

APORTES DO CAMPO DA ANÁLISE DE POLÍTICA EXTERNA (APE) AO TRATAMENTO CIENTÍFICO DE MONITORAMENTO E AVALIAÇÃO DA POLÍTICA EXTERNA (M&A/PE): PEQUENOS AVANÇOS ANTE ENORMES OBSTÁCULOS

Walter Antonio Desiderá Neto¹

No campo da análise de política externa (APE), pesquisas com foco avaliativo são notavelmente pouco representadas entre as diversas áreas temáticas. Essa condição marginal é evidente tanto na limitada quantidade de estudos publicados quanto na quase total ausência dessas ideias nos principais quadros teóricos e conceituais. Para contrariar essa tendência, este artigo exploratório busca identificar e examinar contribuições dentro do campo que visaram direta ou indiretamente fornecer um quadro conceitual, teórico e/ou metodológico para o monitoramento e a avaliação da política externa (M&A/PE). O propósito deste empreendimento é duplo: primeiro, delinear os principais desafios enfrentados por acadêmicos nesta área; e segundo, organizar os progressos mais significativos que eles alcançaram. Apesar da escassez de trabalhos disponíveis, compilar suas percepções é crucial. Esta síntese servirá como uma base para futuras pesquisas voltadas ao desenvolvimento de ferramentas que propiciem a execução dessas tarefas.

Palavras-chave: monitoramento e avaliação; análise de política externa; políticas públicas.

CONTRIBUTIONS FROM THE FIELD OF FOREIGN POLICY ANALYSIS (FPA) TO THE SCIENTIFIC TREATMENT OF FOREIGN POLICY MONITORING AND EVALUATION (M&E/FP): SMALL ADVANCES IN THE FACE OF HUGE OBSTACLES

In the field of foreign policy analysis (FPA), research with an evaluative focus is notably underrepresented among various thematic areas. This marginal status is visible both in the limited number of studies published and in their near-total absence from the discourse of leading theoretical and conceptual frameworks. To counter this trend, this exploratory paper seeks to identify and examine contributions within the field that have directly or indirectly aimed to provide a conceptual, theoretical, and/or methodological framework for monitoring and evaluation of foreign policy (M&E/FP). The purpose of this endeavor is twofold: firstly, to outline the primary challenges that scholars in this area face, and secondly, to organize the most significant progress they have achieved. Even though there is not much research in this area, it is important to combine their findings. This will provide a basis for further research that seeks to create tools that can help these analytical tasks be performed.

Keywords: monitoring and evaluation; foreign policy analysis; public policies.

1. Técnico de planejamento e pesquisa na Diretoria de Estudos Internacionais do Instituto de Pesquisa Econômica Aplicada (Dinte/Ipea). Orcid: 0000-0001-6060-164X. *E-mail:* walter.desidera@ipea.gov.br.

CONTRIBUCIONES DESDE EL CAMPO DE ANÁLISIS DE LA POLÍTICA EXTERIOR (APE) AL TRATAMIENTO CIENTÍFICO DEL MONITOREO Y EVALUACIÓN DE LA POLÍTICA EXTERIOR (M&E/PE): PEQUEÑOS AVANCES FRENTE A ENORMES OBSTÁCULOS

En el campo del análisis de la política exterior (APE), la investigación con un enfoque evaluativo está notablemente mal representada entre las diversas áreas temáticas. Esta condición marginal se manifiesta tanto en el limitado número de estudios publicados como en la ausencia casi total de estas ideas en los principales marcos teóricos y conceptuales. Para contrarrestar esta tendencia, este artículo exploratorio busca identificar y examinar las contribuciones dentro del campo que han apuntado directa o indirectamente a proporcionar un marco conceptual, teórico y/o metodológico para el monitoreo y evaluación de la política exterior (M&E/PE). El propósito de esta investigación es doble: en primer lugar, esbozar los principales retos a los que se enfrentan los estudiosos de este campo y, en segundo lugar, organizar los avances más significativos que se han realizado. A pesar de la escasez de trabajos disponibles, es crucial recopilar sus reflexiones. Esta síntesis servirá de base para futuras investigaciones encaminadas a desarrollar herramientas que permitan llevar a cabo estas tareas.

Palabras clave: monitoreo y evaluación; análisis de política exterior; política pública.

JEL: F50; H83; H87.

DOI: <http://dx.doi.org/10.38116/rtm33art4>

Data de envío do artigo: 5/4/2024. Data de aceite: 16/4/2024.

1 INTRODUÇÃO

Entre as funções de monitoramento e avaliação de políticas públicas (M&A/PP), duas merecem ser destacadas pelo estimado serviço que potencialmente oferecem à sociedade. De um lado, essas atividades produzem um conhecimento científico destinado a garantir melhor desempenho às ações do setor público. De outro lado, elas divulgam informações que servem ao controle, à transparência e à prestação de contas – *accountability* – do governo para com a população. Em última análise, conduzem ao avanço do bem-estar do país e ao fomento da confiança dos cidadãos na democracia.

No que concerne ao acervo de pesquisas que tem se dedicado a essa agenda, a parcela dos trabalhos que tomaram a política externa como objeto se caracteriza como muito e desproporcionalmente baixa. Considerados os benefícios assinalados, o cenário é tão intrigante quanto indesejável. De maneira a investigar suas origens para instigar sua superação, dos fatores relacionados com a produção acadêmico-científica, o primeiro e mais intuitivo passo consiste em buscar explicações no campo da análise de política externa (APE). Embora tenha essa política pública como seu objeto de estudo por excelência, uma consulta geral à literatura especializada confirma como as pesquisas de perfil avaliativo têm relevância mínima entre os eixos temáticos da área. A importância marginal se verifica tanto

pela baixa quantidade de estudos publicados como pela quase absoluta omissão do tema nas reflexões das principais correntes teórico-conceituais.

Na intenção de estimular a reversão dessa tendência, o presente artigo, de perfil exploratório, tem como objetivo identificar e analisar os aportes do campo de APE que tenham se dedicado, de forma mais ou menos direta, a oferecer um tratamento conceitual, teórico e/ou metodológico ao monitoramento e à avaliação da política externa (M&A/PE). A empreitada está orientada pela finalidade de mapear, de um lado, os principais desafios encontrados pelos autores e, de outro, sistematizar os avanços mais robustos que foram capazes de obter. Embora os estudos sejam pouco numerosos, a sistematização de suas contribuições constitui uma tarefa imprescindível para dar direcionamento às pesquisas futuras.

O texto está dividido em quatro seções, contando com esta primeira, introdutória. A segunda seção, por sua vez, conta com uma apresentação prévia de conceitos e ideias fundamentais para a compreensão do tratamento científico geral dado ao M&A/PP. A terceira seção, por seu turno, contém a revisão da literatura na qual se traduz especificamente o propósito definido para este trabalho. A quarta seção abarca as reflexões finais, retomando e sistematizando as ideias apresentadas ao longo do texto que tenham, como traço comum, a utilidade de indicar caminhos para a realização de pesquisas avaliativas sobre a política externa.

2 O TRATAMENTO CIENTÍFICO DO M&A/PP: CONCEITOS E FERRAMENTAS NO DESENHO DA PESQUISA AVALIATIVA

Embora haja uma quantidade razoável de definições conceituais para as ideias de M&A/PP, Cohen e Franco (2008, p. 73-77) sublinham que,

[em qualquer *avaliação* de política pública], a constante é, por um lado, a pretensão de comparar um padrão almejado (imagem-objetivo em direção à qual se orienta uma ação) com a realidade (a medida potencial na qual esta vai ser modificada, ou o que realmente ocorreu como consequência da atividade empreendida) e, por outro lado, a preocupação em alcançar eficazmente os objetivos propostos. (...) Enquanto o acompanhamento ou *monitoramento* é uma atividade gerencial interna que se realiza durante o período de execução e operação [das políticas públicas], a avaliação pode ser realizada tanto antes ou durante a implementação, como ao concluí-la, ou então algum tempo depois, para quando se prevê que o projeto provocou todo o seu impacto.

Colocado de outra forma, conforme defende Garcia (2015, p. 257), a avaliação é uma atividade em que “é julgado o valor de uma [intervenção pública governamental] (...) a partir de um quadro referencial ou padrão comparativo previamente definido (...), tendo como base (...) critérios de aceitabilidade pretendidos”. Por sua vez, o monitoramento é condição ou parte da avaliação, “um processo sistemático e contínuo que, produzindo informações sintéticas e

em tempo eficaz, permite a rápida avaliação situacional e a eventual intervenção que confirma ou corrige as ações monitoradas” (Garcia, 2015, p. 259). Nesse sentido, o monitoramento pode produzir informações para o próprio processo de tomada de decisões e não tem fim em si mesmo.

Nessa direção, em um estudo avaliativo, a pergunta de pesquisa principal e o objetivo geral na qual se traduz correspondem, em última análise, à emissão de um juízo de valor cientificamente orientado. Ao avaliar, pretende-se examinar se as características de uma determinada realidade (os resultados observados) correspondem a uma ou outra categoria qualitativa (bom, regular ou ruim, por exemplo). Do ponto de vista metodológico, esse objetivo corresponde à realização de uma inferência descritiva. De todo modo, não há hipótese previamente estabelecida a ser comprovada.

Três grupos de elementos costumam estar presentes na base de praticamente qualquer estudo que tenha como meta avaliar (ou monitorar) cientificamente uma política pública: i) variáveis e indicadores; ii) critérios; e iii) parâmetros. A definição e a escolha de cada um deles, bem como a garantia da existência de uma interdependência lógica e coerente entre eles, são tarefas teórico-metodológicas fundamentais do M&A/PP. Tomados em conjunto, sugere-se denominá-los *ferramentas de avaliação*. Para esclarecer brevemente, cabe lembrar que as variáveis representam características da realidade que podem ser classificadas em mais de uma categoria, ao passo que os indicadores são os referentes empíricos que dimensionam ou qualificam o seu comportamento. A qualidade das variáveis e dos indicadores depende de sua relação com os atributos do fenômeno avaliado, de sua validação na comunidade científica e da disponibilidade e confiabilidade das fontes de dados. Os critérios são as métricas segundo as quais os juízos de valor são emitidos sobre o objeto avaliado. Eles devem ser adequados à avaliação empreendida e aceitos pela sociedade e pelo meio científico. Alguns exemplos de critérios são eficácia, eficiência, efetividade, sustentabilidade e equidade. Os parâmetros, por fim, são os limites que separam as categorias qualitativas de avaliação, de acordo com cada critério adotado. Eles devem ser definidos com base em estudos comparativos e validados cientificamente. Sua função é prover objetividade e imparcialidade às conclusões.

Contando com ferramentas adequadas, o estudo tem maior potencial para corresponder às expectativas que justificam sua realização, das quais cabe destacar os seguintes pontos: i) atender às demandas práticas dos atores envolvidos; ii) produzir resultados em tempo hábil para que possíveis ações de melhoria sejam planejadas, analisadas e implementadas pela gestão; iii) respeitar valores éticos; e iv) demonstrar rigor científico, procedimental e metodológico, de forma a assegurar confiabilidade e legitimidade aos seus resultados (Garcia, 2015, p. 255).

Atento aos desafios envolvidos na elaboração e obtenção dessas ferramentas, Thoenig (2000) assevera que a busca por níveis elevados de robustez geralmente constitui um objetivo inatingível. Concretamente, essa dificuldade pode se tornar um obstáculo que, ao longo do tempo, conduza ao desestímulo da execução de tais tipos de estudos com regularidade. Nesse sentido, ele sustenta que as avaliações devem ser julgadas com mais flexibilidade. A capacidade para produzir algum conhecimento sobre os programas governamentais sobre os quais se debruçam, com informação útil aos atores envolvidos, especialmente os burocratas e os tomadores de decisão, merece ser examinada positivamente.

O autor chama atenção para a importância e o valor de algumas práticas que são mais frequentemente observadas em diferentes países, para as quais propõe a denominação de quase-avaliações. São esforços da própria administração pública focados na elaboração de variáveis e indicadores que informem, segundo os critérios adotados, o desempenho das atividades executadas para o enfrentamento dos problemas, porém sem a preocupação formal de avaliar os resultados segundo parâmetros de referência, tampouco comprovar a causalidade com as ações efetivamente tomadas. Na visão do especialista, essas práticas devem ser encorajadas, uma vez que, de uma forma geral, entregam resultados de maneira mais ágil e com utilidade prática mais visível aos gestores dos programas governamentais, além de alimentarem de forma progressiva uma cultura de avaliação no setor público. No fim das contas, essas práticas se enquadram mais como monitoramento do que como avaliação, razão adicional pela qual são bastante proveitosas para o aprimoramento das ações governamentais.

3 OS TRABALHOS DEDICADOS AO M&A/PE PRODUZIDOS NO CAMPO DE APE

Conforme argumentado em trabalho anterior (Desiderá, 2022), alguns fatores gerais de ordem teórico-metodológica podem ser responsáveis pela baixa atratividade das pesquisas avaliativas no campo da APE. Eles podem ser resumidos em duas motivações complementares. A primeira está relacionada com o próprio surgimento do campo. Os especialistas se concentraram em explicar como as estratégias de política externa são formuladas e como as decisões são tomadas, porque sua fundação, em grande parte, foi inspirada pela ideia de suprir e fazer um contraponto às teorias sistêmicas das relações internacionais, consideradas insuficientes para prever o comportamento de cada país. A segunda, se refere à complexidade do sistema internacional, no qual vários atores interagem simultaneamente em um ambiente anárquico. Essas características dificultam o estabelecimento de nexos causais robustos entre as iniciativas implementadas por cada um e os resultados obtidos. Logo, a capacidade de extrair aprendizados, aplicar aprimoramentos e fazer recomendações fica comprometida, levando à conclusão plausível de que não valeria a pena levar adiante tais esforços.

Por algumas décadas, essas características fundacionais da área e os obstáculos científicos envolvidos contribuíram para que não fosse despertado interesse dos especialistas na realização do M&A/PE. Em decorrência, não houve uma elaboração de procedimentos e ferramentas que, contando com aportes atentos às especificidades das iniciativas exteriores dos países, dessem suporte científico a essa empreitada. De todo modo, ainda que tenham sido (e de certa forma permaneçam) raros, pouco a pouco alguns pesquisadores se dispuseram a se debruçar sobre o tema.

A primeira obra que se propôs a enfrentar esse desafio de forma sistemática, tendo como objetivo geral mobilizar as contribuições da APE para a discussão do M&A/PE, consiste no livro organizado por Vasquez (1986). Conforme relatado pelo autor, sua discordância com a estratégia agressiva adotada pelo governo do presidente Ronald Reagan (1981-1989) na política externa americana, somada à sua convicção de que as alternativas escolhidas pelo governo eram as piores disponíveis, levaram-no a concluir que os esforços da APE desenvolvidos até então haviam falhado. Para ele, esse fracasso estava relacionado ao princípio básico de que as pesquisas científicas devem apresentar uma justificativa normativa perante a sociedade. Em outras palavras, o campo havia malgrado porque seus aportes não tinham a utilidade prática de informar claramente aos formuladores da política externa quais são as iniciativas mais eficientes e eficazes à sua disposição, ou pelo menos demonstrar quais delas não merecem ser adotadas em hipótese alguma. Com essa frustração mobilizada como argumentação, ele foi capaz de provocar alguns colegas da área a refletirem sobre o tema, os quais aceitaram debatê-lo em dois congressos acadêmicos. Os trabalhos preparados para esses eventos foram consolidados em formato de livro posteriormente nessa obra.

De todo modo, à exceção do organizador da compilação, os demais autores se dedicaram a efetuar inferências avaliativas mediante a realização de estudos de caso. Com isso, seus trabalhos tomaram um perfil analítico, conforme a tradição da área, limitando as chances de se extrair rotinas e ferramentas padronizadas para aplicação em outros casos e, mais importante, a partir dos quais se pudesse produzir resultados sistemáticos e comparáveis. Não obstante, no capítulo introdutório do livro, Vasquez (1986) chamou atenção para os problemas teóricos e metodológicos gerais que rodeiam a empreitada do M&A/PE. Cabe destacar que, na sua visão,

a tarefa mais premente da avaliação da política externa seria examinar se as consequências previstas [pelos formuladores das estratégias] de fato ocorreram. (...) De igual importância para a avaliação são pesquisas que distingam as consequências não intencionais e não previstas [das desejadas] (Vasquez, 1986, p. 9).

Com esse ponto de partida, ele sugeria um caminho para que, em primeiro lugar, os resultados fossem mapeados com rigor científico, para que, em seguida,

a existência de causalidade do comportamento da política externa sobre eles fosse de alguma forma testada, explicitada e pormenorizada. Posto de outra forma, sua preocupação inicial se referia à definição de variáveis e à produção de indicadores que dessem conta de medir os resultados de forma objetiva.

Como exemplo a ser seguido nessa direção, o autor indicou os trabalhos realizados pelos pesquisadores envolvidos no projeto denominado Pesquisa Comparativa sobre Eventos² das Nações (Comparative Research on the Events of Nations – Creon). Juntamente com outras iniciativas similares, essa investigação fez parte de um intenso esforço de construção de bases de dados – em sua maior parte quantitativos – que teve lugar nas décadas de 1960 e 1970. A partir da análise qualitativa de eventos específicos, os pesquisadores buscavam obter valores codificados e padronizados para um conjunto de variáveis. Sua meta era cobrir uma ampla gama de atividades e características, de forma a coletar dados sistemáticos de indicadores sobre o comportamento da política externa dos distintos atores, e ao longo do tempo. Os resultados haviam sido divulgados de maneira robusta alguns anos antes, no livro organizado por Callahan, Brady e Hermann (1982). Nessa obra coletiva da equipe, embora o objetivo geral fosse descrever a conduta internacional individual dos países, foi dedicado um capítulo para a discussão dos efeitos ou das consequências da política externa, contando com a proposição de definições e métodos para mensurá-los de alguma maneira.

Conforme relatam no livro, em suas reflexões iniciais, a equipe do projeto tinha considerado que três variáveis abstratas a respeito das propriedades dos resultados da política externa tinham potencial para o desenvolvimento operacional de dados: i) a geração de situações – definida como a natureza da situação imposta ao ator alvo pela atividade do país executor;³ ii) a amplitude de consequências externas – identificada como a probabilidade de que uma ação estimule a tomada de decisões em um amplo espectro de atores no sistema internacional; e iii) o *feedback* recebido – designado como a proporção de aceitação e de rejeições ao comportamento da política externa de um país. Devido a dificuldades conceituais, metodológicas e operacionais, somente para a última delas o grupo conseguiu levar à frente o desafio de produzir dados (Callahan, Brady e Hermann, 1982, p. 257-258).

Propostas por Hermann (1982), as medidas de *feedback* às iniciativas de política externa tinham como foco estudar parte das consequências efetivas e

2. Os dados de eventos vão além da simples descrição codificada de uma determinada iniciativa de política externa, buscando conter também informações sobre quais atores estavam envolvidos na interação e quando ela ocorreu, entre outros dados adicionais, a depender do esquema de codificação. As fontes utilizadas costumam ser notícias publicadas pelos meios de comunicação (Schrodt, 1995).

3. É comum na literatura de eventos de política externa considerar que, em cada evento, sempre há pelo menos uma diáde de atores: i) o ator A, aquele que executa uma ação; e ii) o ator B, aquele a quem a ação se destina direta ou indiretamente, ou seja, o ator que A pretendia influenciar (mudar seu comportamento ou impedir a mudança).

imediatas do comportamento de cada ator – em contraposição aos resultados potenciais e os pretendidos. A autora seguiu os seguintes procedimentos: para cada ação executada pelos países estudados no projeto Creon, ela examinou no comportamento das demais nações a presença de atitudes de aceitação e de rejeição – aferidas a partir de critérios bem definidos⁴ – ou a ausência de ambas. Com esses dados, para cada país foram calculadas algumas estatísticas, com periodicidade anual: as taxas de aceitação, de rejeição e de *feedback*⁵ relativas ao comportamento geral da política externa, além da taxa de aceitação relativa às rejeições e vice-versa. Mesmo que não sejam medidas específicas de êxito e fracasso, a autora defendia sua aplicabilidade a esforços avaliativos, porque

(...) o *feedback* fornece aos formuladores das estratégias algumas pistas referentes ao desempenho da política externa de seus governos – indícios sobre o que está obtendo sucesso e o que está fracassando. (...) Quando a resposta é positiva ou favorável [aceitação], eles tendem a considerar haver algum grau de êxito. (...) Quando a resposta for negativa [rejeição], eles costumam percebê-la com algum sentido de fracasso (Hermann, 1982, p. 260, tradução nossa).

De toda forma, ao passar para o passo seguinte, nomeadamente, verificar a existência de correlação ou causalidade entre essas medidas e alguns indicadores descritivos do comportamento da política externa desenvolvidos no mesmo projeto,⁶ os testes estatísticos não apresentaram resultados significativos e conclusivos. Apesar de se tratar de uma contribuição promissora, mesmo que seminal, a sensação de frustração com os enormes esforços realizados até ali era generalizada entre os pesquisadores, uma vez que este não era um caso isolado nem inédito em que os dados empíricos elaborados não foram capazes de produzir testes com resultados robustos que embasassem a construção de um corpo teórico em APE.

Em decorrência, conforme indicam Farias e Carmo (2021, p. 5),

com a evolução da disciplina, a análise de [dados de] eventos acabou sendo relegada para o segundo plano [até os anos 2010]. Dois fatores explicam essa marginalização. O primeiro foi o predomínio de teorias sistêmicas, começando com a obra de Waltz (1979). O segundo foi endógeno à própria área de análise de política externa. O trabalho de construção das bases de dados era extremamente subjetivo. A atividade de codificar cada evento exigia amplo conhecimento de outros países, e os autores engajados nessa atividade raramente eram familiarizados com esses contextos. (...) Partia-se da premissa de que a política externa é predominantemente pública, quando tanto os eventos como os próprios atores por trás deles podem ser desconhecidos por décadas em decorrência de sigilo da documentação.

4. No modelo, um exemplo de atitude de aceitação é a assinatura de acordos, enquanto a oposição formal a propostas em organizações internacionais é um exemplo de rejeição.

5. Soma de aceitação e rejeição, considerando que há casos de omissão (países que não se manifestam sobre a ação).

6. Por exemplo, uma maior probabilidade de ações conjuntas de dois ou mais atores serem aceitas do que rejeitadas.

A provocação de Vasquez (1986), nesse contexto, não foi capaz de induzir os especialistas a se envolverem com sua agenda de pesquisa para além daquele fôlego inicial. Passada uma década, as inquietações presentes na argumentação de Clark (1995) revelaram a existência de pesos adicionais para que os pesquisadores permanecessem afastados do M&A/PE. Designadamente, o autor propôs que, para que se efetuasse a avaliação da eficácia da política externa, pelo menos três estágios fundamentais estariam envolvidos: i) a definição dos padrões por meio dos quais os resultados são avaliados⁷ (observação empírica); ii) a determinação do impacto efetivo que as iniciativas dos países têm no mundo real (causalidade); e iii) por fim, a confrontação entre os resultados observados e os critérios de avaliação (escrutínio objetivo).

Estabelecido o raciocínio, Vasquez (1986) se dispôs em seu trabalho a se dedicar exclusivamente à primeira fase da etapa inicial, por considerar um passo imprescindível para que se pudesse avançar para a discussão das demais. O recorte proposto se justificava pelo argumento de que, quando os cidadãos comuns ou os especialistas entram em discussões calorosas sobre suas avaliações da política externa implementada por algum governo, muitas vezes eles não estão de acordo sequer sobre quais eram os objetivos daquela estratégia, um elemento básico para que se possa começar a debater se ela obteve êxito. Com efeito, colocado de forma concisa, o artigo de Clark (1995) revela que, pela complexidade intrínseca dos próprios objetivos da política externa, seu estudo busca efetuar dois esforços muito desafiadores do ponto de vista científico: a identificação e a caracterização.

Com relação à primeira tarefa, entre outras coisas, o autor aponta que com frequência não há uma obrigação jurídica nos regimes políticos de que os objetivos sejam formalmente divulgados pelos governos. Quando ela existe, ou mesmo quando um governo o faz voluntariamente, nada garante que aquilo que é divulgado em documentos oficiais ou expresso nos discursos de autoridades seja verdadeiro. Pode ser um blefe dentro de uma estratégia política não declarada e direcionada a confundir determinados atores com quem compete, por exemplo. Tampouco quando essa divulgação ocorre *a posteriori* a informação é totalmente confiável, pois pode estar enviesada, seja com o objetivo de influenciar positivamente sua avaliação, seja por estar contaminada por interesses políticos presentes.

No caso da segunda tarefa, os argumentos do autor giravam ao redor das dificuldades para a formulação e a escolha de critérios para descrever, classificar e hierarquizar os objetivos da política externa. Em suma, o trabalho de Clark (1995), com uma narrativa bastante sólida, chamou atenção para a existência de enormes desafios na etapa mais preliminar do M&A/PE: a definição do objeto de avaliação. Assim, o autor não apenas adicionava um desafio para sua realização,

7. O objeto de avaliação, as variáveis e seus indicadores, os critérios e os parâmetros.

mas também criava um impasse na agenda de investigação, dado que defendia que sem a superação desta etapa não fazia sentido consumir recursos em esforços para o enfrentamento das demais.

Apesar de não ter se dedicado a explorar essa questão, um pequeno e provocativo artigo de Baldwin (2000), autor influente nos debates teóricos em relações internacionais, caracterizou-se como uma das colaborações explicitamente dedicadas ao tema do M&A/PE que, desde então, rendeu mais resultados dentro do campo da APE. Nesse trabalho, ele propôs a discussão dos conceitos de sucesso e fracasso em política externa. A fundamentação à proposição do debate se dava pelo potencial em contribuir para a formação de um quadro analítico rigoroso e sistemático para avaliar as estratégias internacionais adotadas pelos governos, o qual fosse politicamente relevante para seu país e administrativamente útil à burocracia.

Definindo o sucesso, de forma genérica, como a situação em que os resultados políticos são favoráveis ao país, a colaboração de Baldwin (2000) se concentrou em discutir a multidimensionalidade do conceito. Para ele, ela decorre da existência de uma relação de critérios (eficácia, eficiência, entre outros) que devem ser respeitados necessária e simultaneamente para que determinado caso seja assim categorizado. Posto de outra forma, Baldwin (2000) enumera os requisitos para a identificação dessas experiências na realidade.⁸ Apesar de não trazer avanços substantivos para o tema, deve-se ressaltar que o principal mérito da publicação, do ponto de vista teórico, encontra-se na absorção de aportes do campo de análise de políticas públicas (APP) a uma discussão sobre política externa.

Passados alguns anos, o lançamento de alguns trabalhos demonstrou que o aceno do autor foi capaz de dar uma orientação conceitual ao debate e lançar uma agenda de pesquisa ligeiramente frutífera. A partir de então, em grande parte das publicações no campo da APE, o tema passou a ser trabalhado sob alcunhas que não fazem referência explícita à terminologia conceitual – monitoramento e avaliação – adotada e desenvolvida pela área da APP. Em vez disso, mesmo que de forma incipiente, aos poucos as discussões foram se organizando separadamente ao redor das próprias categorias avaliativas, consideradas como objetos de estudo: i) o sucesso em política externa, conforme sugerido pelo artigo de Baldwin (2000); e ii) sua antinomia, o fracasso em política externa, cujo apelo normativo despertou mais interesse e levou a uma produção ligeiramente maior no campo.

A primeira e mais famosa dessas publicações coube a Johnson e Tierney (2006), cuja obra foi agraciada com o respeitado Prêmio Anual de Melhor Livro,

8. Vale assinalar que um trabalho publicado anteriormente por Kratochwill (1981) trazia uma proposta semelhante, ainda que de maneira mais ensaística.

oferecido pela Associação de Estudos Internacionais (International Studies Association – ISA),⁹ em 2008. O problema de pesquisa que despertou as reflexões desse trabalho foi a perplexidade diante das profundas divergências que facilmente se encontram entre as avaliações realizadas pelos mais distintos observadores (especialistas, líderes políticos, cidadãos comuns, meios de comunicação, entre tantos outros) a respeito do mais básico resultado de um conflito internacional: quem venceu?

Partindo do estudo de um conjunto de cinco casos,¹⁰ os autores buscaram demonstrar como essas discrepâncias se verificam também, no debate público como um todo, na forma de um mesmo observador apreciar cada evento. Para eles, a mudança na forma de priorizar os fatores a serem levados em conta é empreendida sem que haja sequer a tentativa de oferecer justificativas lógicas, além de em muitos casos o próprio espectador não ser nem mesmo capaz de perceber a disparidade nos critérios adotados. Conforme sublinham, tais variações se fazem presentes também caso se comparem as narrativas dominantes sobre cada conflito, tanto na historiografia como em relação aos eventos correntes.

De maneira a descrever a natureza dessas incongruências, ademais de explicar os fatores que induzem à sua ocorrência, a obra conta com uma contribuição teórica que despertou bastante interesse. A sofisticação da tese resulta de sua multidisciplinaridade e da simplicidade da linguagem adotada. Usando a analogia ao mundo dos esportes como estratégia heurística, os autores propõem haver duas abordagens por intermédio das quais o vencedor de uma partida é determinado: pela contagem dos pontos (CdP) ou pela manipulação dos resultados (MdR). Na obra, cada uma delas é tratada teoricamente como um quadro referencial, designando os procedimentos tomados pelo observador para chegar ao julgamento final.

A CdP é caracterizada como um mecanismo no qual a objetividade e a imparcialidade são priorizadas. Nela, o julgamento do observador se baseia em uma métrica bem definida, em que são reunidos os indicadores de algumas variáveis, as quais são selecionadas de acordo com o critério de avaliação pretendido, a fim de que os resultados possam ser esmiuçados de maneira padronizada. Parâmetros previamente determinados, cuja função consiste em pontuar o conjunto de características definidoras das categorias vitória e derrota, são confrontados com as informações coletadas, levando à apuração do caso. Em princípio, esse primeiro

9. A ISA é a principal e mais ampla congregação de especialistas da disciplina de relações internacionais no mundo. Além disso, em parceria com a Universidade de Oxford, a ISA é responsável pela edição de um dos mais importantes periódicos do campo de APE, a revista *Foreign Policy Analysis*. Para mais informações, acessar a página da instituição na internet, disponível em: <https://www.isanet.org/>.

10. Especificamente, os seguintes: i) a Crise dos Mísseis de Cuba (1962); ii) a Ofensiva Tet (1968), no contexto da Guerra do Vietnã (1955-1975); iii) a Guerra do Yom Kippur (1973); iv) a intervenção americana na Somália (1992-1994); e v) a Guerra do Iraque (2003-2011), a qual ainda se encontrava em aberto na data de publicação. Todos eles contam com o envolvimento dos Estados Unidos, embora no terceiro caso, comparado aos demais, a participação tenha se dado de forma bem menos direta.

quadro referencial apresenta as características de um modelo satisfatoriamente representativo das abordagens que se encontram em conformidade com a literatura científica de M&A/PP: a utilização objetiva de ferramentas de avaliação.

Trazendo a ideia para o tema do livro, no caso da avaliação dos resultados de uma guerra ou crise internacional segundo esse quadro referencial, Johnson e Tierney (2006, p. 26) defendem que, na fase de produção do diagnóstico, dois elementos precisam ser levados em consideração: i) os ganhos (e as perdas) materiais aferidos; e ii) os objetivos conquistados (e os inatingidos). Para os autores, além de mapear as informações a respeito desses dois fatores, os dados encontrados devem ter seu valor ponderado de acordo com duas dimensões: i) a importância dada, pelo governo e pela sociedade, aos ganhos e aos objetivos; e ii) o grau de dificuldade envolvido na sua obtenção.

Para finalizar o escrutínio, a execução de uma tarefa preliminar, igualmente imprescindível, é assinalada: a definição da data de encerramento da guerra, que permite estabelecer a partir de quando se considera iniciada a conjuntura pós-conflito. A comparação sistemática das duas realidades, anterior e posterior, baseia a avaliação final. Embora um grau considerável de subjetividade esteja envolvido em todas as etapas desse processo (com destaque para os fatores de ponderação, ancorados em conceitos tão abstratos como importância e dificuldade), os autores sustentam que conclusões baseadas nesse quadro referencial “tendem a se aproximar [umas das outras], de maneira bastante previsível, por se tratar de exames racionais (...) de desempenho” (Johnson e Tierney, 2006, p. 32, tradução nossa). Baseados nessa justificativa, eles defendem que toda avaliação deveria ser realizada dessa forma. Isso se desdobra, ao longo da obra, em uma carga valorativa que conduz à ideia de que esse quadro é justo, correto e melhor.

Diferentemente, quando as percepções sobre os produtos da guerra não seguem esses procedimentos e não possuem sustentação em elementos materiais, os caminhos percorridos para a obtenção desse julgamento constituem o quadro referencial da MdR. Conforme pontuado por eles,

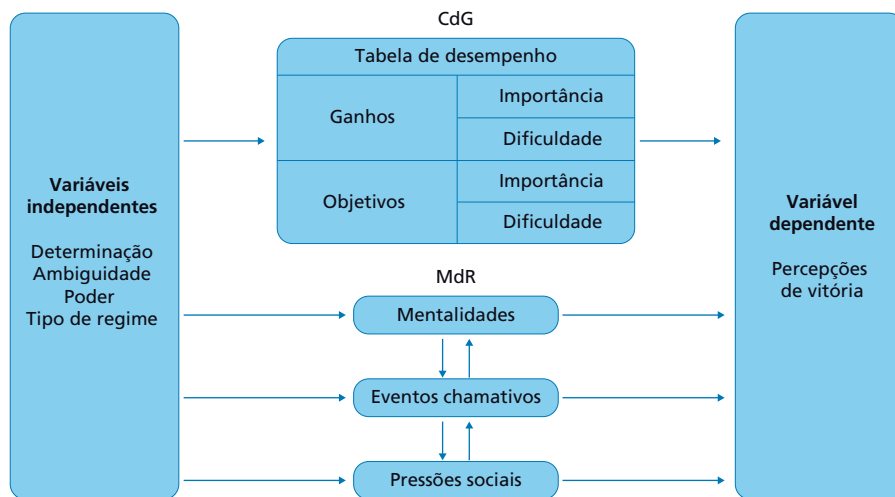
(...) os ganhos e as perdas de cada lado muitas vezes não explicam a percepção de vitória das pessoas, seja em guerras, crises ou outros eventos (...). De fato, ganhos e perdas materiais muitas vezes não explicam três fenômenos distintos: a percepção dominante de vitória e derrota; a diversidade nas percepções de vitória e derrota; e as mudanças nessas percepções ao longo do tempo. (...) [Propõe-se aqui] que as percepções de vitória e derrota não dependem apenas de contar os pontos, mas também são influenciadas por três tipos de fenômenos enraizados na natureza humana: [a diversidade de] mentalidades; [o peso de] eventos chamativos [na memória coletiva]; e [os efeitos de] pressões sociais [sobre os formadores de opinião, bem como o das narrativas dominantes sobre os observadores] (Johnson e Tierney, 2006, p. 38, tradução nossa).

Seja qual for o fator que tenha propiciado essas avaliações que denominam desviantes, a percepção está constituída por vieses de duas naturezas principais: informacional e/ou psicológica. A primeira está relacionada com as limitações que a quantidade e a qualidade dos dados acessados pelo observador lhe impõem. A segunda se refere à maneira pela qual as informações são absorvidas, selecionadas e interpretadas por cada pessoa: processos cognitivos e/ou afetivos, seu conhecimento acumulado, além das influências culturais. Nesse quadro referencial, a subjetividade e a parcialidade ganham terreno. O modelo racional se desvirtua porque os vieses são capazes de gerar confusões ao redor de elementos e procedimentos fundamentais para a aplicação adequada do quadro referencial da CdP. Entre outros pontos, os autores destacam “a ambiguidade do que constitui uma métrica apropriada; (...) a dificuldade de combinar (...) as informações relevantes; e a dificuldade de definir quando uma determinada guerra (...) terminou” (Johnson e Tierney, 2006, p. 42, tradução nossa).

É importante observar que, ao apresentar esse segundo modelo em diálogo com os elementos do primeiro, os autores estabelecem uma oposição normativa das características de cada um. Efetivamente, essa dualidade se traduz como a deturpação, no segundo, das alegadas qualidades do primeiro. Nesse contexto, como as ideias de avaliação e percepção não recebem tratamento conceitual apropriado na obra, elas acabam construindo no trabalho um jogo de palavras com carga valorativa, no qual a primeira é aplicada para a descrição do quadro da CdP, enquanto para a MdR se utiliza a segunda. Essa linguagem exemplifica uma das formas pelas quais se induz no trabalho a ideia de que o segundo modelo é pior, injusto e equivocado.

O último objetivo dos autores consiste em compreender as razões para que haja variação do quadro referencial adotado na avaliação predominante de cada guerra. Em conjunto com os aportes anteriores, a construção teórica da obra se completa com a sugestão de quatro variáveis causais para que, em algum caso, as percepções se ancorem em um modelo ou outro. A cada fator correspondem hipóteses mais ou menos bem definidas para o sentido e a intensidade da causalidade. Em ordem decrescente quanto ao efeito causal, conforme colocado por Johnson e Tierney (2006), as variáveis e os respectivos sentidos tomados pela causalidade são os seguintes: i) contundência dos resultados (quanto mais decisivos e evidentes, predomina o quadro CdP); ii) ambiguidade dos eventos (quanto maior a ambiguidade, predomina a MdR); iii) posição de poder dos envolvidos no conflito (predomina a MdR no sentido de considerar a derrota do mais forte); e iv) tipo de regime político (em democracias, predomina ligeiramente a CdP, enquanto em ditaduras predomina a MdR). A figura 1 contém uma representação esquemática do conjunto da tese na forma de uma cadeia causal.

FIGURA 1
Modelo explicativo da formação de percepções de vitória



Fonte: Johnson e Tierney (2006, p. 78).

Acima de tudo, deve-se ressaltar como o livro apresenta uma virtude indiscutível e muito importante: o papel que desempenhou no estímulo do debate científico sobre o tema. De todo modo, apesar da sua originalidade, bem como da complexidade elevada em que se estabelece a construção do argumento, uma série de limitações tornam muito reduzidas as possibilidades de que as ideias sejam traduzidas em ferramentas avaliativas operacionalizáveis. Sem uma profunda reformulação adaptativa de variados pontos, o construto teórico dificilmente pode se aplicar a áreas temáticas e/ou contextos de atuação da política externa distintos aos dos casos examinados, nos quais predominam questões de segurança. De um modo geral, as variáveis independentes, as hipóteses de pesquisa, o desenho da cadeia causal e a variável dependente apresentam uma quantidade razoável de problemas.

Do ponto de vista mais estrito da metodologia científica, parte das vulnerabilidades prejudica seriamente a validade da proposta teórica. Para se ater aos objetivos deste artigo, cabe chamar atenção para cinco delas: i) a imprecisão conceitual na definição de parte das variáveis, com destaque para a dependente, a respeito da qual não ficam claros até mesmo os valores que ela pode assumir; ii) o elevado grau de subjetividade no processo de atribuição de valores às variáveis, com destaque para as duas primeiras variáveis independentes e para os fatores de ponderação do próprio quadro referencial da CdP; iii) o descompasso na definição dos quadros referenciais, feita a partir de fatores com objetos referentes e níveis de análise completamente distintos; iv) a indefinição da função dos quadros referenciais

na cadeia causal,¹¹ até mesmo de sua relação com a variável dependente, tanto do ponto de vista do mecanismo como do efeito causal; e v) decorrente disso, a desconexão entre as próprias variáveis independentes com a dependente.

Em suma, embora o desenho da cadeia causal conduza à expectativa por conclusões com elevado grau de generalização, o aporte que efetivamente se permite extrair da obra é bem mais limitado. Devido à maneira com que a discussão se desenvolve, bem como ao sentido para o qual é conduzida, do ponto de vista da utilidade geral ao M&A/PE, essa contribuição é ainda mais marginal. Por sua vez, os quadros referenciais de avaliação propriamente ditos, com potencial de aplicabilidade mais direta e evidente em estudos avaliativos, apresentam severos problemas metodológicos. Apenas para um deles é apresentado um desenho lógico geral, contando com ferramentas de avaliação. Entretanto, sua descrição é insuficiente, com escassa clareza na exposição de elementos conceituais e procedimentais. Quanto ao modelo alternativo, os elementos descritos não conformam um quadro referencial, como pretendem os autores. Na verdade, são apresentados fatores que explicam as razões e os mecanismos pelos quais percepções que não se baseiam no outro modelo têm apelo considerável no debate público. Portanto, do ponto de vista prático, a replicação dos quadros para o exame de outros casos é extremamente reduzida. De forma paradoxal, pode-se afirmar que isso se deve precisamente ao elevado grau de subjetividade da proposição teórica, caracterizada pela imprecisão dos seus encadeamentos lógicos.

Fundamentada em uma quantidade robusta de exemplos históricos, a justificativa da pesquisa defendida pelos autores aponta os impactos que as percepções de vitória exercem sobre quatro pontos importantes: os líderes políticos, a sociedade e a democracia, a política externa corrente e a formulação das estratégias futuras. Ao mesmo tempo, conforme relatado, eles defendem que as pessoas deveriam basear suas avaliações em um quadro referencial racional baseado em evidências materiais, devido à sua tendência em conduzir a conclusões mais previsíveis e com menor grau de variabilidade. Para evitar entrar em discussões epistemológicas, mesmo assumindo que tal modelo plenamente objetivo fosse produzido, as virtudes que são apresentadas como motivação para que a sua adoção seja preferível não têm relação explícita, clara ou plausível com a superação dos problemas que são apontados na justificativa do trabalho. Em outras palavras, resulta que o papel das conclusões propriamente ditas de uma avaliação é desprezado pela tese. Sem que explicações sejam apresentadas, a premissa de que quanto maior o grau de objetividade do quadro referencial avaliativo, mais positivos serão os impactos de suas conclusões na realidade, implícita na base de todo o desenho de investigação e da argumentação desenvolvida, não pode ser aceita.

11. Não se sabe se são duas variáveis intervenientes, ou se constituem as duas posições possíveis (ou os extremos de um espectro) de uma só variável interveniente.

Mais do que isso, essa premissa é inadequada ao M&A/PE. Para a ciência, a defesa da objetividade empírica é legítima e tem motivações sólidas, trata-se de um dos próprios elementos fundamentais constituintes do método que a define. Contudo, na avaliação de políticas públicas, com destaque ainda maior no caso de conflitos armados externos, a incorporação de elementos subjetivos – a exemplo de valores, crenças ou preferências sociais de fundo moral e ético – é legítima, necessária e não pode ser sobreposta pela busca por objetividade. Tanto no seu desenho de investigação geral como no desenvolvimento específico das ferramentas de avaliação que serão adotadas, os valores e as preferências devem orientar o exame da adequação dos caminhos alternativos pelos quais se efetuará o escrutínio, posto que são eles que justificam a pesquisa avaliativa.

A condenação do uso de armas de destruição em massa contra civis, sob qualquer circunstância, está ancorada em valores éticos e princípios morais sólidos e legítimos. Contudo, isso não seria considerado em um quadro referencial baseado estritamente em evidências materiais. Numa sociedade na qual tais valores sejam predominantes na formação das preferências, um quadro referencial que os despreza de antemão – pretendendo convencer o público de que, mesmo usando bombas atômicas, um país seja considerado vitorioso – deve ser classificado como um modelo de manipulação dos resultados. Em contrapartida, mesmo no caso de tal sociedade apoiar e justificar tal atitude do ponto de vista ético e moral, é possível que sua adoção seja condenada com base em evidências materiais. A insegurança em nível existencial gerada pela criação e pelo uso dessas armas, principalmente a partir do momento em que se concretiza sua proliferação, acarreta custos materiais em defesa crescentes e imprevisíveis. Não é desprezível o risco de que, em algum momento, eles possam se tornar insustentáveis. Seria racionalmente legítimo que o volume indeterminado das perdas potenciais fosse predominantemente considerado um risco inaceitável em tal sociedade, recebendo importância máxima e prioritária na ponderação ante quaisquer ganhos materiais obtidos. Um quadro referencial em que esses elementos materiais fossem ponderados de outra forma – pretendendo convencer que, mesmo com os custos potenciais resultantes da insegurança sem precedentes e irreversível que foi gerada, um país seja considerado vitorioso – deve ser classificado como um modelo que leva à manipulação dos resultados. Ambos os exemplos apresentam plausibilidade suficiente para que a premissa de Johnson e Tierney (2006), considerando também seu caráter implícito e a fragilidade de suas sustentações, seja questionada e considerada equivocada. Numa visão ortodoxa, como é a valorizada pela própria premissa, todo o desenho de investigação se invalida – desde as perguntas de pesquisa até os resultados obtidos.

Walker e Malici (2011), sob uma abordagem estritamente racionalista, deram seguimento à agenda de investigação com a publicação de um livro cujos objetivos principais consistem em definir conceitualmente o que são erros em política externa e explicar por que eles ocorrem. Do ponto de vista normativo, cabe assinalar que a justificativa apresentada pelos autores para a escolha do tema foi muito similar à de Vasquez (1986): a inquietação diante de escolhas equivocadas e da obtenção de resultados desfavoráveis pela política externa americana. No caso deles, o episódio que catalisou essa angústia foi a invasão do Iraque que se sucedeu aos atentados de 11 de setembro de 2001, no bojo da Guerra Contra o Terror, cujas consequências para a potência hegemônica avaliavam negativamente. O foco da sua investigação se voltou ao caso exclusivo dos Estados Unidos.

No exame do seu aporte teórico-conceitual, cabe indicar que os autores iniciam sua proposta assinalando, em primeiro lugar, a diferença entre erro e fracasso (Walker e Malici, 2011, p. 7-13). Segundo eles, em política externa, erros seriam falhas procedimentais, de fonte cognitiva ou comportamental, ao passo que fracassos seriam falhas substantivas, detectadas somente a partir do exame dos resultados obtidos. Com relação aos primeiros, eles aperfeiçoaram sua definição a partir de uma tipologia baseada no cruzamento de dois eixos de variáveis dicotômicas: (fase do) processo de tomada de decisão (diagnóstico ou prescrição) e comportamento (omissão ou ação). As possíveis combinações resultaram em quatro tipos ideais de erros, conforme o quadro 1.

QUADRO 1
Tipologia geral de erros em política externa

Processo de tomada de decisão	Comportamento	
	Omissão (pragmático)	Ação (dogmático)
Inteligência (diagnóstico)	Erro de detecção (<i>aquém</i>)	Erro de percepção (<i>além</i>)
Política externa (prescrição)	Erro de hesitação (<i>tarde</i>)	Erro de precipitação (<i>cedo</i>)

Fonte: Walker e Malici (2011, p. 11).

Quanto aos fracassos, os autores se concentraram apenas naqueles que são precedidos por erros. Igualmente, eles progrediram sua definição com base em uma tipologia baseada no cruzamento de duas variáveis. Nela, a variável erros – categorizada, então, de forma dicotômica, a partir da agregação em dois blocos dos quatro tipos ideais que propõem (omissão e ação) – é cruzada com a variável situação, a qual assume as posições ameaças (de perdas) e oportunidades (de ganhos). Igualmente, das combinações resultaram quatro tipos ideais de fracassos, conforme o quadro 2.

QUADRO 2

Tipos de situações e tipos de consequências em política externa

Situação	Erro	
	Omissão (aquém/tarde)	Ação (além/cedo)
Ameaças (de perdas)	Fracasso de dissuasão	Fracasso de alarme falso
	↑	↑
	Erro de detecção ou persuasão	Erro de percepção ou precipitação
Oportunidades (de ganhos)	↓	↓
	Fracasso de reafirmação	Fracasso de falsa esperança

Fonte: Walker e Malici (2011, p. 12).

Segundo argumentam esses autores, essas tipologias têm a função analítica de identificar os erros e os fracassos em política externa de forma sistemática, além de prover categorias que forneçam bases comunicativas para sua compreensão e seu exame dentro da comunidade científica. Contudo, conforme bem sublinham e reconhecem, da maneira como estão colocadas as tipologias, não está definido o elemento principal para a detecção dos dois fenômenos na realidade: o padrão moral para definição de certo e errado, ou seja, o parâmetro ou as ferramentas de avaliação que permitam, antes de tipificar erros e fracassos, identificar sua ocorrência. Concentrando-se nos erros, a solução proposta pelos autores baseia-se no modelo da escolha racional. Sinteticamente, depreende-se que um erro em política externa ocorreria quando, numa tomada de decisão, os custos não são minimizados e/ou os benefícios não são maximizados.

Para explicar os motivos que levam os erros a serem observados com considerável frequência, sustentada em uma revisão da literatura de APE, Walker e Malici (2011) assinalam faltar no campo um construto teórico único, ao redor do qual o julgamento normativo a respeito da racionalidade possa ser organizado. Dessa forma, propõem um foco em relações de poder associado a estilos de liderança.¹² Aplicando a teoria dos jogos, os autores tentam demonstrar empiricamente ao longo da obra que

a causa imediata de um erro é a percepção equivocada de um líder (...) das relações de poder [estabelecidas entre o país que ele representa e os outros] (...), a qual leva a cálculos errôneos dos custos e benefícios de cada instrumento de política externa mobilizado pelas estratégias alternativas, incorrendo em fracassos (Walker e Malici, 2011, p. 25, tradução nossa).

Sem deixar de reconhecer os méritos do esforço intelectual, do ponto de vista da sua aplicação em pesquisas avaliativas, duas lacunas devem ser destacadas. Em primeiro lugar, cabe apontar o fato de não se debruçarem sobre o ponto

12. Walker é um dos autores da ferramenta de exame de estilos de liderança denominada análise de códigos operacionais. Para aprofundamento, conferir Schafer e Walker (2006).

elencado por Clark (1995): a identificação do objetivo cujo desempenho será objeto de avaliação. Na obra de Johnson e Tierney (2006), a qual conta com exemplos numerosos e alguns estudos de caso, o objetivo de política externa que se coloca sob avaliação, sobretudo para a condução da argumentação, é bastante claro: vencer uma guerra. Com pretensões mais ambiciosas de generalização, o trabalho de Walker e Malici (2011) toma como premissa a validade do conceito de interesse nacional, do qual derivam as metas de política externa de um país. A partir delas, os autores alegam que seria possível determinar objetivamente quais situações ou estado de coisas constituem custos e quais se caracterizam como benefícios. Entretanto, cabe lembrar que a própria definição daquilo que entra em cada lado da balança é bastante subjetiva. Ela é dependente dos objetivos perpetrados por cada ator, bem como da sua lista de prioridades. Assim, o modelo da escolha racional não garante que, a depender da posição de quem observa, uma mesma estratégia tenha seus resultados avaliados em categorias distintas.

Em segundo lugar, na explicação sugerida pelos autores para a ocorrência dos fracassos em política externa, há uma contradição na argumentação que fragiliza o aporte. Na construção do modelo teórico-conceitual, é defendida a existência de uma relação causal direta dos fracassos com os erros, ou seja, uma presunção de efeito entre a ação individual de um país e o resultado. Contudo, nas reflexões gerais, chama-se atenção para a complexidade do sistema internacional: “os resultados finais, dependem, *entre outras coisas*, da capacidade dos atores envolvidos em uma interação estratégica de tomar novas decisões que cancelem os efeitos de erros cometidos anteriormente” (Walker e Malici, 2011, p. 296, tradução nossa). A contradição tem impactos contundentes sobre a validade do argumento, uma vez que conduz à indefinição do fator causal de um fracasso: um erro prévio, as outras coisas ou uma complexa combinação de ambos. Para além disso, tampouco há uma sugestão objetiva para identificar quando um resultado deve ser considerado o final. A diferença do lapso temporal escolhido por distintos avaliadores pode se caracterizar como fonte adicional de discrepâncias em resultados de estudos avaliativos baseados no quadro teórico que propõem.

Representando outro fruto importante dessa agenda de pesquisa, a revista *Journal of European Public Policy* publicou um número completo, contando com oito artigos, cujo desafio colocado aos autores consistia em propor formas de criar pontes entre os campos de APP e APE. Editado por Oppermann e Spencer (2016a),¹³ o tema eleito como referência para que fosse efetuado o intercâmbio dos aportes acumulados pelas duas áreas consistiu no estudo de

13. Posteriormente, essa coleção foi publicada na forma de livro, sob a organização também de Oppermann e Spencer (2018). Vale mencionar, ainda, que no mesmo ano os autores e um terceiro colega organizaram mais uma obra (Kruck, Oppermann e Spencer, 2018), esta, sobre erros e fracassos nas relações internacionais de um modo geral. Contando com a colaboração de mais quinze autores, são examinados vários casos adicionais, sob óticas muito similares à do outro livro.

iniciativas governamentais implementadas, que se caracterizaram como fiascos – com ênfase na política externa. No conjunto, os artigos abordam, entre outras questões: i) a maneira de definir, identificar, medir e explicar as causas e consequências do fracasso (e, em menor medida, do sucesso) das políticas públicas; ii) na visão pós-positivista, a identificação dos interesses, presentes no jogo político, que estão por trás da construção das narrativas que não apenas rotulam determinadas políticas como fiascos, mas também apontam (um ou poucos) culpados; e iii) os processos de aprendizagem resultantes da experiência de fiascos anteriores, isto é, como, após a fase da avaliação, o ciclo das políticas públicas é retomado nesses casos.

Para os objetivos deste artigo, merecem destaque os aportes de dois trabalhos constantes naquela revista. O primeiro deles, de autoria de McConnell (2016), tem como objetivo, conforme proposto no número da revista, utilizar uma abordagem da APP para entender a natureza e as causas dos fracassos¹⁴ em política externa. Conforme sua definição,

uma política pública fracassa, mesmo que seja bem-sucedida em alguns aspectos mínimos, se fundamentalmente não atingir os objetivos que os proponentes pretendiam alcançar, e a oposição a ela for grande e/ou o apoio for virtualmente inexistente (McConnell, 2016, p. 671, tradução nossa).

Ele considera que as atividades governamentais podem ser classificadas de acordo com três objetivos alternativos: i) orientar processos para produzir políticas públicas; ii) promulgar programas ou decisões que desejam colocar em prática; e iii) gerar impacto na política. A depender da categoria, o autor propõe que os fracassos também assumem formas distintas, podendo ser respectivamente caracterizados como: processuais, programáticos e políticos. Conforme apresentado no quadro 3, para cada tipo de malogro correspondem também critérios de sucesso e características específicas.

QUADRO 3
Formas de fracassos em políticas públicas

Tipo de fracasso	Crítérios de sucesso	Características e sinais de fracasso
Processual: formulação e aprovação	Preservar objetivos e instrumentos de política.	Governo incapaz de produzir os objetivos e instrumentos de política desejados.
	Assegurar a legitimidade.	Processo político ilegítimo.
	Construir uma coalizão político-partidária sustentável.	Ausência de construção de uma coalizão político-partidária sustentável.
	Atrair apoio ao processo.	Oposição ao processo é virtualmente universal e/ou o suporte é virtualmente inexistente.

(Continua)

14. O autor rejeita o uso da palavra fiasco, devido a sua carga valorativa. Em contrapartida, os editores do número da revista, com metodologia pós-positivista, escolheram esse termo exatamente por conta do peso negativo, uma vez que seu trabalho objetiva estudar a construção das narrativas que logram rotular e qualificar no debate público uma política pública como um fiasco (Oppermann e Spencer, 2016b).

(Continuação)

Tipo de fracasso	Crítérios de sucesso	Características e sinais de fracasso
Programático: tomada de decisões e implementação	Implementação em linha com os objetivos.	Apesar de pequenos progressos em direção à implementação de acordo com o pretendido, o programa é afetado por falhas crônicas de implementação.
	Conquistar os resultados desejados.	Algumas pequenas consequências positivas são alcançadas conforme pretendido, mas são oprimidas pela falha em produzir resultados.
	Beneficiar os grupos-alvo.	Pequenos benefícios são acompanhados e ofuscados por danos ao próprio grupo que deveria se beneficiar.
	Satisfazer a critérios altamente estimados na área temática da política pública.	Alguns pequenos sucessos são observados, mas são contaminados pela incapacidade de atender a critérios altamente estimados na área temática da política pública.
	Atrair apoio ao programa.	A oposição aos objetivos do programa, seus valores e os meios de alcançá-los supera os pequenos níveis de apoio.
Político: impacto e repercussões	Colaborar com as perspectivas eleitorais (reputação).	Apesar dos pequenos sinais de benefício, a política pública provoca uma generalizada responsabilização eleitoral e de reputação.
	Facilitar o jogo político.	São observados sinais claros de que a agenda do governo e suas negociações políticas são forçadas a se direcionar a lutar para suprimir uma questão politicamente difícil.
	Promover a trajetória política desejada.	Toda a trajetória do governo corre o risco de ser comprometida.
	Prover benefícios políticos ao governo.	A oposição ao provimento de benefícios políticos ao governo supera os pequenos níveis de apoio.

Fonte: McConnell (2016, p. 673).

Passando para a explicação da sua ocorrência, o autor elenca alguns motivos que tornam muito desafiadora a tarefa de identificar as causas dos fracassos. Entre eles, vale destacar estes quatro: i) muito frequentemente, as explicações dos fracassos são multicausais; ii) existe uma dificuldade metodológica enorme em atribuir o peso de cada uma das múltiplas variáveis independentes sobre o efeito causal; iii) quase sempre, os fracassos são examinados em retrospectiva, fazendo com que a narrativa construída no debate público a respeito deles contamine a escolha das variáveis independentes a serem analisadas numa avaliação científica; e iv) considerado pelo autor o mais importante, atores políticos apoiadores e opositores do governo executor da política pública, mesmo diante das mesmas evidências, constroem seus discursos e fazem suas avaliações individuais com maior e menor grau de complacência com erros, respectivamente, quando o objeto é um programa implementado por seus correligionários ou por seus adversários (McConnell, 2016).

Para que esses problemas não sejam negligenciados ou ocultados, o especialista propõe um quadro referencial que busca exemplificar a maneira pela qual o enquadramento da descrição do fracasso e de sua causalidade é estruturado nas avaliações efetuadas no debate público. Como se pode observar no quadro 4, essas tipificações são resultantes do cruzamento entre três variáveis: i) as perguntas feitas

pelo avaliador ao explicar os fracassos; ii) o ator ao redor do qual está atribuída a parcela principal da culpa; e iii) a afinidade política do avaliador com o governo responsável pela política pública. Com essa proposta, o autor defende que reconhecer a influência das disputas políticas sobre os resultados de uma avaliação, bem como a existência de áreas cinzentas ao longo desse processo, é a chave para avançar na compreensão da própria maneira como as ferramentas de avaliação são construídas.

Apesar de o autor encerrar o texto sem trazer reflexões específicas sobre a aplicabilidade desses aportes ao caso da política externa, de fato suas colaborações são úteis ao M&A/PE. Em primeiro lugar, sua definição de fracasso em políticas públicas, ao incluir como critério, além do exame dos resultados, o coeficiente entre apoio e oposição das forças sociais e políticas ao programa, dialoga com o caso da política externa no sentido de rejeitar a possibilidade de se presumir teoricamente haver um interesse nacional comum em qualquer problema enfrentado por iniciativas da pasta internacional. Em segundo lugar, a especificação das formas pelas quais os fracassos se expressam conformam um quadro analítico de referência muito útil para que, no desenvolvimento de ferramentas de avaliação, analistas se atentem à possibilidade de que, em um caso ou outro, sua aplicabilidade não seja possível de se transportar automaticamente, desenvolvendo estratégias para adaptá-las. Por último, uma vez que as impressões sobre a política externa sofrem grande influência das narrativas colocadas no debate público sobre alguns poucos episódios que são destacados nos meios de comunicação, o quadro referencial para a identificação da forma de enquadramento das causas dos fracassos é muito útil para a construção de ferramentas objetivas de avaliação. Nesse processo, ao efetuarem essa classificação de forma explícita na metodologia, explicando as estratégias adotadas para que, ao mesmo tempo, as ferramentas não sejam contaminadas pelas paixões dos discursos de apoiadores ou de opositores, nem deixem de contemplar os principais argumentos de cada lado, os pesquisadores conferem mais transparência ao seu estudo. Mais do que isso, permitem ao público, seja o geral ou o especializado, apreciar o grau de imparcialidade embutido nas ferramentas construídas.

QUADRO 4
Enquadramentos diferenciados sobre as causas e as consequências de fracassos de políticas públicas

Perguntas realizadas	Ator a quem se atribui a principal parcela de culpa					
	Indivíduo		Instituição ou processo		Sociedade	
	Afinidade do avaliador com o avaliado		Afinidade do avaliador com o avaliado		Afinidade do avaliador com o avaliado	
	Opositor	Apoiador	Opositor	Apoiador	Opositor	Apoiador
O que causou o fracasso?	Interesse individual imprudente, cultivo deliberado do fracasso, negligência, incompetência, julgamento.	Falha no bom senso, erro genuíno, má sorte.	Interesse institucional próprio, arrogância institucional, pontos cegos importantes, capacidade fraca para uma boa tomada de decisão.	Pequena fraqueza em instituições e processos que são sólidos.	Os valores essenciais e/ou os vieses da elite produziram vieses na formulação e fracassos inevitáveis.	A sociedade talvez tenha se distanciado um pouco dos valores e das questões fundamentais.
Os tomadores de decisão deveriam ter previsto e prevenido o fracasso?	Sim, absolutamente, mas eles ignoraram de forma patológica ou foram incapazes de ver os riscos.	Idealmente sim, mas eles enfrentaram circunstâncias difíceis; eles também podem ter tido azar.	Sim, absolutamente, mas a instituição e/ou um processo disfuncional ignorou ou não foi capaz de ver os riscos.	Idealmente sim, mas pequenas deficiências nos processos e procedimentos limitaram a capacidade de prever e prevenir.	Sim, em teoria, mas preconceitos sistemáticos produziram uma cegueira para a possibilidade de fracasso ou mesmo uma preparação para prever (substituída por um preparo para tolerar).	Idealmente sim, mas na prática não se pode prever e prevenir todas as falhas.
Como se pode aprender a partir do fracasso?	Removendo indivíduos ruins do cargo e substituindo-os por outros que sejam mais competentes.	Fazer o possível para treinar melhor os formuladores de políticas, mas não se deve ser muito duro com eles.	Desmontagem drástica ou auditoria da instituição e/ou do processo.	Refinamento ligeiro da política institucional dos procedimentos e/ou dos processos.	As causas do fracasso estão profundamente arraigadas em valores centrais disfuncionais e nos sistemas de poder. A aprendizagem só pode acontecer quando houver uma mudança de paradigma.	Reflexão e talvez refinamento dos valores e das prioridades essenciais e de como são postos em prática, sendo desaconselhado desafiá-los fundamentalmente.

Fonte: McConnell (2016, p. 678).

O segundo artigo que merece destaque no número publicado pela revista é de autoria de Meyer (2016).¹⁵ Aperfeiçoando o trabalho de Walker e Malici (2011), examinado anteriormente, o autor propõe um maior detalhamento dos fracassos em política externa decorrentes dos erros caracterizados como reações que se revelaram estar além ou aquém do necessário. Conforme argumenta Meyer (2016), na tipologia de erros original dos autores em questão, resultam dois tipos ideais adicionais nos quais a variável de diferenciação implícita é a adequação temporal (*timeliness*) do momento da reação (tarde ou cedo, respectivamente, erro de hesitação e erro de precipitação). Em outros termos, trata-se de uma qualidade da resposta dos gestores da pasta internacional aos estímulos externos que não está explicitada pelos autores originais, tampouco tem suas categorias separadas na outra tipologia que eles propõem, isto é, a de consequências de erros.

De forma a corrigir essa imprecisão, delimitar uma tipologia genuinamente de fracassos em política externa e expandir a quantidade de situações contempladas, Meyer (2016) sugere uma reestruturação geral da proposta. Como justificativa, ele defende que a terminologia utilizada originalmente e o grau de detalhamento em que é considerada restringem a aplicabilidade da proposta a casos de defesa e segurança. De fato, sob uma análise menos rigorosa, pelo menos um esforço de adaptação é exigido para que sua adoção dialogue com outras áreas temáticas de atuação de um país. Assim, ele recomenda que, como ponto de partida, os fracassos relacionados com a perda de oportunidades e os associados ao enfrentamento de ameaças devam ser tratados separadamente, resultando em duas tipologias distintas. No trabalho, o pesquisador se dedica exclusivamente ao segundo caso.

Na sua reestruturação, o eixo dos tipos de erros é mantido, mas com alterações. O primeiro ajuste se constitui na formulação de uma dimensão, chamada avaliação da ameaça (gestão de risco), em substituição à categoria denominada originalmente omissão. Ele propõe ainda uma matização dessa dimensão, a qual está baseada na adequação, de um lado, no exame da possibilidade de concretização da ameaça (probabilidade) e, de outro lado, na apreciação da natureza e da severidade das suas possíveis consequências (gravidade). O segundo ajuste, por sua vez, consiste na conversão da categoria oposta, que recebia o nome de ação, em uma dimensão batizada de proporcionalidade da resposta (análise de custo-benefício). Nela, duas nuances são levadas em conta para sua caracterização: a escala (quantidade adequada de recursos mobilizados) e o escopo (limitação do alcance dos efeitos à resolução do problema). Quanto ao eixo situação, que perdeu sua razão de ser com o foco apenas nas situações de ameaça, sua eliminação ensejou ao autor

15. Apesar de não ser pormenorizado neste texto, vale apontar que o artigo de Brummer (2016) também representa uma tentativa de avançar na construção das ideias propostas por Walker e Malici (2011). O autor aplica distintas ferramentas para definir o estilo de liderança de uma amostra de primeiros-ministros britânicos, para então, com esses dados, testar a correlação de cada estilo de liderança com a ocorrência de fracassos em política externa.

singularizar e explicitar a variável *timeliness*, com sua categorização permanecendo a mesma (cedo ou tarde).

O resultado dessa reestruturação pode ser visualizado no quadro 5. Coerente com uma de suas críticas, no lugar de fornecer um rótulo para cada tipo ideal resultante das combinações, ele prefere deixar a terminologia em aberto, apenas indicando exemplos de casos reais pertencentes a domínios temáticos distintos das relações internacionais. Uma vantagem direta dessa estratégia é demonstrar na tabela como um mesmo caso de fracasso pode estar associado a mais de uma característica nas colunas, nas quais se destacam os matizes referentes aos erros. Colocado de outra forma, a grande diferença resultante, em comparação com o quadro de Walker e Malici (2011), está na demonstração de que um fracasso pode derivar, ao mesmo tempo, de problemas na fase de diagnóstico e de prescrição. Além disso, nessa nova versão exclusiva dos casos de situações que representam ameaças, a adequação temporal da resposta (cedo ou tarde) não fica restrita a ter havido erros em uma ou outra fase específica, permitindo constatar respostas tardias, por exemplo, devido a problemas de prescrição no que tange ao escopo ou à escala da ação escolhida como resposta.

QUADRO 5

Tipologia de fracassos em política externa quanto à reação a ameaças transfronteiriças (com casos de exemplo)

Erros	Subdimensões dos erros	Eficiência da reação	
		Aquém do necessário	Além do necessário
		<i>Timeliness</i>	
		Atrasado	Precipitado
Avaliação da ameaça	Probabilidade	Subestimada	Superestimada
		Ataques de 11/9 (2001)	Nuvens de poeira da Islândia (2010)
		Guerra da Crimeia (2014)	
	Severidade	Subestimada	Superestimada
		Crise financeira internacional dos <i>subprimes</i> (2008)	Gripe suína (2009)
		Ataques de 11/9 (2001)	
Proporcionalidade da resposta	Escopo	Insuficiente	Extrapolado
		Crise financeira internacional dos <i>subprimes</i> (2008)	Nuvens de poeira da Islândia (2010)
			Guerra do Iraque (2003)
	Escala	Insuficiente	Extrapolada
		Guerra da Crimeia (2014)	Guerra do Iraque (2003)
			Gripe suína (2009)

Fonte: Meyer (2016, p. 673).
Elaboração do autor.

Essas reflexões não apenas corrigem algumas partes um pouco confusas daquelas tipologias, mas também conduzem a uma conceituação mais refinada para esse conjunto determinado de fracassos em política externa, quais sejam, os que decorrem da maneira de lidar com ameaças externas de toda sorte. Na conceituação do autor, esses fracassos se definem como

falhas em mobilizar os recursos cognitivos e materiais disponíveis para a formulação de políticas que sejam sensíveis ao momento oportuno de agir e/ou proporcionais à gravidade, à probabilidade de ocorrência e/ou à natureza de uma ameaça transfronteiriça (Meyer, 2016, p. 740, tradução nossa).

Publicada no periódico *Review of International Studies*, não poderia deixar de ser mencionada neste levantamento uma publicação recente de Cormac *et al.* (2022), a qual contém uma contribuição construtiva no que diz respeito ao diálogo teórico entre algumas dessas abordagens que se dedicaram ao M&A/PE. O mérito do trabalho reside no esforço de sintetizar, confrontar e integrar aportes cujos procedimentos se baseiam em epistemologias distintas – positivismo e pós-positivismo. Como a análise realizada até aqui evidenciou de maneira superficial, a comunicação entre estudos dessas tradições tem sido bastante incipiente. Para além dessa virtude do trabalho dos especialistas, é inegável reconhecer como a escolha do seu objeto de pesquisa também se configura como um elemento que gera muita curiosidade a respeito do desenho de investigação trilhado pelos autores, sobretudo no que concerne à coleta de dados. A descrição do trabalho e de sua proposta teórica esclarecerá tais apontamentos.

No artigo, os autores têm como objetivo construir uma ferramenta de análise que desempenhe a função de definir parâmetros de sucesso para avaliar um grupo de iniciativas específicas de política externa que, por definição, não têm dados divulgados: as ações do serviço secreto.¹⁶ Na ideia de propor uma definição consensual para essas atividades, como requisitos básicos para sua identificação, os autores indicam que duas características são comumente utilizadas na literatura especializada: i) são ações que têm o objetivo de influenciar eventos fora do território nacional; e ii) cujo patrocinador não está aparente nem é conhecido publicamente. Ainda que proponham uma tipologia geral para fornecer alguns exemplos concretos,¹⁷ os autores reconhecem que o elevado grau de abstração na delimitação desse objeto incorre na impossibilidade de que todas as suas formas de expressão sejam mapeadas. O acesso limitado a informações, fruto do caráter sigiloso das atividades, dificulta sua própria observação empírica (Cormac *et al.*, 2022, p. 112).

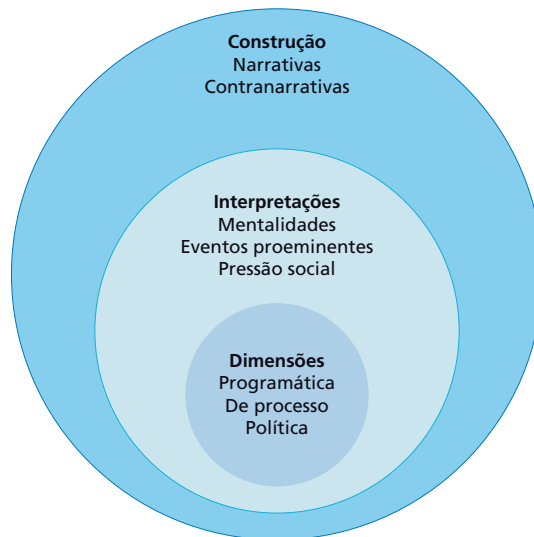
16. O termo utilizado pelo autor, em inglês, é *covert action*.

17. Em um esforço de classificação com fins didáticos, três tipos são sugeridos: "propaganda; ação política, como canalizar dinheiro para um partido político ou fomentar protestos; e ação paramilitar, desde o treinamento de grupos insurgentes até [a autoria propriamente dita de] assassinatos" (Cormac *et al.*, 2022, p. 112, tradução nossa).

A exposição dessas especificidades fundamenta a tese defendida pelos autores, no contexto de uma discussão das formas alternativas de conceber e desenvolver um modelo de avaliação de desempenho quanto à eficácia de tais atividades. Retomando as ideias contidas na obra de Johnson e Tierney (2006), Cormac *et al.* (2022) argumentam que a aplicação do quadro referencial da CdP, definido pela mensuração objetiva de ações e resultados, não é sequer possível no caso do seu objeto de estudo. A partir da apresentação de alguns exemplos, os autores evidenciam como, apesar disso, as ações secretas, bem como sua avaliação em termos de sucesso e fracasso, não deixam de estar presentes no debate público. Consequentemente, tal panorama demonstra como as percepções, se não forem mais importantes, “são tão importantes quanto os resultados e os impactos reais (...), [porque uma ação secreta de sucesso] é aquela que foi assim rotulada pelos observadores mais proeminentes – e cujo rótulo emplacou [no debate público]” (Cormac *et al.*, 2022, p. 112, tradução nossa).

Em um esforço ecumênico de integração das visões, é proposto um quadro analítico, representado esquematicamente na forma de três círculos concêntricos (figura 2), no qual os elementos constitutivos da construção social do sucesso são definidos e têm seus papéis explicitados teoricamente. Este modelo de avaliação, portanto, está baseado em uma abordagem intersubjetiva, cujo foco se direciona ao exame a respeito de como uma iniciativa chega a ser vista como exitosa.

FIGURA 2
A construção social do sucesso



Em tal modelo, chama-se atenção à maneira pela qual, nesse processo, alguns atores proeminentes

interpretam os eventos, e como aspectos cognitivos e culturais moldam a compreensão dos resultados na política internacional. A competição entre narrativas e contranarrativas eventualmente conduz a alguma forma de consenso – a narrativa dominante – na qual a interpretação dos eventos como um sucesso ou um fracasso se torna um fato social. (...) Narrativas dominantes, mais do que realidades materiais, determinam até que ponto as ações secretas são construídas como exitosas por um certo período. Uma ação secreta é bem-sucedida quando observadores proeminentes julgam que uma operação atingiu os objetivos que os proponentes se propuseram a alcançar; quando esses julgamentos emplacam [no debate público]; e quando são mínimas as críticas contrárias à forma como isso foi conquistado e às consequências políticas que acarretaram (Cormac *et al.*, 2022, p. 118, tradução nossa).

No exame dessa contribuição teórica, para os objetivos deste capítulo, devem ser destacados quatro aspectos. Dois são positivos: i) trata-se de um aporte que, de certa forma, supera a polarização presente no debate entre as epistemologias positivista e pós-positivista, uma vez que encontra pontos de convergência para uma sistematização lógica das ideias em um mesmo construto teórico; e ii) que, ao mesmo tempo, dialoga com a literatura, reunindo os principais pontos positivos de alguns trabalhos – com destaque para as propostas de Johnson e Tierney (2006) e de McConnell (2016) – ao redor de um modelo que pode ser aproveitado por desenhos de investigação avaliativos.

O outro par de aspectos, como é de se esperar, é negativo: i) da forma como está colocado, o modelo assume haver e depende da existência (e da identificação) de uma narrativa dominante, reduzindo sua utilidade para os objetivos aqui colocados, uma vez que as rotinas gerais de M&A/PE fariam parte do processo de construção social do sucesso, tendo o papel de delimitar a compreensão sobre o comportamento dos resultados (localizadas no círculo intermediário), em certas situações precedendo e/ou concorrendo para a formação da interpretação que estará embutida na narrativa dominante; e ii) além disso, mesmo havendo uma visão dominante, a disputa de narrativas também pode ser considerada uma premissa do modelo, sendo forçoso reconhecer que, em muitos casos, para os quais não é criada polêmica e/ou não é dada tanta atenção a ponto de que a maneira de perceber o desempenho se torne objeto de questionamentos, a aplicação do modelo se revela inviável e/ou inócua.

No conjunto, de qualquer maneira, a análise desse aporte deve ser apreciada como uma colaboração valiosa. Os elementos constitutivos do modelo podem eventualmente ser utilizados, na comparação entre propostas construídas como rotinas gerais para o M&A/PE, como forma de organizar as características e, com isso, sistematizar critérios para que sejam enumeradas e ponderadas as vantagens

e as fragilidades de cada um. A abrangência dos elementos, com a transparência que carregam, tornam o modelo um bom balizador para a seleção das rotinas que se apresentem como as mais adequadas, conforme cada contexto avaliado.

Para fechar este levantamento, deve ser destacada uma última obra de fôlego que derivou dessa agenda de pesquisa. Trata-se do trabalho de Friedman (2019), livro no qual o autor se debruçou sobre o lado mais preditivo que envolve o reconhecimento antecipado das situações de ameaça e de oportunidade que têm impacto sobre o desempenho da estratégia exterior de um país. Na pesquisa, o autor tem como objetivo mostrar a maneira pela qual, com muita frequência, os burocratas responsáveis pela política externa tentam evitar o desafio envolvido em efetuar raciocínios probabilísticos a respeito de eventos do ambiente global. Adicionalmente, ele busca demonstrar como apreciações sobre as incertezas da política internacional são muito mais valiosas do que convencionalmente se costuma considerar. Como justificativa para a realização desse estudo, o autor defende que “é impossível avaliar as opções de política externa sem apreciar a incerteza de alguma forma” (Friedman, 2019, p. 5, tradução nossa).

Reconhecendo a enorme dificuldade de que essa tarefa seja executada com precisão matemática, ele argumenta que a teoria subjetiva da probabilidade se configura como uma saída útil para que os analistas de política externa¹⁸ deixem de se esquivar diante desse obstáculo. Nessa abordagem, os exames probabilísticos devem se basear nas convicções pessoais do próprio analista, de acordo com sua experiência e o conhecimento acumulado sobre o tema. Mesmo que o rigor científico desse procedimento seja incomparável ao da metodologia estatística, o autor reforça que fazer esses exames de maneira explícita sempre será melhor do que a omissão. Ao serem realizados e apresentados, ao menos algum alerta concreto é dado aos tomadores de decisão sobre os riscos envolvidos por cada alternativa de ação disponível, dando embasamento para a escolha de quantos, como, quando e quais recursos serão alocados pela pasta, tendo em vista as impressões dos analistas.

Do ponto de vista avaliativo, caso fosse adotada pelos formuladores da política externa, a proposta de Friedman (2019) ofereceria elementos mais palpáveis com os quais se poderiam elaborar alguns parâmetros para apreciar as estratégias executadas. O risco aos quais os tomadores de decisão se dispuseram a correr quando da escolha das alternativas de ação é uma informação sugestiva até mesmo dos graus de prioridade diferenciados dados pelo governo aos seus

18. Com esse termo ele se refere a qualquer pessoa que busque informar os debates de política externa, tanto dentro como fora do governo.

objetivos exteriores.¹⁹ Em última análise, explicitar essa apreciação proveria mais transparência para com os cidadãos a respeito das escolhas realizadas em política externa pela administração pública.

4 CONSIDERAÇÕES FINAIS

Seguindo com este fio condutor, é interessante notar que, ao chamar atenção para as análises de impressão de probabilidades, Friedman (2019) se refere implicitamente a duas questões distintas, ensejando dois tipos de exames probabilísticos. Em alguns momentos, o autor faz alusão ao exame das chances de que determinados eventos externos, que tenham impacto doméstico (positivo ou negativo), ocorram na política internacional. Em outras passagens, de maneira distinta, ele se remete às apreciações prévias quanto à possibilidade de êxito que envolve a adoção de cada alternativa disponível em um determinado caso.

Considerados todos os desafios científicos envolvidos na realização do M&A/PE que foram discutidos neste artigo, repassando alguns dos aportes que foram elencados, é interessante notar como essa distinção contribui para que se perceba como a tipologia de erros e fracassos em política externa de Walker e Malici (2011), inclusive na versão aprimorada proposta por Meyer (2016), tem maior aplicabilidade na realização de avaliações de programas (formulação) e processos (implementação). Os exames probabilísticos implicitamente exigidos nesses modelos são essencialmente do primeiro tipo: dizem respeito, de um lado, à precisão do diagnóstico efetuado sobre a conjuntura e, de outro, ao planejamento realizado para obter uma execução eficiente da estratégia adotada. Assim, as construções teóricas propostas por esses autores ajudam a detectar e tipificar os erros quanto a *quando e como/quanto* os instrumentos são empregados, ou ainda quanto a se algo deveria ter sido feito ou não.

Deve-se reconhecer, de todo modo, que eles pouco colaboram com a tarefa de sugerir aos tomadores de decisão *qual* ferramenta de ação se caracteriza como a mais adequada. Essa informação depende de exames probabilísticos do segundo tipo, ou seja, a chance de êxito. A escolha do tipo de instrumento de política externa mais indicado se configura como uma apreciação fundamental nas avaliações de desempenho. É importante que esse raciocínio seja colocado ao levar em consideração que a adoção de uma ferramenta menos eficaz e/ou eficiente em comparação com as demais alternativas disponíveis,²⁰ mesmo que aplicada no

19. Considere-se, por exemplo, que os responsáveis pela política externa tenham avaliado que uma ameaça potencial aos interesses do seu país tinha uma probabilidade baixa de se concretizar. Ainda assim, eles decidiram tomar uma medida enérgica para evitá-la, de modo a garantir que determinado objetivo fosse atingido. Esse contexto revelaria, em princípio, uma alta prioridade dessa questão no conjunto da estratégia do país.

20. Num exemplo extremo: escolher uma intervenção militar, com os custos de toda sorte que pode envolver, no lugar de investir em negociações diplomáticas, em uma situação em que os resultados seriam os mesmos.

momento oportuno e em escopo e escala devidamente proporcionais ao objetivo a ser alcançado, pode ainda caracterizar-se inadequada. Mesmo que tenha resultados satisfatórios, ainda assim tal alternativa teria um desempenho inferior ao da opção ótima.

Nesse sentido, do ponto de vista metodológico, é fundamental observar que um caminho alternativo e/ou complementar para a realização do M&A/PE consiste em tomar como objeto de avaliação, para além da conquista dos objetivos, a adequação do instrumento empregado em uma estratégia de inserção internacional. Com efeito, o desempenho isolado de alguns itens à disposição no cardápio da pasta internacional tem sido um caminho trilhado com maior ou menor efervescência em agendas de pesquisa paralelas ao campo de APE, comumente em caráter multidisciplinar.

Em trabalho anterior (Desiderá e Schmitz, 2023), buscou-se esmiuçar alguns aportes resultantes das pesquisas avaliativas com recortes desse perfil. Os instrumentos selecionados foram as operações de paz, as sanções econômicas, a diplomacia tradicional e a diplomacia pública. Embora fosse legítimo imaginar que, em um esforço de construção de rotinas gerais para o M&A/PE, essa variedade pudesse representar um empecilho, a discussão desenvolvida evidenciou como, ao contrário, a diversidade se revelou uma importante fonte da riqueza de suas contribuições à empreitada. Ao elaborar desenhos de investigação para apreciar o desempenho de cada um deles, os autores dedicados a essas agendas se deparam com algumas das mesmas dificuldades discutidas ao longo deste artigo. Não obstante, é importante ressaltar que a cada instrumento correspondem desafios maiores ou menores com uma ou outra questão devido a especificidades próprias.

Por exemplo, em geral os objetivos de política externa que motivam a adoção de sanções econômicas são bastante claros e amplamente divulgados. Em contrapartida, o mesmo não pode ser dito a respeito das operações de paz, nas quais os interesses das nações envolvidas diretamente nas intervenções geralmente estão ocultos. Ao mesmo tempo, é mais simples definir o momento final para o exame de seus resultados (a retirada das tropas, por exemplo) em comparação com as sanções econômicas – as quais podem ter evidentemente fracassado enquanto os executores se negam a aceitar a derrota e as mantêm em implemento. Do ponto de vista teórico-metodológico, particularidades como essas apresentam a riqueza de oferecer ideias criativas a serem cotejadas de um lado a outro, gerando uma dinâmica construtiva para superar os distintos entraves que acometem o M&A/PE de um modo geral.

Antes de encerrar este trabalho, deve ser mencionada uma questão importante para que se faça justiça aos esforços de uma parte considerável dos especialistas do campo da APE. É impossível ignorar que, em muitos trabalhos, sobretudo em

estudos de caso sobre o comportamento da política externa de governos específicos, os autores acabam desenvolvendo argumentações que podem ser consideradas como avaliações, mesmo que em um sentido amplo. Entretanto, pela natureza de seus desenhos de investigação, com alguma frequência essas pesquisas não estão preocupadas com a produção de análises cujos dados sejam comparáveis com outros casos, dificultando sua colaboração sistemática com a superação de parte das dificuldades de fundo teórico aqui elencadas.

De certo modo, quanto aos requisitos esperados de uma avaliação no campo de APP (a adoção de ferramentas com os três componentes bem estruturados), na grande maioria deles o exame de resultados realizado é distinto. De maneira alguma isso significa que suas contribuições sejam inválidas cientificamente nem irrelevantes ao M&A/PE. Ao contrário, são fontes de informações muito ricas em detalhes que os pesquisadores dedicados a essa tarefa podem e devem consultar ao construir ferramentas de avaliação gerais consistentes com essa tradição de pesquisa diferente. Efetivamente, as inferências produzidas por estudos de caso, por exemplo, são parâmetros valiosos para apontar quais caminhos trilhar no momento da elaboração e definição do desenho de investigação e das ferramentas de avaliação.

REFERÊNCIAS

- BALDWIN, D. A. Success and failure in foreign policy. **Annual Review of Political Science**, v. 3, n. 1, p. 167-82, 2000.
- BRUMMER, K. “Fiasco prime ministers”: leaders’ beliefs and personality traits as possible causes for policy fiascos. **Journal of European Public Policy**, v. 23, n. 5, p. 702-717, 2016.
- CALLAHAN, P.; BRADY, L. P.; HERMANN, M. G. (Ed.). **Describing foreign policy behavior**. Beverly Hills: Sage Publications, 1982.
- CLARK, J. F. Evaluating the efficacy of foreign policy: an essay on the complexity of foreign policy goals. **Southeastern Political Review**, v. 24, n. 4, dez. 1995.
- COHEN, E.; FRANCO, R. **Avaliação de projetos sociais**. Petrópolis: Vozes, 2008.
- CORMAC, R. *et al.* What constitutes successful covert action? Evaluating unacknowledged interventionism in foreign affairs. **Review of International Studies**, v. 48, n. 1, p. 1-18, 2022.
- DESIDERÁ, W. Por que o monitoramento e a avaliação da política externa não se consolidaram como uma agenda de pesquisa produtiva do campo da análise de política externa? Uma explicação geral de cunho teórico-metodológico. **Boletim de Economia e Política Internacional**, n. 33, p. 161-169, maio/ago. 2022.

DESIDERÁ, W.; SCHMITZ, G. O. Como efetuar o monitoramento e a avaliação da política externa? As contribuições oferecidas por estudos focados no desempenho de instrumentos de política externa alternativos. **Boletim de Economia e Política Internacional**, n. 36, p. 9-25, maio/ago. 2023.

FARIAS, R. S.; CARMO, G. F. Atores, eventos e redes da política externa brasileira. **Dados**, Rio de Janeiro, v. 64, n. 1, p. 40, 2021.

FRIEDMAN, J. A. **War and chance**: assessing uncertainty in international politics. Oxford: Oxford University Press, 2019.

GARCIA, R. C. Subsídios para organizar avaliações da ação governamental. *In*: CARDOSO JUNIOR, J. C.; CUNHA, A. S. (Org.). **Planejamento e avaliação de políticas públicas**. Brasília: Ipea, 2015. p. 235-96.

HERMANN, M. G. Feedback: acceptance and rejection. *In*: CALLAHAN, P.; BRADY, L. P.; HERMANN, M. G. (Ed.). **Describing foreign policy behavior**. Beverly Hills: Sage Publications, 1982. cap. 12.

JOHNSON, D.; TIERNEY, D. **Failing to win**: perceptions of victory and defeat in international politics. Boston: Harvard University Press, 2006.

KRATOCHWILL, F. Alternative criteria for evaluating foreign policy. **International Interactions**, v. 8, n. 1-2, p. 105-122, 1981.

KRUCK, A.; OPPERMANN, K.; SPENCER, A. **Political mistakes and policy failures in international relations**. Cham: Springer International Publishing, 2018.

MCCONNELL, A. A public policy approach to understanding the nature and causes of foreign policy failure. **Journal of European Public Policy**, v. 23, n. 5, p. 667-684, 2016.

MEYER, C. O. Over and under-reaction to transboundary threats: two sides of a misprinted coin? **Journal of European Public Policy**, v. 23, n. 5, p. 735-752, 2016.

OPPERMANN, K.; SPENCER, A. Studying fiascos: bringing public and foreign policy together. **Journal of European Public Policy**, v. 23, n. 5, p. 643-652, 2016a.

_____. Telling stories of failure: narrative constructions of foreign policy fiascos. **Journal of European Public Policy**, v. 23, n. 5, p. 685-701, 2016b.

_____. **Fiascos in public policy and foreign policy**. London: Routledge, 2018.

SCHAFER, M.; WALKER, S. G. **Beliefs and leadership in world politics**: methods and applications of operational code analysis. New York: Palgrave Macmillan, 2006.

SCHRODT, P. A. Event data in foreign policy analysis. *In*: NEACK, L.; HEY, J. H. K.; HANEY, P. J. (Ed.). **Foreign policy analysis: continuity and change in its second generation**. Englewood Cliffs: Prentice-Hall, 1995. cap. 9.

THOENIG, J. C. Evaluation as usable knowledge for public management reforms. **Evaluation**, v. 6, n. 2, p. 217-229, Apr. 2000.

VASQUEZ, J. A. (Ed.). **Evaluating US foreign policy**. New York: Praeger Publishers, 1986.

WALKER, S. G.; MALICI, A. **U.S. presidents and foreign policy mistakes**. Stanford: Stanford Security Series, 2011.

WALTZ, K. **Theory of international politics**. New York: McGraw-Hill, 1979.