Tradition and Innovation:
The Historical Contingency of R&D Statistical Classifications

Benoît Godin

3465, rue Durocher
Montréal (QUE)
Canada H2X 2C6

Project on the History and Sociology of S&T Statistics

Working Paper No. 11

2001

Previous papers in the series:

1. B. Godin, *Outlines for a History of Science Measurement*.
2. B. Godin, *The Measure of Science and the Construction of a Statistical Territory: The Case of the National Capital Region (NCR)*.
3. B. Godin, *Measuring Science: Is There Basic Research Without Statistics?*
4. B. Godin, *Neglected Scientific Activities: The (Non) Measurement of Related Scientific Activities*.
5. H. Stead, *The Development of S&T Statistics in Canada: An Informal Account*.
6. B. Godin, *The Disappearance of Statistics on Basic Research in Canada: A Note*.
7. B. Godin, *Defining R&D: Is Research Always Systematic?*
8. B. Godin, *The Emergence of Science and Technology Indicators: Why Did Governments Supplement Statistics With Indicators?*
9. B. Godin, *The Number Makers: A Short History of Official Science and Technology Statistics*.
10. B. Godin, *Metadata: How Footnotes Make for Doubtful Numbers*.

Tradition and Innovation:
The Historical Contingency of R&D Statistical Classifications

In the last few years, the US National Research Council (NRC) has repeatedly criticized the National Science Foundation (NSF) for not measuring appropriate dimensions of science and technology. [1] The research enterprise has changed considerably since the 1950s when the first R&D surveys were conceived, but the scope of the surveys and the concepts measured, according to the NRC, have not really kept pace with this change: "Current science and technology indicators and data fall woefully short of illuminating a major part of the story of American industry in the 1990s (…)". [2]

This was only one of the many recent criticisms that have been voiced in OECD Member countries. In fact, not so long ago, national experts on R&D statistics were frequently complaining that the data collected and published by the OECD were too aggregated and insufficiently detailed. In the last ten years, Australia has been one of the main countries to argue for a better breakdown of R&D data by fields of science and socioeconomic objectives.

The problem has to do with the classifications used for measuring R&D and the way countries have applied them. Classifications are ways of imposing order onto the world, natural or social. [3] They separate the world into different entities that are subsequently

[1] National Research Council (2000), *Measuring the Science and Engineering Enterprise: Priorities for the Division of Science Resources Studies*, Washington; NRC (1999), *Securing America's Industrial Strength*, Washington. NRC (1997), *Industrial Research and Innovation Indicators*, Washington.
[2] NRC (1997), *op. cit*. p. 43.
[3] *C. Perelman (1970),* Réflexions philosophiques sur la classification, in *Le champ de l'argumentation*, Bruxelles: Presses universitaires de Bruxelles, pp. 353-358; N. Goodman (1978), *Ways of Worldmaking*, Indianapolis: Hackett Publishing; M. Douglas and D. Hull (1992), *How Classification Works: Nelson Goodman Among the Social Sciences*, Edinburg: Edinburg University Press; M. Foucault (1966), *Les mots et les choses*, Paris: Gallimard; J. Goody (1977), *The Domestication of the Savage Mind*, Cambridge: Cambridge University Press; A. Desrosières (1998), *The Politics of Large Numbers: A History of Statistical Reasoning*, Cambridge: Harvard University Press.

aggregated to give indicators. In so doing, they simplify reality into specific dimensions, often without consideration for the heterogeneity of the world.

The study of classifications reveals as much about the classifier than the classified. They carry people's representations of the (measured) world and disclose their understanding of the very world they are intended to measure. This paper will focus on the classifications used in R&D statistics.

A peculiar characteristic of R&D classifications is that they were borrowed from statistical series that were not specifically designed for measuring R&D. Official documents generally attributed this situation to the need for comparability with other statistics, namely national accounts, government expenditures, economic indicators, and education statistics. The third NSF industrial R&D survey, for example, was moved from the Department of Labor to the Bureau of Census in 1957 to improve comparability with industrial statistics. [4] Similarly, the first edition of the OECD Frascati manual stated that the classification of R&D data by economic sector "corresponds in most respects to the definitions and classifications employed in other statistics of national income and expenditure, thus facilitating comparison with existing statistical series, such as gross national product, net output, investment in fixed assets and so forth". [5]

This paper proposes a completely different hypothesis, namely that current R&D classifications, at least those in the OECD Frascati manual, were borrowed from othe classifications simply by virtue of their availability. With regard to the NSF's *Science Indicators* series, the General Accounting Office (GAO) of the United States referred to the procedure as "operationalism": the tendency to use existing measurements rather than develop new ones supported by explicit model of science and technology. [6] The OECD was consequently presented with an important problem: how to link R&D data if, although

[4] NSF (1960), *Funds for R&D in Industry: 1957*, Washington: NSF 60-49, pp. 97-100.
[5] OECD (1963), *The Measurement of Scientific and Technical Activities: Propose Standard Practice for Surveys of R&D*, Paris, p. 21.
[6] See: B. Godin (2001), *The Emergence of S&T Indicators: Why Did Governments Supplement Statistics with Indicators*, Montreal: OST.

comparable to other statistics, they remained poorly comparable between themselves? What should be done, for example, if the objectives of government funding of R&D cannot be linked to university performed R&D because the two sectors – government and university – are classified differently?

This paper complements a previous one called *Metadata*, which dealt with some of the limitations of R&D statistics. [7] It divides into two parts. The first presents how current classifications break down R&D data into the main economic sectors of national accounts: government, business, and university. It shows that these classifications came mainly from the United Nations, with slight modifications. Drawing upon the last two revisions of the Frascati manual in 1991 and 2000, which coincided with a decade of increasing policy demands for more detailed R&D statistics, the second part discusses the characteristics and problems of current classifications.

The System of National Accounts (SNA)

The first edition of the Frascati manual suggested classifying R&D by dimensions. One of the central dimensions proposed was the basic versus applied character of research activities. I have already dealt at length with this classification in a previous paper. [8] Another important dimension was concerned with economic sectors. In line with the system of national accounts, and following the practice of the NSF [9] – the first organization to survey all economic sectors systematically – [10] the manual recommended classifying R&D according to the following main economic sectors: business, government, private non-

[7] B. Godin (2001), *Metadata: How Footnotes Make for Doubtful Numbers*, Montreal: OST.
[8] See: B. Godin (2000), *Measuring Science: Is There Basic Research without Statistics?*, Montreal: OST.
[9] K. Arnow (1959), National Accounts on R&D: The National Science Foundation Experience, in NSF, *Methodological Aspects of Statistics on R&D: Costs and Manpower*, Washington: NSF: 57-61; H.E. Stirner (1959), A National Accounting System for Measuring the Intersectoral Flows of R&D Funds in the United States, in NSF, *Methodological Aspects of Statistics on R&D: Costs and Manpower*, Washington: NSF: 31-38.
[10] See: B. Godin (2001), *The Number Makers: A Short History of Science and Technology Statistics*, Montreal: OST.

profit. [11] To the three sectors, however, the OECD added, following the NSF again, a fourth one: higher education. The following rationale was offered for the innovation: [12]

> The definitions of the first three sectors are basically the same as in national accounts, but higher education is included as a separate main sector here because of the concentration of a large part of fundamental research activity in the universities and the crucial importance of these institutions in the formulation of an adequate national policy for R&D.

The system of national accounts, now in its fourth edition, was developed in the early fifties and conventionalized at the world level by the United Nations. [13] At the time, R&D was not recognized as a category of expenditures that deserved a specific mention in the national accounts. [14] The same holds true today: during the last revision of the system of national accounts ten years ago, the United Nations rejected the idea "because it was felt that it opened the door to the whole area of intangible investment". [15] It decided to instead develop a functional classification of expenditures that would make items such as R&D visible in the system of national accounts by way of what was called "satellite accounts". [16]

Despite its alignment to the system of national accounts, the Frascati manual still used a different system of classification in a number of cases, [17] including the coverage of each economic sector for example. [18] Nevertheless, the manual's specifications allowed one to follows the flow of funds between sectors (by way of a matrix), specifically between funder and performer of R&D, and to construct the main R&D indicator: the GERD (Gross Expenditures on R&D), defined as the sum of R&D expenditures of the four previously identified sectors.

[11] Households, as a sector in the SNA, was not considered by the manual.
[12] OECD (1963), *op. cit.*, p. 22.
[13] United Nations, **Title**; OECD (1958*), Standardized System of National Accounts*, Paris.
[14] Only institutions primarily engaged in research are singled out as a separate category.
[15] J.F. Minder (1991), *R&D in National Accounts*, DSTI/STII (91) 11, p. 3.
[16] See annex 11 of the 1993 edition of the Frascati manual.
[17] S. Peleg (2000), *Better Alignment of R&D Expenditures as in Frascati Manual with Existing Accounting Standards*, OECD/EAS/STP/NESTI (2000) 20; OECD (2001), *Better Alignment of the Frascati Manual with the System of National Accounts*, DSTI/EAS/STP/NESTI (2001)14/PART8.
[18] See Godin (2001), Metadata*, op. cit*. Montreal :OST.

The system of R&D classification was peculiar in that each sector had its own kind of classification. Whereas in most official surveys the units are analyzed according to a common system of classifications (every individual of a population, for example, is classified according to the same age structure), here three different kinds of units were distinguished and classified separately. The business sector was classified according to industries, the university (and private non-profit) sector according to fields of science (FOS) or scientific disciplines, and the government sector according to socioeconomic objectives (SEO). The principal recommendations regarding these classifications were made in the first edition of the Frascati manual in 1963.

Business Classification

The 1963 edition of the Frascati manual proposed classifying business R&D according to the International Standard Industrial Classification (ISIC) which has been in existence for more than a decade (1948). [19] Four main divisions were originally suggested:

1. agriculture, forestry, hunting and fishing
2. mining and quarrying
3. manufacturing industry
4. construction, utilities, commerce, transport and services

The manual further specified that it might sometimes be useful to subdivide the last two categories into subcategories, such as the separation of aircraft from transportation equipment because each was in itself particularly research intensive and of special interest from the standpoint of R&D. Today, the recommended classification, in line with the third revision of ISIC (1993), lists twelve main divisions, the more detailed being for the manufacturing industries (Annex 1).

[19] OECD (1963), op.cit. p. 22.

Grouping by industry was only one of the ways Business R&D could be classified. Following the NSF, the OECD also suggested a classification by product fields as early as 1970. If applied, it would allow for more detailed information and permit the R&D of large multi-product enterprises to be properly classified. In fact, with the classification by industry each enterprise was classified according to its principal activity only. This caused significant differences between allocation of R&D expenditures by industry and product field. [20] It resulted in overestimations for certain industries, like manufacturing enterprises, or in underestimations for others, like services enterprises. According to a recent evaluation, the range of discrepancy between the two classifications varied from 5,1% (Sweden) to 8,3% (Australia). [21]

In general, countries considered the product field classification more appropriate that the industrial one, several collected such data, [22] but only 55% used it. [23] The main argument for not collecting the information was that adding such questions in surveys would impose a burden on businesses.

An important question regarding the classification of the Business sector was "whether the R&D should be classified with the content of the R&D activity itself or with the end use. For example, should R&D on an electrical motor for agricultural machinery be classified under electrical or agricultural machinery". [24] In general, the philosophy of R&D statistics had always been to classify data according to the purpose of the research activities. This was the case for the character of research (basic or applied) [25] and was also the case, as we shall later see, for classifying government R&D. In the case of business R&D, the OECD recommended, without being imperative, content rather than purpose. Thus, "R&D for

[20] J. Morgan (2001), *Assess Ways to Improving Product Fields Data*, DSTI/EAS/STP/NESTI (2001) 14/PART10.
[21] J. Morgan (2001), *op. cit*. p. 6.
[22] OECD (1994), *Report on the Mini-Survey on the Availability of Product Field Data*, DSTI/EAS/STP/NESTI (94) 11.
[23] OECD (2000), *Review of the Frascati Manual: Use of Product Field Classification*, DSTI/EAS/STP/NESTI/RD (2000) 5.
[24] J.F. Minder (1991), *Treatment of Industrial R&D Data*, DSTI/STII (91) 17, p. 7.
[25] See B. Godin (2000), Measuring Science, *op. cit*.

rubber tires for aircraft should be classified under rubber tires, and the subsequent flow of embodied R&D to the aircraft industry be captured by I/O [input/output] techniques”. [26]

University Classification

The Frascati manual (1963) proposed to classify university (and private non-profit) R&D by fields of science. Six broad fields (including around thirty sub-fields) were and are still recommended as follows (see Annex 2 for details): [27]

1. natural sciences
2. engineering
3. medical sciences
4. agriculture
5. social sciences
6. humanities and fine arts

The classification came directly from previous OECD work on scientific and technical personnel. [28] In the 1950s, the OECD conducted three international surveys intended to assess the supply and demand of scientific and technical personnel. The classification used was itself based on a UNESCO recommendation dating from 1958. [29] The UNESCO classification was never intended to measure R&D activities however, but rather the qualifications of S&T personnel. In fact, the Frascati manual admitted: [30]

> It may be desirable at a later stage to work out a new classification specially suited to the requirements of R&D statistics and taking into account of the growing importance of inter-

[26] J.F. Minder (1991), *op. cit*.
[27] Social sciences and humanities are covered since 1973 only.
[28] OEEC (1968), *A Study of Resources Devoted to R&D Member Countries in 1963/64: Statistical Tables and Notes*, Paris, p. 25.
[29] The recommendation was concerning the International Standardization of Educational Statistics, now known as ISCED.
[30] OECD (1963), *op.cit.* p. 25.

> disciplinary fields of research. The UNESCO classification is mainly suited to the needs of manpower measurement.

Nevertheless, the Frascati manual conventionalized the classification by fields of science for university sector R&D, *faute de mieux*. [31] It would thus never really classify research projects and their content but rather the department or qualification of the researcher under which the research was conducted. As the UNESCO noted in 1966: "the classification recommended for higher education is not completely satisfactory when used for categories of personnel [or research expenditures] actually engaged in scientific and technical activities. The discipline of study or faculty granting the degree does not always adequately describe the fields of specialization in the actual employment situation outside the academic field". [32] As a result, less than half of OECD Member countries report data by fields of science to the OECD today. [33] As in the following quotation from 1991, delegates at the Frascati manual revision meeting in 2000 apparently expressed dissatisfaction over the absence of details in most classification systems: [34]

> Most of the OECD Member countries give information/data with regard to fields of science and technology at a high level of aggregation" (…). Only a few classification systems include research fields like biotechnology, materials research and information technology. [35]

All in all, as the Australian delegate stated, the classification by fields of science was "based on conceptions of discipline boundaries that were more relevant to the state of science in the first half of the 20th Century". [36] Classifications by fields of science, as most

[31] Better classifications are in fact available in Australia, the Netherlands, the United States, as well as in J. Irvine, B. Martin, and P. Isard, *Investing in the Future: An International Comparison of Government Funding of Academic and Related Research*, Worchester: Billing and Sons Ltd.

[32] Unesco (1966), *Statistical Data on Science and Technology for Publication in the UNESCO Statistical Yearbook*, UNESCO/CS/0666.SS-80/4, pp. 3-4.

[33] K.W. Maus (2000), *R&D Classification*, DSTI/EAS/STP/NESTI/RD (2000) 14.

[34] OECD (2000), *Review of the Frascati Manual: Use of Field of Science Classification* DSTI/EAS/STP/NESTI/RD (2000) 7; OECD (2000), *Outcomes of the Frascati Manual Revision Meeting Held on 13-14 March 2000*, DSTI/EAS/STP/NESTI (2000) 30, p. 3.

[35] E.L. Rost (1991), *Fields of Science and Technology*, DSTI/STII (91) 8, p. 6; OECD (1991), *Exemples de classification par domains scientifiques et techniques*, DSTI/STII (91) 9.

[36] K. Bryant (2000), *An Outline Proposal to Extend the Detail of the SEO and FOS Classifications in the Frascati Manual*, DSTI/EAS/STP/NESTI (2000) 27, p. 3.

classifications, are retrospective and therefore "lag behind the progress of science". [37] Their usefulness was therefore continually put into question: [38]

> A large part of the 1960s push to obtain detailed R&D [classifications] stemmed from the belief that it would come to fill a major role in priority setting at the national level. (…). Instead, the data proved to be highly useful for a wide variety of strategic and analytical purposes at a slightly lower level.

Government Classification

Of the three main economic sectors, it was in the government sector that the OECD departed most from available standards. Included in the United Nations system of national accounts was an international Classification of the Functions of Government (COFOG) that was rejected by the OECD group of experts (NESTI) for reasons to be explained below. This was in contradiction to the principle of comparability used for choosing the two previous classifications and to a draft recommendation made in 1978. [39] Instead, the OECD opted for the European *Nomenclature for the Analysis and Comparison of Science Programmes and Budgets* (NABS).

The OECD began collecting data on socioeconomic objectives of government funded R&D in the early seventies and introduced corresponding standards in the third edition of the Frascati manual (1975). [40] The method was in fact supplied by the European Commission. A task group of European statisticians was set up as early as 1968 by the Working Group on Scientific and Technical Research Policy (CREST) in order to study central government funding of R&D. The purpose was to "indicate the main political goals of government when committing funds to R&D". [41] The implicit goal was to contribute to the "construction" of a European science policy and budget. To this end, the Commission

[37] UNESCO (1969), *List of Scientific Fields and Disciplines Which Will Facilitate the Identification of Scientific and Technical Activities*, UNESCO/COM/CONF.22/10, p. 4.
[38] D. Byars and K. Bryant (2001), *Review of FOS Classification*, DSTI/EAS/STP/NESTI (2001) 14/PART11, p. 3.
[39] OECD (1978), *Draft Guidelines for Reporting Government R&D Funding by Socio-Economic Objectives: Proposed Supplement to the Frascati Manual*, DSTI/SPR/78.40, p. 9.
[40] The first two editions of the Frascati manual included preliminary and experimental classifications.

adopted the *Nomenclature for the Analysis and Comparison of Science Programmes and Budgets* (NABS) in 1969 [42] produced by the group of experts, and published a statistical analysis based on the classification. [43]

In line with the spirit of the Brooks report, which had argued for changes in the objectives of government-funded R&D, [44] the OECD Directorate for Science, Technology and Industry (DSTI) adopted the European Commission's approach for obtaining appropriate statistics. [45] The recommended classification had eleven categories or motives for government-funded R&D (see Annex 3 for details), [46] which allowed, for example, the development of specific statistics on energy in the early seventies. [47]

1. agriculture
2. industrial development
3. energy
4. infrastructures
5. environment
6. health
7. social development and services
8. earth and atmosphere
9. advancement of knowledge
10. space
11. defense

[41] Eurostat (1991), *Background Information on the Revision of the NABS,* Room document to the Expert Conference to Prepare the Revision of the Frascati Manual for R&D Statistics, OECD.
[42] The first NABS was issued in 1969, revised in 1975 (and included in the 1980 edition of the Frascati Manual) and again in 1983 (to include biotechnology and information technology, not as categories but broken down across the whole range of objectives). In 1993, improvements were made in the Environment, Energy, and Industrial Production categories.
[43] CEC (1970), *Research and Development: Public Financing of R&D in the Community Countries, 1967-1970*, BUR 4532, Brussels.
[44] OECD (1971), *Science, Growth, and Society*, Paris.
[45] The first OECD (experimental) analysis of data by socioeconomic objectives was published in 1975: OECD (1975), *Changing Priorities for Government R&D: An Experimental Study of Trends in the Objectives of Government R&D Funding in 12 OECD Member Countries, 1962-1972*, Paris.
[46] Since 1993, the expression "government-financed R&D" has been replaced by GBAORD (Government Budget Appropriations and Outlays for R&D).

The classification of government-funded R&D was determined by the following fundamental fact: few governments actually conducted surveys of government R&D. [48] Most preferred to work with budget documents because, although less detailed and accurate than a survey, the information was easier and cheaper to obtain. [49] Among the methodology's advantages was speed, since the data were extracted directly from budget documents without having to wait for a survey. But it also had several limitations, [50] among them the fact that national data relied on different methodologies and concepts, and different administrative systems. As regard the classification, it reflected intention to spend rather than real expenditures. Moreover, data were difficult to extract from budgets because they lacked the required level of details: "the more detailed the questions are, the less accurate the data become" because it was not always possible to define the specific NABS sublevel in the budget – budget items can be quite broad. [51]

OECD statisticians were also confronted with a wide range of budgetary and national classificatory systems in Member countries over which they had relatively little control: [52]

> The unit classified varied considerably between countries (…) because national budget classification and procedures differ considerably. In some countries, such as Germany, the budget data are available in fine detail and can be attributed accurately between objectives. In others, such as the United Kingdom and Canada, the budgetary data are obtained from a survey of government funding agencies which is already based on an international classification. However, in others again such as France, the original series are mainly votes by ministry or agency.

To better harmonize national practices, a draft supplement to the Frascati manual specifically devoted to the measurement of socioeconomic objectives of government R&D

47

48 Exceptions are Canada and the United Kingdom. Other countries either made text analysis of budgets or estimate appropriations from budget documents. For methodologies used in European countries, see: Eurostat (1995), *Government R&D Appropriations: General University Funds*, DSTI/STP/NESTI/SUR (95) 3, pp. 2-3.

49 Eurostat (2000), *Recommendations for Concepts and Methods of the Collection of Data on Government R&D Appropriations*, DSTI/EAS/STP/NESTI (97) 10, p. 3.

50 Eurostat (2000), *The Frascati Manual and Identification of Some Problems in the Measurement of GBAORD*, DSTI/EAS/STP/NESTI (2000) 31.

51 OECD (2000), *The Adequacy of GBAORD Data*, DSTI/EAS/STP/NESTI (2000) 18, p. 3.

52 OECD (1990), *Improving OECD Data on Environment-Related R&D*, DSTI/IP (90) 25, p. 9.

was completed in 1978, [53] but was never published as a separate publication. These data "play only a modest role in the general battery of S&T indicators and do not merit a separate manual" stated the OECD. [54] Instead of a manual, the specifications were abridged and relegated to a chapter in the fourth edition of the manual, and a *Sources and Methods* document was developed in 1982. [55]

In the meantime, criticisms persisted. Seventy five percent of countries actually used the classification, but only 30% were satisfied with it: [56] "It is not a classification in that it does not have a truly logical structure and it contains overlapping categories and gaps in coverage", claimed the Australians. [57] In fact, the OECD list was a compromise between the NABS, Scandinavian countries (grouped under Nordsfork) and national lists, notably the NSF's.

The NABS and the COFOG had similar categories,[58] but different methodologies for collecting data. In fact, the OECD chose to classify and measure the purposes (or intentions) of government R&D expenditures rather than the content, whether by areas of relevance or end results. [59] The COFOG differed from the NABS in three respects: [60] 1) the COFOB classified expenditures, while the NABS measured appropriations; 2) until recently, the COFOG did not distinguish different types of research when they were an integral part of a government program, and 3) nor did it explicitly separate basic from

[53] OECD (1978), *Draft Guidelines for Reporting Government R&D Funding by Socio-Economic Objectives: Proposed Supplement to the Frascati Manual*, DSTI/SPR/78.40.
[54] OECD (1991), *Classification by Socio-Economic Objectives*, DSTI/STII (91) 19, p. 9.
[55] In 1991, Australia again proposed that there should be a supplement to the Manual dealing with detailed classification by socioeconomic objectives and fields of science. See: OECD (1992), *NESTI Meeting*, DSTI/STII/STP/NESTI/M (92) 1.
[56] OECD (2000), *Review of the Frascati Manual: Use of SEO Classification* DSTI/EAS/STP/NESTI/RD (2000) 8.
[57] K. Bryant and D. Byars (2001), *Review of SEO Classification*, DSTI/EAS/STP/NESTI (2001) 14/PART12.
[58] When the main categories of COFOG did not correspond perfectly to the NABS and the OECD classification, subcategories allowed one to develop links with COFOG. For comparisons between classifications, see: OECD (1978), *Draft Guidelines for Reporting Government R&D Funding by Socio-Economic Objectives: Proposed Supplement to the Frascati Manual*, DSTI/SPR/78.40,. pp. 40-71.
[59] OECD (1972), *The Problems of Comparing National Priorities for Government Funded R&D*, DAS/SPR/72.59. This document became chapter 2 of OECD (1975), *Changing Priorities for Government R&D*, Paris.
[60] OECD (1996), *Revision of COFOG*, DSTI/EAS/STP/NESTI (96) 16.

applied research, putting the former under general public services and the latter under the function concerned. [61] These were the three main reasons why OECD Member countries decided to use the NABS instead of the COFOG for classifying government-funded R&D. The NABS was really no better that the COFOG, however, when it came to the classification of general university funds (GUF): governments cannot orient this money when they fix their budgets, and consequently these funds were never broken down by socioeconomic objectives.

Problems of Classifications

Classifications are rarely static. They evolve in time as new things or issues and as new policy priorities emerge: user demands for R&D data is always changing and the environment is constantly evolving. All three R&D classifications were in fact regularly adjusted over the last forty years, as reflected, for example, in each revised edition of the Frascati manual. Such changes often caused important breaks in the statistical series, however. This was probably one of the most difficult problems to resolve and a lot of time was devoted to this task. In fact, when using classifications that are not specifically developed for the purpose of measuring R&D, the challenges multiply rapidly.

Three major problems have thus limited the usefulness of classifications for R&D purposes. The first concerned the degree of details for policy analysis. As regards the business sector, for example, two industries have received considerable attention: the service industry and the information and communication sector. Until recently, coverage of the service industry was still very poor because of the manufacturing bias of most economic classifications: "the service industries have been thought of mainly as importing technology (…). This was unrealistic in that the services include a number of industries whose main economic activity is technology supply (commercial R&D firms, software houses, etc.) and that even for industries where change is initially brought about by embodied technology, this

[61] OECD (1991), *Classification by Socio-Economic Objectives*, DSTI/STII (91) 19, p. 5.

subsequently leads firms to carry out independent R&D”. [62] In fact, in seven OECD countries, one quarter or more of all R&D expenditures in the business sector is carried out in the service industries.

But few countries actually included all the service industries in their national survey. Only communication and computers were covered, because they were presumably the most R&D intensive sectors in services as a whole. Consequently, “the high discrepancy in the share of service R&D [in data surveys] results [not only from the presumed lesser importance of the service sector but] also from different methods of surveying services R&D data”. [63]

Reliable information was also missing for the “information and communication sector”, a “strategic” sector that every government targeted for support. After many years of complaints, a new classification has now been proposed by two Working Parties of the OECD. [64] It classifies enterprises according to their principal activity and includes seven classes for manufacturing and four classes for services. There is still work to be done, however, before the sector is properly defined through a classification based on products rather than industry.

Besides the lack of detailed classifications, a second problem with classifications was related to the treatment of horizontal projects or products. Some authors have argued that most classifications are generally ill suited for intermediate or transitional states “because transition is neither one state nor the next”. [65] Similarly, generic technologies, multidisciplinary projects, and problem-oriented research have always been poorly classified as R&D categories, if at all. This was the case for the environmental sciences which conduct research in a number of areas that are not confined to the classical category

[62] A. Young (1996), *Measuring R&D in the Services*, OECD/GD (96) 132, p. 7.
[63] C. Grenzmann (2000), *Measurement of R&D in Services*, DSTI/EAS/STP/NESTI (2000) 11, p. 4; OECD (1996), *Replies to the Mini-Survey on R&D in the Services*, DSTI/EAS/STP/NESTI (96) 6; OECD (1997), *Progress Report on Services R&D*, DSTI/EAS/STP/NESTI (97) 11; A. Young (1996), *op. cit*. p. 19.
[64] OECD (1998), *Measuring the ICT Sector*, Paris.
[65] M. Douglas (1966), *Purity and Danger*, p. 97; see also: H. Ritvo (1997), *The Platypus and the Mermaid and Others Figments of the Classifying Imagination*, Cambridge: Harvard University Press.

of the classification by socioeconomic objectives or to the environment category of ISIC. [66] It was also the case for biotechnology for which a definition and a list of technologies for statistical purposes have only recently been proposed. [67] And finally, it was also the case for the health sector which cut across numerous classifications and categories, while actual measures were confined to the health category of socioeconomic objectives in the classification by government R&D, to the medical sciences in the classification by fields of science, or to the pharmaceutical industry in the ISIC classification. [68]

A third problem with R&D classifications was the difficulty of establishing links between data. [69] Since each of the three sectors had its own system of classification, it was always difficult, for example, to relate socioeconomic objectives of government funding to industrial or university R&D activities (Table 1). As a result, some countries preferred to apply the (university) classification by fields of science to the government sector [70] instead of applying a classification by socioeconomic objectives "in order to be able to make comparisons with R&D in the higher education sector". [71] In fact, although there was always was desire for consistency between classifications, [72] no consensus was ever developed for classifying government R&D according to either one of the two

[66] OECD (1998), *Trends in Environmental R&D Expenditures*, DSTI/STP/TIP (98) 10.

[67] OECD (2000), *Reviewing and Refining the Definitions of Biotechnology*, DSTI/EAS/STP/NESTI (2000) 7; OECD (2000), *Discussions and Recommendations*, DSTI/EAS/STP/NESTI (2000) 9; OECD (2001), *Discussion and Recommendations for Future Work*, DSTI/EAS/STP/NESTI (2001) 7; OECD (2001), *Biotechnology Statistics in OECD Member Countries: Compendium of Existing National Statistics*, DSTI/EAS/STP/NESTI (2001) 2.

[68] A. Young (2000), *Proposals for Improving the Availability and International Comparability of the Health-Related R&D Data Overseen by NESTI*, DSTI/EAS/STP/NESTI (2000) 28; A. Young (2000), *Assessment of International Practices for the Compilation of Data on R&D Related to Health and Preparation of Guidelines for Improved Data Collection*, DSTI/EAS/STP/NESTI (2000) 29 and Annex; A. Young (2001), *Develop Methodologies for Better Derivation of R&D Data on Hospitals and Clinical Trials*, DSTI/EAS/STP/NESTI (2001) 14/PART4.

[69] Linkage of various existing data had recently been identified by the OECD as a way to bring forth new indicators, partly because of budget constraints – linking existing data would be far less expensive than developing totally new indicators. See: OECD (1996), *Conference on New S&T Indicators for a Knowledge-Based Economy: Summary Record of the Conference Held on 19-21 June 1996*, DSTI/STP/NESTI/GSS/TIP (96) 5; OECD (1996), *New Indicators for the Knowledge-Based Economy: Proposals for Future Work*, DSTI/STP/NESTI/GSS/TIP (96) 6.

[70] Or to the business sector in the case of UNESCO surveys. See: UNESCO (1969), *Report of the Session Held in Geneva, 2-6 June 1969*, UNESCI/COM/CONF.22/7, p. 9.

[71] OECD (1992), *NESTI Meeting*, DSTI/STII/STP/NESTI/M (92) 1.

classifications (fields of science or socioeconomic objectives) and both are actually recommended by the Frascati manual.

Similarly, problems of comparability affected the business sector. When, at the beginning of the 1990s, the DSTI launched a project to create the STAN (Structural Analysis) database that would link R&D data to economic indicators, it faced the fact that every statistical series – economic as well as R&D – had its own classification: "As most of these databases were created in different environments and at different times, the categories were picked as a function of the needs which each specific database was to serve without particular concern about comparable coverage of categories among different databases". [73]

Table 1.

R&D Classifications – First level

(Frascati manual, 1993)

Business (Industries)	Government (SEO)	University (FOS)
agriculture	agriculture	natural sciences
mining	industrial development	engineering
manufacturing	energy	medical sciences
electricity	infrastructures	agriculture
construction	environment	social sciences
services	health	humanities & fine arts
	social development and services	
	earth and atmosphere	
	advancement of knowledge	
	space	
	defense	

[72] OECD (2000), *Review of the Frascati Manual: Use of SEO Classification* DSTI/EAS/STP/NESTI/RD (2000) 8; K.W. Maus (2000), *R&D Classification*, DSTI/EAS/STP/NESTI/RD (2000) 14; K. Wille-Maus (1991), *Sector Sub-Classification in the Government Sector*, DSTI/STII (91) 21.

[73] *OECD (1991), Choosing Economic Activity Categories Form Revised Classifications for STIID Databases*, Room Document, Expert Conference to Prepare the Revision of the Frascati Manual for R&D Statistics, p. 19.

Linking data is the only way of answering policy demands for measuring the performance and impacts of science and technology. Since at least the mid-1980s, governments have all asked for indicators that would relate inputs to outputs. Statisticians are only now beginning to direct their efforts towards this task.

Conclusion

Agencies are usually proud of developing their own classifications because they are adapted to their own needs and interests. By aligning themselves to the system of national accounts, however, the State statisticians responsible for R&D statistics effectively reduced the amount of control that they could exercise over the data. Although borrowed classifications did undergo some modifications, they had nonetheless been developed for purposes other than classifying R&D and therefore carried with them considerable biases, the consequences of which could not be fully anticipated.

The Frascati manual classified R&D according to three (of the four) sectors of the system of national accounts and a fourth sector of its own – higher education. The sole variable common to each sector was money, that is R&D expenditures. It allowed people to calculate the GERD and to measure financial flows between sectors.

Each sector had its own system of classification. The business sector was classified according to industries. Only the principal activity of the enterprise was classified, however, rather than the whole range of products on which R&D was realized. The university sector was classified according to scientific disciplines. This classification was modeled after the demarcations of traditional university departments and was more useful for scientific and technical personnel than research projects. The government sector was either classified according to scientific disciplines or to socioeconomic funding objectives. It classified intentions rather than real expenditures, was oriented toward industrial motives rather than social problems, and did not deal with the general funds allotted to universities.

Two motives drove the choice of classifications for R&D at the OECD. The first, as disclosed in official documents, was the need for comparing R&D data to other variables, mainly economic ones. The second motive for choosing existing classifications rather than constructing new ones was their availability. The business and university sector classifications were produced by the United Nations, while the government sector classification was developed by the European Union.

Classifications were not without problems, however. Firstly, the level of aggregation prevented detailed analyzes. Secondly, horizontal issues could not be clearly measured. Thirdly, links between dimensions were difficult to establish. There were two methodological choices concerning surveys that contributed to the limitations of classifications. One was the unit surveyed. The questionnaires were addressed to the organizations responsible for R&D with respect to their research activities as a whole. It was therefore not the research projects themselves that were classified but the principal R&D activity of the unit. As the OECD indeed recognized as early as 1968, "it would be difficult to get a good (…) breakdown unless data were available by projects". [74]

The second methodological choice, paradoxically, has nothing to do with the classifications themselves, but rather with the way the OECD collected data. Classifications were, in fact, often more detailed than they might have at first appeared in statistical and analytical reports. But the problem was with the countries that did not collect detailed enough data or that did so with different classifications. Since the OECD relied on national data to construct its statistics, harmonization could only be conducted at a very aggregated level in international tables. The OECD could not really innovate if the countries themselves relied on national traditions.

[74] OECD (1968), *A Study of Resources Devoted to R&D Member Countries in 1963/64: Statistical Tables and Notes*, Paris, p. 24.

Annex 1.

Fields of Science (FOS)

1. NATURAL SCIENCES

1.1 Mathematics and computer sciences (mathematics and other allied fields: computer sciences and other allied subjects (software development only; hardware development should be classified with the engineering fields)
1.2 Physical sciences (astronomy and space sciences, physics, other allied subjects)
1.3 Chemical sciences (chemistry, other allied subjects)
1.4 Earth and related environmental sciences (geology, geophysics, mineralogy, physical geography and other geosciences, meteorology and other atmospheric sciences including climatic research, oceanography, vulcanology, palaeoecology, other allied sciences)
1.5 Biological sciences (biology, botany, bacteriology, microbiology, zoology, entomology, genetics, biochemistry, biophysics, other allied sciences, excluding clinical and veterinary sciences)

2. ENGINEERING AND TECHNOLOGY

2.1 Civil engineering (architecture engineering, building science and engineering, construction engineering, municipal and structural engineering and other allied subjects)
2.2 Electrical engineering, electronics [electrical engineering, electronics, communication engineering and systems, computer engineering (hardware only) and other allied subjects]
2.3 Other engineering sciences (such as chemical, aeronautical and space, mechanical, metallurgical and materials engineering, and their specialized subdivisions; forest products; applied sciences such as geodesy, industry chemistry, etc.; the science and technology of food production; specialized technologies of interdisciplinary fields, e.g. systems analysis, metallurgy, mining, textile technology, other allied subjects)

3. MEDICAL SCIENCES

3.1 Basic medicine (anatomy, cytology, physiology, genetics, pharmacy, pharmacology, toxicology, immunology and immunohaematology, clinical microbiology, pathology)
3.2 Clinical medicine (anaesthesiology, paediatrics, obstetrics and gynaecology, internal medicine, surgery, dentistry, neurology, psychiatry, radiology, therapeutics, otorhinolaryngology, ophthalmology)
3.3 Health sciences (public health services, social medicine, hygiene, nursing, epidemiology)

4. AGRICULTURAL SCIENCES

4.1 Agriculture, forestry, fisheries and allied sciences (agronomy, animal husbandry, fisheries, forestry, horticulture, other allied subjects)
4.2 Veterinary medicine

5. SOCIAL SCIENCES

5.1 Psychology
5.2 Economics
5.3 Educational sciences (education and training and other allied subjects)
5.4 Other social sciences [anthropology (social and cultural) and ethnology, demography, geography (human, economic and social), town and country planning, management, law, linguistics, political sciences, sociology, organisation and methods, miscellaneous social sciences and interdisciplinary), methodological and historical S & T activities relating to subjects in this group. Physical anthropology, physical geography and psychophysiology should normally be classified with the natural sciences]

6. HUMANITIES

6.1 History (history, prehistory and history, together with auxiliary historical disciplines such as archaeology, numismatics, palaeography, genealogy, etc.)
6.2 Languages and literature (ancient and modern languages and literatures)
6.3 Other humanities [philosophy (including the history of science and technology), arts, history of art, art criticism, painting, sculpture, musicology, dramatic art excluding artistic «research» of any kind, religion, theology, other fields and subjects pertaining to the humanities, methodological, historical and other S & T activities relating to the subjects in this group]

Annex 2.

Socio-Economic Objectives (SEO)

OECD categories	NABS categories
1. Development of agriculture, forestry and fishing	6. Agricultural production and technology
2. Promotion of industrial development technology	7. Industrial production and technology
3. Production and rational use of energy	5. Production, distribution and rational utilisation of industry
4. Development of the infrastructure *4.1 Transport and telecommunica-tions* *4.2 Urban and rural planning*	2. Infrastructure and general planning of land use *2.4 Transport systems* *2.5 Telecommunication systems* *2. n.e.c.= general infrastructure and land planning research, construction and planning of buildings, water supplies, infrastructure R & D n.e.c.*
5. Control and care of the environment	
5.1 The prevention of pollution	*(Included in objective concerned)*
5.2 Identification and treatment of pollution	*3. Control of environmental pollution*
6. Health (excluding pollution)	4. Protection and improvement of human health
7. Social development and services	8. Social structures and relationships
8. Exploration and exploitation of Earth and atmosphere	1. Exploration and exploitation of the Earth
9. General advancement of knowledge	
9.1 Advancement of research	*11. Non-oriented research*
9.2 General university funds	*10. Research financed from general university funds*
10. Civil space	9. Exploration and exploitation of space
11. Defence	13. Defence
12. Not specified	12. Other civil research

Annex 3.

Classification of Industries

1. AGRICULTURE, HUNTING AND FORESTRY
2. MINING
3. MANUFACTURING
4. Food, beverages and tobacco
5. Food products and beverages
6. Tobacco products
7. Textiles, wearing apparel, fur and leather
8. Textiles
9. Wearing apparel and fur
10. Leather products and footwear
11. Wood, paper, printing, publishing
12. Wood and cork (not furniture)
13. Pulp, paper and paper products
14. Publishing, printing and reproduction of recorded media
15. Coke, petroleum, nuclear fuel, chemicals and products, rubber and plastics
16. Coke, refined petroleum products and nuclear fuel
17. Chemicals and chemical products
18, Chemicals and chemical products (less pharmaceuticals)
19. Pharmaceuticals
20. Rubber and plastic products
21. Non-metallic mineral products («Stone, clay and glass»)
22. Basic Metals
23. Basic metals, ferrous
24. Basic metals, non-ferrous
25. Fabricated metal products (except machinery and equipment)
26. Machinery equipment, instruments and transport equipment
27. Machinery
28. Office, accounting and computing machinery
29. Electrical machinery
30. Electronic equipment (radio, TV and communications)
32. Television, radio and communications equipment
33. Medical, precision and optical instruments, watches and clocks (instruments)
34. Motor vehicles
35. Other transport equipment
36. Ships
37. Aerospace
38. Other transport
39. Furniture, Other Manufacturing n.e.c.
40. Furniture
41. Other manufacturing
42. Recycling
43. ELECTRICITY, GAS AND WATER SUPPLY (UTILITIES)
44. CONSTRUCTION
45. SERVICE SECTOR
46. Wholesale, retail trade and motor vehicle, etc., repair
47. Hotels and restaurants
48. Transport and storage
49. Communications
50. Post
51. Telecommunications
52. Financial intermediation (including insurance)
53. Real estate, renting and business activities
54. Computer and related activities
55. Software consultancy
56. Other computer services
57. Research and development
58. Other business activities
59. Community, social and personal service activities, etc.