R&D Management and the Stokes Diagram: An Exploratory Study

Willy Hoppe de Sousa[1], Mery Piedad Zamudio Igami[1],

Diógenes de Souza Bido[2]

Abstract

The Stokes Diagram model was developed to replace the linear mode view of a progression of science by a two dimensional plane in which the vertical dimension represents the search of understanding and the horizontal dimension represents the use. Although this diagram is easy to understand, no cases of organizations or government programs were identified in literature which developed a method to put this proposal in practice. Thus, based on the development of a conceptual and operating model, both dimensions were measured and validated through structural equation modeling, using information from 28 dissertations developed in a Postgraduate program. The results suggest that the proposed method is viable and give some examples of its use to analyze and manage R&D activities.

Key words: Pasteur´s quadrant; Stokes diagram; R&D management; knowledge management.

[1] Instituto de Pesquisas Energéticas e Nucleares IPEN - CNEN/SP, Prof. Lineu Prestes 2242, São Paulo, S.P., Brazil, 05508-000.
(Corresponding author) E-mail: whsousa@ipen.br, mery@ipen.br
[2] Universidade Presbiteriana Mackenzie, Rua da Consolação, 930, São Paulo, S.P., Brazil, 01302-907.
E-mail: diogenesbido@yahoo.com.br

1. Introduction

The evolution of humanity has been accompanied by efforts to master knowledge. In the last 600 years, it is the existence of a work- and knowledge-focused society that explains why some countries have become much richer than what could be expected based on their dimensions or military power (Landes, 1998).

Different types of knowledge can be involved in this evolutionary process - popular, scientific, philosophical and religious (Trujillo Ferrari, 1974) but, according to Sáenz and Capote (2002), it is scientific knowledge, more specifically about how to practice economic activities and about forms in which social relations are manifested that make a social system function appropriately.

Although the acceleration of scientific knowledge evolution dates back to the 17th century Scientific Revolution and the Industrial Revolution that started in the 19th and 20th centuries, it is at the end of the Second World War that the countries constituting the main scientific powers started to define policies that influenced the orientation of how science should be conducted in the next years or decades.

One landmark in the definition of these policies is the Bush Report (1945) entitled Science, the Endless Frontier. This report not only set a perspective on how the United States could keep up their scientific research investments when the war was over, but also described basic science and its relation with technological innovation. Half a century later, it became the framework of that country's scientific policy for the post-war decades (Stokes, 1997).

Bush' view can be summarized in two premises: (1) basic research is performed without thinking about practical goals and (2) basic research is a precursor of technological progress (Stokes, 1997). The former supports a static version of the relation between basic and applied research as parts of a one-dimensional space, while the second underlies a dynamic view, with basic research leading to applied research and development and, subsequently, to production or operations, depending on whether it is product- or operating research. This became known as the linear model.

However, half a century after the Second World War, the conceptual split between basic and applied science started to be questioned:

> Indeed, these doubts have appeared in each of the major industrial countries. It is no longer believed that a heavy investment in pure, curiosity-driven basic science will by itself guarantee the technology required to compete in the world economy and meet a full spectrum of other societal needs. (Stokes, 1997).

Stokes clearly manifests his unease with the "paradigm" of the linear model, still very present in the scientific community and the general population, when he states the problem as follows:

> The belief that the goals of understanding and use are inherently in conflict, and that the categories of basic and applied research are necessarily separate, is itself in tension with the actual experience of science. (Stokes, 1997).

According to Stokes, although much research can either be classified as one or another category, highly important studies have shown that studies' successive choices are influenced at the same time by understanding and by use, as observed by Pasteur's microbiological studies, Langmuir's work in physics-chemistry or Noyes and Lewis' research in chemical engineering, among others.

In Pasteur's case, the problem of obtaining alcohol from beet juice is quite illustrative. Pasteur, initially trying to understand racemic acid, discovered micro-organisms responsible for the alcoholic fermentation of beet juice. That is, to the extent that his research became progressively more fundamental, the problems he chose and the research lines became progressively more applied.

Although the Organization for Economic Cooperation and Development (OECD) made efforts to improve the definitions of research categories, according to Stokes, the problem of summarizing the spectrum of basic-applied research in one single dimension persisted. How could Pasteur's work – directed at both understanding and use - be represented in this one-dimensional spectrum? Stokes' conclusion is that, because these works need to be simultaneously represented in the two ends of the basic-applied spectrum, they generate an anomalous

representation in the one-dimensional Euclidean space (Stokes, 1997).

To solve this problem, Stokes suggested the construction of a bidimensional plan, in which one axis represents the extent to which a certain research seeks to go beyond the frontiers of fundamental understanding, while the horizontal axis represents the extent to which the same research is guided by usage considerations. Although there may be many degrees of commitment to these two objectives, for the sake of simplification, Stokes proposes a

structure with two dichotomies in the form of a diagram, composed of four quadrants, as presented in Fig. 1.

To facilitate the understanding of his proposal, Stokes proposes a characterization for each of the quadrants in the diagram. The upper left quadrant includes basic research and the researcher representing this quadrant is Niels Bohr, due to his efforts to develop an atomic model, independently of what this proposal would represent to the world further ahead.

Considerations about the use?

Research inspired by:

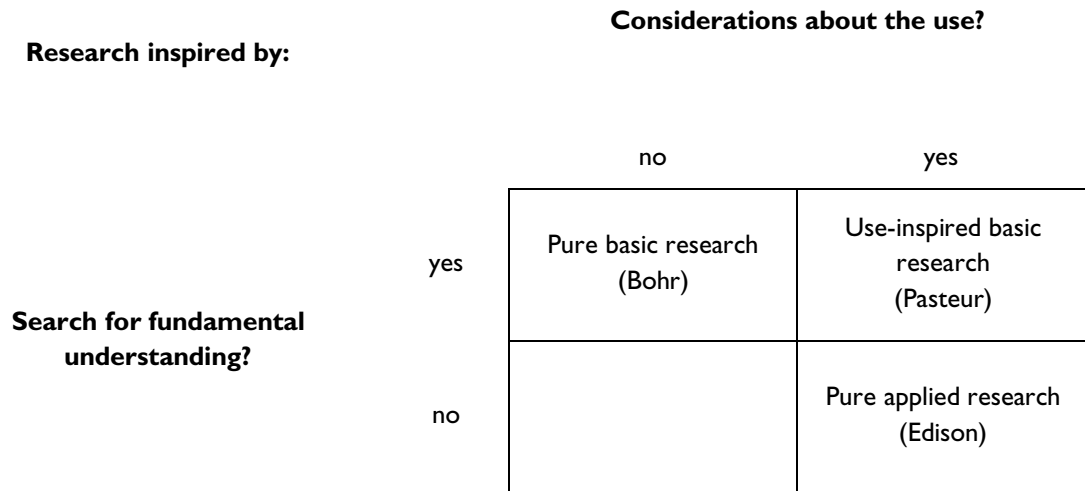|		no	yes
Search for fundamental understanding?	yes	Pure basic research (Bohr)	Use-inspired basic research (Pasteur)
	no		Pure applied research (Edison)

Figure 1. Quadrant model of scientific research. (Source: Stokes, 1997)

The lower right quadrant represents research conducted for applied purposes only, without any concern with understanding more general phenomena underlying the research. Stokes called this quadrant Edison's because of the way Thomas Edison pursued the search for a commercially profitable electric illumination system and impeded his collaborators from seeking a more in-depth understanding.

The upper right quadrant represents research that simultaneously seeks to extent the frontiers of knowledge but is also inspired by use, which could be represented by Pasteur.

The lower left quadrant represents studies that neither take into account fundamental understanding, nor are they use-oriented. This is the case of more specific studies, directed at particular facts, or precursor studies of research for the Bohr quadrant, according to Stokes.

The understanding of Stokes' proposal is quite intuitive and its use can be observed in different contexts: as a proposal to guide the development of a research agenda (Simmons et al., 2005; Price, Behrens, 2003); as part of a theoretical model on knowledge production (Tsao et al., 2008); as part of funding agencies' analyses to formulate public policies in science, technology and innovation (Brito Cruz, 2008), in a discussion presented in the editorial of an academic journal reflecting on – basic and applied – research and its utility (Balaram, 2008) and even as part of a presentation in application processes for a leading function in a research institution (Camara, 2005).

It is remarkable, however, that in databases examined in mid-2008 (Proquest, Scopus, Web of Science and specifically Technological Innovation Management Symposia from 1998 till 2008 and the Latin American Technology Association from 2005 and 2007), using the key words "Quadrant" and "Pasteur", only "Pasteur" and only

97

"Stokes", no paper or study was found that developed some kind of application of this model, as an operating tool for project or research program analysis.

This arouses the following questions: Which analyses could be developed by applying Stokes' Diagram: Could comparative analyses be developed between different research programs? Could any change in research profile be identified, based on strategic decisions made in the management of research programs? Could the dynamics in the profile of the dissertations developed under one single supervisor across a time period be observed? This study aims to put Stokes' diagram in practice and present some examples of these applications.

2. Development of conceptual research model

To start the search for answers to these questions, a conceptual research model was developed (Miles and Huberman, 1994) to represent how Stokes' Diagram would be put in practice.

To start the development of this model, it was taken into account that: a) (scientific) research represents a search for knowledge through an investigation of solutions to scientific and social problems, using objective and systematic analyses (Rajasekar, *et al.*, 2006); b) (scientific) research can be understood as one form of the knowledge

work process, that is, a structure for action – a process with a set of activities executed over time, with defined inputs and outputs (Davenport, *et al.*, 1996) and involves, c) cognitive processes that demand substantial creative efforts or mental activities (Amavaradi, Lee, 2005).

Thus, if scientific research can be understood as a process, then both the content of what is being processed during a research and the way this content is being processed can be represented by variables of the input, processing and output of the research process for each of the two dimensions of the diagram. Thus each of the two axes of Stokes' diagram could be put in practice by unfolding them in these three variables: input, processing and output.

As to the "Considerations about use" axis, the variable *nature of the problem* was defined as the input variable; to represent the result profile as a function of the research process used, the variable *nature of the research* was defined and, to represent the immediate applicability profile of the research results, the variable *perspective of immediate use was defined*. In the "Search for fundamental understanding" axis, the *knowledge requisites* was defined as the input variable to conduct this process; to represent the process in itself, a variable called *knowledge generation process* was defined and, finally, as the output variable, the *knowledge progress* profile generated by the process was defined.
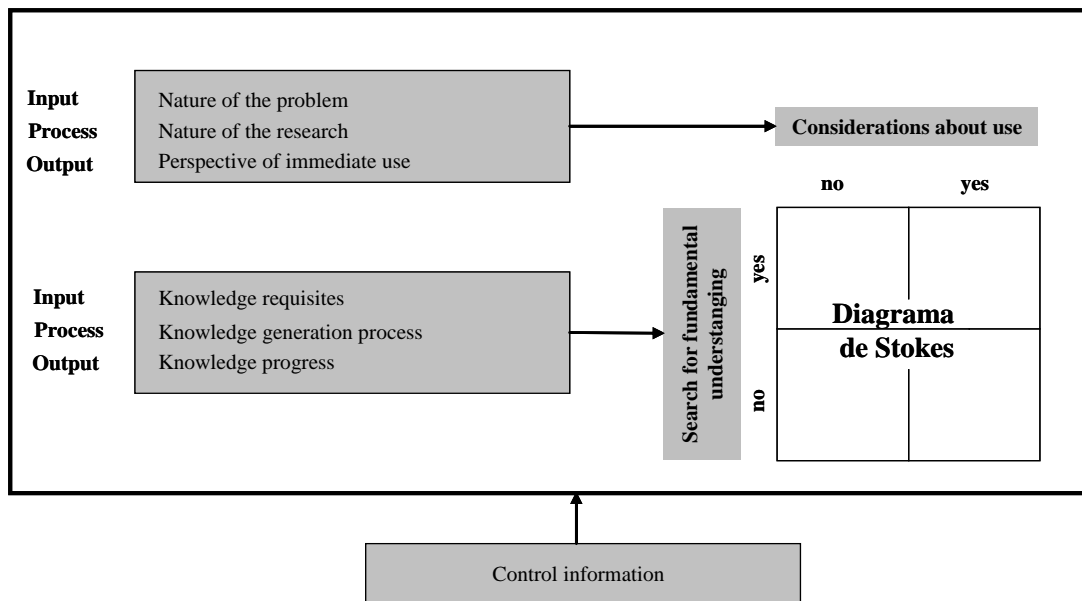

Figure 2. conceptual research model

These definitions and the inclusion of a set of control variables for the research object, defined in the operating part of the research, conclude the conceptual research model, as presented in Fig. 2. Part 3, which discusses the research method, gives details on how these variables are used in practice.

3. Research method

To develop the practical part of the research, the following steps were developed: a) Definition of research material and information sources; b) Development of data collection instrument; c) Definition of analysis method; d) Definition of the organization where the study would be developed and e) Operational research model. Each of these steps is defined next.

3.1. Definition of research material and information sources

The research material consisted of doctoral dissertations because these represent knowledge production processes that follow the traditional mode of knowledge generation which, among other aspects, is characterized by problems and solutions ruled by the academic community, homogeneous production, quality control by peer review and involves emphasis on individual creativity (Gibbons, Limoges, Nowotny, Schwartzmann, Scott and Trow, 1994).

The dissertation authors were defined as the information sources. In the cases were the authors could not be contacted, their respectives supervisors were contacted.

3.2. Development of data collection instrument

The data collection instrument was developed as a questionnaire with closed questions, organized in three blocks: the first block contains six questions, three to measure the "Considerations about use" axis of the conceptual research model and three to measure the "Search for fundamental understanding" axis. Thus, for each axis, three closed questions were elaborated to represent the "input", "process" and "output" variable. The procedures adopted to categorize each of the variables are presented in the next two blocks.

3.2.1 *"Considerations about use" axis*

To represent the *nature of the problem* variable, variations of two situations were considered: (a) problems can be more theoretical, such as studies on natural phenomena or pure science, which are not concerned with an immediate solution to any problem, or (b) problems can be more practical, when they aim for solutions of immediate use (Rajasekar *et al.*, 1996), the following categories were elaborated: (1)."totally theoretical"; (2)."more theoretical than practical"; (3)."balance between theoretical and practical"; (4)."more theoretical than practical" and (5)."totally practical".

The *nature of the research* variable was constructed to represent the association between the nature of the research and the type of result associated with this research, based on the definitions found in the Frascati Manual (OECD, 2002). Thus, the following categories were established for this variable: (1)."pure basic research, without aiming for immediate economic or social benefits or for solutions to practical problems"; (2)."basic research aimed at knowledge construction in an undefined future"; (3). "basic research oriented towards knowledge construction for use in an undefined future"; "basic research oriented towards knowledge construction for use in the near future"; (4)."applied research to solve problems and/or reach goals defined in the present" and (5). "experimental development to obtain new products, processes, materials, equipments, systems or their improvements in the present".

As for the *perspective of immediate use* variable to construct the categories, the classification of the research results could be a function of two aspects: (a) the time horizon considered (Saenz and Garcia Capote, 2002) and (b) the scaling of the experimental development results which can range from the laboratory to the production scale (Saenz and Garcia Capote, 2002). Thus, the following categories of immediate use of research results were elaborated: (1)."theoretical foundations for other strictly theoretical studies"; (2)."theoretical foundations for other theoretical and experimental studies"; (3)."incorporation in technologies at laboratory level"; (4)."incorporation in technologies on a pilot scale" and (5)."incorporation in technologies for commercial use".

The premise underlying the categorization for each of these three variables is as follows: the lower the number of the answer chosen, the lesser the degree of applicability of the material under analysis (doctoral dissertations in this case); the higher the number, the higher the degree of applicability.

3.2.2 "Search for fundamental understanding" axis

To represent the variable *knowledge requisites*, the profile a researcher needs to have to integrate, process and generate new knowledge was considered. Root-Bernstein believes that great discoverers and inventors are people with knowledge in more than one area who discover by questioning premises and by experimenting; on the other hand, these authors also consider that these great inventors can be individuals who contributed within the same disciplinary domain – like in the case of Nobel Prize winner Paul Ehrlich. These can be people without any formal education to become inventors, like in the case of Samuel Morse, Nikola Tesla or Thomas Edison (Root-Bernstein, 1989).

Thus, to represent this spectrum of requisites, the following categories were developed: (1)."profound theoretical-practical knowledge in a specific area"; (2)."good theoretical and practical knowledge in some specific knowledge areas"; (3)."limited theoretical and practical knowledge in some specific knowledge areas"; (4)."limited theoretical and good practical knowledge in few areas" and (5)."no theoretical and in-depth practical knowledge in many knowledge areas".

The *knowledge generation process* variable was elaborated to assess if new knowledge is generated by (a) understanding why things happen (Sim and Duffy, 2004; Rajasekar, *et al.*, 2006), (b) by a more specific or more general understanding (Root-Bernstein, 1989; Stokes, 1997; Sim and Duffy, 2004) or (c) by experimentation (Root-Bernstein, 1989) or (d) as precursor of oriented research or fundamental understanding, (Stokes, 1997) or (e) combination of multiples approaches (Bigwood, 2004). To represent combinations of the different possible knowledge generation forms, the following categories were developed: (1). "deepening and understanding of existing knowledge on the same knowledge base as the initial research problem"; (2)."deepening and understanding of existing knowledge on a broader knowledge base than the initial research problem"; (3)."integration and classification-

systemization of existing knowledge; (4)."experimentation and aggregation of existing knowledge to the same knowledge base as the initial research problem and (5)."experimentation and aggregation of new knowledge, involving a broader knowledge base than the initial research problem".

And, finally, for the third variable in this dimension – *knowledge progress* – the following options were made available: (1)."extraordinary scientific advancement – publication potential in an international journal equivalent to Nature or Science"; (2)."significant scientific advancement – publication potential in international journals equivalent to Qualis A[3]; (3). "moderate scientific advancement – publication potential in national journals equivalent to Qualis A and/or potential to obtain a national patent"; (4). "significant technological-scientific advancement – publication potential in international journals equivalent to Qualis A and/or potential to obtain an international patent" and (5). "extraordinary technological advancement - potential rupture in quality of life".

The following premise supports the categorization of each of these variables: the lower the answer number chosen, the greater the degree of effort to seek a fundamental understanding of the research material and, the higher the number, the higher the practical and technological nature of the research[4].

During the questionnaire development phase, a pilot test was applied among senior researchers at the Energy and Nuclear Research Institute (IPEN), with a view to verifying whether the interviewees understood both the formulated questions and the available answers.

It was observed that, although the questionnaire had gone through different reviews to clarify the meaning of questions, the interviewees still faced difficulties to understand some of the terms used in the questionnaire.

---

[3] Qualis evaluation refers to a Brazilian governmental evaluation process that, among other aspects, evaluates the quality of national and international journals and proceedings. Grade "A" represents the best evaluation. For further details, access http://www.capes.gov.br/avaliacao/qualis.

[4] Obs.: In order to facilitate data interpretation (e.g.: the higher a scale grade, the higher the degree of fundamentality it represents), the scale of these three items was inverted for data analysis (1→ 5, 2 → 4, 3 → 3, 4 → 2 e 5 → 1).

Therefore, it was decided that, for some of the questions, the categories would be accompanied by notes, giving examples of representative situations and illustrating their meaning better.

A sixth option, "I cannot assess", was also offered for each of the six variables, so as to identify possible comprehension difficulties by the interviewee. In those cases when these options were marked, the interviewees were contacted later and their doubts were clarified, so that they could answer the question in one of the five categories that ranked the research variables.

To support the analysis of answers to the research questions presented earlier as well to open the possibility to explore new ones in the future, some control variables were surveyed for each of the dissertations (either in the questionnaire itself or through additional surveys): (1) characterization of research objectives (2) dissertation author's education; (3) year of conclusion (dissertation defense year); and (4) supervisor's education.

First, data were collected by contacting the dissertation authors, through direct interviews as well as by e-mail. In those cases when the author could not be contacted, data were collected through a direct interview with the supervisor. In all cases, together with the questionnaire, the interviewees received a letter explaining the research objectives.

3.3. Definition of analysis method

For data analysis, a multivariate data technique called structural equation modeling (SEM) was chosen. This is due to the fact that this technique has two distinct characteristics of interest for this research: (1) estimation of multiples and inter-related relations of dependence and (2) ability to represent non-observable concepts in these relations and to explain the measurement error in the estimation process (Hair Jr. et al., 2005).

The estimation method of the SEM was component-based (PLS-PM: Partial Least Squares Path Modeling or simply PLS). The following figure among the main reasons for using that method in this study: (1) greater flexibility in terms of requisites of the theory under analysis (that is, it is appropriate for exploratory studies); (2) there is no need to make suppositions on the distribution of data; (3) there may be both formative and reflexive indicators in the

measurement model and (4) small requisite in terms of sample size in comparison with other estimation methods (Zwicker, *et al.*, 2008).

3.4. Choice of the organization

The Energy and Nuclear Research Institute (IPEN), located in São Paulo State, Brazil, was chosen to develop the research. This choice was due to the fact that this organization develops a successful Postgraduate program – the Master's and Doctoral program received CAPES[5] score 6 in recent years – associated with the University of São Paulo, having concluded more than 1000 Master's and Doctoral defense sessions since its creation in 1976. Another factor in this choice was the authors' easy access to the technical staff that constituted the target of this research.

IPEN is a scientific and technological research institution, founded in 1956, which is active in the nuclear and related areas. Its budget mainly comes from the Brazilian federal government. It has a fixed staff of 1029 employees, of whom 219 have a Doctoral and 118 a Master's degree (IPEN, 2007).

Between the mid-1980's and 1990's, the Brazilian Nuclear Program lost strategic importance and IPEN had to reorient its education programs – which used to focus on the Federal Government demands and now focus also on demands from society in general. Its Board especially responded to calls from funding agencies to finance its research activities. These efforts were relatively successful, but led to great diversification in its activities. To try and decrease this problem, strategic planning efforts started in 1996 with the definition of IPEN's mission. In 2000, this culminated in the formalization of the institution's first Master Plan.

To develop the research, Doctoral graduates and supervisors from two Research Programs were invited to participate: Lasers and Biotechnology. These two programs were chosen because their scientific-technological contents are very distinct, so as to compare the implications of fitting dissertations produced in these programs into the

---

[5] CAPES evaluation refers to the postgraduate program evaluation developed by the Coordination for the Improvement of Higher Education Personnel (CAPES), a governmental department of the Brazilian Ministry of Education. The score each postgraduate program can reach ranges from 1 to 7.

Stokes Diagram. One control variable called *Institutional Program* was developed to classify the dissertations in one of these two research programs.

To assess possible alterations in the research profile of the doctoral dissertations in function of the implantation of IPEN's Master Plan, the analyzed dissertations were separated in two groups: before and after 2004. This four-year lag with the year the Master Plan was put in practice is due to the fact that four years is the normal maximum period to elaborate a doctoral dissertation in Brazil. This was meant to guarantee that, after 2004, only dissertations that started after the implantation of the Master Plan would be incorporated. Two control variables were used to support this analysis: *Year of Master Plan* and *Year of conclusion*.

Since the creation of IPEN's Postgraduate program in 1976, 44 doctoral theses were defended in the Lasers area and the same number in the Biotechnology area. In this research, 28 dissertations were analyzed, representing about 34% of each knowledge area. These were selected in function of the facility to contact and get answers from the interviewees.

3.5 Operational research model

To put the research in practice, the conceptual model was expanded to include to representation of latent variables and their inter-relation and detail the control variables as presented in Fig. 3.
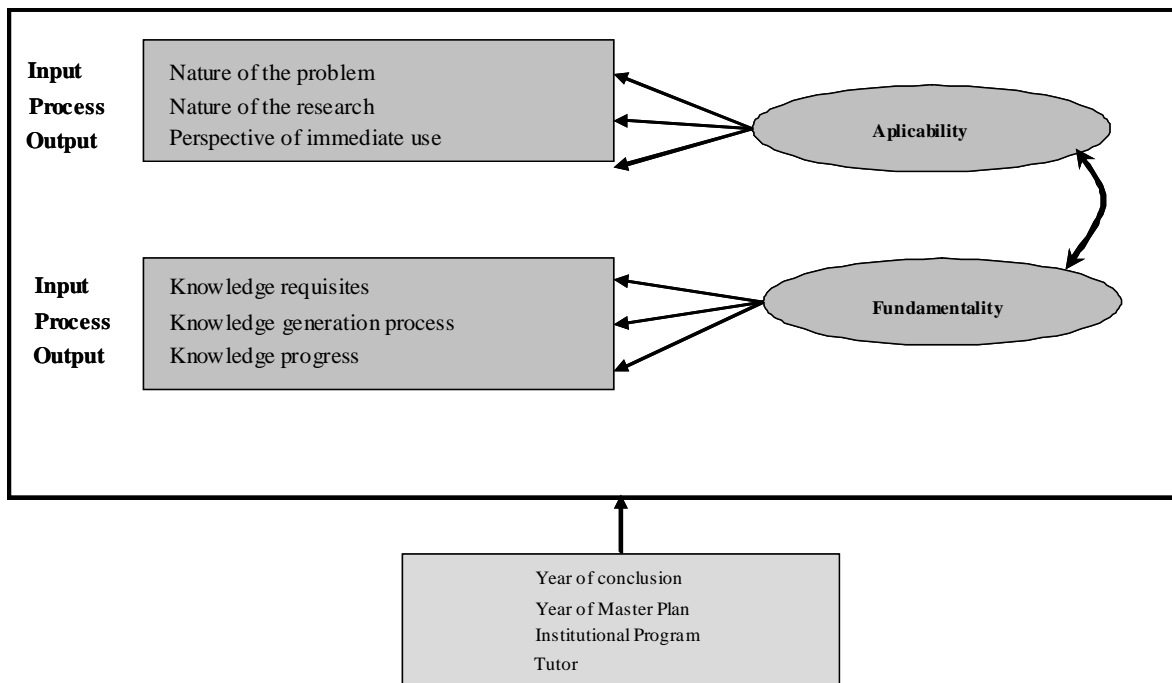

Figure 3. Operational research model

This model represents how the conceptual research model was investigated in order to answer the research questions presented at the end of the introduction, and can be explained as follows:

(1) The latent variables (VL) Applicability and Fundamentality can present a mutual inter-relation. Intuitively, a negative correlation between them is

expected, as more basic research would tend towards less applicability and vice-versa;

(2) As commented in the section about the development of the instrument, the three variables or indicators (input, process and output) were elaborated so as to reflect the construct's behavior or VL to be measured, which is why a correlation between them is expected. Thus, a reflexive measurement model was used;

(3) The control variables (year of conclusion, year of master plan, institutional program and supervisor) are connected with the research questions and will be considered in the graphic analyses of the Stokes Diagram and presented in the next section.

4. Result analysis

4.1 Data about the research and the interviewees

Twenty-eight interviews were held, 14 for the Lasers Program and 14 for the Biotechnology Program. As to the

distribution of the authors' answers, 12 were answered by the dissertation author himself and 16 by the dissertation supervisor. In terms of the authors' education, 13 interviewees had a degree in Exact and Earth Sciences and 15 in Life Sciences.

4.2 Representative model of the analyzed dissertations

The first step of the quantitative analysis was accomplished through univariate and bivariate statistics, summarized in Tab. 1.

	Aplicability			Fundamentality		
	1	**2**	**3**	**4**	**5**	**6**
1 - Nature of the problem	1,000					
2 - Nature of the research	0,666	1,000				
3 - Perspective of immediate use	0,291	0,396	1,000			
4 - Knowledge requisites	-0,468	-0,265	0,036	1,000		
5 - Knowledge generation process	-0,383	-0,391	-0,121	-0,350	1,000	
6 - Knowledge progress	0,006	-0,225	0,124	0,131	0,162	1,000
Mean =	3,8	3,8	2,9	3,7	2,7	3,7
Median =	4	4	3	4	2	4
Standard error =	0,69	0,79	0,66	0,66	1,65	0,61
Frequency distribution						
Answers = 1	0	0	0	0	9	0
Answers = 2	1	1	6	3	8	2
Answers = 3	7	9	19	2	0	5
Answers = 4	17	13	2	23	4	21
Answers = 5	3	5	1	0	7	0

Obs: Correlations equal to or higher than |0.382| were significant at 5% .

Table 1. Pearson's correlation matrix and descriptive statistics of indicators

Some results should be highlighted:

a) The three Applicability indicators are positively correlated, as expected, and main component analysis showed that 64% of the variance could be explained by only one extracted component and the factor loads varied from 0.65 to 0.89, which resulted in a Cronbach's alpha of 0.72;

b) The three Fundamentality indicators presented low and non-significant correlations, except for the

correlation between Knowledge generation process and Knowledge requisites, which presented a negative correlation significant at 10%. In the main components analysis, two components were extracted with self-value higher than 1. In view of these results, the appropriateness of measuring this VL as a reflexive model was questioned, which was reconsidered in subsequent sections;

c) One plausible explanation for the unexpected results related to Fundamentality can be the homogeneity of the analyzed dissertations in terms of Knowledge profile and Knowledge progress, which obtained 23 and 21 answers in the same category, respectively, out of 28 cases. In future research, a larger and more varied sample could solve this problem.

To estimate the model presented in Fig. 3, SmartPLS 2.0.M3 software was used (Ringle, *et al.*, 2005). Given the exploratory nature of this research, the measurement model was changed (Fig. 4) based on the results obtained for this sample, explained next.
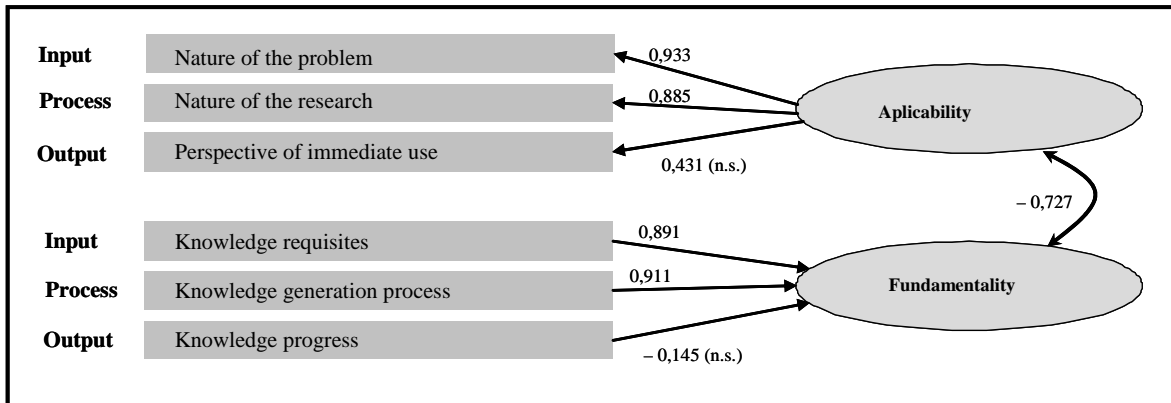
Obs: All coefficients were significant (p < 0.001), except when indicated (n.s.).

Significance rates were calculated by bootstrap, using 28 cases and 400 repetitions.

In *SmartPLS,* the double arrow (correlation between VL) was estimated as a simple regression (one-directional arrow) and results were the same, no matter which of the VL was defined as the dependent or independent variable.

Figure 4: Revised model of laser and biotechnology program dissertations

The original model (reflexive VL) presented adequate results for Applicability, that is, extracted variance of 0.61 and composite reliability of 0.81, higher than Chin's recommendations (Chin, 1998). However, results were inadequate for Fundamentality: extracted variance of 0.27 and composite reliability of 0.53, much below the minimum values of 0.5 and 0.7, respectively.

Considering the results in Tab. 1 and estimation by PLS-PM, it was decided to consider the Fundamentality indicators as formative (Jarvis, *et al.*, 2003), which is more coherent with the low correlations between its indicators.

To assess the consequences of this *ad hoc* decision, correlations between the factor scores were estimated in a formative and reflexive way. Both for Applicability and Fundamentality, they exceeded 0.94, indicating that this decision exerted little effect on the estimation of the factor scores used later in the Stokes' Diagram.

Besides, the correlation between Applicability and Fundamentality remained at an intermediary level between the totally reflexive model and the totally formative model, which is equivalent to a canonical correlation, as presented next:

- Applicability (reflexive) and Fundamentality (reflexive) → r = − 0,69 (p < 0.001)
- Applicability (reflexive) and Fundamentality (formative) → r = − 0,73 (p < 0.001)
- Applicability (formative) and Fundamentality (formative) → r = − 0,79 (p < 0.001)

The revised model with the results obtained in *SmartPLS* is presented in Fig. 4. As *SmartPLS* produces scores for the VL on the same scale as the indicators, in this case from 1 to 5, these scores could be used to represent all dissertations in the Stokes Diagram according to the constructed model. The result is illustrated in Fig. 5 and Tab. 2.
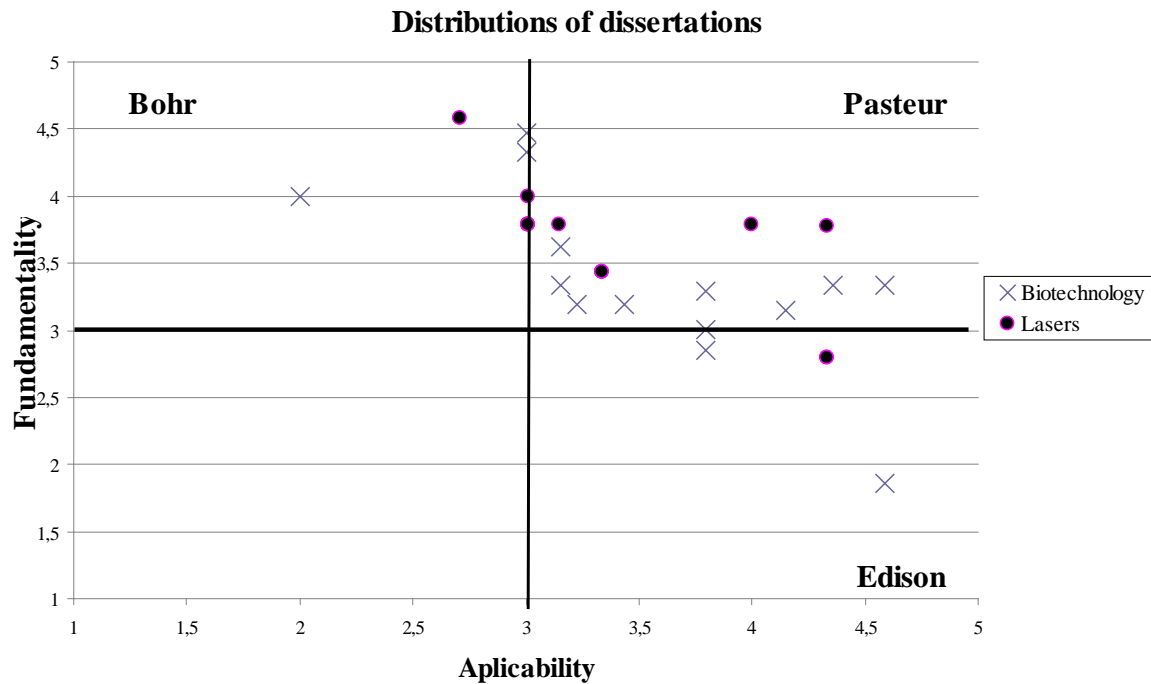
104

Figure 5. Graphic distribution of theses from the Laser and Biotechnology programs in Stokes' Quadrants

Quadrant	Biotechnology		Lasers	
	n	%	n	%
Bohr	1	7,1%	2	14,3%
Edison	2	14,3%	1	7,1%
Pasteur	11	78,6%	11	78,6%
undefined	0	0,0%	0	0,0%
total	14	100,0%	14	100,0%

Table 2. Percentage distribution of theses from the Laser and Biotechnology programs in Stokes' Quadrants

The results presented above suggest that the two Postgraduate programs are quite similar in terms of their dissertations' distribution in the Stokes Diagram. Analyzing the reasons for this similarity lies outside the scope of this paper, but it does answer the first research question, that is, comparative analyses between research programs can be developed by means of the Stokes Diagram.

The second analysis refers to the possibility of using the Stokes Diagram to identify any change in the research profile, based on strategic research program management decisions. Therefore, one strategic decision was identified and a reference year was defined in order to separate the dissertations database in two groups: before and after that year. In this case, the strategic decision considered was the

implantation of the IPEN Master Plan in 2000 but the reference year was 2004 due to the following reasons: (1) only doctoral research started and defended after that year could somehow be influenced by this implantation and (2) on the average, it takes four years to conclude a doctoral dissertation. The results are presented in Fig. 6 and Tab. 3.
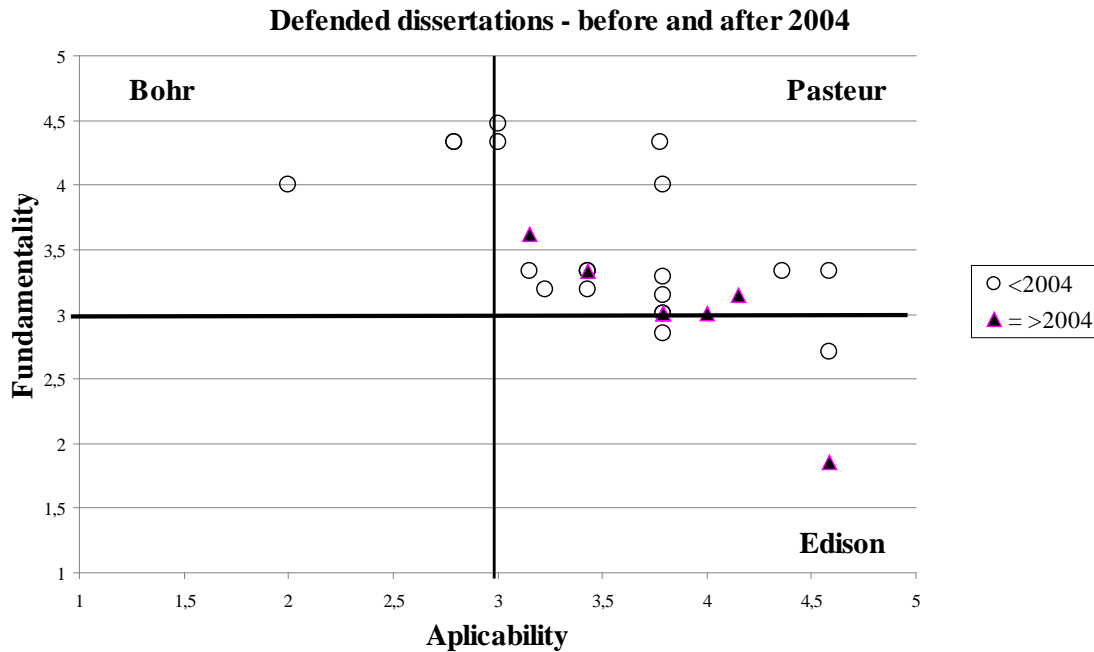

Defended dissertations - before and after 2004

Figure 6: Graphic distribution of theses defended before and after 2004.

Quadrant	< 2004		= > 2004	
	n	**%**	**n**	**%**
Bohr	3	15,0%	0	0,0%
Edison	2	10,0%	1	12,5%
Pasteur	15	75,0%	7	87,5%
undefined	0	0,0%	0	0,0%
total	**20**	**100,0%**	**8**	**100,0%**

Table 3. Percentage distribution of dissertations defended before and after 2004

The small number of cases per cell in Tab. 3 does not permit statistical tests like chi-square. Limited to this sample, however, the results suggest some alteration in the profile of dissertations defended before and after 2004.

The following was observed: (1) before 2004, 15% of defended dissertations fitted into the Bohr Quadrant; as from 2004, none of the analyzed dissertations fitted into this Quadrant; (2) dissertations defended as from 2004 showed a slight increase by 2.5% in the Edison Quadrant and a 12.5% increase in the Pasteur Quadrant.

On the whole, it is concluded that a change occurred in the profile of the analyzed dissertations, moving towards applicable results. Hence, the Stokes Diagram can also be a useful tool for this type of analysis. It should specifically be highlighted that this analysis allows for the identification of a change in the profile of the dissertations defended as from 2004, although it cannot be affirmed that the

introduction of the Master Plan was the cause of this change.

A third data analysis was carried out to check for the possibility of analyzing the dynamics in the profile of the

dissertations developed under the same supervisor during a certain time period. The results are presented in Fig. 7 and summarized in Tab. 4.
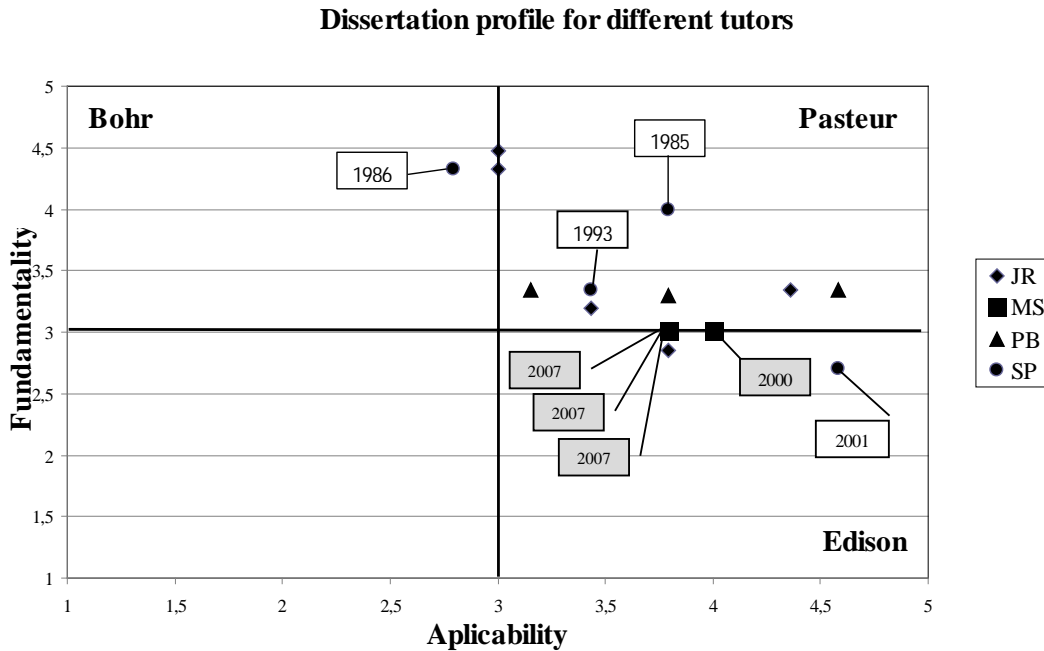
Dissertation profile for different tutors

Figure 7. Profile of dissertations elaborated under the advice of four different supervisors

Quadrant	JR		MS		PB		SP	
	n	%	n	%	n	%	n	%
Bohr	0	0,0%	0	0,0%	0	0,0%	1	25,0%
Edison	1	20,0%	0	0,0%	0	0,0%	1	25,0%
Pasteur	4	80,0%	4	100,0%	4	100,0%	2	50,0%
undefined	0	0,0%	0	0,0%	0	0,0%	0	0,0%
total	5	100,0%	4	100,0%	4	100,0%	4	100,0%

Table 4. Percentage distribution of theses under the advice of four different supervisors

Data in Fig. 7 and Tab. 4 represent the dissertations developed under four supervisors. For the sake of illustration, the dissertation defense years were highlighted for two supervisors.
Tab. 4 shows that the profile of the dissertations linked with the same supervisor can be subdivided in two groups: supervisors active in more than one Quadrant of the

Stokes Diagram (the case of JR and SP) and supervisors who work in one Quadrant only (the case of MS and PB supervisors).

The analysis of Fig. 7, on the other hand, discloses the time sequence of the dissertations developed under SP. Over the years, the dissertations gradually "moved" along the

107

fundamental dimension. At first, the dissertations focused more on scientific knowledge advancement; over time, the focus moved to scientific and technological knowledge advances. In the case of MS supervisor, it seems that he has concentrated his activities in an even more concentrated area of the Stokes Diagram over time.

In short, it is confirmed that the Stokes Diagram can be used as a tool to analyze the dynamics of dissertations developed under the same supervisor.

4.3 Research limitations

As this is an exploratory study, different limitations should be highlighted with a view to future improvements. In the first place, as mentioned earlier, for some research variables, some of the interviewees faced difficulties to understand some questions.

The second problem refers to the need for the interviewee to "go back to the time" when the dissertation was defended and to answer the questionnaire now in view of the time when the dissertation was defended. The effects of these and other difficulties to answer the questionnaire were captured by a reliability test. This test was developed by reapplying the questionnaire to 4 of all 28 interviewees five months after the first data collection. The reliability rate ranged between 50% and 67%, that is, below desirable levels (at least 70%, approximately).

Another problem refers to the low significance of two variables (*knowledge progress and perspective of immediate use*) – this low significance can be associated with the above mentioned difficulties to understand the questionnaire, but also with the small sample size (28 dissertations), although these represent 34% of all dissertation defenses.

And, finally, reflections are needed with a view to better understanding the implications of modeling the fundamentality indicators as formative.

In summary, the specific results achieved here should be considered more as a starting point for future research than as the final results of a modeling process for research program analyses.

Despite these limitations, the study demonstrates the potential use of the Stokes Diagram as an analytic tool to assess R&D programs and to supplement others managerial R&D activities (e.g.: project portfolio category evaluation (Bigwood, 2004); selection of the human resource management practices (UN, 2007) and innovation capability audit tool (Rush, Bessant and Hobday, 2007)).

5. Conclusions

Adequately assessing the profile of Research Programs and, consequently, guiding them so as to appropriately combine knowledge production and application, still represents a challenge for public Science & Technology policy makers and managers.

This paper represents a small but unedited contribution to put in practice a tool that was conceptually idealized by Donald Stokes. Some of its potential uses were presented here, but further research is needed now to improve the method with a view to developing similar studies in different and broader R&D scopes.

References

AMAVARADI, C, S., Lee, I. (2005). The dimensions of process knowledge, *Knowledge and Process Management*, 12, pp. 65-76.

BALARAM, P. (2008), Science, invention and Pasteur´s quadrant, *Current Science*, 94, pp. 961-962. http://www.ias.ac.in/currsci/_apr252008/961.pdf [Retrieved from the Web 22 May 2008]

BIGWOOD, M. P., (2004), Managing the new technology exploitation process. *Research Technology Management*, 47, pp. 38-42.

BRITO Cruz, C. H., (2008), Políticas para C&T&I para o Brasil. http://www.ifi.unicamp.br/~brito [Retrieved from the Web 13 Jun 2008]

CAMARA, G. (2005) Um Instituto, Três Missões, Dez Compromissos. http://www.dpi.inpe.br/gilberto/inpe/apresentacao_comite_busca.ppt [Retrieved from the Web 20 Apr 2008]

CHIN, W. W. (1998). The Partial Least Squares approach to structural equation modeling, in: Marcoulides, G. A. (Ed.), *Modern Methods for Business Research*. Lawrence Erlbaum Associates, Inc., USA.

DAVENPORT, T.H., Jarvenpaa, S. L., Beers, M. (1996). Improving knowledge work, *Sloan Management Review*, 34, pp. 53-65.

GIBBONS, M., Limoges, C., Nowotny, H., Schwarztman, S., Scott, P. and Trow, M. (1994) *The New Production of Knowledge*. Sage Publications, Inc., London.

HAIR Jr., J., Anderson, R. E., Tatham, R. L., Black, W. C., (2005). *Análise Multivariada de Dados*. Bookman, Porto Alegre

IPEN (2007). Relatório de Gestão, São Paulo.

JARVIS, C. B., Mackenzie, S. B., Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research*, 30, pp. 199-218.

MILES, M. B., Huberman, A. M. (1994). *Qualitative Data Analysis: an expanded sourcebook*. California Sage Publications, Inc.

OECD (2002). *Manual Frascati*. OECD publications service, Paris.

LANDES, D. S. (1998). *The Wealth and Poverty of Nations: why some are so rich and some so poor*. W. Norton & Company, New York.

PRICE, R. H., Behrens, T. (2003). Working Pasteur´s quadrant: harnessing science and action for community change. *American Journal of Community Psychology*, 31, pp. 219-223.

RAJASEKAR, S., Philominathan, P., Chinnathambi, V., (2006). Research methodology, *Ar XIV Physics*. http://arxiv.org/abs/physics/0601009v2 [Retrieved from the Web 22 May 2008]

RINGLE, C. M., Wende, S., Will, A. (2005). *SmartPLS 2.0 M3 (beta)*. University of Hamburg, Germany. http://www.smartpls.de [Retrieved from the Web 04 Sep 2009]

ROOT-BERNSTEIN, R. S., (1989). Who discovers and invents. *Research Technology Management*, 32, pp. 43-50.

RUSH, H., Bessant, J. and Hobday, M. (2007). Assessing the technological capabilities of firms: developing a policy tool. *R&D Management*, 37, pp. 221-236.

SÁENZ, T. W. Capote, E. C. (2002). *Ciência, Inovação e Gestão Tecnológica*. CNI/IEL/SENAI ABIPTI, Brasília.

STOKES, D. E. (1997). *Pasteur´s Quadrant: basic science and technological innovation*. The Brookings Institution, Washington.

SIM, S. K., Duffy, A. H. B. (2004). Knowledge transformers: a link between learning and creativity, *Artificial Intelligence for Engineering Design, Analysis and Manufacturing*, 18, pp. 271-279.

SIMMONS, P. E., Brunkhorst, H., Lunetta, V., Penick, J., Peterson, J., Pietrucha, B., Staver, J. (2005). Developing a research agenda in science education, *Journal of Science Education and Technology*, 14, pp. 239-252.

TSAO, J. Y., Boyack, K. W., Coltrin, M. E., Turnley, J. G., Gauster, W. B. (2008). Galileo´s stream: a framework for understanding knowledge production, *Research Policy*, 37, pp. 330-352.

TRUJILLO Ferrari, A. (1974) *Metodologia da Ciência*. 2 and 3rd. ed. Kennedy, Rio de Janeiro.

UN, C. A. (2007). Managing the innovators for exploration and exploitation. *Journal of Technology Management & Innovation*, 2, pp. 4-20.

ZWICKER, R., Souza, C. A., Bido, D. S. (2008). Uma revisão do modelo do grau de informatização de empresas: novas propostas de estimação e modelagem usando PLS (*partial least squares*), *Proceedings of the XXXII Encontro da Associação dos Programas de Pós-Graduação e Pesquisa em Administração*, ANPAD.