

Evaluation policies for research and graduate education in the Netherlands: lessons on self-evaluation and institutional autonomy

Joviles Vitório Trevisol^I 

André Brasil^{II} 

ABSTRACT

This paper presents the results of a study on the evaluation of research and graduate education in the Netherlands. The policies implemented in the country over the last four decades have consolidated a *sui generis* and stable evaluation system, contributing to strengthening the country's scientific leadership on the international scene. The Dutch experience has inspired several countries, and its lessons could also be relevant for the Brazilian reality. In this context, this study aimed to analyse the main characteristics of the Dutch system, understand the role of self-evaluation within that system and present a self-evaluation experience conducted by a noteworthy research centre from Leiden University. From the lessons learned, we conclude by presenting some subsidies for improving the evaluation of graduate education conducted in Brazil.

KEYWORDS

evaluation policies; self-evaluation; institutional autonomy; the Netherlands.

^IUniversidade Federal da Fronteira Sul, Chapecó, RS, Brazil.

^{II}Universiteit Leiden, Leiden, Netherlands.

AS POLÍTICAS DE AVALIAÇÃO DA PESQUISA E DA PÓS-GRADUAÇÃO NOS PAÍSES BAIXOS: LIÇÕES SOBRE AUTOAVALIAÇÃO E AUTONOMIA INSTITUCIONAL

RESUMO

Este artigo apresenta os resultados de um estudo sobre o sistema de avaliação da pesquisa e da pós-graduação dos Países Baixos. As políticas implementadas ao longo das últimas quatro décadas consolidaram um sistema *sui generis* e estável de avaliação, contribuindo para fortalecer a liderança científica do país no cenário internacional. A experiência holandesa tem inspirado diversos países e suas lições também podem ser relevantes para a realidade brasileira. Nesse contexto, este estudo visou analisar as principais características do sistema holandês; compreender o papel da autoavaliação no sistema e apresentar uma experiência de autoavaliação desenvolvida por um notável centro de pesquisa da Universidade de Leiden. Com base nas lições apreendidas, concluímos apresentando alguns subsídios para o aprimoramento do sistema brasileiro de avaliação da pós-graduação.

PALAVRAS-CHAVE

políticas de avaliação; autoavaliação; autonomia institucional; Países Baixos.

POLÍTICAS DE EVALUACIÓN DE LA INVESTIGACIÓN Y DEL POSGRADO EN LOS PAÍSES BAJOS: LECCIONES SOBRE AUTOEVALUACIÓN Y AUTONOMÍA INSTITUCIONAL

RESUMEN

Este artículo presenta los resultados de un estudio sobre el sistema de evaluación de la investigación y de los estudios de posgrado en los Países Bajos. Las políticas implementadas durante las últimas cuatro décadas han consolidado un sistema de evaluación *sui generis* y estable, contribuyendo a fortalecer el liderazgo científico del país en la escena internacional. La experiencia holandesa ha inspirado a varios países, y sus lecciones también podrán ser relevantes para la realidad brasileña. En este contexto, este estudio tiene como objetivos analizar las principales características del sistema holandés; comprender el papel de la autoevaluación dentro del sistema y presentar una experiencia de autoevaluación desarrollada por uno destacado centro de investigación de la Universidad de Leiden. A partir de las lecciones aprendidas, concluimos presentando algunos subsidios para la mejora del sistema brasileño de evaluación de posgrado.

PALABRAS CLAVE

políticas de evaluación; autoevaluación; autonomía institucional; Países Bajos.

INTRODUCTION

In recent decades, the importance of evaluation has been paramount within national systems of basic through higher education, as well as in the realms of science, technology, and innovation. Nevertheless, it is important to recognize that the current relevance of this subject should not lead us to assume that interest in it is of recent origin. As per Vught (1995), Amaral (2009), and Cobban (2017), historical records indicate that evaluation practices in higher education can be traced back to the 13th century. According to these scholars, the first European universities employed rudimentary evaluation processes, which were based on two distinct conceptions. The first, known as the French model, conferred an external authority with the power to grant a license to teach (*licentia ubique docendi*) and the authority to determine the curriculum. Evaluation served as a form of accountability to the superior external authority responsible for granting approval. For example, in the early 13th century, the bishop of Paris and the chancellor of Notre-Dame Cathedral oversaw the University of Paris. The second conception, influenced by the English model, revolved around the principle of self-governance. Universities enjoyed a considerable degree of autonomy and peers conducted evaluations. Collegial bodies possessed the right to evaluate their fellow colleagues and, if necessary, to replace them (Amaral, 2009).

These divergent conceptions have given rise to distinct evaluation cultures, as noted by Hicks (2012), Lepori, Reale, and Spinello (2018), and Ochsner, Kulczycki, and Gedutis (2018). Over the centuries, evaluation practices have undergone transformations, resulting in hybrid and highly heterogeneous models. It can be argued that, until the 1980s, evaluation was an internal, sporadic, and inadequately institutionalized practice. External evaluation, particularly when carried out by government entities, was perceived as a threat to scientific ethos. It was rejected based on the notion that science should be conceived and conducted according to its own assumptions, methodologies, and legitimacy criteria (Meulen, 1998; 2007). Quality assessment was deemed to be the responsibility of academic institutions and peers. External evaluation was viewed as an instrument of interference in institutional autonomy, as well as a means of curbing intellectual freedom and creativity. The evaluation of the merit of scientific endeavors was conducted by peers and not by external parties (Molas-Gallart, 2012).

From the 1980s, the discourse surrounding evaluation took on new dimensions. National agencies entrusted with the regulation and funding of higher education and research began to advocate for the establishment of permanent and standardized evaluation policies. According to these agencies, institutionalizing these processes would allow the assessment of the results resulting from increased public investments, while ensuring regularity, transparency, and consistency of criteria and procedures at the national level (Hicks, 2012; Ochsner, Kulczycki, and Gedutis, 2018). In addition to being an internal and exclusive concern of universities and researchers, evaluation gradually assumed a prominent role with-

in national systems of higher education, science and technology (Amaral, 2009; Verhine and Freitas, 2012). Through their regulatory and funding bodies, nation states progressively undertook the responsibility of conducting periodic evaluations, which not only influenced research and higher education policies but also had an impact on resource allocation. This process led to the so-called “performance-based funding system” (Ochsner, Kulczycki, and Gedutis, 2018, p. 1235). The nexus between performance and funding caused profound changes in the management of universities and research institutions (Castro, 1999; Sobrinho, 2000; 2003; Amaral, 2009; Afonso, 2013).

Over the past few decades, numerous countries have established institutionalized evaluation systems. While some have chosen centralized and standardized models based on metrics, others have implemented decentralized models that are less quantitative, emphasizing the idea that evaluation should be conducted by the institutions themselves to reinforce their institutional missions and guide strategic planning.

In this context, the focus of the present study is on the evaluation system in the Netherlands. The Netherlands was among the first European countries to adopt a formal system for assessing the quality of research and higher education.¹ The Dutch experience has served as a benchmark for other nations and institutions, including Brazil. Their decentralized, bottom-up model is structured around the principles of autonomy, self-evaluation (SE), and institutional planning (Weert and Boezerooy, 2007; Drooge *et al.*, 2013). The recognition this model has gained within the academic and scientific community, as well as its notable position in internationally indexed scientific production rankings (14th place according to the Web of Science, Clarivate, n. d.), highlights the significance and contributions of evaluation policies in consolidating quality.

Specifically, this article aims to:

1. analyze the evolution of research and higher education evaluation policies in the Netherlands,
2. comprehend the main characteristics that the model has acquired in recent years, and
3. understand the role of SE within the system.

The examination of the key characteristics and specificities of the Dutch model is complemented by a case study. This case study pertains to a recent SE initiative conducted between 2021 and 2022 by the Center for Science and Technology Studies (CWTS) at Leiden University. Drawing from this study, we highlight certain aspects that can provide valuable insights for enhancing the higher education evaluation system.

1 Since the 1980s, several European countries have developed their evaluation systems, with emphasis on to England (1986), Netherlands (1993), Spain (1989), Poland (1991), China (1993), Finland (1998), Hong Kong (2002), New Zealand (2003), Belgium (2003), Norway (2006), Denmark (2006), France (2006), Sweden (2009), Italy (2009), Australia (2010) (Hicks, 2012; Drooge *et al.*, 2013).

REGULATORY FRAMEWORKS FROM A HISTORICAL PERSPECTIVE

According to the research conducted, the Netherlands initiated its first efforts to establish norms and guidelines for evaluating the quality of research in the late 1960s. In 1969, the Advisory Council for Science Policy recommended the formation of expert committees to evaluate research programs funded by the government. In the early 1970s, committees were appointed for areas such as education, social policy, and environmental planning, followed by chemistry, biochemistry, biology, and physics in 1976. These committees comprised representatives from three stakeholder groups: researchers, government members, and stakeholders. However, lack of clarity regarding objectives, procedures, and evaluation criteria hindered the success of these initial endeavors (Meulen, 1998; 2007). In 1979, the Ministry of Education and Sciences introduced new guidelines for organizing research in institutions. Expert committees were tasked with evaluating research programs, which became the focal point of assessment (Goedegebuure and Westerheijden, 1991).

Evaluation experiences throughout the 1970s provided valuable insights for the development of the first regulatory framework for evaluation in the Netherlands. The Conditional Financing (CF), published in 1982, laid down general guidelines for establishing the national evaluation system. According to Goedegebuure and Westerheijden (1991), the CF can be seen as the initial attempt to introduce a formal system for assessing the quality of higher education and research. Its objectives were to promote quality, improve transparency, and evaluate the results of government and other funding agency investments. The document solidified research units as central organizing cores within institutions (Rip and Meulen, 1995; 1996; Meulen, 1998; 2007). Additionally, the CF:

- defined criteria and requirements for the creation and evaluation of research units;
- made it mandatory to articulate objectives, goals, work plans, budgets, and research teams within these units;
- implemented an accreditation process conducted by external bodies; and
- introduced the requirement of evaluating results after five years of activity (Goedegebuure and Westerheijden, 1991; Meulen, 1998).

The CF introduced a mandatory external evaluation, conducted through the peer review system. As described in the document, the researchers were assigned the responsibility of evaluating academic quality and research relevance due to their extensive and specialized training. The presence of external and independent evaluators was considered crucial to ensure legitimacy, improve transparency, and provide reliable information to the academic community and the institutions involved (Weert and Boezeroy, 2007; Drooge *et al.*, 2013).

The appointment of external committees was the responsibility of the Royal Netherlands Academy of Arts and Sciences (KNAW), a national scientific association. The peer review system served to protect the evaluation process from external interference, particularly from funding agencies. According to the docu-

ment, evaluation should not be conducted by funding agencies or the institutions being evaluated (Rip and Meulen, 1995; 1996; Meulen, 1998).

Furthermore, the CF brought about changes in funding policies. The sole criterion of enrollment numbers was no longer used to determine institutional funding (Goedegebuure and Westerheijden, 1991). Evaluation results in terms of research quality and societal relevance began to influence institutional budget allocation (Meulen, 1998; 2007; Weert and Boezeroy, 2007).

During the 1980s, the Ministry of Education, Culture and Science published several supplementary documents aimed at refining and enhancing the policies established by the CF. The principal document, titled *Hoger Onderwijs: Autonomie en Kwaliteit* (OCW, 1985)²—commonly known as the Hoak Document—reinforced institutional autonomy and introduced a retrospective quality assessment system. Instead of imposing interventionist policies, the document encouraged the development of institutional evaluation policies and the strengthening of external evaluation processes. Universities were directed to implement more efficient and professional institutional management practices (Goedegebuure and Westerheijden, 1991). Alongside the promotion of autonomy, the Hoak Document established new guidelines for funding and institutional planning. Institutions were challenged to operate in a more professional manner and align with societal needs. In exchange for increased autonomy, universities were required to establish institutional evaluation policies and ensure transparency of results (Goedegebuure and Westerheijden, 1991; Goedegebuure *et al.*, 1994; Kaiser and Weert, 2006; Droog *et al.*, 2013).

In 1992, the government enacted a new law to regulate the relationships between the government, higher education institutions, and research institutes. The *Hoger Onderwijs en Wetenschappelijk Onderzoek* law (OCW, 1992)³ further solidified autonomy and the role of institutions in the evaluation process. The law reduced state control and limited government interference in the operational aspects of evaluation. State intervention was restricted to specific circumstances. According to the law, the government should only intervene in cases of identified deficiencies, misallocation of resources, or discrepancies between objectives and results. Rather than conducting the evaluation itself, the government was responsible for ensuring the macro-efficiency of the system.

According to Weert and Boezeroy (2007, p. 14), the law established four general principles to guide government action:

1. The government should intervene only when self-management by institutions could lead to unacceptable outcomes.
2. Government intervention should primarily focus on addressing system imperfections after they arise.
3. The instruments available to the government should be characterized by a minimum of detailed regulation.

2 Higher Education: Autonomy and Quality.

3 Higher Education and Scientific Research.

4. Institutions should establish norms to ensure legal certainty, reasonableness, and proper administration.

The document emphasizes that evaluation is a component of institutional planning. Institutional policies should incorporate the participation of external and independent evaluators, preferably affiliated with other institutions (OCW, 1992, sec. 1.18). Evaluation is not an end in itself, and should not be used as a tool for external control. Instead, it should be seen as a means to enable the higher education system to respond more effectively and decisively to the needs and changes of society (Weert and Boezeroy, 2007, p. 13). The document provides the legal framework for institutions to develop their own evaluation policies.

In 1993, the Association of Dutch Universities (*Vereniging van Universiteiten*—VSNU),⁴ in collaboration with the Netherlands Organization for Scientific Research (*Nederlandse Organisatie voor Wetenschappelijk Onderzoek*—NWO)⁵ and the Royal Netherlands Academy of Arts and Sciences (*Koninklijke Nederlandse Akademie van Wetenschappen*—KNAW),⁶ formulated the initial general protocol for evaluating research quality. After testing and refinements, the protocol was adopted by all universities and research institutions in the Netherlands. The coordination of the evaluation process was entrusted to the VSNU.

In 1998, the second evaluation protocol, covering the period 1998–2003, was approved (VSNU, 1998). The following year, in 1999, the three main regulatory and funding agencies for higher education and research in the Netherlands (VSNU, NWO and KNAW) established a working group tasked with reviewing and improving evaluation policies. The final report, completed in 2000, served as a basis for the development of the third protocol. Subsequently, four additional regulatory frameworks were developed between 2003 and 2021, corresponding to the periods 2003–2009, 2009–2015, 2015–2021 and 2021–2027. From 2003 onward, the protocols were renamed the Standard Evaluation Protocol (SEP). In 2021, the term “standard” was replaced by “strategy” to reinforce a key principle of the Dutch system. The essence of the evaluation lies in the objectives and goals that the unit (institution) has set for itself. The evaluation is based on what the unit aspires to achieve. Therefore, the evaluation process is considered a form of SE. The strategy precedes the evaluation. The mission and objectives should be clearly formulated, known, shared, and consistently pursued (VSNU, 1993; 1998; KNAW, 2008; Drooge, 2021d).

The series of regulatory frameworks discussed above have established the foundation of what is commonly referred to as the Dutch model of evaluation.

4 The VSNU was created in 1985. It is an entity formed by fourteen research universities, four special universities and one open university.

5 The NWO is the most important research agency in the country. It receives public funding, mainly from the Ministry of Education, Culture and Science, and represents all the research areas in the Netherlands.

6 The KNAW is an independent scientific association formed by professors and researchers linked to public and private research institutes and laboratories.

These frameworks have defined the principles and guidelines for the institutionalization of periodic evaluation policies and processes at the national level. While each protocol reflects the context and requirements of its respective period, there are clear threads of continuity that run through them. The fundamental characteristics of the current evaluation model were already present in the SEP 2003–2009 (VSNU, KNAW, and NWO, 2003). The subsequent protocols, covering the periods 2009–2015, 2015–2021, and 2021–2027, have primarily introduced incremental changes, further solidifying the system. As noted by Drooge *et al.* (2013), these successive protocols have contributed to the institutionalization of a robust and lasting quality evaluation system.

Examining the evolution of the protocols over the past four decades allows for the identification of the key principles and guidelines that have shaped the policies and processes of research and higher education evaluation in the Netherlands. The presence of these lines of continuity enables us to analyze the similarities between the protocols and to understand the changes and improvements implemented throughout the years.

THE INSTITUTIONAL DESIGN OF THE EVALUATION SYSTEM

The ongoing changes and continuous improvements in the Dutch evaluation system have solidified a bottom-up model of evaluation. Since the publication of the Hoak Document (OCW, 1992), all subsequent regulatory frameworks have reinforced the principle that evaluation should be conducted by the institutions themselves, which engage in research and offer graduate programs. Evaluation is considered an inherent responsibility of institutional autonomy and is closely related to institutional planning and quality improvement policies. Quality evaluation cannot be delegated to funding agencies or external institutions. Therefore, in the Netherlands, there is no national agency, public or private, tasked with conducting quality evaluation processes for research and graduate education. Evaluation is mandatory every six years, but institutions and units are not required to undergo evaluation in the same year, and there is no unified national schedule.

The system is decentralized, giving institutions and units the autonomy and competence to establish their own internal evaluation policies and procedures. However, this decentralization and exercise of autonomy occur within a framework of nationally established guidelines provided by protocols published every six years. These protocols, known as the Strategy Evaluation Protocol (SEP), serve as the primary regulatory frameworks for evaluation.⁷ According to the current SEP (VSNU, KNAW, and NWO, 2020), the main objective of

⁷ In the period between 1993 and 2021, seven protocols were published, corresponding to the periods 1993–1998, 1998–2003, 2003–2009, 2009–2015, 2015–2021 and 2021–2027. From 2003 onwards, the protocols began to be called Standard Evaluation Protocol (SEP). In 2021, the term “standard” was replaced by “strategy”.

the protocols is to improve the quality and societal relevance of research, while fostering a continuous dialogue about the actions needed to improve quality and increase transparency and accountability to society, funding agencies, and the government. In addition to setting general guidelines, the protocols provide detailed procedures to guide the evaluation process. They are intended for all stakeholders involved in the evaluation, including researchers, unit directors, university managers, council members, science and technology managers, evaluation committee members, committee secretaries, and graduate students, among others.

As mentioned earlier, a significant change was introduced in the legal framework in 1992 to minimize external interference and control. With the publication of the document *Hoger Onderwijs en Wetenschappelijk Onderzoek* [Higher Education and Scientific Research] (OCW, 1992), the Ministry of Education, Culture, and Science ceased to issue guidelines and norms regarding the evaluation process. The regulatory authority was transferred to the associations representing the institutions involved in the evaluated activities. The protocols began to be jointly developed by the three major scientific and academic associations in the Netherlands: VSNU, NWO, and KNAW (Drooge *et al.*, 2013). The changes implemented over the past four decades have solidified the principle that evaluation should be conducted by peers, based on the assumptions and criteria that guide academic ethos. Funding institutions should not be responsible for the evaluation. Therefore, the separation between funding and evaluation is another crucial characteristic of the Dutch model.

The SEP undergoes a revision every six years. These periodic revisions aim to adapt the protocols to the needs of several types of higher education institutions, research institutions, and units. Consequently, the SEP is not the only evaluation protocol in effect in the country. There are specific protocols for research universities, universities of applied sciences, and academic and nonacademic research institutes (Drooge *et al.*, 2013). It is important to note that the Dutch higher education system encompasses fourteen research universities (*Universiteiten-WO*), four small, specialized universities, numerous universities of applied sciences (*Hogescholen-HBO*), and an open university (*Open Universiteit*) (Drooge, 2021a; VSNU, 2022).

An important characteristic of the evaluation protocols in the Netherlands is the absence of a distinct separation between research and graduate education. Doctoral programs are considered integral components of research programs and are evaluated as such. Therefore, the Strategy Evaluation Protocol provides guidelines for assessing the quality of research without the need for separate regulatory frameworks for evaluating graduate education.

Furthermore, it is worth noting that master's degree programs are approached differently in the current Dutch model compared to Brazil. In the Netherlands, as well as in many European countries, master's programs are part of the sequential educational process following undergraduate studies. This approach is in line with the Bologna Process and the adoption of the European Credit Transfer System, which promotes harmonization and standardization of higher education across Europe.

PRINCIPLES AND OBJECTIVES OF EVALUATION

The organization of the Dutch evaluation system, as described above, is founded on principles and overarching objectives that have guided evaluation policies since the publication of *Hoger Onderwijs en Wetenschappelijk Onderzoek* (OCW, 1992) and the development of the first protocol in 1994. The regulatory frameworks developed over the years have defined and solidified the principles shaping the system’s organization and functioning. Evaluation has been conceptualized based on key assumptions, including university autonomy, external evaluation through the peer review system, SE, accountability, quality improvement, and separation of funding and evaluation.

Based on these principles, the protocols have established a clear distinction from performative evaluation conceptions that view evaluation as a means of performance evaluation, outcome control, accountability, ranking generation, and the development of indicators for the allocation of rewards and incentives, such as scholarships and financial resources.

The mentioned principles are summarized in Chart 1, serving as the foundation for the objectives set forth by the protocols over the decades. These objectives reflect the system’s commitment to improving quality, ensuring accountability, and enhancing the societal relevance of research (Drooge, Jong, and Smit, 2022).

Chart 1 – General evaluation objectives established by the protocols since 2003.

Protocols	Objectives
SEP 2003–2009	“The evaluation system aims to achieve three objectives regarding research quality and management: (i) to enhance the quality of research through an evaluation process conducted according to international standards of quality and relevance; (ii) to improve research management and leadership; (iii) to be accountable to research organizations, funding agencies, government, and society at large” (VSNU, NWO, KNAW, and 2003, p. 5).
SEP 2009–2015	“The SEP 2009–2015 aims to achieve two objectives regarding research evaluation (including doctoral training) and management: (i) improve the quality of research through peer review evaluation, including academic and social relevance of research, research policies, and management; (ii) be accountable to research organizations, funding agencies, government, and society at large” (VSNU, NWO, and KNAW, 2009, p. 4).
SEP 2015–2021	“The main objective of the evaluation is to reveal and confirm the quality and relevance of research to society and improve it when necessary” (VSNU, NWO, and KNAW, 2016, p. 5).
SEP 2021–2027	“The main objective of the SEP is to evaluate the research unit based on its own objectives and strategies and to maintain and improve the quality and societal relevance of the research, as well as facilitate ongoing dialogues on quality, societal relevance, and feasibility within the context of research quality assurance” (VSNU, NWO, and KNAW, 2020, p. 6).

Source: VSNU, KNAW, and NWO (2003; 2009; 2016; 2020).

Chart 1 highlights not only the similarities and continuity in the objectives but also the qualitative changes that have occurred over time. The SEP 2021–2027 places a stronger emphasis on SE. In this context, the core of the evaluation process is based on the objectives, goals, and strategies established by the units⁸ themselves over the six-year period. Evaluation is not an externally imposed exercise, but rather a contextual and formative process. As emphasized by Drooge (2021a; b; c; d), it is an integral part of a continuous cycle of quality assurance and improvement. The evaluation process fosters an ongoing dialogue between the evaluated unit and various governing bodies, particularly regarding strengths, weaknesses, and areas for improvement that need to be addressed.

DIMENSIONS AND CRITERIA OF EVALUATION

The evaluation criteria, along with the principles and objectives discussed above, reflect the evolution and transformations of higher education, science, and technology policies in the Netherlands. These changes are a response to the evolving demands of society, the scientific community, and institutions. Therefore, the evaluation criteria express what society expects from science, researchers and their academic and scientific institutions (Spaapen and Drooge, 2011).

Chart 2 highlights the incremental changes introduced over time. A significant change was the reduction in the number of evaluation criteria from four to three. In 2015, the criterion of “productivity” was removed. This change was based on a proposition that had been advocated by the scientific community for several years (Drooge, Jong, and Smit, 2022). Additionally, the criterion of “relevance” underwent improvements over the decades. Starting from the SEP 2009–2015, it was expanded and became known as “societal relevance”. This change broadened the categories that guide the analysis of research impacts on society, encompassing not only scientific and socioeconomic dimensions but also cultural, social, and political impacts (VSNU, KNAW, and NWO, 2009). In 2015, the SEP 2015–2021 introduced further elements. It required that evaluation committees assess the quality and relevance of research based on the specific areas of operation and target audiences of each unit. Evaluation should consider the contributions of research to society, particularly in areas such as economic development, innovation, education, culture, health, sustainability, and more (VSNU, KNAW, and NWO, 2009; Drooge *et al.*, 2013).

8 The protocols use the term “unit” to designate research institutes, research centers, research groups, multi and interdisciplinary research centers, etc. Units are evaluated that meet the following conditions: be recognized, internally and externally, as an entity of search; have objectives, goals and strategies clearly defined and shared by the group; have, at least, the equivalent of ten researchers in its permanent academic staff, not counting doctoral students and post-doctoral fellows (in the Netherlands the concept of full-time equivalent — FTE is used, so that the required ten FTE can be covered by a larger number of researchers); and have at least three years of operation (VSNU, KNAW, and NWO, 2020).

Chart 2 – Dimensions and evaluation criteria established by the protocols since 2003.

Protocols	Criteria	Main aspects
SEP 2003–2009	Quality	- International recognition and innovative potential
	Productivity	- Scientific production
	Relevance	- Scientific and socioeconomic impact
	Vitality and viability	- Flexibility, management and leadership
SEP 2009–2015	Quality	- Quality and scientific relevance - Leadership - Academic reputation - Human and financial resources - PhD development
	Productivity	Productivity strategy - Productivity
	Relevance	Societal relevance
	Vitality and viability	Strategy - SWOT analysis - Robustness and stability
SEP 2015–2021	Research quality	- Scientific quality and relevance - Results: publications, products, infrastructure, and other contributions to science - PhD development policy - Scientific integrity - Diversity
	Relevance to society	- Scientific, economic, social, or cultural impact - Quality, scale, and relevance of contributions to specific groups (economic, social, cultural sectors); advice for the elaboration of public policies, reports, and other documents; contributions to the public debate, etc.
	Viability	- Strategy - Governance - Skills and leadership in the management of the research unit
SEP 2021–2027	Research quality	- Quality and contributions to scientific development (international, national, or regional) - Reputation and leadership - Open science - PhD development policy - Academic culture - Scientific integrity - Human resources policy: diversity and talent management
	Relevance to society	- Impact and public participation of the unit in areas related to the economy, social development, culture, education, and others that may be relevant
	Viability	- Viability and relevance of the objectives (the extent to which they remain scientifically and socially relevant) - Adequacy of management and resources to achieve objectives - Viability of the unit in relation to the foreseen objectives - Leadership

SWOT: Strengths, Weaknesses, Opportunities and Threats.
Source: VSNU, KNAW, and NWO (2003; 2009; 2016; 2020).

The evaluation criteria used in the Dutch system cover both quantitative and qualitative aspects, allowing for a comprehensive assessment of research and higher education. It is important to note that these criteria should not be applied rigidly. Institutions, councils, and committees have the autonomy to adapt and adjust according to the characteristics and objectives of each unit being evaluated. According to the SEP 2021–2027 (VSNU, KNAW, and NWO, 2020), the criteria are flexible tools; they should be interpreted according to the goals, strategies, and SE policies of each unit.

GUIDING PRINCIPLES OF SELF-EVALUATION

As previously highlighted, SE is a fundamental aspect of the Dutch evaluation system and distinguishes it internationally. It permeates and articulates the principles, objectives, criteria, and procedures of the evaluation process. It serves a dual role as both a principle and a method, defining the objectives and guiding the organization and procedures of the evaluation.

Based on the principles of autonomy, institutional planning, and accountability, the Dutch evaluation system has institutionalized evaluation as a practice of self-management conducted from within and from the bottom up. The units being evaluated are simultaneously “subjects” and “objects” of the process. External evaluation does not overshadow internal evaluation. As emphasized in the SEP 2021–2027 (VSNU, KNAW, NWO, and 2020), the core of the evaluation lies in the objectives, goals, and strategies set by the unit itself to maintain and improve quality, societal relevance, and viability. Goals and strategies are central elements of SE, allowing the unit to analyze the results achieved based on the previously established objectives and assess performance considering the mission and purposes it has established for itself. According to the SEP 2021–2027, SE enables reflection “[...] on the strategies adopted by the unit, as well as the effects they have produced” (VSNU, KNAW, and NWO, 2020, p. 19).

SE is, in this sense, a process related to the planning and management of the unit. As emphasized by Drooge (2021a; b), it is contextual and formative. In addition to considering the stages of development of the units, SE allows for the identification of progress, potential, weaknesses, and threats. In the words of Drooge (2021b), SE “[...] is not focused on research itself, but on the unit’s strategy regarding research.” For this reason, the SEP 2021–2027 (VSNU, KNAW, and NWO, 2020, p. 19) states that SE begins when the unit “[...] explicitly articulates its strategy and objectives.”

As highlighted earlier, the unit is both the subject and the object; it is both the starting point and the end point of the evaluation process. The results of the evaluation are returned to the unit in the form of a report to guide the planning and strategic actions of the next cycle. SE is, therefore, both an end and a means. Instead of being a specific, isolated, and independent stage, it is the structuring dimension around which the distinct stages are articulated and fed into one another.

SELF-EVALUATION AS A PROCESS: THE CENTER FOR SCIENCE AND TECHNOLOGY STUDIES EXPERIENCE

The experience described here provides valuable insights into the Dutch model of research and graduate evaluation, as implemented by the Center for Science and Technology Studies (CWTS) at Leiden University.

This experience is particularly relevant, as the CWTS is a renowned research center in the fields of research evaluation and scientometrics. As a result, SE provided opportunities for CWTS researchers to explore new methods, techniques, and differentiated approaches. These efforts to construct a robust evaluation process, based on the combination of quantitative and qualitative methods, were reported by the working group that supported the center's management in producing the report and related activities (CWTS, 2022), which also formed part of the final evaluation report completed in 2022 (*ibidem*).

The main stages of the SE process conducted by the CWTS between 2021 and 2022 can be summarized as follows, based on the reports and analysis of procedures outlined in the Strategy Evaluation Protocol 2021–2027 (VSNU, KNAW, and NWO, 2020).

DEFINITION OF THE UNIT'S STRATEGY, OBJECTIVES, AND GOALS

This is the starting point. The strategy and objectives of each research unit are formulated through a series of regular meetings between the units and the university council, which plays a similar role to that of the research and graduate studies councils in Brazil. The definition of objectives is part of an institutional quality assurance policy, which is realized through a set of strategies and goals that each unit must pursue during the evaluation cycle (VSNU, KNAW, and NWO, 2020).

As described in the section on the institutional design of the system, the evaluation cycle in the Netherlands is six years, and the current protocol covers the period 2021–2027. Each research unit conducts its evaluation process at its own pace, based on the previous six years. CWTS was one of the first units in the country to conduct its evaluation based on the guidelines established by the SEP 2021–2027.

The Dutch model is in fact quite different from the one adopted in Brazil, where the evaluation cycle is of four years, and all graduate programs are evaluated simultaneously. For example, the Quadrennial Evaluation in 2017 assessed the performance of graduate programs from 2013–2016, and the Quadrennial Evaluation conducted in 2022 evaluated the period from 2017–2020. An argument for synchronous evaluation in Brazil stems from the comparative approach adopted in the system, which has a significant impact on the distribution of funding in the country (Brasil, Trevisol, and Drooge, 2022).

ELABORATION OF THE TERMS OF REFERENCE AND THE EVALUATION ACTION

Based on the guidelines established by the SEP, the university council and the research unit jointly develop Terms of Reference (ToR), which is finalized approximately one year before the evaluation committee's visit. The ToR includes, among other aspects:

- basic information about the unit to be evaluated;
- the strategy, objectives, and key characteristics of the unit;
- the purposes and evaluation criteria;
- the key aspects that the external evaluation committee should consider; and
- the public nature of the final report (VSNU, KNAW, and NWO, 2020).

Although the ToR are straightforward documents, with a template included in the annexes of the SEP (VSNU, KNAW, and NWO, 2020), they are vital in ensuring that a research unit is evaluated according to its mission and objectives, playing an essential role in a truly multidimensional evaluation.

As mentioned in Brasil (2022), multidimensional evaluation is one in which the mission or vocation of a research program can be valued through a non-reductionist analysis, distinct from the traditionally adopted evaluation process in Brazil. In evaluations conducted by the Coordination of Improvement of Higher Education Personnel (CAPES) or Brazil's National Institute for Space Research (INEP), postgraduate programs or higher education institutions receive a grade based on the results of their evaluations. These grades reduce multiple dimensions of analysis to a single indicator that is incapable of demonstrating how each entity performs in terms of its different impacts: economic, social, scientific, etc.

In the Dutch evaluation model, research units can focus their analysis on dimensions that align with their objectives. For example, the CWTS — despite having a doctoral program — does not prioritize personnel training and does not offer undergraduate or master's degree programs. However, the institute focuses on research (with the majority of its academic staff consisting of researchers rather than professors), transdisciplinarity (maintaining strong ties with funding agencies and science policy makers), and societal impact (including through its own company, providing various services related to bibliometrics, scientometrics, and science policy consultancy). In a reductionist evaluation, the choices made by the CWTS could potentially hinder their results.

Once the ToR is validated, reflecting the missions and objectives of the unit that should have centrality in the evaluation, a series of parallel steps occur to inform the construction of an action plan produced by the research unit. Such a plan needs to include essential elements, such as the indication of members to integrate the external evaluation committee, the design of activities and key dates of the process, planning of bibliometric and scientometric analyses, intention to conduct comparative analyses with other units, and budgeting (hiring a secretary to support the work of the external committee, travel and accommodation costs, committee remuneration, etc.). The produced document should be submitted to the university council for approval.

In early 2021, CWTS appointed a working group responsible for assisting the center's board of directors throughout the evaluation process, including the development of the action plan. Based on a detailed planning of all evaluation stages, the following decisions were integrated into the document produced by the center, called *CWTS self-evaluation report* (CWTS, 2022):

1. Due to the unique profile of CWTS in the Netherlands, the planning did not include a comparative evaluation, and the center's performance would be analyzed in relation to the previous cycle.
2. The use of some traditional evaluation indicators were frowned upon (such as the H-index), and others, such as the impact factor, were prohibited (VSNU, KNAW, and NWO, 2020). However, research units can select sets of indicators and other evidence to support their self-evaluation. Thus, CWTS designed the desired analyses and selected preferred databases (Dimensions, ORCID, Altmetric), justifying their decision in the action plan.
3. The name of a secretary for the evaluation committee was indicated, with a proposed job description attached to the plan. It should be noted that the responsibilities of this professional go beyond traditional secretarial tasks, since the individual hired by CWTS was also a member of the committee responsible for the development of the current SEP.
4. An external evaluation committee conducted an on-site visit in May 2022.

Parallel to the activities of preparing the action plan, activities were conducted to define the members of the external evaluation committee, as described below.

COMPOSITION OF THE EXTERNAL EVALUATION COMMITTEE

The selection of evaluators is a crucial phase of the evaluation cycle. Each institution establishes procedures for composing the committee, considering the diversity of academic and scientific expertise. The committee should consist of impartial specialists, with international representation, including a doctoral candidate, an early and/or mid-career researcher, and, when necessary, a non-academic specialist. It is the responsibility of the committee to assess the quality, relevance to society and feasibility of the research conducted by the unit over the past six years. The unit should be evaluated based on the strategy and goals it has set for itself, considering the international, national, and, when necessary, regional context (VSNU, KNAW, and NWO, 2020).

Unlike the Brazilian model, where evaluation committees are formed based on the recommendations of the CAPES area coordinators, in the Dutch model, each department selects its own committee based on criteria defined in the SEP. At CWTS, this nomination process was participatory and followed the steps listed next:

1. The center's management and the self-evaluation working group held a meeting with the entire staff, presenting the evaluation process and the requirements for selecting committee members.
2. An open consultation was initiated, inviting all center staff to submit suggestions for committee members, resulting in fifty-five nominated names.
3. The head of management conducted an analysis of the nominations and the nominees' *curricula vitae*. Twelve experts were selected, six as primary members and six as substitutes.

4. The list was discussed with the entire center for validation and additional adjustments (e.g., one member had to be replaced because they had just become involved in a research project in collaboration with CWTS).

Nominated names were included in the action plan and submitted to the university council for approval. The document also proposed the appointment of a committee chairperson.

ELABORATION OF THE SELF-EVALUATION REPORT

The self-evaluation report is a central element of the evaluation process and is drafted by the research unit. It should be concise and well-founded, including:

- a description of the key characteristics of the unit;
- the mission, strategy and objectives of the unit (both current and future);
- evidence, quantitative and qualitative indicators, case studies, benchmarking, and other relevant information to demonstrate the main results and achievements;
- strengths and weaknesses; and finally,
- the strategic actions that the unit intends to implement in the next evaluation cycle (VSNU, KNAW, and NWO, 2020).

The process of producing the CWTS report involved the participation of the entire center and was conducted according to the following steps:

1. **Survey:** At the beginning of 2021, a survey form was sent to the entire CWTS staff, consisting of a set of questions grouped into two categories: one related to the respondent's activities, research conducted, topics of interest, and participation in projects; and the other related to the perception of the center's mission, objectives, and motivations, including perspectives or adjustments needed for the new cycle.
2. **Definition of report themes:** Units have the autonomy to define the structure and content of the report. As a result of analyzing the team's contributions, the document was structured based on themes of interest rather than the center's organizational or research structure. Therefore, the CWTS research groups and thematic axes were reorganized to better represent the center's identity.
3. **Thematic groups:** Under the coordination of GT members, *ad hoc* groups were created to discuss the central themes to be included in the self-evaluation report. Participants were selected to represent the center's expertise and different professional profiles (including faculty, researchers, doctoral students, and technical staff). Each group produced documents that helped management in constructing the final report.
4. **Bibliometric and scientometric analyses:** Research units have the option to hire external services to conduct bibliometric and scientometric studies. In the case of CWTS, as it is an area of expertise within the center itself, these analyses were carried out internally. During the process, sets of indicators were selected and the team developed a series

of dashboards and networks to meet the various categories of evidence required by the SEP, as presented in Chart 3.

Chart 3 – Categories of evidence for the domains of research quality and relevance to society.

		Quality domains	
		Research quality	Relevance to society
Evaluation dimensions	Scientific production	Academic scientific production	Scientific/technical/ technological production
	Production impact	Academic impact measures	Impact measures on society
	Evidence of recognition	Evidence of academic recognition	Evidence of recognition by society

Source: Adapted by the authors from the dimensions of the Standard Evaluation Protocol (VSNU, KNAW, and NWO, 2020, p. 34).

5. **SWOT Analysis.** The process of analyzing the strengths, weaknesses, opportunities, and threats of CWTS was conducted based on the responses from the center’s team to the survey form, combined with dedicated activities in internal meetings and with selected stakeholders.
6. **Appendices.** Additional evidence and additional information can be included in the appendices of the SE report. The SE report of CWTS included organizational information about the center, selected indicators with corresponding methodological descriptions, dashboards with networks and visualizations, research program, financial data, information about doctoral students, etc.

Based on the guidelines established in the most recent protocol for the period 2021–2027, it can be affirmed that the SE report fulfills four fundamental objectives. It challenges the units to:

- analyze the adequacy and relevance of the strategies and goals established by the unit in the last six years;
- clearly articulate the strategies and objectives that will guide the actions and activities of the unit in the next evaluation cycle;
- present the main achievements, weaknesses, and challenges in each of the evaluation criteria and subcriteria established by the SEP and the unit; and
- highlight the key results in the last six years for each of the evaluation criteria and subcriteria (VSNU, KNAW, and NWO, 2020).

Approximately two months before the scheduled date for the on-site visit, the SE report of the CWTS was submitted for approval by the university council and subsequently sent to the external evaluation committee (VSNU, KNAW, and NWO, 2020).

ON-SITE VISIT AND REPORT FROM THE EXTERNAL COMMITTEE

After the final report is sent to the external committee, the preparation process for the on-site visit begins. The evaluated unit organizes this visit, in direct contact with the committee members, and with the support of the secretary hired to assist in the evaluation activities. Through the on-site visit, it is possible to familiarize yourself with the infrastructure of the unit, conduct interviews with directors, research leaders, and other relevant individuals. During the on-site visit, the committee may request additional information not covered in the SE report (VSNU, KNAW, and NWO, 2020).

The evaluation committee's visit to CWTS was organized with a three-day schedule. The first day was dedicated to the committee's internal work. This activity was conducted without interference or participation from center representatives and was coordinated by the secretary and the chair of the committee. One of the initial activities included a brief presentation of the *Strategy Evaluation Protocol*, conducted by the group's secretary, highlighting its unique role in the process. This presentation was necessary due to the international profile of the committee, as some evaluators may not be familiar with the Dutch evaluation model.

On the second day of work, activities included interviews with various members of the CWTS team and discussions of the results, indicators, and other evidence presented in the SE report. Analysis of the recommendations made by the evaluation committee in 2015, derived results, and changes in the profiles of strengths, weaknesses, opportunities, and threats (SWOT) were also central to the work. On the third day, the committee met with the center members to present their initial impressions and recommendations.

After the on-site visit, the committee submitted an evaluation report to the CWTS. The evaluation protocol established that the report aims to communicate the main findings, proposals, and recommendations of the committee and should be clear and consistent. The protocol established a 30 day deadline to produce the report, starting from the end of the visit. The report is initially submitted to the evaluated unit and then to the institution's management (VSNU, KNAW, and NWO, 2020).

FINAL REPORT, RESPONSE DOCUMENT, AND FOLLOW-UP OF RESULTS

The institution's management is responsible for issuing a response document regarding the external committee's report, although in practice the evaluated research unit conducts the task itself. This document can clarify any doubts expressed in the report of the committee, present additional arguments, and even contest or reject the received recommendations. Both the response document and the evaluation report should be published on the institution's website (VSNU, KNAW, and NWO, 2020).

In the current *Strategy Evaluation Protocol*, the evaluation results are formative and it is up to the institution itself to reflect on the necessary actions to reward satisfactory performance, mitigate risks in the research unit, or take more drastic measures if the results have shown insufficient quality. There is no assignment of a

grade or direct impact on the unit's funding, whether it comes from the institution or other funding agencies. Therefore, the evaluation process aims to deepen the culture of evaluation, transparency, and strategic planning of institutions and units (VSNU, KNAW, and NWO, 2020).

Part of the formative process materializes shortly after the completion of the external evaluation through discussions of the results among leaders, the research unit, and the work team. Follow-up is conducted according to each institution's quality assurance policies, which may, if necessary, monitor the performance of the units, recommend a mid-term evaluation (VSNU, KNAW, and NWO, 2020).

In the case of CWTS, the final part of the evaluation process coincided with the beginning of the construction of a knowledge agenda—an expanded research program—effective from the beginning of 2023. This agenda was developed based on lessons learned throughout the evaluation process, including recommendations and discussions with the external committee and the University of Leiden itself. Furthermore, the new agenda aims to address the objectives and priority strategies for the 2023–2028 period and is constructed collaboratively through thematic group meetings, internal seminars, and the organization of two research retreats held in April and September 2022.

CONCLUSION

The Netherlands was one of the first European countries to implement a formal system for evaluating the quality of research and graduate education. The policies and changes implemented over the past four decades have established a *sui generis*, enduring, and stable system. Like all national evaluation systems, the Dutch model was conceived and improved on the basis of the country's characteristics and needs. It is part of a broad and complex political, cultural, social, and academic process. In addition to reflecting evaluation concepts, it also expresses the guidelines that the country has established for organizing higher education and the national science and technology system. Unlike other nations that have implemented centralized, standardized, performance-based evaluation systems, the Netherlands has consolidated a decentralized and heterogeneous model based on the principle that evaluation should be conducted based on the mission, goals, and strategies that the units being evaluated establish for themselves. Evaluation is therefore an internal, formative, contextual and participatory practice.

The Dutch system has been a reference and inspiration for numerous institutions and countries, including Brazil. Researchers and evaluation agencies have studied the model to understand its specificities and identify contributions to improving national evaluation systems. As this purpose also motivated the present study, we consider it important to conclude this article by highlighting some aspects that can serve as subsidies for improving the postgraduate evaluation system in Brazil.

The first aspect concerns the protagonism exercised by the institutions and the units being evaluated. Evaluation is an ongoing institutionalized practice of self-awareness and self-management. It is related to the exercise of institutional autonomy and strategic planning. It is not an instrument of state regulation, external control, or the production of rankings, performance indicators, and comparisons.

There is no national institution responsible for organizing, coordinating, and implementing the process at a national level. National evaluation protocols invariably attribute the competence and responsibility for evaluation to the institutions themselves. This trust has been reaffirmed and strengthened over the decades.

Second, the Dutch system is based on the principle that evaluation should prioritize the mission, goals, and strategies established by the unit being evaluated. The evaluation objectives should not be disconnected from strategic planning. Instead of assessing whether the unit has fully or partially achieved nationally established performance standards, the evaluation process seeks to analyze the results considering the goals and strategies previously established. Units are both the starting point and the destination. Through self-evaluation, they identify their strengths, weaknesses, and areas for improvement that need to be implemented.

Third, the evaluation aims to respect and strengthen the identity of each unit through a multidimensional approach. Units have the freedom to define the dimensions they wish to consider in their own evaluation, adopting criteria and indicators tailored to support each individual narrative, without generating standardized results for the entire system. Instead of promoting homogenization, the process preserves diversity and differentiation, even allowing for optional comparisons between units.

The fourth important aspect concerns the periodicity of the evaluation. Since 2003, the cycles have been set at six years. Longer periods facilitate the development and monitoring of units' strategic planning.

Indeed, while not perfect, the Dutch model seeks to integrate complex and challenging dimensions. Through SE, the system aims to promote balance and complementarity between qualitative and quantitative dimensions, internal and external evaluation processes, formative and summative dimensions, institutional autonomy and accountability, and processes of standardization and differentiation.

REFERENCES

- AFONSO, A.J. Mudanças no Estado-avaliador: comparativismo internacional e teoria da modernização revisitada. **Revista Brasileira de Educação**, v. 18, n. 53, p. 267-284, 2013. <https://doi.org/10.1590/S1413-24782013000200002>
- AMARAL, A. Avaliação e qualidade do ensino superior: as muitas racionalidades da qualidade. *In*: LEITE, D. (org). **Avaliação participativa e qualidade: os atores locais em foco**. Porto Alegre: Sulina, 2009. p. 11-32.
- BRASIL, A. Multidimensionality through self-evaluation: From theory to practice in the Brazilian graduate system. **Fteval Journal for Research and Technology Policy Evaluation**, n. 53, p. 97-106, 2022. <http://doi.org/10.22163/fteval.2022.546>
- BRASIL, A.; TREVISOL, J. T.; DROOGE, L. Research evaluation in Brazil and the Netherlands: a comparative study. International Conference of the European Forum for Studies of Policies for Research and Innovation. *In*: EUROPEAN FORUM FOR STUDIES OF POLICIES FOR RESEARCH AND INNOVATION, 2022, Utrecht. **Extended abstracts** [...]. Utrecht: Utrecht University, 2022. p. 258-261. Available at: <https://bit.ly/3gRMNUo>. Accessed on: Mar. 10, 2022.

CASTRO, M. H. G. **A educação para o século XXI**: o desafio da qualidade e da equidade. Brasília: INEP, 1999.

CLARIVATE. **Web of Science Core Collection**. [s.l.: s.n., s.d.]. Available at: <https://bit.ly/3rHmCJT>. Accessed on: Mar. 16, 2022.

COBBAN, A. B. **The Medieval English Universities**: Oxford and Cambridge to c. 1500. London: Routledge, 2017.

CWTS. **CWTS self-evaluation report: SEP 2021-2027**. Leiden, NL: Centre for Science and Technology Studies, Leiden University, 2022. Available at: <https://sep.cwts.nl>. Accessed on: Mar. 15, 2022.

DROOGE, L.; JONG, S.; FABER, M.; WETERHEIJDEN. **Twenty years of research evaluation**: Facts & Figures. The Hague: Rathenau Instituut, 2013. Available at: <http://bit.ly/3cXp4mK>. Accessed on: Dec. 10, 2021.

DROOGE, L. Research evaluation in context 1: Introducing research evaluation in the Netherlands. **Leiden Madtrics**, 12 fev. 2021a. Available at: <https://bit.ly/3IsQZJE>. Accessed on: Jan. 7, 2022.

DROOGE, L. Research evaluation in context 2: One joint protocol, three criteria and four aspects. **Leiden Madtrics**, 17 fev. 2021b. Available at: <https://bit.ly/3Atb67G>. Accessed on: Jan. 12, 2022.

DROOGE, L. Research evaluation in context 3. The board, the research unit and the assessment committee. **Leiden Madtrics**, 4 mar. 2021c. Available at: <https://bit.ly/3fMAoE4>. Accessed on: Jan. 12, 2022.

DROOGE, L. Research evaluation in context 4. The practice of research evaluation. **Leiden Madtrics**, 12 maio 2021d. Available at: <https://bit.ly/3IqzLfL>. Accessed on: Jan. 13, 2022.

DROOGE, L.; JONG, S.; SMIT, J. P. **What makes societal impact 'Excellent'?** A qualitative comparative analysis of evaluation registers in Dutch Research Assessment. Leiden: Center for Science and Technology Studies, 2022.

GOEDEGEBUURE, L.; KAISER, F.; MAASSEN, P.; WEERT, E. Higher education policy in The Netherlands. *In*: GOEDEGEBUURE, L.; KAISER, F.; MAASSEN, P.; MEEK, L; VUGHT, F.; WEERT, E. (ed.). **Higher education policy**: an international comparative perspective. Issues in Higher Education. Oxford: Pergamon Press, 1994. p. 188-213. <https://doi.org/10.1016/B978-0-08-042393-7.50015-8>

GOEDEGEBUURE, L. C. J.; WESTERHEIJDEN, D. F. Changing balances in Dutch higher education. **Higher Education**, v. 21, n. 4, p. 495-520, 1991. <https://doi.org/10.1007/bf00134986>

HICKS, D. Performance-based university research funding systems. **Research Policy**, v. 41, n. 2, p. 251-261, 2012. <https://doi.org/10.1016/j.respol.2011.09.007>

KAISER, F.; WEERT, E. Retoriek of rationeel beleid: deelname en deelnamebeleid in internationaal perspectief. **Tijdschrift voor hoger onderwijs**, v. 24, n. 3, p. 128-151, 2006.

KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW. **From SEP to CEP: Balancing fairness and simplicity**: Quality assurance in scientific research. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 2008. Available at: <https://bit.ly/3DmzOYY>. Accessed on: Mar. 14, 2022.

- LEPORI, B.; REALE, E.; SPINELLO, A. O. Conceptualizing and measuring performance orientation of research funding systems. **Research Evaluation**, v. 27, n. 3, p. 171-183, 2018. <https://doi.org/10.1093/reseval/rvy007>
- MEULEN, B. V. D. Science policies as principal-agent games Institutionalization and path dependency in the relation between government and science. **Research Policy**, v. 27, n. 4, p. 397-414, 1998. [https://doi.org/10.1016/s0048-7333\(98\)00049-3](https://doi.org/10.1016/s0048-7333(98)00049-3)
- MEULEN, B. V. D. Interfering governance and emerging centres of control. *In*: WHITLEY, R.; GLÄSER, J. (ed). **The changing governance of the sciences: the advent of research evaluation systems**. Sociology of the Sciences Yearbook. Dordrecht: Springer, 2007. p. 191-203. https://doi.org/10.1007/978-1-4020-6746-4_9
- MINISTERIE VAN ONDERWIJS, CULTUUR EN WETENSCHAP — OCW. **Hoger onderwijs: autonomie en kwaliteit**. Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschap, 1985. Available at: <https://bit.ly/3GXeid7>. Accessed on: Dec. 30, 2021.
- MINISTERIE VAN ONDERWIJS, CULTUUR EN WETENSCHAP — OCW. **Hoger onderwijs en wetenschappelijk onderzoek**. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap, 1992. Available at: <https://bit.ly/33YmT11>. Accessed on: Dec. 22, 2021.
- MOLAS-GALLART, J. Research governance and the role of evaluation: a comparative study. **American Journal of Evaluation**, v. 33, n. 4, p. 583-598, 2012. <https://doi.org/10.1177/1098214012450938>
- OCHSNER, M.; KULCZYCKI, E.; GEDUTIS, A. The diversity of european research evaluation systems. *In*: INTERNATIONAL CONFERENCE ON SCIENCE AND TECHNOLOGY INDICATORS, 23., 2018, Leiden. **STI 2018 Conference Proceedings** [...]. Leiden, NL: STI, 2018. Available at: <https://bit.ly/3SR11sI>. Accessed on: Jan. 5, 2022.
- RIP, A.; MEULEN, B. V. D. The patchwork of the Dutch evaluation system. **Research Evaluation**, v. 5, n. 1, p. 45-53, 1995. <https://doi.org/10.1093/rev/5.1.45>
- RIP, A.; MEULEN, B. V. D. The post-modern research system. **Science and Public Policy**, v. 23, n. 6, p. 343-352, 1996. <https://doi.org/10.1093/spp/23.6.343>
- SOBRINHO, J. D. **Avaliação da educação superior**. Petrópolis: Vozes, 2000.
- SOBRINHO, J. D. Avaliação da educação superior: Regulação e emancipação. **Avaliação: Revista da Avaliação da Educação Superior**, v. 8, n. 2, p. 31-47, 2003. Available at: <https://bit.ly/3I5MR2m>. Accessed on: Nov. 26, 2021.
- SPAAPEN, J.; DROOGE, L. V. Introducing “productive interactions” in social impact assessment. **Research Evaluation**, v. 20, n. 3, p. 211-218, 2011. <https://doi.org/10.3152/095820211x12941371876742>
- VERHINE, R. E.; FREITAS, A. A. S. M. A avaliação da educação superior: modalidades e tendências no cenário internacional. **Revista Ensino Superior Unicamp**, n. 7, p. 16-39, 2012. Available at: <https://bit.ly/2C9herj>. Accessed on: Nov. 25, 2021.
- VERENIGING VAN UNIVERSITEITEN – VSNU. **Quality Assessment of Research**: Protocol 1993. Utrecht: VSNU, 1993.

VERENIGING VAN UNIVERSITEITEN – VSNU. **Protocol 1998**: Series Assessment of Research Quality. Utrecht: VSNU, 1998.

VERENIGING VAN UNIVERSITEITEN – VSNU. **Universiteiten van Nederland**. 2022. Available at: <https://www.universiteitenvannederland.nl/>. Accessed on: Jan. 21, 2022.

VERENIGING VAN UNIVERSITEITEN – VSNU; KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW; NEDERLANDSE ORGANISATIE VOOR WETENSCHAPPELIJK ONDERZOEK — NWO. **Standard Evaluation Protocol (2003-2009)**. The Hague, 2003. Available at: <https://bit.ly/3rbJtzR>. Accessed on: Dec. 16, 2021.

VERENIGING VAN UNIVERSITEITEN – VSNU; KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW; NEDERLANDSE ORGANISATIE VOOR WETENSCHAPPELIJK ONDERZOEK — NWO. **Standard Evaluation Protocol (2009-2015)**. The Hague, 2009. Available at: <https://bit.ly/3t0B02X>. Accessed on: Dec. 27, 2021.

VERENIGING VAN UNIVERSITEITEN – VSNU; KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW; NEDERLANDSE ORGANISATIE VOOR WETENSCHAPPELIJK ONDERZOEK — NWO. **Standard Evaluation Protocol (2015-2021)**. [Rep. Amended version]. The Hague, 2016. Available at: <https://bit.ly/45TIpzr>. Accessed on: Jan. 16, 2022.

VERENIGING VAN UNIVERSITEITEN – VSNU; KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW; NEDERLANDSE ORGANISATIE VOOR WETENSCHAPPELIJK ONDERZOEK — NWO. **Strategy Evaluation Protocol (2021-2027)**. The Hague, 2020. Available at: <https://bit.ly/3wAFbzi>. Accessed on: Jan. 16, 2022.

VUGHT, F. V. The new context for academic quality. *In*: DILL, D. D.; SPORN, B. (ed.). **Emerging patterns of social demand and university reform**: Through a glass darkly. Kidlington, UK: Emerald Publishing, 1995. p. 194-211.

WEERT, E.; BOEZEROOY, P. **Higher education in the Netherlands**. Enschede, NL: Center for Higher Education Policy Studies (CHEPS), 2007.

ABOUT THE AUTHORS

JOVILES VITÓRIO TREVISOL has a doctorate in Sociology from the Universidade de São Paulo (USP). He is a professor at the Unversidade Federal da Fronteira Sul (UFFS).

E-mail: joviles.trevisol@uffs.edu.br

ANDRÉ BRASIL is a doctoral student in Assessment Studies and Scientometrics from the Leiden University (Netherlands). He is a server of the Evaluation Board at the CAPES.

E-mail: a.l.brasil@cwts.leidenuniv.nl

Conflicts of interest: The authors declare they don't have any commercial or associative interest that represents conflict of interests in relation to the manuscript.

Funding: This study was carried out with the support of the National Council for Scientific and Technological Development (CNPq), within the scope of the Post Doctorate Abroad modality (Public Call 08/2019) and the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES) — Financing Code 001.

Authors' contribution: Conceptualization, Data Curation, Formal Analysis, Methodology, Writing — Original Draft, Writing — Review & Editing.

Received on April 22, 2022

Approved on November 4, 2022

