

As políticas de avaliação da pesquisa e da pós-graduação nos Países Baixos: lições sobre autoavaliação e autonomia institucional

Joviles Vitório Trevisol^I 

André Brasil^{II} 

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.

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

^{II}Universiteit Leiden, Leiden, Países Baixos.

EVALUATION POLICIES FOR RESEARCH AND GRADUATE EDUCATION IN THE NETHERLANDS: LESSONS ON SELF-EVALUATION AND INSTITUTIONAL AUTONOMY

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.

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.

INTRODUÇÃO

Nas últimas décadas a avaliação assumiu enorme centralidade nos sistemas nacionais de educação básica e superior, ciência, tecnologia e inovação. A relevância atual não pode, no entanto, nos levar a supor que o interesse pelo tema seja recente. De acordo com Vught (1995), Amaral (2009) e Cobban (2017), há registro de práticas de avaliação na educação superior desde o século XIII. Segundo os autores, as primeiras universidades europeias realizavam embrionários processos de avaliação baseados em duas concepções bastante distintas. A primeira, conhecida como modelo francês, conferia a uma autoridade externa o poder de conceder a licença para ensinar (*licentia ubique docendi*) e o direito de decidir sobre o conteúdo dos estudos. A avaliação era uma espécie de prestação de contas à autoridade superior externa responsável pela chancela. A título de ilustração, no início do século XIII a Universidade de Paris era tutelada pelo bispo de Paris e pelo chanceler da catedral de Notre-Dame. A segunda concepção, de inspiração inglesa, tinha como principal característica o autogoverno. As universidades gozavam de elevado grau de autonomia e a avaliação era realizada pelos pares. Os colegiados tinham o direito de avaliar os colegas e, quando necessário, substituí-los (Amaral, 2009).

Essas diferentes concepções deram origem a distintas culturas de avaliação (Hicks, 2012; Lepori, Reale e Spinello, 2018; Ochsner, Kulczycki e Gedutis, 2018). As práticas transformaram-se ao longo dos séculos, dando origem a modelos híbridos e bastante heterogêneos. Pode-se afirmar que até década de 1980 a avaliação era predominantemente uma prática interna, esporádica e pouco institucionalizada. A avaliação externa — especialmente a realizada pelos governos — era vista como uma ameaça ao *ethos* científico. Ela era rechaçada sob o argumento de que ciência deveria ser concebida e realizada segundo seus próprios pressupostos, métodos e critérios de legitimidade (Meulen, 1998; 2007). A avaliação da qualidade devia ser feita pelas instituições acadêmicas e pelos pares. A avaliação externa era compreendida como um instrumento de interferência na autonomia das instituições e uma forma de limitar da liberdade de pensamento e de criação. A avaliação do mérito da atividade científica devia ser feita pelos pares e não por *outsiders* (Molas-Gallart, 2012).

A partir dos anos 1980, os debates sobre a avaliação assumiram novas perspectivas. As agências nacionais de regulação e de financiamento da educação superior e da pesquisa passaram a defender a necessidade de políticas permanentes e padronizadas de avaliação. Por meio da institucionalização desses processos seria possível, de acordo com as agências, avaliar os resultados dos crescentes investimentos públicos, conferindo regularidade, transparência e uniformidade de critérios e procedimentos em âmbito nacional (Hicks, 2012; Ochsner, Kulczycki e Gedutis, 2018). Além de ser uma dimensão interna e exclusiva das universidades e dos pesquisadores, a avaliação foi, gradativamente, se consolidando como uma das principais dimensões dos sistemas nacionais de educação superior, ciência e tecnologia (Amaral, 2009; Verhine e Freitas, 2012). Por meio de suas agências de regulação e de financiamento, os Estados nacionais foram, aos poucos, assumindo o papel de conduzir as avaliações periódicas, influenciando não apenas as políticas de pesquisa e educação superior, mas também a distribuição dos recursos. Desse processo surgiu o chamado “sistema de financia-

mento baseado no desempenho” (Ochsner, Kulczycki e Gedutis, 2018, p. 1235). A vinculação entre desempenho e financiamento desencadeou mudanças profundas na gestão das universidades e das instituições de pesquisa (Castro, 1999; Sobrinho, 2000; 2003; Amaral, 2009; Afonso, 2013).

Nas últimas décadas, diversos países institucionalizaram os seus sistemas de avaliação. Enquanto alguns optaram por modelos centralizados, padronizados e baseados em métricas, outras nações implementaram modelos descentralizados, menos quantitativos e alicerçados no princípio de que a avaliação deve ser realizada pelas próprias instituições a fim de fortalecer as suas missões institucionais e orientar os seus planejamentos estratégicos.

Esse contexto motivou a realização do presente estudo, cujo objeto de análise é o sistema de avaliação holandês. Os Países Baixos foram um dos primeiros países europeus a implementar um sistema formal de avaliação da qualidade da pesquisa e da pós-graduação.¹ A experiência holandesa tem sido referência para outras nações e instituições, entre elas o Brasil. Trata-se de um modelo descentralizado, *bottom-up* e estruturado sobre os princípios da autonomia, da autoavaliação (AA) e do planejamento institucional (Weert e Boezeroy, 2007; Drooge *et al.*, 2013). A legitimidade que o modelo adquiriu para a comunidade acadêmica e científica e, principalmente, a posição que o país ocupa no *ranking* da produção científica internacional indexada — 14ª posição de acordo com a *Web of Science* (Clarivate, n. d.) — indicam, entre outros aspectos, a pertinência e as contribuições das políticas de avaliação para a consolidação da qualidade.

De forma mais objetiva, o presente artigo pretende:

1. analisar a evolução das políticas de avaliação da pesquisa e da pós-graduação nos Países Baixos;
2. compreender as principais características que o modelo assumiu nos últimos anos; e
3. entender o papel da AA no sistema.

A pesquisa sobre as principais características e especificidades do modelo holandês será complementada por um estudo de caso. Trata-se de uma experiência recente de AA, desenvolvida entre 2021 e 2022 pelo *Center for Science and Technology Studies* (CWTS) da Universidade de Leiden. Com base nesse estudo destacaremos, por fim, alguns aspectos que podem servir de subsídios para o aprimoramento do sistema brasileiro de avaliação da pós-graduação.

OS MARCOS REGULATÓRIOS EM PERSPECTIVA HISTÓRICA

De acordo com a pesquisa realizada, as primeiras iniciativas no sentido de introduzir normas e diretrizes para a avaliação da qualidade da pesquisa nos Países

1 A partir dos anos 1980 vários países europeus criaram os seus sistemas de avaliação, cabendo destaque para a Inglaterra (1986), Países Baixos (1993), Espanha (1989), Polônia (1991), China (1993), Finlândia (1998), Hong Kong (2002), Nova Zelândia (2003), Bélgica (2003), Noruega (2006), Dinamarca (2006), França (2006), Suécia (2009), Itália (2009), Austrália (2010) (Hicks, 2012; Drooge *et al.*, 2013).

Baixos foram implementadas no final dos anos 1960. Em 1969, o *Advisory Council for Science Policy* recomendou a criação de comitês de especialistas para a avaliação dos programas de pesquisa financiados pelo governo. No início dos anos 1970 foram nomeados os comitês das áreas de educação, política social e planejamento ambiental, seguidos, em 1976, pelas áreas de química, bioquímica, biologia e física. Esses grupos tinham uma composição tripartite, sendo integrados por pesquisadores, membros do governo e *stakeholders*. Contudo, a falta de clareza sobre os objetivos, procedimentos e critérios da avaliação comprometeram o sucesso dessas primeiras experiências (Meulen, 1998; 2007). Em 1979, o Ministério da Educação e Ciências estabeleceu novas diretrizes para a organização da pesquisa nas instituições. Os programas de pesquisa passaram a ser o núcleo central da avaliação realizada pelos comitês de especialistas (Goedegebuure e Westerheijden, 1991).

As experiências de avaliação realizadas ao longo da década de 1970 ofereceram elementos para a elaboração do primeiro marco regulatório da avaliação dos Países Baixos. O *Conditional Financing* (CF), publicado em 1982, estabeleceu as diretrizes gerais para a organização do sistema nacional de avaliação. De acordo com Goedegebuure e Westerheijden (1991), o CF pode ser considerado a primeira tentativa de introduzir um sistema formal de avaliação da qualidade da educação superior e da pesquisa. Ele foi elaborado com o propósito de promover a qualidade, aumentar a transparência e avaliar os resultados dos investimentos feitos pelo governo e demais agências de fomento. O documento consolidou as unidades de pesquisa como os núcleos centrais da organização da pesquisa nas instituições (Rip e Meulen, 1995; 1996; Meulen, 1998; 2007). Além disso, o CF:

- estabeleceu critérios e exigências para a criação e a avaliação das unidades de pesquisa;
- tornou obrigatória a definição dos objetivos, metas, plano de trabalho, orçamento e equipe de pesquisadores que integram as unidades;
- implantou um processo de acreditação conduzido por organismos externos; e
- introduziu a obrigatoriedade da avaliação de resultados após cinco anos de atividades (Goedegebuure e Westerheijden, 1991; Meulen, 1998).

O CF tornou obrigatória a avaliação externa, realizada por meio do sistema *peer review*. De acordo com o documento, os pesquisadores deviam ser — por sua longa e especializada formação — os responsáveis pela avaliação da qualidade acadêmica e da relevância da pesquisa. A presença de avaliadores externos e independentes foi considerada fundamental para assegurar a legitimidade, elevar o nível de transparência e oferecer informações confiáveis à comunidade acadêmica e ao conjunto das instituições envolvidas (Weert e Boezeroy, 2007; Drooge *et al.*, 2013).

A nomeação dos comitês externos ficou sob a responsabilidade de uma associação científica nacional, a *Royal Netherlands of Arts and Sciences* (KNAW). Por meio do sistema *peer review* seria possível proteger a avaliação das interferências externas, especialmente do controle das agências financiadoras. De acordo com o documento, a avaliação não devia ser realizada pelas agências financiadoras, tampouco pelas instituições avaliadas (Rip e Meulen, 1995; 1996; Meulen, 1998).

O documento também introduziu mudanças nas políticas de financiamento. O número de matrículas deixou de ser o único critério para a definição do financiamento das instituições (Goedegebuure e Westerheijden, 1991). Os resultados da avaliação da qualidade e da relevância societal da pesquisa passaram a ter impacto na definição do orçamento institucional (Meulen, 1998; 2007; Weert e Boezeroy, 2007).

Ao longo dos anos 1980, o Ministério da Educação, Cultura e Ciência publicou vários documentos complementares com o propósito de detalhar e aprimorar as políticas estabelecidas pelo CF. O principal deles, intitulado *Hoger Onderwijs: Autonomie en Kwaliteit* (Ministerie van Onderwijs, Cultuur en Wetenschap – OCW, 1985)² — conhecido também como *Hoak Document* — fortaleceu a autonomia das instituições e introduziu um sistema retrospectivo de avaliação da qualidade. Em vez de estabelecer políticas intervencionistas, o documento recomendou a elaboração de políticas institucionais de avaliação e o fortalecimento dos processos de avaliação externa. As universidades foram orientadas a introduzir processos de gestão institucional mais eficientes e profissionais (Goedegebuure e Westerheijden, 1991). Além de fortalecer a autonomia, o *Hoak Document* estabeleceu novas diretrizes para o financiamento e o planejamento institucional. As instituições foram desafiadas a operar de forma mais profissional e em sintonia com as necessidades da sociedade. Em troca de maior autonomia, as universidades tiveram que criar políticas institucionais de avaliação e de transparência dos resultados (Goedegebuure e Westerheijden, 1991; Goedegebuure *et al.*, 1994; Kaiser e Weert, 2006; Drooge *et al.*, 2013).

Alguns anos depois, em 1992, o governo publicou uma nova lei com o propósito de regulamentar as relações entre o governo, as instituições de ensino superior e os institutos de pesquisa. A *Hoger Onderwijs en Wetenschappelijk Onderzoek* (OCW, 1992)³ fortaleceu ainda mais a autonomia e o papel das instituições no processo avaliativo. A lei reduziu o controle do Estado e limitou a interferência do governo sobre as dimensões operacionais da avaliação. A intervenção estatal foi limitada a algumas situações específicas. De acordo com a lei, o governo devia intervir apenas quando constatadas deficiências, desvios na aplicação dos recursos ou discrepâncias entre os objetivos e os resultados. Em vez de realizar a avaliação, o governo devia zelar pela macroeficiência do sistema.

De acordo com Weert e Boezeroy (2007, p. 14), a lei estabeleceu quatro princípios gerais que deviam orientar a atuação governamental:

1. O governo deve intervir para evitar desenvolvimentos indesejáveis somente onde a autogestão pelas instituições possa ter resultados inaceitáveis.
2. A intervenção do governo deve ter como objetivo principal remediar imperfeições no sistema *ex post*.
3. Os instrumentos à disposição do governo devem ser caracterizados por um mínimo de regulamentação detalhada.
4. As instituições devem estabelecer normas para garantir segurança jurídica, razoabilidade e administração adequada.

2 Ensino Superior: Autonomia e Qualidade.

3 Ensino Superior e Pesquisa Científica.

De acordo com o documento, a avaliação é uma dimensão do planejamento institucional. As políticas institucionais devem prever a participação de avaliadores externos e independentes, preferencialmente vinculados a outras instituições (OCW, 1992, sec. 1.18). A avaliação não é um fim em si. Ela não pode ser um instrumento de controle externo. Ela deve ser concebida como “[...] uma maneira que permite ao sistema de educação superior responder de forma mais eficaz e decisiva às necessidades e às mudanças da sociedade” (Weert e Boezerooy, 2007, p. 13). O referido documento forneceu a base legal para as instituições elaborarem as suas próprias políticas.

Em 1993, a Associação das Universidades Holandesas (*Vereniging van Universiteiten — VSNU*)⁴ elaborou, em conjunto com o Conselho Holandês de Pesquisa (*Nederlandse Organisatie voor Wetenschappelijk Onderzoek — NWO*)⁵ e a Academia Real Holandesa de Artes e Ciências (*Koninklijke Nederlandse Akademie van Wetenschappen — KNAW*),⁶ o primeiro protocolo geral de avaliação da qualidade da pesquisa. Após alguns testes e ajustes, o protocolo foi adotado por todas as universidades e instituições de pesquisa do país. A coordenação do processo de avaliação ficou sob a responsabilidade da VSNU.

Cinco anos depois, em 1998, foi aprovado o segundo protocolo de avaliação, correspondente ao período 1998–2003 (VSNU, 1998). No ano seguinte, em 1999, as três principais agências de regulação e de financiamento da educação superior e da pesquisa do país (a VSNU, o NWO e a KNAW) designaram um grupo de trabalho responsável pela revisão e o aprimoramento das políticas de avaliação. O relatório final, concluído em 2000, serviu de referência para a elaboração do terceiro protocolo. Entre 2003 e 2021 foram elaborados quatro novos marcos regulatórios, correspondentes aos períodos 2003–2009, 2009–2015, 2015–2021 e 2021–2027. A partir de 2003, os protocolos passaram a se chamar *Standard Evaluation Protocol* (SEP). Em 2021, o termo *standard* foi substituído por *strategy*. A palavra “estratégia” foi incorporada com o propósito de fortalecer um dos principais princípios do sistema holandês. O núcleo central da avaliação são os objetivos e as metas que a unidade estabeleceu para si mesma. A unidade é avaliada com base no que ela se propõe a ser. A avaliação é, neste sentido, essencialmente AA. A estratégia precede a avaliação. Além de claramente formulados, a missão e os propósitos devem ser conhecidos, partilhados e perseguidos continuamente (VSNU, 1993; 1998; KNAW, 2008; Drooge, 2021d).

O conjunto dos marcos regulatórios destacados anteriormente consolidou o que podemos denominar de modelo holandês de avaliação. Estabeleceu os princípios e as diretrizes para a institucionalização das políticas e dos processos de avaliação

4 A VSNU foi criada em 1985. É uma entidade formada por 14 universidades de pesquisa, quatro universidades especiais e uma universidade aberta.

5 O NWO é a mais importante agência de pesquisa do país. Recebe financiamento público, especialmente do Ministério da Educação, Cultura e Ciência, e representa todas as áreas de pesquisa desenvolvidas nos Países Baixos.

6 A KNAW é uma associação científica independente formada por professores universitários e pesquisadores vinculados a institutos e laboratórios públicos e privados de pesquisa.

periódica em âmbito nacional. Cada protocolo reflete o contexto e as necessidades de cada período. Apesar das especificidades de cada um, há claras linhas de continuidade entre eles. As principais características do atual modelo de avaliação já estavam presentes no SÉP 2003–2009 (VSNU, KNAW e NWO, 2003). As mudanças implementadas a partir de 2003 foram essencialmente incrementais. Os aprimoramentos introduzidos pelos últimos quatro protocolos, referentes ao período 2003–2021, consolidaram o sistema. Como destacam Drooge *et al.* (2013), os diferentes protocolos contribuíram para a institucionalização de um longo e estável sistema de avaliação da qualidade.

As linhas de continuidade permitem-nos analisar as similaridades entre os protocolos. Por meio da análise das mudanças e dos aprimoramentos introduzidos ao longo dos últimos 40 anos é possível identificar os princípios e as diretrizes que vêm orientando as políticas e os processos de avaliação da pesquisa e da pós-graduação nos Países Baixos.

O DESENHO INSTITUCIONAL DO SISTEMA DE AVALIAÇÃO

As mudanças e os contínuos aprimoramentos consolidaram no país um modelo *bottom-up* de avaliação. A partir da publicação do *Hoak Document* (OCW, 1992) todos os marcos regulatórios partilharam e fortaleceram o princípio de que a avaliação deve ser realizada pelas próprias instituições que desenvolvem a pesquisa e ofertam os programas de pós-graduação. A avaliação é uma competência inerente ao princípio da autonomia. Ela deve estar estreitamente vinculada ao planejamento e às políticas institucionais de melhoria da qualidade. A avaliação da qualidade não pode ser atribuída às agências financiadoras ou a qualquer outra instituição externa. Por essa razão, nos Países Baixos não há uma agência nacional — pública ou privada — responsável pela realização dos processos de avaliação da qualidade da pesquisa e da pós-graduação. A avaliação é obrigatória a cada seis anos, mas as instituições e as unidades não precisam realizá-la no mesmo ano. Não há um calendário nacional unificado.

O sistema é descentralizado. As instituições e as unidades têm autonomia e gozam de competência para estabelecer as políticas e os procedimentos internos de avaliação. A descentralização e o exercício da autonomia, no entanto, ocorrem apoiados em um conjunto de diretrizes estabelecido nacionalmente por meio dos protocolos publicados a cada seis anos. Os protocolos, conhecidos como *Strategy Evaluation Protocol* (SEP), consolidaram-se como os principais marcos regulatórios da avaliação.⁷ De acordo com o SEP em vigência (VSNU, KNAW e NWO, 2020), o principal objetivo dos protocolos é aprimorar a qualidade e a relevância societal da pesquisa, assim como promover um contínuo diálogo

7 No período entre 1993 e 2021 foram publicados sete protocolos, correspondentes aos períodos 1993–1998, 1998–2003, 2003–2009, 2009–2015, 2015–2021 e 2021–2027. A partir de 2003 os protocolos passaram a se chamar *Standard Evaluation Protocol* (SEP). Em 2021, o termo “*standard*” foi substituído por “*strategy*”.

sobre as ações que precisam ser implementadas a fim de melhorar a qualidade e ampliar a transparência e a prestação de conta perante a sociedade, as agências de fomento e o governo. Além de estabelecer as diretrizes gerais, os protocolos detalham os procedimentos que devem orientar a avaliação. Eles se destinam a todos os atores que participam do processo avaliativo, especialmente aos pesquisadores, diretores das unidades, gestores universitários, membros dos conselhos, gestores de ciência e tecnologia, membros das comissões de avaliação, secretários dos comitês, estudantes de pós-graduação etc.

Como destacado anteriormente, os protocolos não são elaborados pelas agências governamentais. Com o propósito de reduzir ao mínimo a interferência e o controle externo, em 1992 foi introduzida uma mudança substantiva no arcabouço legal que define as competências e as atribuições das diferentes instituições que integram o sistema. A partir da publicação do documento *Hoger Onderwijs en Wetenschappelijk Onderzoek* (OCW, 1992), o Ministério da Educação, Cultura e Ciência deixou de emitir diretrizes e normas referentes ao processo de avaliação. A competência da regulação passou a ser exercida pelas associações que representam as instituições que realizam as atividades fim avaliadas. Os protocolos passaram a ser elaborados, de forma conjunta, pelas três mais importantes associações científicas e acadêmicas do país: VSNU, NWO e KNAW (Drooge *et al.*, 2013). As mudanças introduzidas nos últimos 40 anos consolidaram o princípio de que a ciência deve ser avaliada pelos pares, conforme os pressupostos e critérios que orientam o *ethos* acadêmico. As instituições que financiam não devem ser responsáveis pela avaliação. A separação entre financiamento e avaliação é, desse modo, outra importante característica do modelo holandês.

Os SEP são revisados a cada seis anos. As revisões periódicas visam, entre outros objetivos, adequar os protocolos às necessidades das diferentes instituições de ensino superior e instituições e unidades de pesquisa. Por essa razão, o SEP não é o único protocolo de avaliação em vigor no país, existindo protocolos específicos para as universidades de pesquisa, para as universidades de ciências aplicadas e para os institutos de pesquisa acadêmica e não acadêmica (Drooge *et al.*, 2013). Cabe dizer que o sistema de educação superior holandês é formado por 14 universidades de pesquisa (*Universiteiten-WO*); quatro pequenas universidades especiais; dezenas de universidades de ciências aplicadas (*Hogescholen-HBO*); e uma universidade aberta (*Open Universiteit*) (Drooge, 2021a; VSNU, 2022).

Um detalhe relevante dos protocolos de avaliação na Holanda é que não existe separação entre pesquisa e pós-graduação. Os programas de doutorado são considerados elementos integrantes dos programas de pesquisa e são avaliados como tal. Nesse sentido, o *Strategy Evaluation Protocol* estabelece diretrizes para a avaliação da qualidade da pesquisa, sem que sejam necessários marcos regulatórios específicos para a avaliação da pós-graduação.

Finalmente, cabe dizer que os cursos de mestrado são abordados de forma diferente do modelo corrente no Brasil, pois na Holanda, assim como a maior parte dos países europeus, tais cursos integram a estrutura do processo formativo sequencial dos cursos de graduação, um resultado do Processo de Bolonha e da aprovação do *European Credit Transfer System*.

PRINCÍPIOS E OBJETIVOS DA AVALIAÇÃO

A organização do sistema descrita anteriormente reflete os princípios e os objetivos gerais que vêm orientando as políticas de avaliação desde a publicação *Hoger Onderwijs in Wetenschappelijk Onderzoek* (OCW, 1992) e da elaboração do primeiro protocolo em 1994. Os marcos regulatórios elaborados ao longo das décadas definiram e consolidaram os princípios que modelaram a organização e o funcionamento do sistema. A avaliação foi concebida de acordo com alguns pressupostos fundamentais, como autonomia universitária, avaliação externa (por meio do sistema *peer review*), AA, *accountability*, melhoria da qualidade e separação entre financiamento e avaliação.

Com base nesses princípios, os protocolos consolidaram um claro distanciamento em relação às concepções performáticas de avaliação e às abordagens que concebem e realizam os processos avaliativos como instrumentos de aferição de desempenho, controle de resultados, prestação de contas, produção de *rankings* e elaboração de indicadores para a distribuição de recompensas e premiações, especialmente bolsas e recursos financeiros.

Os princípios anteriormente referidos podem ser observados no Quadro 1. Eles fundamentam a grade dos objetivos estabelecidos pelos protocolos ao longo das décadas. Os objetivos refletem, em linhas gerais, os compromissos do sistema com a melhoria da qualidade, com a *accountability* e com o fortalecimento da relevância societal da pesquisa (Drooge, Jong e Smit, 2022).

Quadro 1 – Objetivos gerais da avaliação estabelecidos pelos protocolos a partir de 2003.

| Protocolos | Objetivos |
|---------------|--|
| SEP 2003–2009 | “O sistema de avaliação visa a três objetivos no que tange à qualidade da pesquisa e à gestão: (i) melhorar a qualidade da investigação por meio de um processo de avaliação realizado de acordo com os padrões internacionais de qualidade e relevância; (ii) melhorar a gestão e a liderança da pesquisa; (iii) prestar contas para as organizações de pesquisa, agências de fomento, governo e sociedade em geral” (VSNU, KNAW e NWO, 2003, p. 5). |
| SEP 2009–2015 | “O SEP 2009–2015 visa a dois objetivos no que diz respeito à avaliação da pesquisa (incluindo a formação de doutores) e à gestão: (i) melhorar a qualidade da pesquisa por meio da avaliação realizada pelos pares (<i>peer review</i>), incluindo a relevância acadêmica e social da pesquisa, políticas de pesquisa e gestão; (ii) prestar contas para as organizações de pesquisa, agências de fomento, governo e sociedade em geral” (VSNU, KNAW e NWO, 2009, p. 4). |
| SEP 2015–2021 | “O principal objetivo da avaliação é revelar e confirmar a qualidade e a relevância da pesquisa para a sociedade e aprimorá-la quando necessário” (VSNU, KNAW e NWO, 2016, p. 5). |
| SEP 2021–2027 | “O principal objetivo do SEP é avaliar a unidade de pesquisa com base em seus próprios objetivos e estratégias e [...] manter e aprimorar a qualidade e a relevância societal da pesquisa, assim como facilitar contínuos diálogos sobre a qualidade, relevância societal e viabilidade no contexto da garantia da qualidade da pesquisa” (VSNU, KNAW e NWO, 2020, p. 6). |

Fonte: VSNU, KNAW e NWO (2003; 2009; 2016; 2020).

O Quadro 1 revela — além das similaridades e linhas de continuidade entre os objetivos — algumas mudanças qualitativas. O SEP 2021–2027, particularmente, reforçou a AA. O núcleo central da avaliação é composto, neste sentido, dos objetivos, metas e estratégias que as unidades avaliadas⁸ estabelecem para si mesmas ao longo do período de seis anos. A avaliação não é um exercício de fora para dentro. Ela é contextual e tem uma dimensão claramente formativa. Como destaca Drooge (2021a; b; c; d), ela é parte de um contínuo processo de garantia e de aprimoramento da qualidade. A avaliação promove um permanente diálogo entre a unidade avaliada e os diferentes conselhos superiores, especialmente sobre os pontos positivos, as fragilidades e as melhorias que precisam ser implementadas.

DIMENSÕES E CRITÉRIOS DE AVALIAÇÃO

Os critérios de avaliação, assim como os princípios e os objetivos descritos anteriormente, refletem a evolução e as transformações das políticas de educação superior, ciência e tecnologia do país. As mudanças traduzem as novas e contínuas demandas colocadas pela sociedade, pela comunidade científica e pelo conjunto das instituições. Os critérios de avaliação expressam, nesse sentido, o que a sociedade espera da ciência, dos pesquisadores e de suas instituições acadêmicas e científicas (Spaapen e Drooge, 2011).

O Quadro 2 mostra algumas mudanças incrementais introduzidas no período. A primeira delas foi a redução do número de critérios de avaliação (de quatro para três). Em 2015, o item “produtividade” foi excluído. De acordo com Drooge, Jong e Smit (2022), o SEP 2015–2021 incorporou uma proposição que vinha sendo formulada há anos pela comunidade científica. O critério “relevância” também foi aprimorado ao longo das décadas. A partir do SEP 2009–2015, o item passou a se chamar “relevância societal”. Essa alteração ampliou as categorias que integram e orientam a análise dos impactos da pesquisa na sociedade. Além das dimensões científicas e socioeconômicas, a avaliação deve considerar outras modalidades de impactos, especialmente as de natureza cultural, social e política (VSNU, KNAW e NWO, 2009). Em 2015, o SEP 2015–2021 introduziu novos elementos. Ele estabeleceu que os comitês deveriam avaliar a qualidade e a relevância da pesquisa de acordo com as áreas de atuação e os públicos-alvo de cada unidade. A avaliação devia considerar as contribuições da pesquisa para o conjunto da sociedade, especialmente as contribuições para o desenvolvimento econômico, inovação, educação, cultura, saúde, sustentabilidade etc. (VSNU, KNAW e NWO, 2009; Drooge *et al.*, 2013).

8 Os protocolos utilizam o termo “unidade” para designar os institutos de pesquisa, os núcleos de pesquisa, os grupos de pesquisa, os centros multi e interdisciplinares de pesquisa etc. São avaliadas as unidades que atendem às seguintes condições: ser reconhecida, interna e externamente, como uma entidade de pesquisa; ter objetivos, metas e estratégias claramente definidos e compartilhados pelo grupo; ter, ao menos, o equivalente a dez pesquisadores em seu *staff* acadêmico permanente, sem contar os doutorandos e os pós-doutorandos (nos Países Baixos se utiliza o conceito de *full-time equivalent* — FTE, de forma que os dez FTE exigidos podem ser cobertos por um número maior de pesquisadores); e ter, ao menos, três anos de funcionamento (VSNU, KNAW e NWO, 2020).

Quadro 2 – Dimensões e critérios da avaliação estabelecidos pelos protocolos a partir de 2003.

| Protocolos | Crítérios | Principais aspectos |
|----------------------|-----------------------------|---|
| SEP 2003–2009 | Qualidade | - Reconhecimento internacional e potencial inovador |
| | Produtividade | - Produção científica |
| | Relevância | - Impacto científico e socioeconômico |
| | Vitalidade e viabilidade | - Flexibilidade, gestão e liderança |
| SEP 2009–2015 | Qualidade | - Qualidade e relevância científica - Liderança - Reputação acadêmica - Recursos humanos e financeiros - Formação de doutores |
| | Produtividade | - Estratégia de produtividade - Produtividade |
| | Relevância | - Relevância societal |
| | Vitalidade e viabilidade | - Estratégia - Análise SWOT - Robustez e estabilidade |
| SEP 2015–2021 | Qualidade da pesquisa | - Qualidade e relevância científica - Resultados: publicações, produtos, infraestrutura e outras contribuições para a ciência - Política de formação de doutores - Integridade científica - Diversidade |
| | Relevância para a sociedade | - Impacto científico, econômico, social ou cultural - Qualidade, escala e relevância das contribuições para grupos específicos (setores econômicos, sociais, culturais etc.); assessoria para a elaboração de políticas públicas, relatórios e outros documentos; contribuições para o debate público etc. |
| | Viabilidade | - Estratégia - Governança - Competências e liderança na gestão da unidade de pesquisa |
| SEP 2021–2027 | Qualidade da pesquisa | - Qualidade e contribuições para o desenvolvimento científico (internacional, nacional ou regional) - Reputação e liderança - Ciência aberta - Política de formação de doutores - Cultura acadêmica - Integridade científica - Política de recursos humanos: diversidade e gestão de talentos |
| | Relevância para a sociedade | - Impacto e engajamento público da unidade em áreas relacionadas à economia, desenvolvimento social, cultura, educação e outras que possam ser relevantes |
| | Viabilidade | - Viabilidade e pertinência dos objetivos (até que ponto eles permanecem científica e socialmente relevantes) - Adequação da gestão e dos recursos para atingir os objetivos - Viabilidade da unidade em relação aos objetivos previstos - Liderança |

SWOT: *Strengths, Weaknesses, Opportunities and Threats.*

Fonte: VSNU, KNAW e NWO (2003; 2009; 2016; 2020).

Os critérios contemplam tanto os aspectos quantitativos quanto os qualitativos. Além disso, eles não devem ser apropriados de forma rígida. As instituições, os conselhos e os comitês têm autonomia para fazer ajustes e adequações de acordo com as características e os objetivos de cada unidade avaliada. De acordo com o SEP 2021–2027 (VSNU, KNAW e NWO, 2020), os critérios são instrumentos flexíveis; eles devem ser interpretados com base nas metas, estratégias e políticas de AA de cada unidade.

PRINCÍPIOS ORIENTADORES DA AUTOAVALIAÇÃO

Como destacado anteriormente, a AA é a espinha dorsal do sistema de avaliação holandês. Ela é uma característica marcante, que confere especificidades ao modelo e o distingue em termos internacionais. A AA perpassa pelos princípios, objetivos, critérios e *modus operandi* do processo avaliativo, articulando-os. Pode-se afirmar, nesse sentido, que a AA cumpre um duplo papel. Além de ser um princípio, ela é também um método. Além de definir os fins, ela orienta a organização e os procedimentos da avaliação.

Com base nos princípios da autonomia, planejamento institucional e *accountability*, os diferentes protocolos institucionalizaram a avaliação como prática de autogestão, realizada de dentro para fora e de baixo para cima. As unidades avaliadas são, simultaneamente, “sujeitos” e “objetos” do processo. A avaliação externa não se sobrepõe à avaliação interna. Como destaca o SEP 2021–2027 (VSNU, KNAW e NWO, 2020), o núcleo central da avaliação são os objetivos, as metas e as estratégias que a unidade estabelece para si mesma a fim de manter e aprimorar a qualidade, a relevância societal e a viabilidade. As metas e as estratégias são elementos centrais da AA, permitindo que a unidade analise os resultados obtidos conforme os objetivos previamente estabelecidos e avalie o desempenho à luz da missão e dos propósitos que estabeleceu para si mesma. De acordo com o SEP 2021–2027, a AA possibilita uma reflexão “[...] sobre as estratégias que a unidade adotou, assim como os efeitos que elas produziram” (VSNU, KNAW e NWO, 2020, p. 19).

A AA é, nesse sentido, um processo estreitamente vinculado ao planejamento e à gestão da unidade. Como destaca Drooge (2021a; b), ela é contextual e formativa. Além de levar em consideração os estágios de desenvolvimento das unidades, a AA permite identificar os avanços, as potencialidades, as fragilidades e as ameaças. Nas palavras de Drooge (2021b, p. 2), ela “[...] não está voltada para a pesquisa em si, mas para a estratégia da unidade em relação à pesquisa”. Por essa razão, o SEP 2021–2027 (VSNU, KNAW e NWO, 2020, p. 19) afirma que a AA começa no momento que a unidade “[...] explicita a sua estratégia e os seus objetivos”.

Como destacado anteriormente, a unidade é o sujeito e o objeto; ela é tanto o ponto de partida quanto o ponto de chegada do processo avaliativo. Os resultados finais da avaliação regressam à unidade — na forma de relatório — a fim de orientar o planejamento e as ações estratégicas do próximo ciclo. AA, dessa forma, é um fim e um meio. Em vez de ser uma etapa específica, isolada e independente, ela é a dimensão estruturante em torno da qual as diferentes etapas se articulam e se retroalimentam.

A AUTOAVALIAÇÃO COMO PROCESSO: A EXPERIÊNCIA DO CENTER FOR SCIENCE AND TECHNOLOGY STUDIES

De forma a apresentar os principais procedimentos relacionados ao modelo holandês de avaliação da pesquisa e da pós-graduação, apresentaremos a seguir um relato da experiência de AA realizada pelo *Center for Science and Technology Studies* (CWTS) da Faculdade de Ciências Sociais e Comportamentais da Universidade de Leiden.

A experiência ora relatada é especialmente relevante visto que o CWTS é um renomado centro de investigação nas áreas de avaliação de pesquisa e cienciométrica. Como consequência, a realização da AA criou oportunidades para que os pesquisadores do CWTS pudessem explorar novos métodos, técnicas e abordagens diferenciadas. Tais esforços de construção de um processo robusto de avaliação baseado na combinação de métodos quantitativos e qualitativos foram relatados pelo grupo de trabalho que apoiou a diretoria do centro na produção do relatório e demais atividades relacionadas (CWTS, 2022), em parte também compondo o relatório final de avaliação finalizado em 2022 (*ibidem*).

Os tópicos apresentados a seguir sintetizam as principais etapas do processo de AA realizado pelo CWTS entre 2021 e 2022, tendo sido construídos com base nos relatos mencionados, assim como na análise de procedimentos definidos no *Strategy Evaluation Protocol 2021–2027* (VSNU, KNAW e NWO, 2020).

DEFINIÇÃO DA ESTRATÉGIA, DOS OBJETIVOS E DAS METAS DA UNIDADE

Este é o ponto de partida. A estratégia e os objetivos de cada unidade de pesquisa são formulados com base em um conjunto de reuniões periódicas realizadas entre as unidades e o conselho universitário, que desempenha papel similar ao das pró-reitorias de pesquisa e pós-graduação no Brasil. A definição dos objetivos é parte de uma política institucional de garantia da qualidade, que se materializa pautada por um conjunto de estratégias e metas que cada unidade deve perseguir ao longo do ciclo avaliativo (VSNU, KNAW e NWO, 2020).

Conforme descrito na seção sobre o desenho institucional do sistema, o ciclo avaliativo é de seis anos, e o protocolo vigente corresponde ao período 2021–2027. Cada unidade de pesquisa conduz seu processo de avaliação no seu próprio tempo, tendo como base os seis anos anteriores. O CWTS foi uma das primeiras unidades do país a realizar a avaliação com base nas diretrizes estabelecidas pelo SEP 2021–2027.

O modelo holandês é bastante diferente do adotado no Brasil, onde o ciclo avaliativo é de quatro anos, com todos os programas de pós-graduação sendo avaliados simultaneamente no mesmo período. Por exemplo, a Avaliação Quadrienal de 2017 avaliou o desempenho dos PPG referente a 2013–2016, e a Quadrienal realizada em 2022 avaliou o período 2017–2020. Parte do argumento para a avaliação síncrona no Brasil vem da abordagem comparativa adotada no sistema, com grande impacto na distribuição do fomento no país (Brasil, Trevisol e Drooge, 2022).

ELABORAÇÃO DO TERMO DE REFERÊNCIA E PLANO DE AÇÃO DA AVALIAÇÃO

Apoiados nas diretrizes estabelecidas pelo SEP, o conselho universitário e a própria unidade de pesquisa elaboram juntos um Termo de Referência (TR), finalizado cerca de um ano antes da visita do comitê de avaliação. O TR apresenta, entre outros aspectos,

- as informações básicas sobre a unidade a ser avaliada;
- a estratégia, os objetivos e as principais características da unidade;
- os propósitos e os critérios da avaliação;
- os principais aspectos que a comissão externa de avaliação deve considerar; e
- a natureza pública do relatório final (VSNU, KNAW e NWO, 2020).

Ainda que os TR sejam documentos simples, com um *template* incluído entre os anexos do SEP (VSNU, KNAW e NWO, 2020), eles são vitais para permitir que uma unidade de pesquisa seja avaliada de acordo com sua missão e objetivos, desempenhando assim um papel essencial para uma avaliação realmente multidimensional.

Conforme descrito em Brasil (2022), a avaliação multidimensional é aquela em que a missão ou vocação de um programa de pesquisa pode ser valorizada com base em uma análise não reducionista, de forma distinta do processo de avaliação tradicionalmente adotado no Brasil. Nas avaliações realizadas pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) ou pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), programas de pós-graduação ou instituições de ensino superior recebem uma nota consoante ao resultado das suas avaliações. Tais notas reduzem múltiplas dimensões de análise a um único indicador, que é incapaz de demonstrar como cada avaliado desempenha em termos de seus diferentes impactos: econômico, social, científico etc.

No modelo holandês de avaliação, as unidades de pesquisa podem focar sua análise naquelas dimensões que se alinham com seus objetivos. Por exemplo, o CWTS — ainda que conte com um programa de doutorado — não tem a formação de pessoal entre suas prioridades e não oferece cursos de graduação ou de mestrado. Por outro lado, o instituto foca em pesquisa (com a maior parte do seu *staff* acadêmico composto de pesquisadores e não professores), transdisciplinaridade (mantendo fortes laços com agências de fomento e gestores de políticas científicas) e impacto societal (até mesmo por intermédio de uma empresa própria, prestando serviços diversos relacionados à bibliometria, cienciométrica e consultoria em políticas científicas). Em uma avaliação reducionista, as escolhas feitas pelo CWTS poderiam prejudicar seus resultados.

Uma vez que o TR é validado, refletindo as missões e objetivos da unidade que devem ter centralidade na avaliação, uma série de etapas ocorrem paralelamente de forma a informar a construção de um plano de ação que é produzido pela unidade de pesquisa. Tal plano precisa incluir elementos primordiais, como a indicação dos membros que devem integrar a comissão externa de avaliação, o projeto das atividades e as principais datas do processo, planejamento de análises bibliométricas e cienciométricas, intenção de conduzir análises comparativas com outras unidades e orçamento (contratação de um secretário para apoiar os trabalhos da comissão

externa, custos de viagem e hospedagem, remuneração da comissão etc.). O documento produzido deve ser submetido ao conselho universitário para aprovação.

No início de 2021, o CWTS nomeou um grupo de trabalho com a responsabilidade de auxiliar o corpo de diretores do centro ao longo do processo de avaliação, incluindo a elaboração do plano de ação. Com base em um planejamento detalhado de todas as etapas da avaliação, as seguintes decisões integraram o documento produzido pelo centro, denominado *CWTS self-evaluation report* (CWTS, 2022):

1. Em virtude do perfil singular do CWTS nos Países Baixos, o planejamento não incluiu uma avaliação comparativa, e o desempenho do centro seria analisado em relação ao ciclo anterior.
2. Alguns indicadores tradicionais em processo de avaliação não são bem-vistos (como o índice H). Alguns outros, como o fator de impacto, são proibidos (VSNU, KNAW e NWO, 2020). Contudo, as unidades de pesquisa podem selecionar conjuntos de indicadores e outras evidências para fundamentar sua AA. Assim, o CWTS projetou as análises desejadas e selecionou as bases de dados de preferência (Dimensions, ORCID, Altmetric), justificando sua decisão no plano de ação.
3. O nome de um secretário para a comissão de avaliação foi indicado, com uma proposta de trabalho anexada ao plano. Cabe destacar que a atividade desse profissional vai muito além da expectativa de atividades secretariais, e o contratado pelo CWTS foi, também, membro da comissão responsável pela elaboração do SEP corrente.
4. Uma comissão externa de avaliação realizou uma visita *in loco* em maio de 2022.

Paralelamente às atividades de preparação do plano de ação, foram implementadas as atividades para a definição dos membros da comissão de avaliação, conforme descrito a seguir.

COMPOSIÇÃO DO COMITÊ EXTERNO DE AVALIAÇÃO

A escolha dos avaliadores é uma fase crucial do ciclo de avaliação. Cada instituição define os procedimentos para a composição do comitê, considerando uma diversidade de competências acadêmicas e científicas. A comissão deve ser formada por especialistas imparciais, com composição internacional e incluindo um doutorando, um pesquisador em início e/ou no meio de carreira e, quando necessário, um especialista não acadêmico. Cabe ao comitê avaliar a qualidade, a relevância para a sociedade e a viabilidade da pesquisa desenvolvida pela unidade nos últimos seis anos. A unidade deve ser avaliada em relação à estratégia e às metas que estabeleceu para si mesma, levando em consideração o contexto internacional, nacional e, quando necessário, regional (VSNU, KNAW e NWO, 2020).

Diferentemente do modelo brasileiro — em que as comissões de avaliação são formadas conforme a indicação dos coordenadores das áreas da CAPES —, no modelo holandês cada departamento seleciona a sua própria comissão pautado por critérios definidos no SEP. No CWTS, esse processo de indicação foi participativo, seguindo as etapas listadas a seguir:

1. A diretoria do centro e o GT da AA organizaram uma reunião com toda a equipe de trabalho, apresentando o processo de avaliação e os requisitos necessários para a seleção dos integrantes da comissão.
2. Uma consulta aberta foi iniciada, e toda a equipe de trabalho do centro foi convidada a encaminhar sugestões para compor a comissão, resultando em 55 nomes indicados.
3. A direção conduziu uma análise das indicações e dos seus currículos. Doze especialistas foram selecionados, sendo seis titulares e seis suplentes.
4. A lista foi discutida com todo o centro para validação e ajustes adicionais (por exemplo, um dos integrantes precisou ser alterado, visto que acabara de se envolver em um projeto de pesquisa em colaboração com o CWTS).

Os nomes indicados foram incluídos no plano de ação e submetidos ao conselho universitário para aprovação. O documento também propôs a indicação de um membro para presidir a comissão.

ELABORAÇÃO DO RELATÓRIO DE AUTOAVALIAÇÃO

O relatório de AA é um elemento central do processo avaliativo, sendo redigido pela unidade de pesquisa. Além de conciso e bem fundamentado, ele deve contemplar

- uma descrição das principais características da unidade;
- a missão, a estratégia e os objetivos da unidade (tanto os atuais quanto os futuros);
- evidências, indicadores quantitativos e qualitativos, estudos de caso, *benchmarking* e outras informações relevantes para evidenciar os principais resultados e conquistas;
- os pontos fortes e as fragilidades; e, por fim,
- as ações estratégicas que a unidade pretende implementar ao longo do ciclo avaliativo seguinte (VSNU, KNAW e NWO, 2020).

O processo de produção do relatório do CWTS contou com o engajamento de todo o centro, sendo conduzido de acordo com as etapas descritas a seguir:

1. **Consulta.** No início de 2021, um formulário de pesquisa foi encaminhado para toda a equipe do CWTS, incluindo um conjunto de questões agrupadas em dois grupos: um sobre as atividades do respondente, pesquisas realizadas, temas de interesse, participação em projetos; e outro sobre a percepção com relação à missão, objetivos e motivações do centro, incluindo perspectivas ou necessidade de ajustes para o novo ciclo.
2. **Definição dos temas do relatório.** As unidades têm autonomia para definir a estrutura e o conteúdo do relatório. Dessa forma, um dos resultados da análise das contribuições da equipe foi a estruturação do documento em temas de interesse, em vez de baseado na estrutura organizacional ou de pesquisa do centro. Dessa forma, os grupos de pesquisa e eixos temáticos do CWTS foram reorganizados para melhor representar a identidade do centro.

3. **Grupos temáticos.** Sob a coordenação dos membros do GT, grupos *ad hoc* foram criados com o propósito de discutir os temas centrais a serem contemplados no relatório de AA. Os participantes foram selecionados para representar a *expertise* do centro e os diferentes perfis profissionais (incluindo professores, pesquisadores, doutorandos e corpo técnico). Cada grupo produziu documentos que auxiliaram a diretoria a construir o relatório final.
4. **Análises bibliométricas e cienciométricas.** As unidades de pesquisa podem contratar serviços de externos para a realização de estudos bibliométricos e cienciométricos. No caso do CWTS, por se tratar de área de conhecimento do próprio centro, tais análises foram desenvolvidas internamente. No processo, os conjuntos de indicadores foram selecionados e a equipe desenvolveu uma série de *dashboards* e *networks* em atendimento às diferentes categorias de evidências requeridas pelo SEP, de acordo com o apresentado no Quadro 3.

Quadro 3 – Categorias de evidências para os domínios de qualidade da pesquisa e relevância para a sociedade.

| | | Domínios de qualidade | |
|------------------------|------------------------------|--|---|
| | | Qualidade da pesquisa | Relevância para a sociedade |
| Dimensões de Avaliação | Produção científica | Produção científica de caráter acadêmico | Produção científica técnica/tecnológica |
| | Impacto da produção | Medidas de impacto acadêmico | Medidas de impacto na sociedade |
| | Evidências de reconhecimento | Evidências de reconhecimento acadêmico | Evidências de reconhecimento pela sociedade |

Fonte: Adaptado pelos autores com base nas dimensões do Standard Evaluation Protocol (VSNU, KNAW e NWO, 2020, p. 34).

5. **Análise SWOT.** O processo de análise de forças, fraquezas, oportunidades e ameaças do CWTS foi conduzido com base nas respostas da equipe do centro ao formulário de consulta, em combinação com atividades dedicadas em reuniões internas e com *stakeholders* selecionados.
6. **Anexos.** As evidências e as informações complementares podem ser inseridas nos anexos do relatório de AA. No relatório de AA do CWTS foram incluídas informações organizacionais sobre o centro, indicadores selecionados com descrições metodológicas correspondentes, *dashboards* com redes e visualizações, programa de pesquisa, dados financeiros, informações sobre doutorandos etc.

Com base no que estabelece o mais recente protocolo, correspondente ao período 2021–2027, é possível afirmar que o relatório de AA cumpre quatro objetivos fundamentais. Ele desafia as unidades a:

- analisar a adequação e a pertinência das estratégias e das metas estabelecidas pela unidade nos últimos seis anos;
- explicitar claramente as estratégias e os objetivos que orientarão as ações e as atividades da unidade no próximo ciclo avaliativo;
- apresentar as principais conquistas, fragilidades e desafios em cada um dos critérios e subcritérios de avaliação estabelecidos pelo SEP e pela unidade; e
- evidenciar os principais resultados nos últimos seis anos em cada um dos critérios e subcritérios da avaliação (VSNU, KNAW e NWO, 2020).

Cerca de dois meses antes da data prevista para a visita *in loco*, o relatório de AA do CWTS foi submetido à aprovação pelo conselho universitário e, em seguida, remetido ao comitê externo de avaliação (VSNU, KNAW e NWO, 2020).

VISITA *IN LOCO* E RELATÓRIO DO COMITÊ EXTERNO

Após o envio do relatório final para a comissão externa, deu-se início ao processo de preparação da visita *in loco*. Tal visita é organizada pela unidade avaliada, em contato direto com os membros da comissão e apoio do secretário contratado para as atividades avaliativas. Por meio da visita é possível conhecer a infraestrutura da unidade, realizar entrevistas com os diretores, líderes de pesquisa e pessoas consideradas relevantes. Durante a visita *in loco*, a comissão pode solicitar informações não contempladas pelo relatório de AA (VSNU, KNAW e NWO, 2020).

A visita da comissão de avaliação do CWTS foi organizada com uma programação de três dias. O primeiro deles foi dedicado ao trabalho interno da comissão. Essa atividade foi conduzida sem interferência ou participação de representantes do centro, sendo coordenada pelo secretário e pelo presidente da comissão. Entre as atividades iniciais realizadas constou uma breve apresentação do *Strategy Evaluation Protocol*, conduzida pelo secretário do grupo, ressaltando seu papel diferenciado no processo. A apresentação faz-se necessária em virtude do perfil internacional da comissão, uma vez que parte dos avaliadores não tem familiaridade com o modelo de avaliação holandês.

O segundo dia de trabalho incluiu atividades como entrevistas com diversos membros da equipe e discussão dos resultados, indicadores e outras evidências apresentadas no relatório de AA. Análises das recomendações feitas pela comissão avaliadora em 2015, resultados derivados e alterações nos perfis de forças, fraquezas, oportunidades e ameaças (SWOT) também foram centrais no trabalho. No terceiro dia, a comissão reuniu-se com os integrantes do centro para apresentar suas impressões e recomendações iniciais.

Após a visita *in loco*, o comitê submeteu ao CWTS um relatório de avaliação. O protocolo de avaliação estabeleceu que o relatório tem como objetivo comunicar as principais conclusões, propostas e recomendações do comitê, devendo ser claro e consistente. O protocolo estabeleceu o prazo de 30 dias para a produção do relatório, contados a partir da conclusão da visita, sendo ele inicialmente submetido à unidade avaliada e posteriormente à direção da instituição (VSNU, KNAW e NWO, 2020).

RELATÓRIO FINAL, DOCUMENTO DE RESPOSTA E ACOMPANHAMENTO DOS RESULTADOS

Cabe à direção da instituição emitir um documento de resposta em relação ao relatório da comissão externa, ainda que a tarefa seja, na prática, conduzida pela própria unidade de pesquisa avaliada. Tal documento pode esclarecer dúvidas explicitadas no relatório da comissão, apresentar argumentos adicionais e, até mesmo, contestar ou rejeitar recomendações recebidas. Tanto o documento de resposta quanto o relatório da avaliação devem ser publicados no *site* da instituição (VSNU, KNAW e NWO, 2020).

No atual *Strategy Evaluation Protocol*, o resultado da avaliação é formativo e cabe à própria instituição refletir sobre ações necessárias para recompensar o bom desempenho, mitigar os riscos da unidade de pesquisa ou tomar medidas mais drásticas caso os resultados tenham revelado qualidade insuficiente. Não há atribuição de nota nem impacto direto no financiamento da unidade, seja ele proveniente da instituição, seja de outras agências de fomento. Assim, o processo de avaliação visa, essencialmente, aprofundar a cultura da avaliação, da transparência e do planejamento estratégico das instituições e das unidades (VSNU, KNAW e NWO, 2020).

Parte do processo formativo materializa-se logo após a conclusão da avaliação externa, por meio da discussão dos resultados entre os dirigentes, a unidade de pesquisa e a equipe de trabalho. O acompanhamento é realizado de acordo com as políticas de garantia de qualidade de cada instituição, que pode, se necessário para monitorar o desempenho das unidades, recomendar a realização da avaliação de meio-termo (VSNU, KNAW e NWO, 2020).

No caso do CWTS, a parte final do processo de avaliação coincidiu com o início da construção de uma agenda de conhecimento — um programa de pesquisa expandido — vigente a partir do início de 2023. A referida agenda foi elaborada com recurso às lições aprendidas ao longo do processo avaliativo, incluindo as recomendações e discussões com a comissão externa e com a própria Universidade de Leiden. Adicionalmente, a nova agenda visa contemplar os objetivos e as estratégias prioritárias para o período 2023–2028, sendo construída colaborativamente por intermédio de reuniões de grupos temáticos, realização de seminários internos e da organização de dois retiros de pesquisa realizados em abril e setembro de 2022.

CONSIDERAÇÕES FINAIS

Os Países Baixos foram um dos primeiros países europeus a implementar um sistema formal de avaliação da qualidade da pesquisa e da pós-graduação. As políticas e as mudanças implementadas ao longo das últimas quatro décadas consolidaram um sistema *sui generis*, duradouro e estável. Assim como todos os sistemas nacionais de avaliação, o modelo holandês foi concebido e aprimorado com base nas características e necessidades do país. É parte, portanto, de um amplo e complexo processo político, cultural, social e acadêmico. Além de refletir as concepções de avaliação, ele expressa também as diretrizes que o país estabeleceu para a organização da educação superior e do sistema nacional de ciência e tecnologia. Os Países Baixos — diferentemente de outras nações que implementaram sistemas de avaliação centralizados, padronizados, quantitativos e baseados em desempe-

nho — consolidaram um modelo descentrado e heterogêneo, fundamentado no princípio de que a avaliação deve ser realizada tendo em vista a missão, as metas e as estratégias que as unidades avaliadas estabelecem para si mesmas. A avaliação é, nesse sentido, uma prática interna, formativa, contextual e participativa.

O sistema holandês tem sido referência e inspiração para inúmeras instituições e países, inclusive o Brasil. Os pesquisadores e as agências de avaliação têm estudado o modelo a fim de compreender as suas especificidades e identificar possíveis contribuições para o aprimoramento dos sistemas nacionais de avaliação. Como esse propósito também motivou a realização do presente estudo, julgamos importante concluir este artigo destacando alguns aspectos que podem servir de subsídios para a melhoria do sistema de avaliação da pós-graduação no Brasil.

O primeiro aspecto diz respeito ao protagonismo exercido pelas instituições e unidades avaliadas. A avaliação é, acima de tudo, uma prática institucionalizada e permanente de autoconhecimento e de autogestão. É uma dimensão estreitamente vinculada ao exercício da autonomia institucional e do planejamento estratégico. Não é um instrumento de regulação estatal, de controle externo e de produção de *rankings*, indicadores de desempenho e comparações. Não há uma instituição nacional responsável pela organização, coordenação e execução do processo em âmbito nacional. Os protocolos nacionais de avaliação têm, invariavelmente, atribuído às instituições a competência e a responsabilidade pela avaliação. Essa confiança tem sido reafirmada e fortalecida ao longo das décadas.

Em segundo lugar, o sistema holandês está baseado no princípio de que a avaliação deve priorizar a missão, as metas e as estratégias que a unidade avaliada estabelece para si mesma. Os objetivos da avaliação não podem estar desassociados do planejamento estratégico. Em lugar de aferir se a unidade atingiu integral ou parcialmente os padrões de desempenho estabelecidos nacionalmente, o processo avaliativo procura analisar os resultados à luz das metas e das estratégias previamente estabelecidas. As unidades são, nesse sentido, tanto o ponto de partida quando o ponto de chegada. Por meio da AA elas identificam os pontos fortes, as fragilidades e os aprimoramentos que precisam ser implementados.

A avaliação procura, em terceiro lugar, respeitar e fortalecer a identidade de cada unidade por intermédio de uma abordagem multidimensional. As unidades têm liberdade para definir as dimensões que desejam considerar na sua própria avaliação, adotando critérios e indicadores pensados para embasar cada narrativa individual sem, por conseguinte, gerar resultados padronizados para todo o sistema. Ao invés de promover a homogeneização, o processo preserva a diversidade e a diferenciação, permitindo até que comparações entre unidades sejam opcionais.

O quarto aspecto importante diz respeito à periodicidade da avaliação. Desde 2003, os ciclos têm sido de seis anos. Os períodos mais longos favorecem a elaboração e o acompanhamento do planejamento estratégico das unidades.

Embora não seja perfeito, o modelo holandês procura, por fim, integrar dimensões bastante complexas e de difícil equação. Por meio da AA o sistema procura promover equilíbrio e complementar entre as dimensões qualitativas e quantitativas; os processos internos e externos de avaliação; as dimensões formativas e somativas; a autonomia institucional e *accountability*; e os processos de homogeneização e de diferenciação.

REFERÊNCIAS

- 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. Disponível em: <https://bit.ly/3gRMNuo>. Acesso em: 10 mar. 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., n. d.]. Disponível em: <https://bit.ly/3rHmCJT>. Acesso em: 16 mar. 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. Disponível em: <https://sep.cwts.nl>. Acesso em: 15 mar. 2022.
- DROOGE, L.; JONG, S.; FABER, M.; WETERHEIJDEN. **Twenty years of research evaluation: Facts & Figures**. The Hague: Rathenau Instituut, 2013. Disponível em: <http://bit.ly/3cXp4mK>. Acesso em: 10 dez. 2021.
- DROOGE, L. Research evaluation in context 1: Introducing research evaluation in the Netherlands. **Leiden Madtrics**, 12 fev. 2021a. Disponível em: <https://bit.ly/3IsQZJE>. Acesso em: 7 jan. 2022.
- DROOGE, L. Research evaluation in context 2: One joint protocol, three criteria and four aspects. **Leiden Madtrics**, 17 fev. 2021b. Disponível em: <https://bit.ly/3Atb67G>. Acesso em: 12 jan. 2022.
- DROOGE, L. Research evaluation in context 3. The board, the research unit and the assessment committee. **Leiden Madtrics**, 4 mar. 2021c. Disponível em: <https://bit.ly/3fMAoE4>. Acesso em: 12 jan. 2022.
- DROOGE, L. Research evaluation in context 4. The practice of research evaluation. **Leiden Madtrics**, 12 maio 2021d. Disponível em: <https://bit.ly/3IqzLfL>. Acesso em: 13 jan. 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. Disponível em: <https://bit.ly/3DmzOYY>. Acesso em: 14 mar. 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. Disponível em: <https://bit.ly/3GXeid7>. Acesso em: 30 dez. 2021.

MINISTERIE VAN ONDERWIJS, CULTUUR EN WETENSCHAP — OCW. **Hoger onderwijs en wetenschappelijk onderzoek.** Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap, 1992. Disponível em: <https://bit.ly/33YmT1l>. Acesso em: 22 dez. 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. Disponível em: <https://bit.ly/3SR11sI>. Acesso em: 5 jan. 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. Disponível em: <https://bit.ly/3I5MR2m>. Acesso em: 26 nov. 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. Disponível em: <https://bit.ly/2C9herj>. Acesso em: 25 nov. 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. Disponível em: <https://www.universiteitenvannederland.nl/>. Acesso em: 21 jan. 2022.
- VERENIGING VAN UNIVERSITEITEN – VSNU; KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN — KNAW; NEDERLANDSE ORGANISATIE VOOR WETENSCHAPPELIJK ONDERZOEK — NWO. **Standard Evaluation Protocol (2003-2009)**. [Report]. The Hague, 2003. Disponível em: <https://bit.ly/3rbJtzR>. Acesso em: 16 dez. 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. Disponível em: <https://bit.ly/3t0B02X>. Acesso em: 27 dez. 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. Disponível em: <https://bit.ly/45TIpzr>. Acesso em: 16 jan. 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. Disponível em: <https://bit.ly/3wAFbzi>. Acesso em: 16 jan. 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.

SOBRE OS AUTORES

JOVILES VITÓRIO TREVISOL é doutor em Sociologia pela Universidade de São Paulo (USP). Professor da Universidade Federal da Fronteira Sul (UFFS).
E-mail: joviles.trevisol@uffrs.edu.br

ANDRÉ BRASIL é doutorando em Estudos de Avaliação e Cienciometria pela Leiden University (Países Baixos). Servidor da Diretoria de Avaliação da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).
E-mail: a.l.brasil@cwts.leidenuniv.nl

Conflitos de interesse: Os autores declaram não possuir nenhum interesse comercial ou associativo que represente conflito de interesses em relação ao manuscrito.

Financiamento: O presente estudo foi realizado com o apoio do Conselho Nacional de Pesquisa e Desenvolvimento Tecnológico (CNPq), no âmbito da modalidade Pós-Doutorado no Exterior — PDE (Chamada Pública 08/2019) e da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior — Brasil (CAPES) — Código de Financiamento 001.

Contribuições dos autores: Conceituação, Curadoria de Dados, Análise Formal, Metodologia, Escrita – Primeira Redação, Escrita – Revisão e Edição: Trevisol, J. V.; Brasil, A.

Recebido em 22 de abril de 2022
Aprovado em 4 de novembro de 2022

