SciELO Preprints

Estado da publicação: O preprint não foi publicado em outro meio.

ASSIMETRIA AVALIATIVA DA ÁREA 27 DA CAPES NOS ESTRATOS DE EXCELÊNCIA (2010–2025)

Emerson Maccari

https://doi.org/10.1590/1982-7849rac2027260184.por

Submetido em: 2026-11-17
Postado em: 2026-11-18 (versão 1)
(AAAA-MM-DD)

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

ASSIMETRIA AVALIATIVA DA ÁREA 27 DA CAPES NOS ESTRATOS DE EXCELÊNCIA (2010–2025)

EVALUATION ASYMMETRY OF CAPES AREA 27 IN THE EXCELLENCE TIERS (2010–2025)

Emerson Maccari[1]
ORCID: https://orcid.org/0000-0001-7085-224X
maccari@uninove.br

[1] Universidade Nove de Julho, São Paulo, SP, Brasil

RESUMO

Objetivo: explicar por que a Área 27 (Administração Pública e de Empresas, Ciências Contábeis e Turismo) da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), que adota os critérios de avaliação mais exigentes das Ciências Sociais Aplicadas, é a que menos alcança os estratos de excelência dessa Grande Área. **Marco teórico:** desacoplamento institucional, reatividade das métricas e validade consequencial. **Método:** pesquisa documental que acompanha a trajetória da Área 27 desde 2010 e compara as fichas de avaliação das sete áreas das Ciências Sociais Aplicadas em dois ciclos avaliativos. **Resultados:** a Área 27 cresceu nos estratos de excelência, mas segue entre as últimas do sistema, enquanto suas fichas são as mais exigentes da Grande Área em todas as especificações analisadas. A diferença não está nas regras comuns, que todas as áreas cumprem, mas no modo como cada área as calibra, sobretudo na classificação de periódicos e nas métricas de citação, que passam a depender de bases internacionais pouco representativas da produção da área. **Conclusões:** a mesma nota pode corresponder a níveis de exigência diferentes entre áreas, o que fragiliza o princípio de comparabilidade que sustenta o sistema. Sem afirmar causalidade, o estudo propõe o conceito de desacoplamento por calibração discricionária: quando o regulador fixa critérios comuns e delega às áreas a decisão sobre o quanto exigir, a conformidade formal convive com réguas desiguais. Na prática,

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

indica que a CAPES precisa verificar a proporcionalidade dos critérios entre áreas, sem reduzir o rigor de nenhuma delas.

Palavras-chave: Educação de Pós-Graduação; Avaliação Educacional; Validade de Testes; Desacoplamento Institucional; Avaliação CAPES.

ABSTRACT

Objective: to explain why Area 27 (Public and Business Administration, Accounting and Tourism) of the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES), which applies the most demanding evaluation criteria in the Applied Social Sciences, is the area that least attains the excellence strata of this broad area. **Theoretical framework:** institutional decoupling, reactivity of metrics and consequential validity. **Method:** documentary research that traces the trajectory of Area 27 since 2010 and compares the evaluation forms of the seven Applied Social Sciences areas across two evaluation cycles. **Results:** Area 27 has grown in the excellence strata (grades 6 and 7) but remains among the lowest-ranked areas, while its evaluation forms are the most demanding in the broad area on every specification analyzed. The difference lies not in the common rules, which all areas meet, but in how each area calibrates them, particularly in journal classification and citation metrics, which depend on international databases that poorly represent the area's output. **Conclusions:** the same grade may correspond to different levels of stringency across areas, which weakens the comparability principle on which the system rests. Without claiming causality, the study proposes the concept of *decoupling through discretionary calibration*: when the regulator sets common criteria and delegates to each area how much to require, formal compliance coexists with unequal standards. In practical terms, the study indicates that CAPES needs to verify the proportionality of criteria across areas without lowering the rigor of any of them.

Keywords: Postgraduate Education; Educational Assessment; Test Validity; Institutional Decoupling; CAPES Evaluation.

1. Introdução

A qualidade da pós-graduação stricto sensu é reconhecida como vetor central do desenvolvimento científico e tecnológico de um país (Balbachevsky, 2005; Dantas, 2004; Maccari et al., 2014; Schwartzman, 2022). No Brasil, esse papel é desempenhado por mais de 4.500 programas de pós-graduação distribuídos em 50 áreas de avaliação, cuja qualidade é periodicamente aferida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). Dentro desse sistema, os estratos de excelência reúnem os conceitos 6 e 7, associados a padrão de qualidade internacional, e conferem acesso ao Programa de Excelência Acadêmica da CAPES (PROEX).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Maccari e Nishimura (2014) mapearam a ocupação dos estratos 6 e 7 pela Área 27 da CAPES (Administração Pública e de Empresas, Ciências Contábeis e Turismo), revelando um paradoxo persistente nas avaliações trienais de 2010 e 2013: a área apresentava crescimento expressivo de produção científica qualificada, de 274% em artigos A1 e 361% em artigos A2, e da formação de doutores (+62%), mas permanecia nas últimas posições do ranking de excelência, com apenas 4,2% de seus programas nos estratos superiores, posição inferior à de áreas com produção per capita menor, como Serviço Social (20%), Direito (10,7%) e Economia (9,1%).

Os autores questionaram se os critérios avaliativos conseguiam captar a qualidade diferenciada de uma área predominantemente orientada à produção nacional e ao impacto aplicado. Após três ciclos quadrienais (2017, 2021 e 2025) e uma reformulação do modelo de avaliação, a questão permanece em aberto: a Área 27 alcançou 9,4% de programas nos estratos de excelência em 2025, posicionando-se na 47ª colocação entre 50 áreas, enquanto a média geral do sistema atingiu 18,2%.

A sub-representação tem implicações institucionais diretas: os conceitos 6 e 7 são condição de elegibilidade ao PROEX, cuja concessão depende ainda de procedimentos próprios de adesão e fomento, não decorrendo de forma automática da nota. Com 9,4% de excelência, a Área 27 acessa proporcionalmente menos esse reconhecimento do que as áreas vizinhas. Este estudo examina se essa diferença coexiste com critérios de acesso mais exigentes do que os das demais áreas da Grande Área, situação que estaria em tensão com o princípio de comparabilidade do próprio sistema CAPES.

A produção em periódicos recebe atenção analítica privilegiada por duas razões: é a dimensão em que a assimetria entre as fichas é diretamente mensurável no texto normativo, e é a que se conecta ao indicador de maior peso diferencial da Área 27, a mediana do H-index Scopus e do índice H Spell do corpo docente, que na ficha projetada para o ciclo 2025–2028 passou a indicador quantitativo do item que responde por metade do quesito Impacto (CAPES, 2025c). A análise não desconsidera as demais dimensões, mas prioriza aquelas em que as assimetrias se manifestam de forma mais evidente e produzem maiores consequências.

Nesse contexto, três conceitos orientam a investigação. O primeiro, nível de exigência, refere-se aos critérios estabelecidos nas fichas de avaliação de cada área para alcançar os estratos superiores, considerando aspectos como indexação de periódicos, indicadores bibliométricos e quantidade e profundidade das evidências requeridas. A comparação desses critérios entre as áreas é realizada por meio de um índice específico, apresentado na seção metodológica. O segundo, sub-representação, caracteriza a participação proporcional de programas nos estratos 6 e 7 inferior tanto à média da Grande Área de Ciências Sociais Aplicadas quanto à média das 50 áreas avaliadas pela CAPES. O terceiro, mandato, corresponde às determinações normativas formalmente estabelecidas nos regulamentos, parâmetros e diretrizes da CAPES. Para assegurar a precisão conceitual, o termo exigência operacional é reservado à dimensão mensurada pelo índice, enquanto rigor é utilizado em sentido mais amplo, associado às expectativas e aos objetivos declarados pelas áreas de avaliação.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Apesar da relevância do tema, não há estudos que integrem: (a) a análise longitudinal da trajetória da Área 27 nos estratos de excelência desde 2010; (b) a caracterização do paradoxo da sub-representação e seus determinantes estruturais; e (c) a comparação sistemática das fichas de avaliação entre todas as áreas pertencentes à Grande Área de Ciências Sociais Aplicadas. Esta lacuna é particularmente relevante à luz do próprio mandato normativo da CAPES: desde o Regulamento da Avaliação Trienal 2013 (CAPES, 2013b) e reiterado nos Parâmetros dos ciclos quadrienais subsequentes (CAPES, 2017b; CAPES, 2021b; CAPES, 2025a), o sistema determina explicitamente que notas iguais entre áreas devem representar igual valor e que a avaliação não pode se restringir à perspectiva intraárea.

A comparação entre áreas da mesma Grande Área é, portanto, princípio estruturante do sistema, detalhado na seção 2, e não escolha metodológica deste estudo.

A questão central que orienta este estudo é: como se caracteriza a assimetria entre a exigência operacional das fichas de avaliação da Área 27 e sua ocupação dos estratos de excelência, quando comparada às demais áreas da Grande Área de Ciências Sociais Aplicadas, à luz do princípio de comparabilidade interáreas estabelecido pelo próprio sistema? As implicações para a calibração dos parâmetros são derivadas dos achados e não constituem objeto de verificação empírica.

Com base em pesquisa documental e análise de dados da CAPES, os resultados são consistentes com o paradoxo apontado por Maccari e Nishimura (2014), indicam que ele se aprofundou nos ciclos quadrienais subsequentes e discutem cinco fatores estruturais associados a ele. As contribuições centrais são: (i) análise longitudinal completa do desempenho da Área 27 nos estratos de excelência de 2010 a 2025, com inclusão da média geral do sistema como linha de referência em todos os ciclos; (ii) comparação sistemática das fichas de avaliação das sete áreas da Grande Área de Ciências Sociais Aplicadas no ciclo avaliativo de 2021–2024 e projetado para o ciclo 2025–2028, operacionalizada por matriz de codificação, índice de exigência operacional e análises de sensibilidade, indicando que a Área 27 registra o maior índice do conjunto nos dois ciclos; e (iii) a especificação do mecanismo de desacoplamento por calibração discricionária e de suas condições de fronteira, com a discussão, ancorada nos próprios dispositivos da CAPES sobre comparabilidade interáreas, das implicações desses achados para a calibração dos parâmetros.

2. O Sistema de Avaliação da CAPES: do Modelo Trienal ao Modelo Quadrienal

O sistema de avaliação da pós-graduação da CAPES passou por transformações estruturais entre 2010 e 2025: do modelo trienal quantitativo para um modelo quadrienal com ênfase em impacto, inovação e inserção social (Maccari et al., 2014; Schwartzman, 2022; Verhine & Dantas, 2012). A avaliação da CAPES orienta a alocação de recursos públicos para formação, produção científica e pesquisa, em instituições públicas e em programas privados que recebem bolsas, auxílios e projetos. Uma régua descalibrada, nesse contexto, repercute sobre a distribuição de recursos, as estratégias dos programas e as trajetórias dos pesquisadores, o que amplia a relevância de examinar os critérios que

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

definem excelência. Nos ciclos de avaliação anteriores, o índice H Scopus era empregado na classificação de periódicos da área, como critério de acesso aos estratos superiores do Qualis (CAPES, 2013a). Na ficha que produziu as notas da Quadrienal 2021-2024 (CAPES, 2025d), o H-index Scopus figura entre quatro métricas de citação examinadas de forma qualitativa; já na ficha projetada para o ciclo 2025–2028 (CAPES, 2025c), a mediana do H-index Scopus e do índice H Spell do corpo docente passou a indicador quantitativo do item 3.3 do sistema de avaliação, que responde por metade do quesito Impacto, e o item 1.3 do quesito Programa passou a exigir planejamento estratégico articulado ao Plano de Desenvolvimento Institucional (PDI). Comparar ciclos exige cautela, pois critérios e pesos foram substancialmente alterados ao longo do período.

O princípio de comparabilidade entre programas é estruturante do sistema desde o primeiro PNPG (1975) e encontrou sua primeira expressão normativa explícita no Regulamento da Avaliação Trienal 2013 (CAPES, 2013b), que determinava que notas iguais entre áreas devem representar igual valor e que a avaliação deve promover a equivalência de desempenho entre programas com as mesmas notas nas distintas áreas.

A partir da transição para o modelo quadrienal, os Parâmetros de todos os ciclos repetiram, com redação praticamente idêntica, os mesmos imperativos: ciclo 2013–2016 (CAPES, 2017b), ciclo 2017–2020 (CAPES, 2021b) e ciclo 2025–2028 (CAPES, 2025a). Em todos eles, a CAPES determina que a avaliação não pode se restringir a uma perspectiva intraárea; que os referenciais devem evitar a perda progressiva da capacidade discriminatória da escala por concentração de programas nos níveis superiores; e que as áreas devem adotar padrões internacionais como referência. Trata-se, portanto, de uma constante normativa que atravessa quatro ciclos, do Trienal 2013 ao Quadrienal projetado para o ciclo 2025–2028, conferindo ao argumento deste estudo amparo documental consistente.

Esse terceiro imperativo merece atenção especial: ao orientar que as comissões adotem padrões internacionais, a CAPES legitima a estratégia da Área 27. O que os dados revelam, porém, é que a área internacionalizou seus critérios de forma tão concentrada que apenas 22,65% dos artigos publicados pela área em 2017–2020 estavam em periódicos indexados no Scopus ou na Web of Science (CAPES, 2022), bases das quais depende, em boa medida, o acesso ao estrato máximo na classificação adotada para 2025–2028. Essa concentração pode produzir o efeito que os parâmetros buscam prevenir: a perda da capacidade discriminatória da escala. A repetição do alerta ao longo de todos os ciclos oferece base normativa para examinar a calibração das fichas.

Maccari e Nishimura (2014) documentaram que, mesmo diante desse mandato, a Área 27 permanecia sub-representada nos estratos de excelência, sugerindo, já naquele ano, tensão entre o princípio orientado pelo regulamento e a ocupação dos estratos de excelência. A Grande Área de Ciências Sociais Aplicadas constitui o recorte natural para verificar se esse mandato é observado: áreas inseridas no mesmo arcabouço regulatório, avaliadas pela mesma instância, são, por definição do sistema, comparáveis entre si.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

3. Fundamentação Teórica

A análise da assimetria nos critérios de excelência entre áreas de avaliação da CAPES demanda um quadro teórico capaz de dar conta de três dimensões: por que organizações inseridas na mesma estrutura regulatória produzem normas operacionais heterogêneas? como sistemas de mensuração constroem e apagam qualidades ao transformá-las em números? e, em que condições essa heterogeneidade compromete a validade consequencial da avaliação. Para responder a essas questões, o estudo articula três contribuições teóricas complementares: o mecanismo de desacoplamento, da teoria neo-institucional (Bromley & Powell, 2012); os conceitos de comensurabilidade e reatividade, da sociologia das métricas (Espeland & Stevens, 1998; Espeland & Sauder, 2007); e a teoria da validade consequencial (Messick, 1989; Kane, 2013).

Da teoria neo-institucional, o isomorfismo normativo (DiMaggio & Powell, 1983) explica que organizações em um mesmo campo convergem quando submetidas a pressões regulatórias comuns, e é exatamente o que se observa na CAPES: todas as áreas adotam o mesmo Sistema de avaliações, com os quesitos gerais, escala de notas e possuem o mesmo órgão deliberativo (o Conselho Técnico-Científico da Educação Superior, CTC-ES, instância colegiada da CAPES responsável por homologar os resultados da avaliação e deliberar sobre os conceitos atribuídos aos programas). Contudo, DiMaggio e Powell (1983) também identificaram que esse isomorfismo se manifesta frequentemente no plano formal sem correspondente uniformidade operacional. Bromley e Powell (2012) sistematizaram esse mecanismo como desacoplamento (decoupling): organizações que partilham estruturas formais idênticas desenvolvem práticas operacionais divergentes quando a implementação é delegada a subunidades com tradições cognitivas distintas. É o que ocorre nas fichas da CAPES: a estrutura formal é isomorfa (mesmos quesitos), mas os parâmetros operacionais que a concretizam são substantivamente heterogêneos, porque cada comissão traduz os critérios genéricos segundo os esquemas cognitivos da própria comunidade científica (Scott, 2014).

O mandato de comparabilidade do sistema (CAPES, 2013b; CAPES, 2025a) determina que notas iguais entre áreas devem representar igual valor. O que os documentos normativos não endereçam é exatamente o que o desacoplamento produz: parâmetros de excelência entre áreas de uma mesma Grande Área que guardam proporcionalidade apenas formal, não substantiva. Nesse sentido, a exigência operacional da Área 27 pode refletir maturidade científica. A questão que este estudo levanta não é se ela deve ser menos exigente, mas se os indicadores e pesos de suas fichas guardam proporcionalidade com os das áreas que o Sistema de Avaliação declara comparáveis.

A segunda dimensão é provida pela sociologia das métricas, especificamente pelos conceitos de reatividade e comensurabilidade. Sistemas de mensuração não são neutros: ao classificar e comparar objetos, retroalimentam os comportamentos das entidades avaliadas, produzindo reatividade, entendida como a tendência de que atores modifiquem suas práticas para melhorar sua posição nos rankings, independentemente de essa mudança representar melhoria substantiva (Espeland & Sauder, 2007).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

No caso da Área 27, a adoção do H-index Scopus como indicador formal induz programas a redirecionar esforços para periódicos internacionais indexados, com impacto potencial sobre a relevância local da pesquisa produzida. Complementarmente, Espeland e Stevens (1998) introduziram o conceito de comensurabilidade (commensuration): o processo pelo qual qualidades heterogêneas são transformadas em uma escala numérica comum, apagando diferenças qualitativas que podem ser substancialmente relevantes. No contexto deste estudo, a CAPES comensurou produção com perfis estruturalmente distintos (nacional e internacional) numa escala única de notas, sem ajustar os parâmetros para refletir essa diferença estrutural. O risco é que parte da qualidade de uma produção predominantemente nacional se torne menos visível no processo de comensuração, quando a escala privilegia bases de cobertura internacional. Limitações de cobertura das bases afetam de forma desproporcional as ciências sociais em países não anglófonos (Moed, 2008), padrão compatível com o perfil de indexação da Área 27 descrito na seção 5.

A terceira dimensão ancora-se na teoria da validade consequencial (Messick, 1989). Messick propôs que a validade de uma métrica não se esgota na correspondência entre o que ela mede e o que pretende medir (validade interna): inclui também as consequências sociais e institucionais de seu uso. Uma métrica tem validade consequencial comprometida quando seu emprego sistemático produz resultados injustos para grupos que, por razões estruturais legítimas, apresentam perfis distintos daquele para o qual o instrumento foi originalmente calibrado. Kane (2013) desenvolveu essa ideia na forma de argumento de validade como cadeia de inferências: cada passo que liga o desempenho observado a uma conclusão sobre qualidade precisa ser justificado.

No contexto deste estudo, o elo central da argumentação situa-se entre os parâmetros de avaliação adotados pela Área 27, listas internacionais de periódicos, índice H da Scopus e acompanhamento longitudinal de egressos, e as conclusões sobre a densidade de excelência da área. Essa relação exige uma análise criteriosa, uma vez que a adoção de parâmetros mais exigentes não permite concluir, por si só, que a área apresente menor densidade de excelência. Alternativamente, pode indicar que os instrumentos utilizados captam apenas parte da qualidade acadêmica efetivamente produzida. Comparações entre grupos estruturalmente distintos podem gerar resultados sistematicamente enviesados quando as métricas empregadas não consideram essas diferenças (Shadish et al., 2002). Nesse contexto, a predominância de indicadores de inserção internacional na avaliação de uma área cuja produção científica qualificada se concentra majoritariamente no âmbito nacional pode comprometer a validade consequencial da avaliação. Trata-se, contudo, de uma hipótese cuja comprovação exigiria medidas independentes de qualidade, uma vez que o presente estudo se concentra na análise dos instrumentos avaliativos.

Cabe esclarecer que essa argumentação não pressupõe a redução do rigor avaliativo nem a substituição dos critérios internacionais desenvolvidos e aperfeiçoados pela Área 27 no período de quinze anos abrangido por este estudo. Defende-se, antes, que a definição e a aplicação desses parâmetros observem o princípio da proporcionalidade e da comparabilidade entre as áreas que integram

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

a mesma Grande Área, de modo que notas equivalentes expressem níveis efetivamente comparáveis de excelência acadêmica, em consonância com as diretrizes de avaliação da CAPES.

Convém situar o estudo em relação a dois debates. O primeiro questiona a própria lógica da avaliação por rankings e classificações, e não apenas os parâmetros que a operacionalizam. Este estudo dialoga com essa corrente sem aderir integralmente a ela: examina a régua e suas consequências distributivas, não propõe abandoná-la. O segundo debate problematiza a adoção de referenciais do Norte Global como padrão privilegiado de qualidade. A escolha de bases e listas internacionais de indexação como critério de excelência não é neutra: define o que se torna visível como produção de qualidade e tende a subvalorizar circuitos de publicação nacionais e regionais (Bispo, 2024). No caso da Área 27, essa escolha conecta-se diretamente ao dado de que apenas 22,65% dos artigos da área em 2017–2020 estavam em periódicos indexados no Scopus ou na WoS (CAPES, 2022), ordem de grandeza da produção com acesso potencial ao estrato máximo. Por qualidade e excelência, este estudo entende o reconhecimento, pela comunidade avaliadora, de contribuições relevantes para o avanço do conhecimento e para o enfrentamento de problemas científicos e sociais, reconhecimento que pode e deve ser construído também a partir da realidade e dos problemas brasileiros, e não apenas por proxies de circulação internacional.

Esses três quadros teóricos encontram respaldo na literatura internacional sobre avaliação da pesquisa, que convergiu, nas últimas duas décadas, para um princípio central: indicadores de excelência precisam ser calibrados ao contexto disciplinar para que consigam capturar a qualidade que efetivamente existe nos programas avaliados. O Leiden Manifesto (Hicks et al., 2015) estabelece que métricas devem ser ajustadas ao contexto de cada área, evitando que parâmetros excessivamente estreitos deixem invisível a produção de alta qualidade. O Metric Tide (Wilsdon et al., 2015) reforça que a cobertura das bases bibliométricas varia estruturalmente entre áreas, o que exige calibração dos pesos para que os indicadores reflitam com fidelidade o desempenho real. A DORA (2012) alerta que o uso de métricas sem ajuste ao perfil da área pode comprometer a validade da avaliação comparativa.

No contexto brasileiro, esse debate tem tradição consolidada. Schwartzman (2022) argumenta que o Sistema de Avaliação da CAPES, embora reconhecidamente bem-sucedido por décadas, tende a produzir homogeneização excessiva ao aplicar critérios uniformes a programas com perfis e vocações distintas, com o risco de levar áreas inteiras a se transformarem em "fotocópias de qualidade inferior" dos modelos de melhor desempenho. Balbachevsky (2005), em análise sistemática da pós-graduação brasileira, já apontava como desafio central a necessidade de o sistema avaliativo reconhecer a diversidade interna do sistema sem comprometer a comparabilidade entre programas. No mesmo sentido, Barata (2016) e Bispo (2020) documentaram que o descompasso entre parâmetros e perfil de produção gera subcaptura de excelência em áreas com vocação nacional. Esse descompasso de cobertura é técnico e estrutural, não reflexo de ausência de qualidade (Larivière & Gingras, 2010; Alperin et al., 2019). Para a Área 27, a implicação é que a proporcionalidade de seus parâmetros em relação aos das áreas vizinhas merece exame sistemático.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

A articulação desses três referenciais teóricos fundamenta a formulação de três expectativas analíticas que direcionam o exame da relação entre o nível de exigência dos critérios estabelecidos nas fichas de avaliação e a distribuição dos programas nos estratos de excelência, sem pressupor previamente a direção ou a natureza dessa relação.

E1 (desacoplamento institucional com heterogeneidade operacional): derivada do mecanismo de decoupling (DiMaggio & Powell, 1983; Bromley & Powell, 2012), parte de uma base normativa explícita do próprio sistema CAPES. O princípio de comparabilidade entre programas, intraárea e interáreas, é estruturante desde o primeiro PNPG (1975) e foi formalizado no Regulamento da Avaliação Trienal 2013 (CAPES, 2013b), reiterado, com redação praticamente idêntica, nos Parâmetros dos ciclos avaliativos de 2013–2016 (CAPES, 2017b), 2017–2020 (CAPES, 2021b) e 2025–2028 (CAPES, 2025a), todos consignando que a avaliação não pode se restringir à perspectiva intraárea. A comparação entre áreas da mesma Grande Área é, portanto, não apenas legítima: é exigida pelo próprio sistema. O que a expectativa prevê é que, embora todas as áreas operem sob os mesmos quesitos gerais, a mesma escala de notas e o mesmo órgão deliberativo (CTC-ES), as comissões de cada área calibram de forma autônoma os parâmetros operacionais (pesos, exigências de produção qualificada, listas de classificação de periódicos) e essa calibração varia substantivamente entre as sete áreas das Ciências Sociais Aplicadas. A heterogeneidade operacional resultante é o objeto empírico do estudo: a expectativa é que ela seja sistemática, mensurável nas fichas de avaliação e associada à ocupação diferenciada dos estratos de excelência entre áreas que o sistema declara comparáveis.

E2 (subcaptura estrutural por incomensurabilidade de perfil): ancorada nos conceitos de commensuration e reactivity (Espeland & Stevens, 1998; Espeland & Sauder, 2007), prevê que, ao comensurar produção nacional e internacional em uma única escala sem ajuste de parâmetros, o Sistema de Avaliação pode induzir reatividade nos programas e produzir subcaptura: parte da qualidade de programas com perfil predominantemente nacional tenderia a não alcançar o topo das dimensões condicionadas a bases internacionais. Como o estudo não dispõe de medida independente de qualidade, E2 é examinada apenas no plano da estrutura dos instrumentos.

E3 (validade consequencial e agenda de calibração interna): fundamentada na teoria da validade consequencial (Messick, 1989; Kane, 2013), parte da constatação empírica de que outras áreas da mesma Grande Área, operando com parâmetros substantivamente menos exigentes, ocupam proporcionalmente mais os estratos de excelência. Esse contraste sugere que a maior exigência operacional da Área 27 não se acompanha de reconhecimento institucional equivalente. A expectativa é que, se a calibração contribui para o padrão observado, a revisão da proporcionalidade dos parâmetros tende a aproximar o significado das notas entre áreas comparáveis, sem exigir renúncia aos critérios. Essa expectativa orienta a discussão das implicações e não é objeto de verificação empírica.

Articuladas, essas três expectativas configuram um mecanismo que este estudo nomeia como desacoplamento por calibração discricionária em sistemas federados de avaliação: quando o ente regulador define quesitos comuns, mas delega às comissões de área a calibração operacional dos

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

parâmetros, a aderência formal aos quesitos coexiste com variação substantiva nas exigências práticas, produzindo assimetria horizontal entre unidades formalmente comparáveis. Esta é a contribuição teórica que o caso da Área 27 da CAPES permite formular, com condições de fronteira especificadas na seção 6.1: uma maneira específica de desacoplamento que opera não no plano do cumprimento (pois todas as áreas cumprem todos os quesitos), mas no plano da calibração operacional. A seção de resultados examina as três expectativas com base nos dados oficiais da CAPES (2026) e na análise comparativa das fichas das sete áreas da Grande Área de Ciências Sociais Aplicadas.

4. Metodologia

Este estudo adota um design de pesquisa documental com análise de dados secundários (Bowen, 2009; Morgan, 2022), abordagem amplamente reconhecida nas ciências sociais como adequada para examinar políticas públicas, instrumentos regulatórios e sistemas de avaliação institucional, justamente por permitir o exame retrospectivo e sistemático de processos cuja materialidade reside em documentos oficiais (Dalglish, Khalid, & McMahon, 2020). A escolha metodológica é particularmente adequada ao objeto deste estudo: o sistema de avaliação da CAPES é, por desenho, um sistema documental: suas regras, parâmetros, critérios e resultados materializam-se em fichas de avaliação, regulamentos, documentos de área e planilhas oficiais que constituem registros públicos disponíveis para escrutínio independente.

Estudos contemporâneos sobre reformas em sistemas nacionais de avaliação de pesquisa têm adotado abordagem semelhante: Yang e Su (2024) examinaram, por análise documental em profundidade, as respostas institucionais das universidades chinesas à reforma da avaliação de pesquisa daquele país; Cabral (2025) conduziu estudo comparativo entre os sistemas brasileiro e espanhol de avaliação da pós-graduação a partir do exame sistemático de documentos normativos; e Verhine, Freitas e Robl (2022), em estudo comparado entre os sistemas brasileiro e holandês, mostraram a fecundidade da análise documental para revelar as concepções implícitas que estruturam diferentes modelos avaliativos. Esses trabalhos consolidam a análise documental como método central para a investigação de sistemas de avaliação de pós-graduação.

A pesquisa documental organiza-se em dois eixos analíticos complementares. O eixo longitudinal examina o desempenho da Área 27 nos estratos de excelência ao longo de cinco ciclos avaliativos consecutivos (Trienais 2010 e 2013; Quadrienais 2017, 2021 e 2025), reconstituindo a trajetória da área a partir das planilhas oficiais de resultados publicadas pela CAPES em cada ciclo.

O eixo comparativo examina as fichas de avaliação de dois ciclos das sete áreas que compõem a Grande Área de Ciências Sociais Aplicadas: Direito (Área 26), Administração Pública e de Empresas, Ciências Contábeis e Turismo (Área 27), Economia (Área 28), Arquitetura, Urbanismo e Design (Área 29), Planejamento Urbano e Regional/Demografia (Área 30), Comunicação, Informação e Museologia (Área 31) e Serviço Social (Área 32), identificando, para cada uma, os parâmetros operacionais utilizados nos quesitos Programa, Formação e Impacto. São examinados dois conjuntos de fichas: as do

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

quadriênio 2021–2024, que produziram as notas da Quadrienal 2025 analisadas no eixo longitudinal (para a Área 27, CAPES, 2025d), e o projetado para o quadriênio 2025–2028, que indicam a direção das mudanças em curso (para a Área 27, CAPES, 2025c). Para a dimensão de internacionalização (D6), que as áreas regulam também nos Documentos de Área, foram examinados ainda os Documentos de Área de cada ciclo. As fichas e os Documentos de Área das demais seis áreas estão listados, com endereço eletrônico, na Tabela S3 do material suplementar.

Os procedimentos seguiram a sequência consolidada para análise documental qualitativa (Bowen, 2009; Morgan, 2022): (i) identificação e seleção das fontes primárias: Regulamentos, Diretrizes Comuns, Parâmetros, Documentos de Área (para a Área 27, CAPES, 2025b), Fichas de Avaliação e Relatórios de Avaliação publicados pela Diretoria de Avaliação (DAV) da CAPES nos cinco ciclos; (ii) leitura preliminar para mapeamento dos dispositivos relevantes; (iii) leitura analítica com codificação formal dos parâmetros operacionais em seis dimensões, conforme o protocolo detalhado a seguir; e (iv) síntese interpretativa por triangulação entre os dois eixos, articulando os achados quantitativos sobre a distribuição de notas com os achados qualitativos sobre a estrutura das fichas. A triangulação entre dados quantitativos e documentos normativos é prática reconhecida para reforçar a validade interna em estudos de política pública (Dalglish et al., 2020; Wesley, 2010).

O eixo comparativo foi operacionalizado por meio de um protocolo formal de codificação. Seis dimensões de exigência operacional foram definidas a partir dos quesitos comuns às fichas: (D1) sistema de classificação da produção em periódicos, com foco nas condições de acesso ao estrato máximo; (D2) métricas bibliométricas individuais de docentes; (D3) avaliação de teses e dissertações, considerando o volume examinado e a exigência de publicação derivada; (D4) produtos técnicos e tecnológicos, considerando cotas obrigatórias e restrição de tipos aceitos; (D5) acompanhamento de egressos, considerando a profundidade temporal exigida; e (D6) existência de gradação formal de internacionalização entre as notas 6 e 7. As seis dimensões correspondem aos pontos em que os quesitos comuns definidos pelo CTC-ES deixam às comissões de área a escolha de parâmetros operacionais, isto é, aos lugares de calibração discricionária descritos pelo mecanismo examinado; elementos fixados de modo uniforme pela CAPES para todas as áreas, como a estrutura de quesitos e a escala de notas, foram excluídos por não admitirem variação entre comissões.

Cada dimensão foi codificada em escala ordinal de três níveis: 0 quando a exigência é ausente ou a avaliação é qualitativa aberta, sem parâmetro restritivo; 1 quando há parâmetro quantitativo ou estrutura formal de exigência, sem condicionamento a bases ou listas internacionais e sem caráter único na Grande Área; e 2 quando o parâmetro é condicionado a bases ou listas internacionais, ou constitui a exigência estrutural mais alta ou única entre as sete áreas. Os pontos de corte distinguem três graus de restrição ao acesso ao topo de cada dimensão: ausência de parâmetro, parâmetro formal que não depende da cobertura de bases e parâmetro que vincula o topo a bases ou listas internacionais, cuja cobertura da produção da área é limitada. O critério relativo do nível 2 (exigência mais alta ou única)

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

pode favorecer mecanicamente a área com mais exigências; por isso, uma das análises de sensibilidade o elimina.

O somatório das seis dimensões forma o índice de exigência operacional, que varia de 0 a 12 pontos por área. A matriz foi codificada a partir do texto integral das fichas e dos Documentos de Área e conferida por uma segunda codificação independente; as divergências, concentradas na codificação do Qualis Periódicos, do volume de teses examinadas, das cotas de produtos técnicos e da gradação de internacionalização, foram resolvidas pela explicitação das regras correspondentes no Quadro S1.

As regras de atribuição, a comparação detalhada dos parâmetros de cada ficha (Tabela S1), a matriz com a evidência textual de cada código (Tabela S2) e as planilhas de codificação dos dois codificadores estão depositadas no repositório da pesquisa. O índice não mede a qualidade das fichas nem as hierarquiza normativamente: descreve, em métrica comum, o grau de restrição operacional imposto ao acesso aos estratos superiores em cada dimensão. As dimensões recebem pesos iguais por duas razões: não há base teórica para ponderá-las de forma diferenciada, situação em que a ponderação unitária tende a produzir ordenações tão estáveis quanto as de pesos estimados (Dawes, 1979); e os pesos dos itens variam entre as fichas das áreas, que não informam o peso de cada quesito na nota final, o que impede uma ponderação comum derivada dos próprios documentos. As fichas acadêmica e profissional de cada área foram codificadas em conjunto; quando as modalidades diferiam em uma dimensão, prevaleceu o parâmetro mais restritivo, com registro da diferença na Tabela S2, e as diferenças concentraram-se na dimensão de produtos técnicos e tecnológicos. A concordância entre codificadores indica a reprodutibilidade das regras, não a validade de construto. Para examinar se a posição relativa das áreas depende das escolhas do protocolo, três análises de sensibilidade foram conduzidas: (a) codificação binária, que registra apenas a presença de exigência formal; (b) regra absoluta, em que o nível 2 é atribuído somente a parâmetros condicionados a bases ou listas internacionais; e (c) retirada de uma dimensão por vez. Verificou-se ainda se alguma área supera a Área 27 em qualquer dimensão, condição para que a ordenação se inverta sob algum conjunto de pesos não negativos.

Quanto às escolhas e limitações metodológicas: (a) a transição do modelo trienal para o quadrienal limita a comparabilidade direta entre os períodos no eixo longitudinal, razão pela qual as comparações intertemporais são apresentadas com a devida cautela e os achados centrais apoiam-se prioritariamente nos três ciclos quadrienais (2017, 2021 e 2025); (b) os dados detalhados sobre produção qualificada e titulação de doutores referentes ao quadriênio 2021–2024 ainda não estavam disponíveis em nível de programa ao tempo da pesquisa, o que impede verificar se diferenças de desempenho explicam parte do padrão, e por isso o eixo comparativo restringe-se aos instrumentos, examinados nos dois ciclos com equivalência temporal entre as sete áreas; (c) o estudo examina as normas formalmente vigentes nas fichas de avaliação, e não a operacionalização empírica de cada avaliação por sua comissão, escolha justificada pelo objeto da pesquisa, que é a calibração formal das réguas e não as deliberações específicas sobre programas individuais; e (d) o argumento opera no plano descritivo, não

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

causal, documentando a tensão entre a estrutura das fichas e o princípio de comparabilidade interáreas das Diretrizes da CAPES, sem pretender estabelecer relação causal unidirecional entre exigência das fichas e densidade de excelência.

Todos os números deste estudo referem-se ao resultado definitivo da Quadrienal 2025, publicado após a deliberação da 241ª reunião do CTC-ES, em 26/05/2026, já consolidados os pedidos de reconsideração de todas as áreas. Nesse resultado, o déficit da Área 27 em relação à média do sistema ampliou-se em relação ao resultado preliminar, e a área passou da 46ª para a 47ª posição no conjunto das áreas, ultrapassada por Educação.

Em conformidade com a política de ciência aberta, todos os dados utilizados neste estudo foram depositados em repositório público e estão disponíveis para verificação independente. Isso inclui: a planilha de resultados oficiais da CAPES (2026), já consolidada com os resultados dos pedidos de reconsideração de todas as áreas; as tabelas comparativas das fichas de avaliação das sete áreas analisadas; as matrizes de codificação dos dois ciclos, com o trecho de cada ficha que sustenta cada código, e as análises de sensibilidade; e as séries históricas completas de desempenho da Área 27 e das demais áreas nos cinco ciclos avaliativos considerados (2010, 2013, 2017, 2021 e 2025). O repositório encontra-se disponível em: <https://osf.io/je4cw/?view_only=cbaab02127c849f6be62a0cca5f0e38b>.

Uso de inteligência artificial - IA. Em conformidade com as diretrizes da RAC para o uso de IA generativa, os autores declaram que os modelos Claude (Anthropic) e ChatGPT (OpenAI) foram utilizados exclusivamente na revisão linguística e estilística do manuscrito. A concepção do estudo, a definição do referencial teórico, a coleta e análise dos dados, a interpretação dos resultados e as conclusões são de responsabilidade integral dos autores.

5. Resultados

Os resultados estão organizados em quatro subseções: panorama geral da quadrienal 2025 (5.1), evolução histórica dos estratos de excelência entre 2010 e 2025 (5.2), análise do paradoxo da sub-representação e de suas explicações estruturais (5.3), e posicionamento comparativo entre as 50 áreas avaliadas (5.4). A discussão teórica dos achados é apresentada na seção 6. Ao longo das seções empíricas, o texto distingue três planos: a evidência documental, aquilo que fichas, planilhas e regulamentos registram; a inferência interpretativa, aquilo que os padrões observados sugerem; e a recomendação normativa, aquilo que se propõe à CAPES. As afirmações de cada plano são sinalizadas quando a distinção for pertinente.

5.1 Panorama Geral da Avaliação Quadrienal 2025

A avaliação quadrienal de 2025 contemplou 4.555 programas de pós-graduação em 50 áreas, aumento de 36,5% em relação aos 3.337 programas da trienal de 2013. No total, 827 programas (18,2%) obtiveram conceito 6 ou 7, sendo 338 (7,4%) com conceito 7 e 489 (10,7%) com conceito 6.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Para a Área 27, foram avaliados 181 programas, sendo 72 de mestrado/doutorado acadêmico, 41 de mestrado acadêmico, 48 de mestrado profissional, 19 de mestrado/doutorado profissional e 1 de doutorado. A distribuição das notas é apresentada na Tabela 1:

Tabela 1 – Distribuição das notas dos programas da Área 27 na Avaliação Quadrienal 2025.

Nota 1	Nota 2	Nota 3	Nota 4	Nota 5	Nota 6	Nota 7
1 (0,6%)*	3 (1,7%)*	40 (22,1%)	70 (38,7%)	50 (27,6%)	9 (5,0%)	8 (4,4%)

*Notas 1 e 2 correspondem a programas descredenciados; os percentuais usam como base o total de 181 programas avaliados.

Fonte: CAPES (2026).

A concentração massiva nas notas intermediárias é notável: **88% dos programas concentram-se nos estratos 3, 4 e 5.** Apenas 17 programas (9,4%) alcançam os estratos de excelência: 9 com nota 6, o patamar de qualidade internacional, e 8 com nota 7, reservado a programas comparáveis aos melhores do mundo, contra uma média de 18,2% do sistema de avaliação como um todo. O déficit de 8,8 pontos percentuais em relação à média geral será analisado em detalhe nas seções seguintes.

5.2 Evolução Histórica do Povoamento dos Estratos 6 e 7 (2010–2025)

A Tabela 2 apresenta a evolução comparativa completa do povoamento dos estratos superiores incluindo, pela primeira vez com verificação direta nas fontes primárias da CAPES, a média geral do sistema como linha de referência em todos os ciclos:

Tabela 2 – Evolução histórica do povoamento dos estratos 6 e 7 pela Área 27 (2010–2025).

Indicador	Trienal 2010	Trienal 2013	Quad. 2017	Quad. 2021	Quad. 2025	Var. 2013→2025
Total de programas avaliados	100	119	188	187	181	+52%
Programas com nota 7	2 (2,0%)	2 (1,7%)	3 (1,6%)	5 (2,7%)	8 (4,4%)	+300%
Programas com nota 6	3 (3,0%)	3 (2,5%)	4 (2,1%)	9 (4,8%)	9 (5,0%)	+200%
TOTAL nota 6 ou 7	**5 (5,0%)**	**5 (4,2%)**	**7 (3,7%)**	**14 (7,5%)**	**17 (9,4%)**	**+240%**
Média geral do sistema CAPES	**11,8%**	**~14%**	**~14%**	**~15%**	**18,2%**	**DÉFICIT PERSISTENTE: −6,8 p.p. em 2010; −8,8 p.p. em 2025**

Fontes: CAPES (2010a); CAPES (2010b); CAPES (2013a); CAPES (2017a); CAPES (2022); CAPES (2026). Nota: comparação entre modelos trienal e quadrienal deve ser feita com cautela pela mudança de critérios. Em 2017, os números referem-se ao resultado após os recursos, conforme o Documento de Área 2019 (CAPES, 2019). Para referência comparativa, os percentuais de excelência das demais áreas da Grande Área de Ciências Sociais Aplicadas na quadrienal 2021 foram: Serviço Social 25,0% (9/36),

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Direito 12,6% (17/135), Economia 13,2% (10/76), Comunicação 11,6% (11/95), Arquitetura 9,7% (6/62) e Planejamento Urbano 10,2% (5/49), calculados a partir de CAPES (2021a), valores que já em 2021 situavam a Área 27 (7,5%) consistentemente abaixo de todas as demais áreas da Grande Área.

Os dados revelam uma trajetória que combina avanço absoluto inegável com deterioração relativa persistente, o núcleo do paradoxo. A Área 27 triplicou o número de programas nos estratos de excelência (de 5 em 2010 para 17 em 2025, +240%), mas o déficit em relação à média geral ampliou-se de 6,8 p.p. (5,0% vs. 11,8% em 2010) para 8,8 p.p. (9,4% vs. 18,2% em 2025). O pior momento foi a quadrienal 2017 (déficit de 10,3 p.p. e único ciclo com recuo do percentual de excelência). A inflexão positiva de 2021 (7,5%) e sua consolidação em 2025 (9,4%) são reais, mas não alteram a posição estrutural: em 2010 a área ocupava as últimas posições de um ranking com 48 áreas; em 2025 ocupa a 47ª entre 50, com déficit maior do que quando o paradoxo foi identificado pela primeira vez. O padrão é de distância relativa persistente, apesar do avanço absoluto dos programas.

Contudo, o dado mais revelador não é o crescimento absoluto, mas o distanciamento persistente da média do sistema e, principalmente, das áreas vizinhas da Grande Área. Esse resultado impõe a pergunta central que orienta a seção seguinte: **por que o crescimento absoluto consistente da Área 27, acompanhado de exigências operacionais crescentes, não se traduz em maior densidade nos estratos de excelência?** A seção 5.3 examina explicações endógenas e uma explicação institucional, sem pretender ordená-las por importância causal.

5.3 O Paradoxo da Sub-representação: Explicações Endógenas e Explicação Institucional

A análise do paradoxo exige distinguir duas perspectivas. A primeira contempla fatores endógenos, relacionados ao perfil, à trajetória e à estrutura interna da própria área. A segunda considera fatores exógenos de natureza institucional, particularmente as possíveis assimetrias na calibração dos critérios de excelência adotados pelas comissões de área. Essa distinção é fundamental, pois cada perspectiva apresenta implicações distintas para o aprimoramento da avaliação: enquanto os fatores endógenos remetem a estratégias de desenvolvimento e aperfeiçoamento que podem ser implementadas pelos próprios programas; os fatores exógenos envolvem aspectos estruturais e normativos do sistema de avaliação, cuja revisão compete à CAPES e ao Conselho Técnico-Científico da Educação Superior (CTC-ES). A Tabela 3 apresenta os elementos constitutivos do paradoxo:

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Tabela 3 – Exigência operacional crescente e sub-representação persistente da Área 27 nos estratos de excelência (2010–2025).

Ciclo avaliativo / Dimensão	Exigência operacional da Área 27	Posicionamento nos estratos de excelência
Trienal 2010: Início da série histórica	Qualis baseado em periódicos internacionais indexados; produção nacional estruturalmente excluída do estrato máximo	5,0% nos estratos 6–7 (47ª/48 áreas); déficit de 6,8 p.p. em relação à média do sistema (11,8%)
Trienal 2013: Consolidação dos critérios internacionalistas	Exigência de periódicos Scopus/WoS para estrato A1; ausência de periódicos nacionais no topo do Qualis da área	4,2% nos estratos 6–7 (47ª/48 áreas); déficit de ~9,8 p.p.; posição inalterada apesar do crescimento documentado da produção
Quadrienal 2017: Pior resultado relativo da série	Qualis Periódicos da área	3,7% nos estratos 6–7; pior resultado relativo da série; déficit de 10,3 p.p.; único ciclo com recuo do percentual
Quadrienal 2021: Qualis Referência	Qualis Referência apoiado no índice H do Google Acadêmico, com ajuste da classificação por JCR, CiteScore e SPELL (CAPES, 2022)	7,5% nos estratos 6–7; melhora absoluta, mas déficit relativo persiste; H-index mediano dos programas nota 6–7 varia de 2,50 a 9,78, com alta heterogeneidade
Quadrienal 2025: Maior índice de exigência operacional das CSA	Ficha 2021–2024: H-index Scopus entre quatro métricas de citação, em análise qualitativa; até 15 relatos de egressos em três faixas históricas, estrutura compartilhada com Economia e Comunicação; não elegibilidade aos conceitos 6 e 7 de programas sem internacionalização	9,4% nos estratos 6–7 (47ª/50 áreas); déficit de 8,8 p.p. em relação à média do sistema (18,2%); maior do que em 2010

Fonte: CAPES (2010b); CAPES (2013a); CAPES (2017a); CAPES (2022); CAPES (2026).

5.3.1 Explicações Endógenas: Análise e Decomposição

Quatro fatores endógenos podem contribuir para a sub-representação da Área 27. Esses fatores não são causas independentes do paradoxo, mas características estruturais que operam ao lado da explicação institucional examinada na seção 5.3.2.

O primeiro fator é o efeito de diluição pela expansão acelerada de programas. De 119 programas na trienal 2013 para 181 em 2025 (+52%), a maior parte dos ingressantes concentra-se nos estratos 3 e 4, rebaixando o percentual relativo de excelência. Contudo, a expansão por si só não explica o paradoxo: Ciências Agrárias I, com 213 programas, a maior área do sistema de avaliação, alcança 22,5% de excelência na 17ª posição.

O segundo fator é a concentração geográfica e institucional da excelência. Dos 17 programas nos estratos 6 e 7 em 2025, 12 (70,6%) estão no Sudeste, com 9 apenas no estado de São Paulo. Esse padrão é estruturalmente persistente: em 2016, todos os 5 programas de excelência já estavam no Sudeste (CAPES, 2016), e os três ciclos quadrienais subsequentes não alteraram substancialmente essa concentração, sugerindo que há barreiras institucionais e de infraestrutura que transcendem os critérios avaliativos.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

O terceiro fator é a incompatibilidade entre critérios de excelência e o perfil de produção real da área. A produção brasileira da Área 27 é majoritariamente nacional: no período 2017–2020, a área produziu 31.303 artigos em 1.395 periódicos, dos quais apenas 7.092 (22,65%) estavam indexados no Scopus ou WoS, e apenas 2.758 (8,81%) no Web of Science (CAPES, 2022). Esse perfil não é uma deficiência, mas característica estrutural de uma área cuja tradição científica se construiu predominantemente em torno de problemas e periódicos brasileiros (Moed, 2008).

A relação entre o perfil da produção científica da Área 27 e seu reconhecimento no sistema de avaliação depende dos critérios de classificação estabelecidos nas fichas de avaliação elaboradas pela própria área para cada ciclo avaliativo. Na ficha referente ao quadriênio 2021–2024, documento que definiu os quesitos, itens e critérios utilizados para avaliar os programas nesse período, a classificação da produção científica baseava-se no Qualis Periódicos, conforme o modelo adotado pelas demais áreas da CAPES. Para o ciclo 2025–2028, a nova ficha de avaliação estabelece critérios distintos, segundo os quais o enquadramento dos periódicos no estrato máximo (MB – Muito Bom) passa a exigir classificação A ou A* na lista ABDC, pontuação igual ou superior a 2 na lista ABS ou posicionamento no primeiro quartil (Q1) do JCR ou do SJR. Nesse modelo, os periódicos classificados exclusivamente pelo SPELL, sistema de indexação e avaliação de periódicos científicos que contempla a produção acadêmica das áreas de Administração, Ciências Contábeis e Turismo, permanecem limitados ao estrato B (Bom), mesmo quando considerado o ajuste decorrente de sua presença na coleção SciELO Brasil (CAPES, 2025c). Cabe destacar que as listas ABS e ABDC também incluem periódicos não indexados nas bases Scopus e Web of Science (WoS), assim como nem todos os periódicos indexados nessas bases alcançam o primeiro quartil. Portanto, o percentual de 22,65% representa uma estimativa da ordem de grandeza da produção científica potencialmente elegível ao estrato máximo, e não uma mensuração exata dessa proporção.

O quarto fator é a heterogeneidade interna da Área 27. A área reúne subáreas com perfis muito distintos: Administração de Empresas tem maior tradição de publicação em periódicos internacionais; Administração Pública e Ciências Contábeis têm produção predominantemente nacional; Turismo tem base internacional ainda emergente. Essa heterogeneidade dificulta a definição de patamares únicos de excelência e beneficia as subáreas com maior inserção internacional.

Uma simulação ilustrativa ajuda a dimensionar o efeito de diluição: mantido o denominador de 2013 (119 programas), os 17 programas de excelência de 2025 corresponderiam a 14,3%, 3,9 p.p. abaixo da média do sistema de avaliação, ante os 8,8 p.p. observados atualmente. O exercício não constitui decomposição nem controle estatístico, pois supõe que o numerador observado em 2025 seria o mesmo sem a expansão, e parte dos programas de excelência pode ter relação direta ou indireta com o próprio crescimento e amadurecimento do sistema. A diferença remanescente, portanto, não é tratada como evidência da explicação institucional; indica apenas que a expansão, isoladamente, não parece dar conta de toda a distância observada. Em 2021, Serviço Social (25,0%), Direito (12,6%) e Economia (13,2%) já superavam os 7,5% da Área 27.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

5.3.2 A Explicação Institucional: Assimetria na Calibração das Fichas de Avaliação

Os quatro fatores endógenos não esgotam o padrão observado. Áreas com produção de vocação predominantemente nacional, como Direito (135 programas, com 14,1% de excelência) e Serviço Social (36 programas, 30,6% nos estratos de excelência), alcançam densidades superiores às da Área 27, e Direito tem porte comparável ao da Área 27. Esse contraste não identifica a causa da diferença, mas justifica examinar uma explicação adicional: a calibração diferenciada das fichas de avaliação.

Tabela 4 – Matriz de codificação e índice de exigência operacional das fichas de avaliação: sete áreas da Grande Área de Ciências Sociais Aplicadas (Ciclos 2021–2024 e 2025–2028)

Dimensão codificada (código 2021–2024/2025–2028)	**Área 27**	**Serv. Soc.**	**Plan. Urb.**	**Eco.**	**Dir.**	**Com.**	**Arq.**
D1. Classificação de periódicos (acesso ao estrato máximo)	1/2	1/1	1/1	1/2	1/1	1/1	1/0
D2. Métricas bibliométricas individuais de docentes	1/2	0/0	0/0	0/0	0/0	1/1	0/0
D3. Teses e dissertações (volume e publicação derivada)	1/2	1/1	1/1	1/1	1/1	1/1	1/1
D4. Produtos técnicos e tecnológicos (cotas e tipos)	1/1	0/1	1/0	1/1	1/1	1/1	1/0
D5. Acompanhamento de egressos (profundidade temporal)	2/2	1/1	1/1	2/2	1/1	2/2	1/1
D6. Gradação formal de internacionalização (notas 6 e 7)	2/1	1/1	0/1	0/1	0/0	1/1	1/1
Índice de exigência operacional (0 a 12)	**8/10**	**4/5**	**4/4**	**5/7**	**4/4**	**7/7**	**5/3**
Programas nos estratos 6 e 7 em 2025 (%)	9,4	30,6	18,4	15,8	14,1	15,8	16,1

Legenda: Área 27: Administração Pública e de Empresas, Ciências Contábeis e Turismo; Serv. Soc.: Serviço Social; Plan. Urb.: Planejamento Urbano e Regional/Demografia; Eco.: Economia; Dir.: Direito; Com. Comunicação, Informação e Museologia; Arq. Arquitetura, Urbanismo e Design.

Nota: escala por dimensão: 0 = exigência ausente ou avaliação qualitativa aberta; 1 = parâmetro quantitativo ou estrutura formal presente, sem condicionamento a bases ou listas internacionais e sem caráter único na Grande Área de Ciências Sociais e Aplicadas; 2 = parâmetro condicionado a bases ou

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

listas internacionais, ou exigência estrutural mais alta ou única entre as sete áreas. As regras de atribuição, a comparação detalhada dos parâmetros de cada ficha (Tabela S1) e a matriz com a evidência textual de cada código (Tabela S2) estão no material suplementar, depositado no repositório da pesquisa: <https://osf.io/je4cw/?view_only=cbaab02127c849f6be62a0cca5f0e38b.>. Em cada célula, o primeiro código refere-se aos instrumentos do quadriênio 2021–2024, que produziram as notas de 2025, e o segundo aos projetados para o quadriênio 2025–2028; D6 considera também os Documentos de Área. A densidade de excelência refere-se à Quadrienal 2025 após os pedidos de reconsideração (CAPES, 2026). Fonte: elaboração própria com base nas fichas de avaliação acadêmico e profissional e nos Documentos de Área das sete áreas nos dois ciclos (para a Área 27, CAPES, 2025b, 2025c, 2025d; para as demais áreas, Tabela S3).

A Tabela 4 apresenta a matriz de codificação e o índice resultante nos dois ciclos. Nos instrumentos que produziram as notas de 2025, a Área 27 alcança 8 dos 12 pontos, ante índices entre 4 e 7 nas demais áreas (média de 4,8); nos projetados para o ciclo 2025–2028, alcança 10 pontos, ante índices entre 3 e 7 (média de 5,0). Em nenhuma dimensão, nos dois ciclos, outra área recebe código superior ao da Área 27, mas a distância é desigual. Comunicação, Informação e Museologia (índice 7 nos dois ciclos) compartilha a estrutura de acompanhamento de egressos em três faixas, adota índice h docente como indicador e exige produção técnica na modalidade profissional; Economia compartilha a estrutura de egressos e, para o ciclo de 2025–2028, também vincula a classificação de periódicos a bases de impacto.

Entre os ciclos, a exigência da Área 27 aumentou na classificação de periódicos, nas métricas individuais e no volume de teses e dissertações examinadas e, a regra que vinculava a internacionalização aos conceitos 6 e 7 passou da ficha ao Documento de Área, que diferencia os conceitos 6 e 7 pelo grau de maturidade da internacionalização (CAPES, 2025b), gradação que, no ciclo 2025–2028, também consta dos documentos de Serviço Social, Planejamento Urbano, Economia, Comunicação e Arquitetura. Essa heterogeneidade resulta das escolhas das comissões ao operacionalizar, com autonomia normativa, os mesmos quesitos gerais, padrão consistente com E1 (DiMaggio & Powell, 1983). As Diretrizes Comuns para o ciclo avaliativo de 2025–2028 (CAPES, 2025a) consagram a comparabilidade como princípio formal, mas sua operacionalização se dá no eixo intraárea, sem instrumento que verifique a proporcionalidade das réguas entre áreas da mesma Grande Área.

A Tabela 5 apresenta as análises de sensibilidade. A posição da Área 27 no topo do índice se mantém nas três especificações alternativas e na retirada de qualquer dimensão, nos dois ciclos. Como nenhuma área supera a Área 27 em qualquer dimensão, a ordenação não se inverte sob nenhum conjunto de pesos não negativos, inclusive pesos proporcionais aos dos itens das fichas. A distância, contudo, depende da especificação: nas fichas de 2021–2024, a Área 27 empata com Comunicação na codificação binária, na regra absoluta e quando se retira a dimensão de internacionalização; em 2025–2028, empata

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

com Comunicação apenas na codificação binária. O achado que os dados sustentam é, portanto, que a Área 27 opera com o maior índice de exigência operacional do conjunto, e não que opera em patamar isolado das demais áreas.

Tabela 5 – Análises de sensibilidade do índice de exigência operacional (Ciclos 2021–2024 e 2025–2028).

Especificação	2021–2024: Área 27	2021–2024: maior vizinha	2021–2024: média das vizinhas	2025–2028: Área 27	2025–2028: maior vizinha	2025–2028: média das vizinhas
Original (0 a 2, pesos iguais)	8	7 (Com.)	4,8	10	7 (Com., Eco.)	5,0
Codificação binária (presença de exigência)	6 (empate)	6 (Com.)	4,5	6 (empate)	6 (Com.)	4,5
Regra absoluta (nível 2 só para bases ou listas internacionais)	6 (empate)	6 (Com.)	4,5	8	6 (Com., Eco.)	4,7
Retirada de uma dimensão por vez (intervalo)	6 a 7	5 a 6 (Com.)	n.c.	8 a 9	5 a 7 (Eco.)	n.c.

Legenda: Área 27: Administração Pública e de Empresas, Ciências Contábeis e Turismo; Eco.: Economia; Com.: Comunicação, Informação e Museologia; n.c.: não calculado. Os intervalos representam a variação do índice de exigência operacional após a exclusão individual de cada dimensão. A indicação de empate corresponde à igualdade de pontuação entre a Área 27 e a área vizinha com maior índice.

Nota: A coluna "maior vizinha" apresenta o maior índice entre as seis áreas utilizadas na comparação. Em ambas as avaliações e em todas as especificações analisadas, nenhuma área vizinha apresenta, em qualquer dimensão, pontuação superior à da Área 27. As informações relativas aos critérios de excelência nos Documentos de Área 2025–2028 de Direito e Serviço Social não estavam integralmente legíveis nos arquivos disponibilizados. Caso Direito adote gradação equivalente, seu índice passaria de 4 para 5, sem alterar a posição relativa da Área 27; o índice de Serviço Social permaneceria inalterado. Os resultados desagregados por área e especificação encontram-se no material suplementar.

Fonte: elaboração própria.

A primeira dimensão de assimetria é o sistema de classificação de periódicos. Nas fichas de 2021–2024, todas as sete áreas, inclusive a Área 27, classificavam a produção pelo Qualis Periódicos. Nas fichas de 2025–2028, o Qualis deixa de ser a referência e cada área adota procedimento próprio: Direito, Planejamento Urbano e Regional/Demografia e Serviço Social reconhecem a indexação no SciELO ou valorizam explicitamente periódicos nacionais; Economia e Comunicação, Informação e Museologia classificam por indicadores de impacto; a Área 27 atribui o estrato máximo (MB) a periódicos com A ou A* na lista ABDC, ABS igual ou superior a 2 ou primeiro quartil no JCR ou no SJR, e limita ao estrato B os periódicos classificados pelo SPELL, mesmo com o ajuste pela presença no SciELO Brasil (CAPES, 2025c).

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Periódicos nacionais que não alcancem essas posições nas listas internacionais ficam, por regra, fora do estrato máximo, o que é consistente com E2 no plano dos instrumentos. Convém delimitar o alcance desse dado: a restrição opera sobre a classificação da produção em periódicos, uma entre as dimensões da ficha, e não sobre a arquitetura avaliativa completa, que inclui ainda Formação, Impacto e Programa. Seu efeito pode ser ampliado pelo acoplamento entre essa dimensão e o indicador de H-index do corpo docente, parcialmente apoiado na mesma base de menor cobertura, e pela produção derivada de teses e dissertações, avaliada pela mesma classificação. A produção nacional não desaparece da avaliação, mas, pela regra da ficha 2025–2028, não contribui para o topo da dimensão de classificação de periódicos.

A segunda dimensão são as métricas bibliométricas individuais de docentes. Na ficha de 2021–2024, a Área 27 examinava o H-index Scopus como uma de quatro métricas de citação, ao lado do H-index e do i10-index do Google Acadêmico e do índice H do SPELL, em análise qualitativa (CAPES, 2025d); Comunicação adotava a proporção de docentes com índice h superior à mediana da área como componente de um item. Na ficha projetada para o ciclo 2025–2028, a mediana do H-index Scopus e do índice H Spell do corpo docente torna-se indicador quantitativo do item 3.3 da ficha (CAPES, 2025c).

A terceira dimensão é a profundidade exigida no acompanhamento de egressos. A Área 27 demanda até 15 relatos de trajetória em três faixas históricas de cinco anos, cobrindo trajetórias de até quinze anos, exigência compartilhada por Economia (Área 28) e por Comunicação, Informação e Museologia (Área 31), que adotam estrutura equivalente nos dois ciclos. As quatro demais áreas adotam exigências menores: Serviço Social faculta a escolha de apenas um período; Direito exigia 5 egressos dos últimos cinco anos e, em 2025–2028, adota janela de até dez anos; Planejamento Urbano e Regional/Demografia e Arquitetura, Urbanismo e Design avaliam egressos de janela única, sem a estrutura de três faixas. Nessa dimensão, portanto, a diferença envolve quatro das seis áreas vizinhas.

O gradiente capturado pelo índice indica que a assimetria não é binária. Direito, Planejamento Urbano, Arquitetura e Serviço Social têm índices entre 3 e 5; Economia (5 e 7) compartilha o acompanhamento de egressos e, para o ciclo 2025–2028, a classificação por bases de impacto; Comunicação (7 nos dois ciclos) aproxima-se da Área 27 em quase todas as dimensões; a Área 27 (8 e 10) é a única a combinar, no ciclo avaliativo de 2025–2028, classificação de periódicos vinculada a listas internacionais, indicador quantitativo de H-index e o maior volume de teses e dissertações examinadas. Essa hierarquia de exigência não encontra, nos documentos examinados, justificativa explícita em diferencial equivalente de desempenho histórico ou de internacionalização entre as áreas.

A análise conjunta do índice de exigência operacional e da densidade de excelência em 2025 possui caráter descritivo, não causal. A Área 27, com o maior índice de exigência nos instrumentos avaliativos de 2021–2024 (8), registra a menor densidade de excelência da Grande Área (9,4%), enquanto as demais áreas, com índices entre 4 e 7, alcançam percentuais entre 14,1% e 30,6% nos estratos superiores. Essa relação, contudo, é fraca e não monotônica. Comunicação, cujo índice (7) se

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

aproxima do observado na Área 27, alcança densidade de 15,8%, enquanto Serviço Social, Planejamento Urbano e Direito, embora compartilhem o mesmo índice (4), registram densidades distintas, de 30,6%, 18,4% e 14,1%, respectivamente. Esses resultados são compatíveis com a expectativa analítica E1, mas evidenciam que o nível de exigência dos critérios avaliativos, considerado isoladamente, não é suficiente para explicar as diferenças observadas na densidade de excelência. Isso sugere a influência de fatores endógenos, discutidos na seção 5.3.1, bem como de outros elementos não contemplados neste estudo.

A distinção entre diagnóstico e recomendação é fundamental para a interpretação dos resultados. A legitimidade dos critérios adotados pela Área 27 não é questionada, e o crescimento de 240% na ocupação dos estratos de excelência entre 2013 e 2025 ocorreu em um período de progressiva elevação das exigências avaliativas. Contudo, o desenho metodológico deste estudo não permite estabelecer uma relação causal entre esses fenômenos. Os resultados demonstram que a Área 27 reúne, em todas as especificações analisadas, os critérios de maior exigência operacional da Grande Área, isoladamente ou empatada com Comunicação, ao mesmo tempo que registra a menor proporção de programas nos estratos 6 e 7. Essa combinação suscita a hipótese de que os parâmetros adotados possam limitar o reconhecimento de programas cuja qualidade seria suficiente para alcançar esses estratos sob critérios comparáveis aos das áreas vizinhas. A confirmação dessa hipótese, entretanto, dependeria de medidas independentes de qualidade, não contempladas nesta pesquisa. Assim, enquanto a sub-representação da Área 27 nos estratos de excelência constitui um resultado empiricamente documentado, a possibilidade de sub-reconhecimento da qualidade permanece uma interpretação plausível, cuja validação exige evidências adicionais.

5.4 Posição Comparativa com as Demais 50 Áreas (Quadrienal 2025)

A Tabela 6 apresenta o posicionamento da Área 27 no ranking das 50 áreas avaliadas pela CAPES, destacando as áreas da Grande Área de Ciências Sociais Aplicadas (marcadas com ★):

Tabela 6 – Posicionamento comparativo da Área 27 no ranking das 50 áreas avaliadas pela CAPES (Quadrienal 2025).

Pos.	Área	Progs.	Nota 6	% 6	Nota 7	% 7	% 6+7
1ª	Ciências Biológicas III	37	5	13,5%	10	27,0%	40,5%
2ª	Ciências Biológicas I	60	13	21,7%	8	13,3%	35,0%
3ª	Astronomia / Física	63	9	14,3%	12	19,0%	33,3%
4ª	Química	76	12	15,8%	12	15,8%	31,6%
5ª	Serviço Social (Área 32) ★	36	7	19,4%	4	11,1%	30,6%
...	...	...	...	...	...	...	...

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Pos.	Área	Progs.	Nota 6	% 6	Nota 7	% 7	% 6+7
28ª	Plan. Urbano e Regional / Dem. ★	49	6	12,2%	3	6,1%	18,4%
35ª	Arquitetura, Urbanismo e Design ★	62	6	9,7%	4	6,5%	16,1%
36ª	Comunicação, Informação e Museologia ★	95	9	9,5%	6	6,3%	15,8%
37ª	Economia ★	76	6	7,9%	6	7,9%	15,8%
41ª	Direito ★	135	13	9,6%	6	4,4%	14,1%
46ª	Educação	190	12	6,3%	6	3,2%	9,5%
47ª	**ADM Pública, Ciências Cont. e Turismo ★**	**181**	**9**	**5,0%**	**8**	**4,4%**	**9,4%**
48ª	Ensino	177	6	3,4%	5	2,8%	6,2%
49ª	Ciências Ambientais	135	2	1,5%	4	3,0%	4,4%
50ª	C. e Humanidades para Educação Básica	13	0	0,0%	0	0,0%	0,0%
MÉDIA	**SISTEMA CAPES**	**91**	**489**	**10,7%**	**338**	**7,4%**	**18,2%**

Fonte: CAPES (2026). Cálculo próprio. ★ = áreas da Grande Área de Ciências Sociais Aplicadas. Área 27 destacada em azul. Tabelas completas disponíveis em repositório aberto: https://osf.io/je4cw/?view_only=cbaab02127c849f6be62a0cca5f0e38b

A Área 27 ocupa a 47ª posição entre as 50 áreas avaliadas pela CAPES, situando-se entre as últimas colocações do ranking nacional. Essa posição evidencia que sua menor densidade de excelência não se restringe à comparação com áreas das Ciências Exatas e Biológicas, mas também se manifesta de forma expressiva em relação às demais áreas da Grande Área de Ciências Sociais Aplicadas. O contraste com Serviço Social é particularmente significativo: enquanto essa área ocupa a 5ª posição geral, com 30,6% dos programas nos estratos 6 e 7, a Área 27 registra apenas 9,4%, uma diferença de 21,2 pontos percentuais entre áreas pertencentes à mesma Grande Área. Essa disparidade ocorre simultaneamente às diferenças nos níveis de exigência operacional dos critérios avaliativos, documentadas na seção 5.3.2, reforçando a pertinência de examinar a comparabilidade dos parâmetros adotados pelas respectivas comissões de área.

Os resultados evidenciam um contraste relevante: a Área 27 registra o maior índice de exigência operacional nos dois ciclos avaliativos examinados e, simultaneamente, a menor proporção de programas nos estratos de excelência entre as áreas da Grande Área de Ciências Sociais Aplicadas.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Portanto, a maior exigência dos critérios avaliativos não corresponde, necessariamente, a uma maior ocupação proporcional dos estratos 6 e 7.

Esse contraste suscita uma questão institucional relacionada ao princípio da comparabilidade entre áreas no Sistema de Avaliação da CAPES. A adoção de parâmetros e pesos substantivamente distintos para o reconhecimento da excelência pode comprometer a equivalência das classificações atribuídas a programas de diferentes áreas. Quando os requisitos para alcançar uma mesma nota apresentam níveis de exigência desproporcionais, torna-se necessário examinar em que medida essas classificações expressam padrões efetivamente comparáveis de excelência acadêmica. Essa questão não pressupõe a redução do rigor avaliativo, mas aponta para a importância de uma calibração dos critérios que assegure maior consistência e proporcionalidade ao reconhecimento da excelência no sistema.

6. Discussão

Os resultados indicam que a menor presença da Área 27 nos estratos de excelência não decorre do descumprimento das regras comuns, mas do modo como a área as calibra, e esta seção examina o alcance dessa constatação. A seção 6.1 especifica o mecanismo de desacoplamento por calibração discricionária e o diferencia do desacoplamento institucional clássico, da reatividade das métricas e da implementação descentralizada de políticas. A 6.2 analisa os efeitos performativos da régua e a governança de sua calibração. A 6.3 confronta a interpretação proposta com duas explicações rivais: a adaptação disciplinar legítima e a elevação deliberada dos padrões de qualidade. A 6.4 discute a possível ocorrência do mecanismo em outros sistemas nacionais de avaliação, e a 6.5 apresenta as limitações do estudo.

6.1 Desacoplamento por Calibração Discricionária: Especificação do Mecanismo

O desacoplamento clássico descreve duas configurações: a separação entre política e prática, em que organizações adotam formalmente regras que não implementam, e a separação entre meios e fins, em que regras são implementadas sem conexão avaliável com os resultados pretendidos (Bromley & Powell, 2012). Em ambas, o fenômeno ocorre dentro da organização que responde à pressão institucional, e o núcleo do conceito é alguma forma de descumprimento ou de implementação simbólica. O mecanismo identificado neste estudo difere em três aspectos. O primeiro é o lócus: o desacoplamento não ocorre dentro de cada área, mas entre unidades pares de um sistema federado, formalmente isomórficas e submetidas ao mesmo regulador. O segundo é a natureza: não há descumprimento nem implementação simbólica; todas as comissões cumprem integralmente os quesitos e os operacionalizam com rigor, e o que diverge é a calibração dos parâmetros que traduzem os quesitos em exigências práticas. O terceiro é o resultado: a consequência não é a lacuna entre discurso e prática de uma organização, mas a assimetria horizontal entre unidades que o próprio sistema declara comparáveis, com efeitos distributivos sobre reconhecimento e recursos.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

O mecanismo também não se reduz a conceitos vizinhos. A reatividade descreve a resposta dos avaliados às métricas (Espeland & Sauder, 2007); o desacoplamento por calibração discricionária descreve a heterogeneidade dos avaliadores na construção das métricas, operando a montante da reatividade. A literatura sobre implementação descentralizada de políticas trata da tradução de diretrizes gerais por agentes locais, mas tipicamente sem o mandato explícito de comparabilidade que caracteriza sistemas de avaliação com escala única de notas; é esse mandato que converte a variação de calibração em problema de validade, e não em mera diversidade administrativa. Em síntese, propõe-se a seguinte definição: desacoplamento por calibração discricionária é a configuração em que um regulador central define quesitos e escala comuns e delega a subunidades autônomas a calibração operacional dos parâmetros, de modo que a aderência formal uniforme coexiste com exigências práticas substantivamente heterogêneas, produzindo assimetria horizontal entre unidades formalmente comparáveis.

Três condições permitem distinguir o desacoplamento por calibração discricionária de formas legítimas de adaptação descentralizada. A primeira é a existência de um mandato formal de comparabilidade entre as unidades avaliadas. Na ausência desse compromisso institucional, diferenças na calibração dos parâmetros podem representar adaptações legítimas às especificidades de cada área. A segunda é a consequência distributiva, caracterizada pela possibilidade de as diferenças de calibração afetarem o acesso ao reconhecimento institucional ou a recursos vinculados à escala comum de avaliação, como os disponibilizados pelo PROEX. A terceira consiste na ausência de justificativas documentadas que fundamentem essas diferenças em características específicas dos objetos avaliados, especialmente nos perfis de produção científica e impacto das respectivas comunidades acadêmicas. A presença simultânea dessas três condições caracteriza uma situação em que a uniformidade formal dos critérios convive com exigências substantivamente distintas, sem que o sistema assegure sua compatibilidade com o princípio de equivalência das notas. É essa desconexão entre a estrutura normativa comum e a operacionalização dos parâmetros avaliativos que fundamenta o conceito de desacoplamento por calibração discricionária, em diálogo com Bromley e Powell (2012). Na ausência de qualquer dessas condições, conceitos como implementação diferenciada, adaptação local ou heterogeneidade de calibração oferecem interpretações mais adequadas. No caso analisado, as duas primeiras condições encontram respaldo documental, enquanto a terceira apresenta variações entre as dimensões avaliativas, conforme discutido na seção 6.3.

6.2 Efeitos Performativos da Régua e Governança da Calibração

Os conceitos de comensuração e reatividade ajudam a compreender não apenas o que a régua mede, mas o que ela produz (Espeland & Stevens, 1998; Espeland & Sauder, 2007). Rankings e métricas são performativos: redefinem o que conta como qualidade, redistribuem atenção e recursos dentro das organizações avaliadas e alteram as próprias práticas que pretendiam apenas registrar. No caso da Área 27, a estrutura da ficha cria incentivos para o redirecionamento da produção a periódicos internacionais

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

indexados, para a seleção de temas com maior probabilidade de publicação nessas bases e para estratégias de formação orientadas aos indicadores, com possíveis custos para a agenda de problemas nacionais que constitui a vocação histórica de parte relevante dos programas. Esses efeitos independem de qualquer juízo sobre as intenções da comissão: decorrem da mecânica da mensuração.

Essa perspectiva também contribui para compreender a governança dos critérios avaliativos. As comissões de área são constituídas por integrantes da própria comunidade acadêmica avaliada, selecionados por processos que contam com a participação das associações científicas. Nesse contexto, os parâmetros de avaliação tendem a evoluir por um processo de acumulação progressiva de exigências, no qual critérios estabelecidos em ciclos anteriores são incorporados, aperfeiçoados e, eventualmente, ampliados nos ciclos seguintes. Essa trajetória pode ser observada na utilização do índice H da Scopus, empregado na classificação de periódicos em ciclos anteriores (CAPES, 2013a), examinado como uma das quatro métricas de citação na ficha do quadriênio 2021–2024 (CAPES, 2025d) e, finalmente, convertido em indicador quantitativo com peso definido na ficha 2025–2028 (CAPES, 2025c).

Essa evolução não implica atribuir responsabilidades individuais às comissões, mas reconhecer que os instrumentos avaliativos são construções institucionais que se consolidam ao longo do tempo e cuja revisão demanda mecanismos deliberativos capazes de ultrapassar os limites de cada área. Nesse sentido, a contribuição central deste estudo concentra-se na análise dos parâmetros de avaliação, examinando seu nível de exigência, sua comparabilidade com os critérios adotados pelas áreas vizinhas e suas potenciais consequências distributivas.

6.3 Adaptação Disciplinar Legítima ou Assimetria de Calibração?

Uma explicação rival merece resposta direta: as diferenças entre as fichas poderiam refletir adaptação legítima às especificidades epistemológicas de cada disciplina, e não assimetria disfuncional. Culturas de publicação, orientações profissionais e padrões de internacionalização variam entre as sete áreas, e o próprio sistema atribui às comissões a tarefa de traduzir os quesitos gerais para seus contextos. Este estudo acolhe a premissa: parte da heterogeneidade documentada é adaptação esperada e desejável. O critério que permite distinguir adaptação de assimetria combina dois elementos. O primeiro é o mandato de comparabilidade: quando o sistema determina que notas iguais representem igual valor, a variação de calibração deixa de ser questão interna de cada área e passa a ser objeto legítimo de exame interáreas. O segundo é a relação entre a régua e o perfil da própria área que a adotou: uma adaptação legítima aproxima os parâmetros do perfil de produção e impacto da comunidade avaliada, e os dados indicam o movimento oposto na Área 27, uma vez que os critérios de classificação no estrato máximo se distanciam do perfil da produção científica da própria área, na qual aproximadamente três quartos dos artigos são publicados em periódicos não indexados na Scopus ou na Web of Science (WoS). A combinação entre o maior nível de exigência operacional entre as sete áreas e o descompasso com o perfil da produção científica, especialmente nas dimensões de classificação de periódicos e métricas individuais, sustenta a interpretação de assimetria de calibração. Essa constatação, contudo, não exclui

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

a possibilidade de que diferenças observadas nas demais dimensões decorram de adaptações legítimas às especificidades disciplinares.

A aplicação desse critério às seis dimensões avaliativas permite distinguir diferentes situações. No acompanhamento de egressos, a estrutura de classificação em três faixas também é adotada por Economia e Comunicação, sugerindo que as diferenças em relação às demais áreas podem refletir concepções disciplinares distintas sobre os resultados da formação. Na avaliação de teses e dissertações e de produtos técnicos e tecnológicos, as variações nos volumes exigidos e nas modalidades reconhecidas são compatíveis com diferentes tradições acadêmicas e formas de inserção profissional. Quanto à internacionalização, os critérios adotados pela Área 27 no ciclo 2021–2024 encontravam correspondência, embora com menor nível de exigência, em Arquitetura, Comunicação e Serviço Social. No ciclo 2025–2028, a diferenciação entre os conceitos 6 e 7 com base na maturidade da internacionalização passou a integrar os critérios de todas as áreas vizinhas, exceto Direito, cujo Documento de Área não permitiu verificação integral. Em contraste, as dimensões de classificação de periódicos e métricas individuais concentram as evidências mais relevantes de possível assimetria, uma vez que o acesso aos níveis superiores de desempenho depende de listas e bases internacionais cuja cobertura da produção científica da Área 27 é limitada. Assim, embora as diferenças epistemológicas e os perfis de produção justifiquem parte da heterogeneidade identificada, a interpretação de assimetria de calibração concentra-se nas dimensões em que os parâmetros avaliativos se mostram menos alinhados às características da própria comunidade acadêmica avaliada.

Uma segunda objeção diz respeito à trajetória histórica da Área 27: a adoção de critérios mais exigentes não poderia representar uma estratégia legítima de elevação dos padrões de qualidade, em vez de uma disfunção do sistema avaliativo? Essa interpretação é pertinente e reconhecida neste estudo. A orientação internacionalista adotada pela área ao longo dos ciclos avaliativos constituiu uma escolha deliberada, fundamentada no propósito de fortalecer a qualidade e a inserção internacional da pós-graduação brasileira. Nesse período, o número de programas nos estratos de excelência cresceu 240% entre 2013 e 2025, de 5 para 17, acompanhado pela consolidação de programas brasileiros em rankings internacionais, embora o desenho metodológico desta pesquisa não permita atribuir esses avanços diretamente à estratégia adotada. O argumento, portanto, não questiona a legitimidade dessa trajetória, mas a adequação dos parâmetros de excelência ao perfil da produção científica da própria área, uma vez que a predominância de indicadores internacionais pode limitar o reconhecimento de outras manifestações de qualidade acadêmica. As diretrizes da CAPES, ao estabelecerem que os referenciais avaliativos devem considerar o desenvolvimento efetivo dos programas e as especificidades da área (CAPES, 2021b, item b), reforçam a necessidade de compatibilizar padrões internacionais de excelência com a diversidade da produção científica nacional. Assim, a questão central não reside na orientação internacionalista, mas no possível descompasso entre os critérios adotados e as características da comunidade acadêmica avaliada, conforme evidenciado pela análise dos instrumentos.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

6.4 Implicações para Outros Sistemas de Avaliação

O mecanismo proposto é candidato a operar em qualquer sistema nacional de avaliação que combine três características: escala única de notas com pretensão de comparabilidade, painéis disciplinares autônomos e delegação da calibração operacional a esses painéis. O Research Excellence Framework britânico, a Valutazione della Qualità della Ricerca italiana e o Excellence in Research for Australia compartilham essa arquitetura, com painéis que definem critérios de qualidade e regras de uso de indicadores por área sob um quadro comum, e a literatura registra variação substantiva entre painéis no uso de métricas e na valorização de tipos de produto (Wilsdon et al., 2015). Esses sistemas constituem, portanto, contextos relevantes para investigar a possível ocorrência do desacoplamento por calibração discricionária em diferentes ambientes institucionais. A questão central é verificar se as assimetrias horizontais identificadas neste estudo representam uma característica estrutural de sistemas de avaliação que combinam critérios comuns e autonomia disciplinar na definição dos parâmetros, ou se decorrem de particularidades do modelo brasileiro. Essa distinção demanda investigações comparativas internacionais, contempladas na agenda de pesquisas futuras apresentada nas considerações finais.

6.5 Limitações

Cinco limitações metodológicas delimitam o alcance dos resultados e das conclusões deste estudo. A primeira decorre de seu caráter documental e descritivo-normativo, que não permite estabelecer relações causais. Embora a análise demonstre que a Área 27 adota os critérios de maior exigência operacional entre as sete áreas examinadas, não é possível determinar em que medida essa característica influencia sua menor participação nos estratos de excelência, admitindo-se, inclusive, a possibilidade de uma relação bidirecional. Também não foram identificadas alterações específicas nos critérios avaliativos que possam ser associadas à menor densidade relativa de excelência registrada pela Área 27 em 2017 (3,7%), enquanto a simulação do efeito de diluição possui caráter ilustrativo e não permite isolar a influência de outros fatores. A segunda limitação refere-se à indisponibilidade de dados de produtividade desagregados por programa para o quadriênio 2021–2024, o que impossibilitou verificar em que medida diferenças recentes de desempenho acadêmico poderiam contribuir para os resultados observados. A terceira diz respeito ao escopo da análise, restrito aos critérios formalmente estabelecidos nas fichas, sem examinar sua aplicação concreta pelas comissões de avaliação. Desse modo, eventuais práticas deliberativas capazes de atenuar ou ampliar as assimetrias identificadas permanecem fora do alcance da investigação. A quarta limitação relaciona-se ao número de áreas examinadas. Embora as sete áreas constituam um conjunto adequado à comparação documental estruturada, não oferecem base suficiente para inferências estatísticas robustas. Além disso, a associação observada entre o índice de exigência operacional e a densidade de excelência é descritiva e não monotônica, indicando que fatores endógenos e outras variáveis não observadas também devem ser considerados. Por fim, a quinta limitação decorre da sensibilidade do índice às escolhas de codificação.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

As análises de sensibilidade confirmam que a Área 27 mantém o maior nível de exigência operacional em todas as especificações testadas, embora em algumas delas compartilhe essa posição com Comunicação e a magnitude da diferença varie conforme os critérios de codificação. Acrescenta-se que os instrumentos referentes ao ciclo 2025–2028 ainda não produziram resultados avaliativos, de modo que sua comparação com a densidade de excelência observada em 2025 permite identificar mudanças nos níveis de exigência, mas não avaliar seus efeitos sobre a classificação futura dos programas.

7. Considerações Finais

Este estudo examinou a sub-representação da Área 27 da CAPES nos estratos de excelência, ampliando até a Avaliação Quadrienal de 2025 o período investigado por Maccari e Nishimura (2014) e incorporando uma análise comparativa dos critérios estabelecidos nas fichas de avaliação das sete áreas de Ciências Sociais Aplicadas em dois ciclos avaliativos. Os resultados evidenciam um contraste institucional relevante: a Área 27 apresenta o maior índice de exigência operacional entre as áreas examinadas, tanto nos instrumentos utilizados na avaliação de 2025 quanto naqueles definidos para o ciclo 2025–2028, enquanto registra a menor proporção de programas nos estratos de excelência. Essa posição de maior exigência se mantém em todas as análises de sensibilidade, embora seja compartilhada com a área de Comunicação, Informação e Museologia em algumas especificações, confirmando a robustez do resultado diante de diferentes critérios de codificação. No plano teórico, o estudo propõe o conceito de desacoplamento por calibração discricionária, caracterizado pela coexistência de diretrizes institucionais comuns e parâmetros operacionais substantivamente distintos entre unidades formalmente comparáveis. Ao especificar as condições que diferenciam esse mecanismo de formas legítimas de adaptação disciplinar, a pesquisa contribui para a compreensão das tensões entre autonomia avaliativa, comparabilidade institucional e reconhecimento da excelência acadêmica.

Entre 2013 e 2025, a Área 27 ampliou de 5 para 17 o número de programas nos estratos de excelência, mas permanece na 47ª posição entre as 50 áreas avaliadas pela CAPES, com uma densidade de excelência de 9,4%, inferior à média das seis áreas vizinhas (16,8%) e à do conjunto do sistema (18,2%). Esse contraste encontra respaldo nas evidências documentais relacionadas às expectativas analíticas do estudo, embora com diferentes níveis de sustentação. A expectativa E1 é corroborada pela heterogeneidade sistemática dos critérios estabelecidos nas fichas de avaliação para quesitos comuns às sete áreas (Tabela 4). A expectativa E2 encontra sustentação no plano dos instrumentos avaliativos, uma vez que os critérios de classificação de periódicos para o estrato máximo, no ciclo 2025–2028, privilegiam listas e bases internacionais cuja cobertura corresponde a aproximadamente um quarto da produção científica da Área 27, sem que isso permita inferir diferenças na qualidade dos programas. Por sua vez, a expectativa E3 não foi submetida à verificação empírica, mas orienta a discussão das implicações institucionais desenvolvidas a seguir.

A compreensão desse cenário envolve cinco fatores estruturais: o possível efeito de diluição associado à expansão do número de programas; a concentração geográfica e institucional dos programas

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

de excelência; o descompasso entre o perfil da produção científica da Área 27 e os critérios utilizados para sua classificação; a heterogeneidade entre as subáreas que a compõem; e as diferenças na calibração dos parâmetros avaliativos entre áreas submetidas a diretrizes comuns. Embora esses fatores contribuam para interpretar o padrão observado, o desenho metodológico do estudo não permite determinar sua importância relativa nem estabelecer relações causais entre eles e a distribuição dos programas nos estratos de excelência.

No plano normativo, o princípio da comparabilidade interáreas, estabelecido no Regulamento da Avaliação Trienal de 2013 (CAPES, 2013b) e reafirmado nos parâmetros dos ciclos quadrienais subsequentes (CAPES, 2017b; CAPES, 2021b; CAPES, 2025a), pressupõe que notas equivalentes expressem padrões comparáveis de excelência acadêmica. Os resultados deste estudo evidenciam a necessidade de mecanismos institucionais capazes de verificar a proporcionalidade dos critérios adotados pelas diferentes áreas de uma mesma Grande Área, de modo a assegurar a efetividade desse princípio.

Essa necessidade, contudo, não implica estabelecer metas de distribuição de notas ou uniformizar os resultados das avaliações. A densidade média de excelência das seis áreas vizinhas (16,8%), calculada com base no total de seus programas, constitui uma referência descritiva, e não um parâmetro a ser alcançado, uma vez que diferenças de maturidade, estrutura institucional, desempenho, internacionalização e trajetória podem justificar distribuições distintas. O que se busca preservar é a equivalência de significado das classificações, assegurando que programas avaliados com a mesma nota, ainda que pertencentes a áreas diferentes, sejam submetidos a padrões de exigência de magnitude comparável. Nesse sentido, as Diretrizes Comuns 2025–2028 (CAPES, 2025a), ao preverem a definição *a posteriori* dos pontos de corte dos indicadores com base nos dados do conjunto de programas e reafirmarem a comparabilidade como princípio do sistema, oferecem fundamento normativo para o desenvolvimento de mecanismos de verificação da proporcionalidade entre os critérios avaliativos.

A partir dessas constatações, propõe-se uma agenda de verificação e eventual recalibração dos critérios avaliativos estruturada em três horizontes. No curto prazo, recomenda-se examinar ajustes nos parâmetros da própria ficha da Área 27, especialmente a ponderação do índice H da Scopus ou sua complementação por indicadores que contemplem de forma mais abrangente o perfil da produção científica da área. Adicionalmente, sugere-se estabelecer uma possibilidade de reconhecimento qualitativo, fundamentado em evidências documentadas, para que periódicos nacionais de referência possam alcançar o estrato máximo, preservando os critérios internacionais de excelência. No médio prazo, recomenda-se a realização de simulações prévias a eventuais alterações, mediante a aplicação contrafactual dos critérios de áreas vizinhas aos dados dos programas da Área 27 e a análise de sensibilidade dos pontos de corte, cuja definição *a posteriori*, com base no conjunto de programas avaliados, está prevista nas Diretrizes Comuns 2025–2028 (CAPES, 2025a). No longo prazo, propõe-se institucionalizar, no âmbito do Conselho Técnico-Científico da Educação Superior (CTC-ES), mecanismos periódicos de verificação da proporcionalidade dos critérios entre áreas de uma mesma

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Grande Área, acompanhados do monitoramento de possíveis efeitos não intencionais, como a inflação de notas e a redução da capacidade discriminatória da escala avaliativa. Essas medidas não pressupõem flexibilização dos padrões de excelência, mas o aperfeiçoamento da calibração dos parâmetros e de seus respectivos pesos, visando fortalecer a consistência e a comparabilidade do sistema de avaliação.

Como desdobramento, a agenda de pesquisas futuras compreende quatro frentes complementares: (a) ampliar a análise comparativa para outras Grandes Áreas da CAPES, a fim de verificar se as assimetrias de calibração identificadas em Ciências Sociais Aplicadas também se manifestam em outros campos do conhecimento; (b) investigar historicamente os momentos em que os critérios da Área 27 passaram a divergir daqueles adotados pelas áreas vizinhas, especialmente quanto à valorização de indicadores internacionais, buscando identificar os processos institucionais associados a essas trajetórias e, mediante desenhos de pesquisa apropriados, suas possíveis implicações causais; (c) realizar simulações contrafactuais com a aplicação dos critérios de áreas como Direito e Serviço Social aos dados dos programas da Área 27, associadas a medidas independentes de qualidade acadêmica, para examinar se a sub-representação nos estratos de excelência corresponde a um possível sub-reconhecimento da qualidade e estimar os efeitos atribuíveis à calibração; e (d) investigar a ocorrência do desacoplamento por calibração discricionária em outros sistemas nacionais de avaliação, como o Research Excellence Framework (REF), a Valutazione della Qualità della Ricerca (VQR) e o Excellence in Research for Australia (ERA), avaliando o alcance e as condições de aplicação do mecanismo teórico proposto.

Os resultados oferecem à comissão da Área 27, ao CTC-ES e à comunidade acadêmica uma base documental para examinar a comparabilidade interáreas. Mais do que questionar o rigor dos critérios avaliativos, indicam a importância de que a excelência acadêmica seja reconhecida por parâmetros proporcionais e comparáveis, em consonância com os princípios que fundamentam o Sistema de Avaliação da CAPES.

Declaração de conflito de interesse

Os autores declaram que não há conflito de interesse.

Declaração de contribuição do autor

Emerson Antonio Maccari
conceitualização (líder), curadoria de dados (líder), análise formal (líder), investigação (líder), metodologia (líder) administração do projeto (principal), redação - rascunho original (líder), redação - revisão e edição (líder).

Declaração de disponibilidade de dados da pesquisa

Todo o conjunto de dados de apoio aos resultados deste estudo foi disponibilizado em: https://osf.io/je4cw/?view_only=cbaab02127c849f6be62a0cca5f0e38b

Declaração de uso de IA

A ferramenta de inteligência artificial ChatGPT foi utilizada para auxiliar na revisão técnico-gramatical do texto e normalização de referências.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Referências

Alperin, J. P., Muñoz Nieves, C., Schimanski, L., Fischman, G. E., Niles, M. T., & McKiernan, E. C. (2019). How significant are the public dimensions of faculty work in review, promotion and tenure documents? eLife, 8, e42254. https://doi.org/10.7554/eLife.42254

Association of Business Schools (ABS). (2021). Academic Journal Guide 2021. https://charteredabs.org/academic-journal-guide-2021/

Balbachevsky, E. (2005). A pós-graduação no Brasil: novos desafios para uma política bem-sucedida. In C. Brock & S. Schwartzman (Eds.), Os desafios da educação no Brasil (pp. 285–314). Nova Fronteira.

Barata, R. C. B. (2016). Dez coisas que você deveria saber sobre o Qualis. Revista Brasileira de Pós-Graduação, 13(30), 13–40.

Bispo, M. S. (2020). Contradições da pós-graduação em administração brasileira. Revista Eletrônica de Ciência Administrativa, 19(2), 173–200.

Bispo, M. de S. (2024). Implicações da colonização acadêmica nos periódicos científicos brasileiros em administração. Revista de Administração Contemporânea, 28(1), e240049. https://doi.org/10.1590/1982-7849rac2024240049.por

Bowen, G. A. (2009). Document analysis as a qualitative research method. Qualitative Research Journal, 9(2), 27–40.

Bromley, P., & Powell, W. W. (2012). From smoke and mirrors to walking the talk: Decoupling in the contemporary world. Academy of Management Annals, 6(1), 483–530.

Cabral, A. (2025). Avaliação da pós-graduação: uma análise comparativa entre Brasil e Espanha. Revista Ibero-Americana de Estudos em Educação, 20, e19158.

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES). (2010a). Relatório de Avaliação 2007–2009, Trienal 2010: Área de Avaliação 027: Administração, Ciências Contábeis e Turismo. CAPES/MEC.

CAPES. (2010b). Resultados Finais: Avaliação Trienal 2010 [planilha eletrônica]. CAPES/MEC. https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/avaliacao/avaliacao-trienal-2010/resultados-finais

CAPES. (2013a). Relatório de Avaliação 2010–2012, Trienal 2013: Área de Avaliação 027: Administração, Ciências Contábeis e Turismo. CAPES/MEC.

CAPES. (2013b). Regulamento para a Avaliação Trienal 2013 (2010–2012). CAPES/MEC. https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/avaliacao/avaliacao-trienal-2013/02022022_REGULAMENTOTRIENAL2013.pdf

CAPES. (2016). Documento de Área 2016: Administração Pública e de Empresas, Ciências Contábeis e Turismo. CAPES/MEC.

CAPES. (2017a). Relatório de Avaliação Quadrienal 2017, Área 27. CAPES/MEC.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

CAPES. (2017b). Parâmetros da Avaliação Quadrienal 2017 (ciclo 2013–2016). CAPES/MEC. https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/avaliacao/avaliacao-quadrienal-2017/sobre-a-quadrienal/parametros

CAPES. (2019). Documento de Área: Administração Pública e de Empresas, Ciências Contábeis e Turismo (Área 27). CAPES/MEC.

CAPES. (2021a). Resultados da Avaliação Quadrienal 2021. CAPES/MEC.

CAPES. (2021b). Parâmetros da Avaliação Quadrienal 2017–2020. CAPES/MEC. https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/avaliacao/avaliacao-quadrienal/avaliacao-quadrienal-2017-2020/parametros

CAPES. (2022). Relatório de Avaliação, Área 27, Avaliação Quadrienal 2021. CAPES/MEC.

CAPES. (2025a). Diretrizes Comuns da Avaliação Quadrienal 2025–2028. CAPES/MEC. https://www.gov.br/capes/pt-br/acesso-a-informacao/acoes-e-programas/avaliacao/avaliacao-quadrienal/avaliacao-quadrienal-2025-2028

CAPES. (2025b). Documento de Área 2025–2028: Administração Pública e de Empresas, Ciências Contábeis e Turismo (Área 27). CAPES/MEC.

CAPES. (2025c). Ficha de Avaliação Acadêmico e Profissional, Área 27, Quadriênio 2025–2028. CAPES/MEC.

CAPES. (2025d). Ficha de Avaliação 2025: Administração Pública e de Empresas, Ciências Contábeis e Turismo, Área 27 (quadriênio 2021–2024). CAPES/MEC.

CAPES. (2026). Resultados da Avaliação Quadrienal 2025 – Programas Acadêmicos e Profissionais (238ª, 239ª, 240ª e 241ª Reuniões do CTC-ES) [Data set]. CAPES/MEC. https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/avaliacao/27052026_Resultado_2826627_Resultados_da_Avaliacao_Quadrienal_2025_para_publicacao_25_5___senha.xlsx

Dalglish, S. L., Khalid, H., & McMahon, S. A. (2020). Document analysis in health policy research: The READ approach. Health Policy and Planning, 35(10), 1424–1431. https://doi.org/10.1093/heapol/czaa064

Dantas, F. (2004). Responsabilidade social e pós-graduação no Brasil: ideias para (avali)ação. Revista Brasileira de Pós-Graduação, 1(2), 160–172.

Dawes, R. M. (1979). The robust beauty of improper linear models in decision making. American Psychologist, 34(7), 571–582. https://doi.org/10.1037/0003-066X.34.7.571

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizations. American Sociological Review, 48(2), 147–160.

DORA, San Francisco Declaration on Research Assessment. (2012). San Francisco Declaration on Research Assessment. DORA. https://sfdora.org

Espeland, W. N., & Sauder, M. (2007). Rankings and reactivity: How public measures recreate social worlds. American Journal of Sociology, 113(1), 1–40.

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Espeland, W. N., & Stevens, M. L. (1998). Commensuration as a social process. Annual Review of Sociology, 24, 313–343.

Hicks, D., Wouters, P., Waltman, L., De Rijcke, S., & Rafols, I. (2015). Bibliometrics: The Leiden Manifesto for research metrics. Nature, 520, 429–431.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. Journal of Educational Measurement, 50(1), 1–73.

Larivière, V., & Gingras, Y. (2010). The impact factor's Matthew Effect: A natural experiment in bibliometrics. Journal of the American Society for Information Science and Technology, 61(2), 424–427.

Maccari, E. A., Almeida, M. I. R., Riccio, E. L., & Alejandro, T. B. (2014). Proposta de um modelo de gestão de programas de pós-graduação na área de Administração. RAUSP Management Journal, 49(2), 369–383.

Maccari, E. A., & Nishimura, A. T. (2014). Povoamento dos estratos conceitos 6 e 7 no sistema de avaliação da CAPES pela área de Administração, Ciências Contábeis e Turismo. REAd: Revista Eletrônica de Administração, 20(3), 601–624.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13–103). American Council on Education.

Moed, H. F. (2008). UK Research Assessment Exercises: Informed judgments on research quality or quantity? Scientometrics, 74(1), 153–161.

Morgan, H. (2022). Conducting a qualitative document analysis. The Qualitative Report, 27(1), 64–77. https://doi.org/10.46743/2160-3715/2022.5044

Schwartzman, S. (2022). Pesquisa e pós-graduação no Brasil: duas faces da mesma moeda? Estudos Avançados, 36(104), 227–254. https://doi.org/10.1590/s0103-4014.2022.36104.013

Scott, W. R. (2014). Institutions and organizations: Ideas, interests and identities (4th ed.). SAGE.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Verhine, R. E., & Dantas, L. M. V. (2012). Reflexões sobre o sistema de avaliação da CAPES. Revista de Educação Pública, 18(37), 295–310.

Verhine, R. E., Freitas, A. A. da F., & Robl, F. (2022). Avaliação da pós-graduação no Brasil e nos Países Baixos: um estudo comparado. Avaliação: Revista da Avaliação da Educação Superior, 27(3), 729–752.

Wesley, J. J. (2010). Qualitative document analysis in political science. T2PP Workshop, Vrije Universiteit Amsterdam.

Wilsdon, J., Allen, L., Belfiore, E., Campbell, P., Curry, S., Hill, S., Jones, R., Kain, R., Kerridge, S., Thelwall, M., Tinkler, J., Viney, I., Wouters, P., Hill, J., & Johnson, B. (2015). The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management. HEFCE. https://doi.org/10.13140/RG.2.1.4929.1363

SciELO Preprints - Este documento é um preprint e sua situação atual está disponível em: https://doi.org/10.1590/1982-7849rac2027260184.por

Yang, X., & Su, X. (2024). Responding to the new research assessment reform in China: The universities' institutional hybrid actions. Journal of Higher Education Policy and Management, 47(4), 473–489. https://doi.org/10.1080/1360080X.2024.2446372

Este preprint foi submetido sob as seguintes condições:

- Os autores declaram que os necessários Termos de Consentimento Livre e Esclarecido de participantes ou pacientes na pesquisa foram obtidos e estão descritos no manuscrito, quando aplicável.
- Os autores declaram que a elaboração do manuscrito seguiu as normas éticas de comunicação científica.
- Os autores declaram que estão cientes que são os únicos responsáveis pelo conteúdo do preprint e que o depósito no SciELO Preprints não significa nenhum compromisso de parte do SciELO, exceto sua preservação e disseminação.
- Os autores declaram que os dados, aplicativos e outros conteúdos subjacentes ao manuscrito estão referenciados.
- O manuscrito depositado está no formato PDF.
- Os autores declaram que a pesquisa que deu origem ao manuscrito seguiu as boas práticas éticas e que as necessárias aprovações de comitês de ética de pesquisa, quando aplicável, estão descritas no manuscrito.
- Os autores declaram que uma vez que um manuscrito é postado no servidor SciELO Preprints, o mesmo só poderá ser retirado mediante pedido à Secretaria Editorial do SciELO Preprints, que afixará um aviso de retratação no seu lugar.
- Os autores concordam que o manuscrito aprovado será disponibilizado sob licença Creative Commons CC-BY.
- O autor submissor declara que as contribuições de todos os autores e declaração de conflito de interesses estão incluídas de maneira explícita e em seções específicas do manuscrito.
- Os autores declaram que o manuscrito não foi depositado e/ou disponibilizado previamente em outro servidor de preprints ou publicado em um periódico.
- Caso o manuscrito esteja em processo de avaliação ou sendo preparado para publicação mas ainda não publicado por um periódico, os autores declaram que receberam autorização do periódico para realizar este depósito.
- O autor submissor declara que todos os autores do manuscrito concordam com a submissão ao SciELO Preprints.