vol. 60, No. 2 (2004) 000–000

Publication and citation patterns of the Mexican contribution to a "Big Science" discipline: Elementary particle physics

FRANCISCO COLLAZO REYES,[a] ELENA LUNA MORALES,[b] JANE M. RUSSELL[c]

[a] *Centro de Investigación y de Estudios Avanzados del IPN, Biblioteca de Ciencias Exactas, Mexico City (Mexico)*
[b] *Centro de Investigación y de Estudios Avanzados del IPN, Unidad de Servicios Bibliográficos, Mexico City (Mexico)*
[c] *Universidad Nacional Autónoma de México, Centro Universitario de Investigaciones Bibliotecológicas, Mexico City (Mexico)*

The publication and citation patterns of the Mexican community in elementary particle physics (MEPP) were determined by bibliometric analysis of the scientific production and citations registered in the SPIRES-HEP system from 1971 to 2000. All papers, both citing and cited, were classified as theoretical, phenomenological or experimental according to the type of study carried out and citing papers as local (Mexican) or foreign. The growth dynamics of the citation patterns over the thirty-year period was also studied. Results show that the Mexican scientific community in EPP follow the pre-publication and pre-citation communication patterns typical of a Big Science field.

Introduction

Elementary particle physics (EPP), as a Big Science discipline, is highly dependent on a number of factors: technology provided by the large experimental centres; international research collaboration; multi-institutional funding,[1,2] as well as a communication system based on prepublication, initially in print form and from the beginning of the 90s via an open archive system of e-prints developed by Ginsparg[3] which intensified the exchange of original research results between physicists working in EPP.[4]

Preprints, as papers soon to be published,[5,6] have always played an important role in the communication between physicists.[7] Their use and effectiveness in the dissemination of results between researchers and institutions increased considerably after the Second World War. An example of this is the implementation of the CERN (European Organisation for Nuclear Research) web which was soon to become the most

Received January 14, 2004

Address for correspondence:
FRANCISCO COLLAZO REYES
Centro de Investigación y de Estudios Avanzados del IPN, Biblioteca de Ciencias Exactas
Av. IPN 2508, Col. San Pedro Zacatenco, c.p. 07000, Mexico City, Mexico
E-mail: collazo@linux.bfm.cinvestav.mx

0138–9130/2004/US $ 20.00

important interface for publication, dissemination and remote access to research results in the area,[8–10] supported by the development of communication and information transfer protocols.[11] The web has served as a platform for the development of different applications among which is the information retrieval system we use in the present study: Stanford Public Information Retrieval System – High Energy Physics (SPIRES-HEP), implemented in the Stanford Linear Accelerator Centre (SLAC) for the management of documents in their library, including the valuable collection of preprints.[12,13] The system today brings together the archives of distinct centres and specialised libraries in the field such as the Deutsches Electronen Synchrotron (DESY), CERN and the Fermi National Accelerator Laboratory (FERMILAB) which is one of the most successful e-print servers internationally, positioning prepublication, electronic formats and the Internet at the centre of the scientific communication system in EPP and leading to new patterns of publication, dissemination and citation in science.[14–16] This new communication system has also generated controversy within groups responsible for evaluation of scientific activity because citations to non-commercial literature quantifiable in systems such as SPIRES-HEP, are not considered by the *Science Citation Index* (SCI).[17]

Previous studies[18] have shown that the maturing process of the Mexican EPP scientific community under local conditions which started three decades ago, is presently consolidating efforts in terms of production and scientific impact, research focus, diversification of programmes, training of human resources and the integration of research groups.[19] One important aspect is the efforts being made by the Mexican specialists in EPP towards incorporation into the production, collaboration and communication methods characteristic of Big Science experimental research.[20,21]

The objective of the present study is to document and quantify the influences exerted by the ways of Big Science on local communication processes by determining regularities in publication and citation patterns during the major growth phase in the maturing process of the Mexican EPP community. A strong interaction between different types of research (theoretical, phenomenological, experimental) is considered healthy for all scientific disciplines.[22,23]

Material and methods

The SPIRES-HEP was used to identify citations from 1971–2000 to papers authored during the same period by Mexican specialists in EPP and registered in a local database (Mexican Physics of Elementary Particle – MEPP). The MEPP database was constructed by downloading records from the SPIRES-HEP that included Mexico as the country of authorship. The records were subsequently checked manually and anomalies rectified. Cited and citing papers were classified as theoretical, phenomenological or experimental according to descriptions found in SPIRES-HEP, CERN-Library, INSPEC

and in the subject classifications of PACS (Physics and Astronomy Classification Scheme). Citing papers were also classified as local (with the participation of Mexican institutions) or foreign (with the sole participation of foreign institutions). Catalogues published by the Mexican Society of Physics (1987–2000)[24] helped with the identification of scientists working in Mexico. We then looked at the following variables: annual growth patterns of cited and citing papers; the annual ratios of theoretical, phenomenological and experimental source papers; the relationship between cited and citing papers with respect to these three research categories; and their division between local and foreign citations.

Statistical analysis

Citation growth patterns were analysed using time series analysis, correlation and adjustments to different growth tendencies. Growth dynamics were determined by dividing each of the values of the annual series by the arithmetic average (taken as the basic unit of division) of the citations from 1971–2000. The same procedure was used to determine the growth dynamics for each decade.

Results

Table 1 shows the overall results of production, impact and citation growth dynamics of the Mexican contribution to EPP from 1971–2000. Of the total number of 2,301 papers produced, 1,006 were entered into SPIRES-HEP as e-prints, 788 as journal articles and 507 as proceedings, technical reports, pre-prints or theses. SPIRES-HEP registers the change from e-print status to published article but keeps the same code for the bibliographical record which allowed us to determine that of the 1,006 e-prints, 645 have been accepted for journal publication which, with the 788 originally entered as journal articles, makes a total of 1,433 published papers when we carried out our analysis. Of the 1,039 cited papers, 58% were to the 603 papers that entered the system as e-prints, 41% (426) as journal articles and 1% (10) as proceedings. Another important characteristic of SPIRES-HEP is that the citation levels of cited papers can be monitored before and after publication.

According to the SPIRES-HEP classification scheme 1 paper of the Mexican EPP scientific community was in the "renowned" category with > 500 citations, 6 were classified as "famous" with >100 citations, 19 as "well-known" with >50 citations, 202 as "known", with >10 citations and 811 as "less known" as they were cited between one and nine times.

Table 1. Production, impact and citation growth dynamics

Years	Published papers	Cited papers	Citations per year of publication	Citations per year of citation	General growth dynamics	Growth dynamics per decade
1971	2	1	0	0	0.00	0.00
1972	1	0	2	0	0.01	0.02
1973	3	2	44	0	0.13	0.48
1974	10	5	22	2	0.06	0.24
1975	27	16	218	10	0.64	2.36
1976	25	11	54	22	0.16	0.59
1977	48	21	181	31	0.53	1.96
1978	34	17	161	82	0.47	1.74
1979	30	12	66	94	0.19	0.72
1980	31	11	175	69	0.52	1.90
1981	42	15	115	69	0.34	0.82
1982	38	16	198	85	0.58	1.41
1983	31	22	73	78	0.22	0.52
1984	31	12	57	83	0.17	0.41
1985	28	8	66	74	0.19	0.47
1986	33	9	19	57	0.06	0.14
1987	26	14	124	78	0.37	0.88
1988	36	14	77	64	0.23	0.55
1989	39	25	416	58	1.23	2.97
1990	52	18	257	112	0.76	1.83
1991	83	24	301	197	0.89	0.38
1992	75	35	233	195	0.69	0.30
1993	108	33	237	168	0.70	0.30
1994	125	60	677	346	1.99	0.86
1995	162	98	1871	658	5.51	2.38
1996	199	105	1171	1045	3.45	1.49
1997	215	109	968	1191	2.85	1.23
1998	227	118	1056	1251	3.11	1.34
1999	247	112	847	1614	2.49	1.08
2000	293	96	499	1950	1.47	0.63
Totals	2 301	1 039	10 185			

Source. Local database on MEPP (Mexican Elementary Particle Physics)

Taking into consideration that the citation index of SPIRES-HEP is constructed using the references contained in the documents as initially entered into the system and that no updating occurs with respect to possible changes in the reference lists between the e-print and the final article, the majority of the citations to MEPP were made in the pre-publication period. The 1,039 cited papers received a total of 10,185 citations (Table 1) from 6,307 different citing papers: 69% (7,058) were from 4,409 e-prints,

16% (1,637 citations) from 1,060 papers entered into the system as journal articles and 15% (1,490 citations) from 838 documents classified as proceedings, technical reports, pre-prints and theses. Of the 4,409 citing e-prints, 2,976 changed their status to published journal articles which, with the 1,060 that entered the system already published, gives a total of 4,036 articles in 119 journal titles, 93 of which are indexed in the SCI and which account for 97.5% of citations from journal articles. The remaining citations from journal articles were scattered over a total of 26 journal titles not included in the SCI.

Figure 1 shows the trends in the publication patterns of the MEPP according to type of paper. Apart from a general increase in production during the period studied we note a change over time with respect to the different publication types used. The first two decades were characterised by a constant proportional use of the more traditional dissemination media such as journal articles and, to a lesser extent, proceedings, technical reports and pre-prints. During this time the level of scientific production remained basically static.

However, in the 90s we see radical changes. During this period the overall volume of papers produced increased dramatically and e-prints became the most important form for the dissemination of results for the MEPP community, as can be seen in Figure 1.

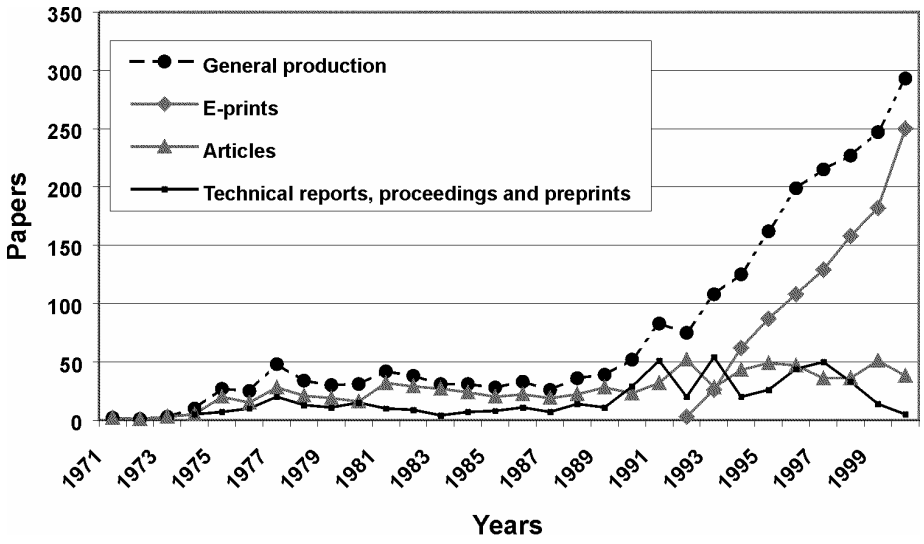

Figure 1. Publication patterns of MEPP

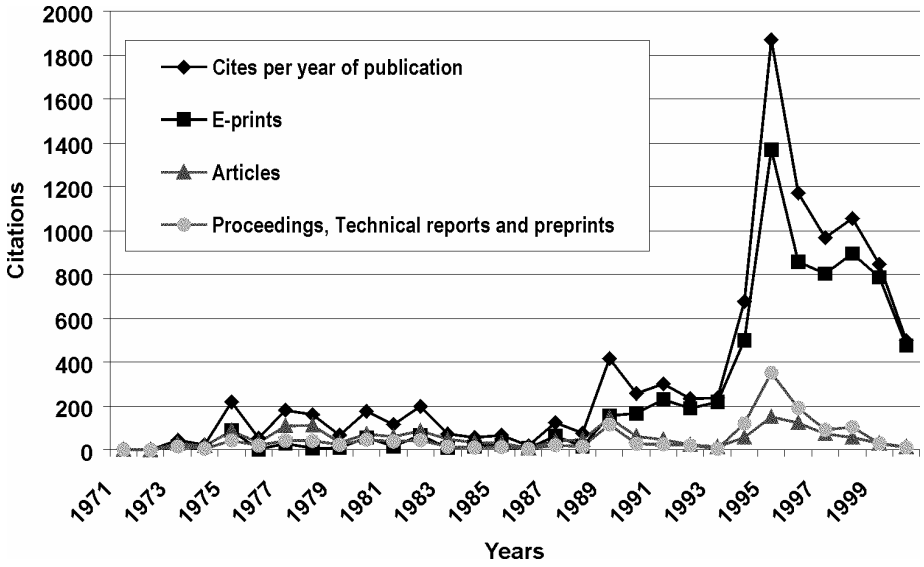

Figure 2. Citation sources to MEPP

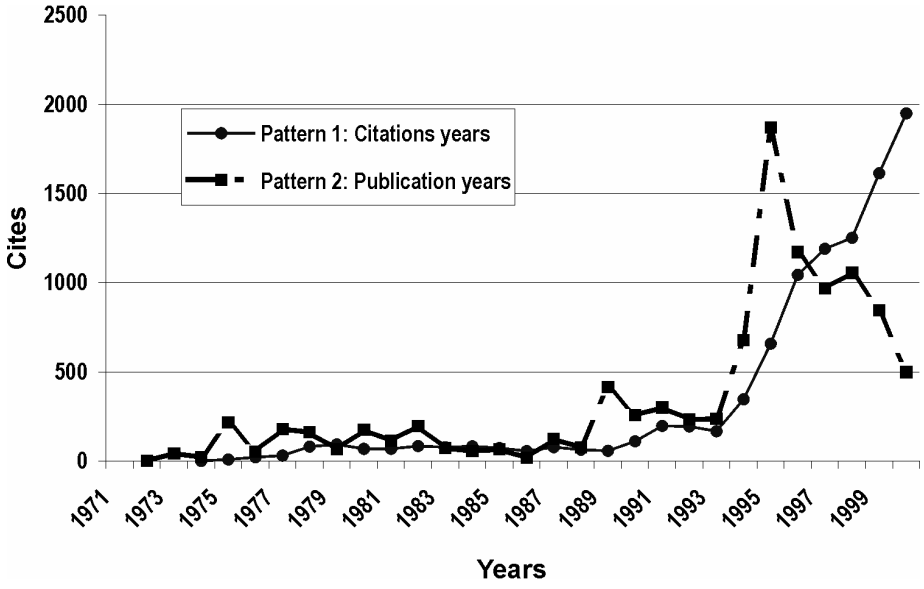

Figure 3. Citation patterns of MEPP

The citations made to the MEPP literature by the three principal types of documents are shown in Figure 2. As in the previous figure e-prints emerge around the beginning of the 90s from which time they increase their importance dramatically so that by 1999 they are responsible for 95% of all citations. Simultaneously we see a gradual decline in the journal article as the major type of citing document.

Figure 3 shows the general growth of citations using the data of Table 1: the broken line gives the data for publication years (column 4) and the continuous line, the accumulated increases per citation year (column 5). The distributions given by patterns 1 and 2 magnify the periods of greatest concentration of citations. Pattern 1 consists of two distinct citation phases: a static phase during the first two decades followed by an unexpected rate of growth, based on the trend of the previous years, during the final decade. Pattern 2 shows atypical behaviour in the sense that it does not conform to any regular growth tendency. Nonetheless, it shows the same contrast as pattern 1 between the first two decades and the last as well as indicating the years when relevant papers were produced.

Given that Pattern 2, publication year of the cited documents, better characterizes the citation distribution of MEPP, we have used this in Figure 4 to analyze growth with respect to two criteria that allow comparison between the general growth dynamics and those obtained during each one of the decades. The range of values on the vertical axis correspond to the scale for the citation growth dynamics, as seen in Table 1. The values in column 6 give rise to the continuous line representing the general dynamics and those in column 7 to the broken lines for each of the decades, resulting from the division of the annual series by the arithmetic average (taken as the basic unit of division) of the citations for 1971–2000, for each decade.

The growth dynamics by decade show a more balanced redistribution of the citations than does the overall pattern, weighting in relative terms the existence of equally important growth in the citations throughout the period, as can be observed in 1975, 1977, 1978, 1980, 1989 and 1995. When examining separately each one of these years we find that growth in 1975 and in 1980, are due to the publication of two important studies, one in the field of relativity and gravitation and the other in elementary particles, both classified in SPIRES-HEP as “famous papers” cited between 100 and 499 times and both theoretical studies developed locally by a single author. The first concerns equation solving and the second is a review of the Yang-Mills quantum theory. Unlike these examples, the growth peaks in 1977 and 1978 are the result of the collective effort of several publications each with 20 or more citations that together accumulate more than 100 citations. These are mainly theoretical studies from individual scientists.

The end of the 80’s saw the highest proportional growth between the annual series of the three decades as can be seen in Figure 4, due to the accumulation of citations to a highly cited paper as well as the collective contribution of 10 studies with an average of

20 citations each one. Of these 11 studies, written mainly in coauthorship, five are theoretical, five phenomenological, and one experimental in international collaboration. Unlike earlier decades where the most highly cited papers were theoretical or phenomenological, carried out locally and written either individually or by two or three authors, in the 90s the most notable growth, occurring in 1995 was the result of eight experimental studies, one with 816 citations and seven others with an average of 71 citations. All were multi-institutional and multi-author international collaboration projects. Among these is the paper on the discovery of the top quark which is the most cited of the studies analysed during this period.

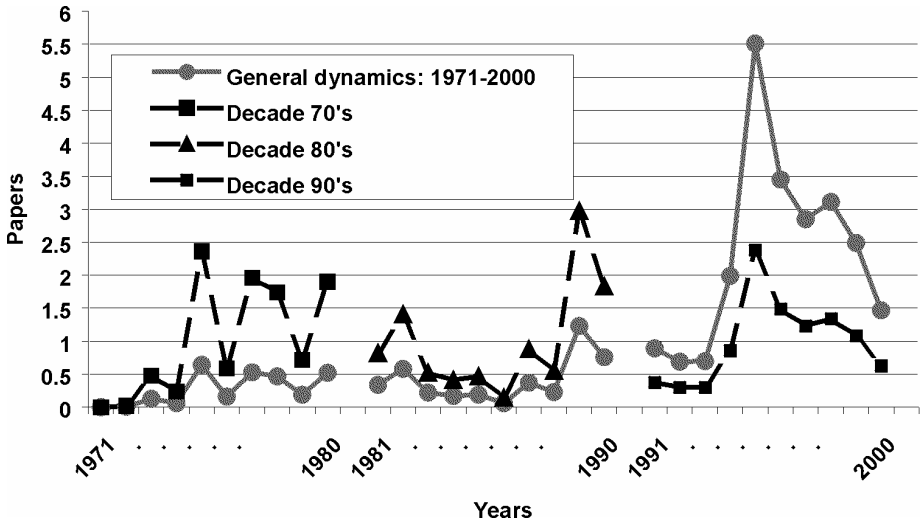

Figure 4. Comparison of citations growth dynamics

Table 2 shows that citation in MEPP occurs mainly between the same types of research, e.g. theoretical to theoretical and experimental to experimental. Cross citing occurs mainly by phenomenological papers to experimental and theoretical studies. Experimental studies influence phenomenological papers while theoretical studies produce little echo in experimental papers and vice-versa. Table 2 also shows that the MEPP papers are cited mainly by phenomenological and theoretical studies, the most cited papers being theoretical and experimental. A considerable difference exists between the average number of citations received by the three types of studies, experimental, phenomenological and theoretical, 29.8, 6.7 and 6.6, respectively.

When looking at differences in the citation patterns between the three types of research with respect to MEPP papers involving national institutions (Table 3) we find

the same situation in that the most common citing/cited relationships is between papers of the same type and especially theoretical to theoretical and experimental to experimental.

No evidence was found that experimental papers were influencing theoretical studies. The relationship of the theoretical and the phenomenological papers towards the experimental ones was the least frequent. Cross citing was most frequent from phenomenological to theoretical papers, followed by theoretical documents towards phenomenological papers. MEPP papers are cited mainly by theoretical and experimental studies, the source papers in the same categories being the most cited. Of the total number of 10,185 citations, a quarter of which (2,567) involve researchers from Mexican institutions, 998 correspond to experimental studies involving international Big Science type collaboration, 551 to theoretical and phenomenological papers written with foreign coauthors, and 1,018 to papers from Mexican institutions either in coauthorship or individually.

Table 2. Relationships between cited and citing papers by types of research

	Cited papers	Theoretical citations	Phenomenological citations	Experimental citations	Totals: cited study
Theoretical	658	3298	960	97	4355
Phenomenological	240	302	1194	125	1621
Experimental	141	226	1909	2074	4209
Totals: Citing		3 826	4 063	2 296	10 185

Source. Local database on MEPP

Table 3. Local citations. Relationships between cited and citing papers by types of research

	Theoretical citations	Phenomenological citations	Experimental citations	Totals: cited studies
Theoretical	1007	179	0	1186
Phenomenological	72	291	31	394
Experimental	3	20	964	987
Totals: Citing studies	1 082	490	995	2 567

Source. Local data base on MPEP

Discussion

The publication and citation patterns of MEPP and the SPIRES-HEP information system, that made their characterisation possible, are part of a new expression of the traditional pre-publication culture of EPP that, in the process of transition from pre-print to e-print, has created an alternative scientific communication and information transfer structure that monitors the use, dynamics and influences of scientific literature prior to publication as previously noted by Manuel[16] Ramalho-Correia & Castro[4] and Langer.[14] Traditional information systems such as Physics Abstracts, INSPEC, *Science Citation Index* (SCI) do not share this characteristic. Consequently, citation analysis of the MEPP using SPIRES-HEP is a novel alternative to the traditional procedure carried out using SCI, that offers a wider coverage of types of cited and citing documents and the identification of a greater number of citations.

Our results show an increasing preference, in quantitative terms, of the MEPP community for publishing and distributing e-prints and their use as cited and citing documents, together with a gradual decline in journal articles in these roles. This indicates that the e-print culture has penetrated the scientific communication structure of the MEPP community and is affecting local publication and citation processes.

Results suggest a maturing process of MEPP in two contrasting stages: the first static and the other showing unexpected growth during which we see distinct publication and citation patterns. The first stage is characterised by a constant level in the scientific production, a dominance of theoretical studies, a predominance of the traditional model of scientific communication centred on commercial print publications and an absence of experimental studies. The second stage presents very different characteristics that reflect a new way of doing science in the MEPP community, based on a more productive organisational structure, diversification of types of research and a new role for the experimental research groups in the development of the discipline. The resulting publication and citation patterns are more representative of the second stage dynamics.

The development of the line corresponding to Pattern 1 shows the effect of the concentration of citations towards the most recent years, linked to the accumulation of studies which is also greater during this period. This phenomenon, according to previous studies,[19,21] coincides with periods where more resources were available, with the growth of research groups and kinds of research in MEPP, minimising what occurred during the initial years, associated with lower production and less infrastructure and resources.

The citation distribution seen in Pattern 2, is clearly concentrated around the mid 90s, owing principally to the participation of Mexican physicists at the most crucial moment of the Collaboration DO project which led to the discovery of the top quark in

1995. The papers reporting this discovery were immediately and highly cited, occurring during the same year of publication, which is shown graphically as a sudden and ephemeral effect lasting only a year, which modified the citation growth tendency for 1995. This situation affected not only MEPP but Mexican physics in general.

Unlike the immediate citation experienced by experimental studies, the period of greater citation to theoretical and phenomenological publications during the 70s did not occur during the same period but were distributed over the three decades analyzed. This situation is clear when comparing the two lines corresponding to citation patterns by year of publication and year of citation. From these data we can conclude that theoretical papers have a longer citing life than the experimental ones.

Unlike the cumulative effects shown in the citation patterns 1 and 2, the distribution of the citations in relation to the growth dynamics of each of the decades shows a novel redistribution with respect to the earlier ones. In relative terms this appears more balanced, redirecting attention towards the citation processes occurring in each one of the periods which we believe require explications different from those for the cumulative phenomena and more appropriate to the conditions and contexts of each decade which give rise to the cited works. These aspects merit further study.

The relationships found between cited and citing works with respect to the three different types of research focus (theoretical, phenomenological and experimental) in MEPP indicate the following important characteristics: a) theoretical and experimental studies are carried out independently; b) the closest relations between physicists carrying out the different kinds of research is the influence of the experimentalists on the phenomenologists resulting mainly from the discovery of "top quark". This situation is not reflected within the local community where the "top quark" experimental work has passed unnoticed, seemingly, by the theoretical and phenomenological researchers.

Attracted by one of the major periods of discovery in EPP (1950–1975), specialists in the study of science carried out a citation analysis during this period of one of its most dynamic fields, weak interactions. They found that the interrelationship between theoretical, phenomenological, and experimental research in moments of greatest progress show changing patterns. For example, between 1953 and 1961,[25,26] three different patterns were found 1) from 1953–1954 phenomenology was highly dependent on experimental work, 2) from 1955–1956 the interrelationship was between theoretical and phenomenological studies and 3) 1957–1961 phenomenological and experimental research showed dependence on general theory. The moment of greatest dependence of theory on experimental results was identified in the area of weak-electromagnetic unification[27,28] during the 70s.

The growth of MEPP, occurring mainly in the decades of the 80s and 90s, coincides with the consolidation period of the standard model at international level as one of the most solid theoretical frameworks within the history of the discipline. Experimental discoveries including the top quark, were not unexpected events but rather were widely

foreseen , the results of which helped to strengthen the scientific paradigm dominant in EPP. In this sense, our analysis corresponds to the moment of predominance of theoretical knowledge over experimental findings and our results show that this period is not the best moment for the interaction between the different types of research and that the scant relationship between theoretical and experimental aspects is by means of phenomenological research. These circumstances help to explain the isolation of theoretical and experimental physicists at local level.

*

This paper has been partially supported by project of Physics Department of the CINVESTAV

References

1. D. CRANE, *Invisible College: Diffusion of Knowledge in Scientific Communities,* University of Chicago Press, Chicago, 1972, pp. 22–40.
2. M. J. MORAVCSIK, The contemporaneous assessment of a big science discipline. In: *The Evaluation of Scientific Research*, Conference, Wiley, Ciba Foundation, London Chichester, 1989, pp.188–200.
3. P. GINSPARG, *Electronic Publishing in Science, Winners and Losers in The Global Research Village*, CIUC/UNESCO, Paris, 1996. [Online].
   Available: http://www.epub.org.br/papers/ginsparg.htm (May, 2002).
4. A. M. RAMALHO-CORREIA, M. CASTRO-NET, The role of eprint archive in the access to, and dissemination of, scientific gray literature: LIZA a case study by the National Library of Portugal, *Journal of Information Science,* 28 (2002) 231–241.
5. C. M. BROWN, The E-volution of preprints in the scholarly communication of physicists and astronomers. *Journal of the American Society for Information Science and Technology*, 52 (2001) 187–200.
6. G. K. YOUNGEN, Citation patterns to electronic preprints in the astronomy and astrophysics literature, *College & Research Libraries*, September (1998) 448–456.
7. L. LUNIN, J. M. HURD, The changing communication system of science: behavioural and organizational aspects. *Journal of the American Society for Information Science*, 51 (2000) 1276–1278.
8. B. White, The World Wide Web and high-energy physics, *Physics Today*, (11) (1998), 30–36.
9. S. R. TOMPSON, A tale of two worlds: high-energy preprints in the 1990s, *Science and Technology Libraries*, 19 (2) (2001) 43–51.
10. J. GILLIES, R. CAILLIAU, *How the Web Was Born: The Story of the World Wide Web*, Oxford University Press, New York, 2000, p. 372.
11. C. VIZCAÍNO-SAHAGÚN, Redes, integración corporativa. *Pc/Tips*, (6) (1991) 26–41.
12. H. B. O´CONNELL, Physicists thriving with paperless publishing. ArXiv:physics/0007040 (2000) 2 1–17, [Online], Available: http://www.stanford.edu/spires/hep, (September, 2002)
13. P. A. KREITZ, L. ADDIS, H. GALIC, T. JOHNSON, The virtual library action: collaborative international control of higher-energy physics pre-prints, *Publishing Research Quarterly*, 13 (2) (1997) 24.
14. J. LANGER, Physicists in the new era of electronic publishing: new electronic opportunities for scientific communication, and changes in how scientists use journals, are prompting a complete revaluation, *Physics Today* [Online], (2002), Available: http://www.aip.org/pt/vol-53/iss-8/p35.html, (May, 2002).
15. J. M. RUSSELL, Scientific communication at the beginning of the twenty-first century. *International Social Science Journal*, 168 (2001) 273–282.
16. K. MANUEL, The place of e-prints in the publication patterns of physical scientists, *Science & Technology Libraries*, 20 (1) (2001) 59–85.

17. N. R. KAPLAN, M. L. NELSON, Determining the publication impact of a digital library, *Journal of the American Society for Information Science*, 51 (1993) 324–339.
18. F. COLLAZO-REYES, Participación del CINVESTAV en disciplinas de gran ciencia: física de partículas elementales, *Avance y Perspectiva*, 21 (2002) 361–374.
19. M. E. LUNA-MORALES, F. COLLAZO-REYES, El síndrome Big Science y su influencia en el proceso de maduración de la física mexicana de partículas elementales, *Revista Española de Documentación Científica*, 24 (4) (2002) 186–197.
20. G. HERRERA-CORRAL, Física experimental de altas energías. *Avance y Perspectiva*, 12 (1993) 81–85.
21. F. COLLAZO-REYES, M. E. LUNA-MORALES, Física mexicana de campos y partículas: organización, producción y crecimiento. *Interciencia*, 27 (2002) 347–353.
22. M. J. MORAVCSIK, The crisis in particles physics. *Research Policy*, 6 (1977) 78–107.
23. S. S. SCHWEBER, Physics community and the crisis in physical theory, *Physics Today*, (4) (1993) 34–40.
24. SOCIEDAD MEXICANA DE FÍSICA, FEDERACIÓN LATINOAMERICANA DE SOCIEDADES DE FÍSICA, *Catálogo Latinoamericano de Programas y Recursos Humanos en Física*, SMF: FLSF, México, (1987–2000).
25. D. H. WHITE, D. SULLIVAN, E. J. BARBONI, The interdependence of theory and experiment in revolutionary science: the case of Parity Violation, *Social Studies of Science*, 9 (1979) 303–327.
26. D. H. WHITE, D. SULLIVAN, Social currents in Weak Interactions: The controversy surrounding possible violation of the ^S = ^Q selection rule yields useful insights into the rewards system of a scientific field, *Physics Today*, April (1979) 40–47.
27. D. SULLIVAN, D. KOESTER, D. H. WHITE, R. KERN, Understanding rapid theoretical change in particle physics: A month-by-month co-citation analysis, *Scientometrics,* 2 (1980) 309–319.
28. D. KOESTER, D. SULLIVAN, D. H. WHITE, Theory selection in particle physics: A quantitative case study of the evolution of weak-electromagnetic unification theory, *Social Studies of Science*, 12 (1982) 73–100.