







Artigo original

Metaavaliação: quais padrões de qualidade, porquê, de quem e para quê?

Meta-Evaluation: Which quality standards, why, to whom and for what?

Elizabeth Moreira dos Santos^{1,2*} [branca] , Andreia Ferreira de Oliveira¹ [branca] ,
Patrícia Pássaro da Silva Toledo¹ [branca] , Egléubia Andrade de Oliveira³ [parda] ,
Gisela Cordeiro Pereira Cardoso² [branca] , Paulo de Martino Jannuzzi^{1,4,5} [branco] 

¹Fundação Cesgranrio, Rio de Janeiro, RJ, Brasil

²Fundação Oswaldo Cruz (FIOCRUZ), Escola Nacional de Saúde Pública Sérgio Arouca (ENSP), Departamento de Endemias Samuel Pessoa, Laboratório de Avaliação de Situações Endêmicas Regionais (LASER), Rio de Janeiro, RJ, Brasil

³Universidade Federal do Rio de Janeiro (UFRJ), Instituto de Estudos de Saúde Coletiva (IESC), Rio de Janeiro, RJ, Brasil

⁴Instituto Brasileiro de Geografia e Estatística (IBGE), Escola Nacional de Ciências Estatísticas (ENCE/IBGE), Rio de Janeiro, RJ, Brasil

⁵Fundação Getúlio Vargas (FGV), Escola Brasileira de Administração Pública e de Empresas (EBAPE/FGV), Rio de Janeiro, RJ, Brasil

COMO CITAR: Santos, Elizabeth Moreira dos, Oliveira, Andreia Ferreira de, Toledo, Patrícia Pássaro da Silva, Oliveira, Egléubia Andrade de, Cardoso, Gisela Cordeiro Pereira, & Jannuzzi, Paulo de Martino (2021). Metaavaliação: quais padrões de qualidade, porquê, de quem e para quê? *Revista Brasileira de Avaliação*, 10(2), e101721. <https://doi.org/10.4322/rbaval202110017>

Resumo

Neste artigo problematiza-se dificuldades com o ensino da qualidade de avaliações na formação de avaliadores. A experiência evidencia que a problematização deve ultrapassar as diferenças semânticas, embora as considerem para a clareza dos processos de avaliação da avaliação. Os profissionais, ao mobilizarem as categorias utilizadas em meta-avaliação, devem estar atentos à estas equivalências e às suas mudanças no tempo. Aborda os padrões de qualidade de quatro documentos representativos da influência anglo-saxônica, europeia, latino-americana, incluindo as Diretrizes para a prática da avaliação no Brasil. A comparação baseou-se nos pontos críticos dos processos avaliativos, que explicitam a interação entre competência técnica, investigação sistemática e habilidades de negociação envolvendo uma pedagogia e uma prática política. Toma-se como pressuposto para a comparação entre os documentos sete perguntas norteadoras. Os achados apontam a qualidade da avaliação pautada por evidências científicas. São incipientes a prática avaliativa profissional e o valor agregado como aprendizado e melhoria social.

Palavras-chave: Meta-avaliação. Avaliação. Padrões de referência. Terminologia.

Abstract

This article discusses difficulties in teaching the quality of evaluation in the training of evaluators. Experience shows that this problematization might consider beyond semantic differences, although they should be taken in account for the clarity of the evaluation of evaluations. When mobilizing the categories used in meta-evaluation, professionals should be aware of these differences, equivalences and possible changes over time. The text addresses the quality standards from four documents representing Anglo-Saxon, European, Latin American influence and the Guidelines for the practice of evaluation in Brazil. The comparison highlights the critical points of the evaluation processes. They involves the relationships among technical competence, systematic investigation and negotiation skills encompassing a pedagogy and a political practice. Seven guiding questions were applied to contrast the documents. Findings indicate scientific evidence as the main criteria of quality in evaluations. As such, learning and social improvement are still incipient.

Keywords: Metaevaluation. Evaluation. Reference standards. Terminology.

A RBAVAL apoia os esforços relativos à visibilidade dos autores negros na produção científica. Assim, nossas publicações solicitam a autodeclaração de cor/etnia dos autores dos textos para tornar visível tal informação nos artigos.

Recebido: Agosto 26, 2021

Aceito: Agosto 31, 2021

***Autor correspondente:**

Elizabeth Moreira dos Santos

E-mail: betuca51@gmail.com



Este é um artigo publicado em acesso aberto (Open Access) sob a licença Creative Commons Attribution, que permite uso, distribuição e reprodução em qualquer meio, sem restrições desde que o trabalho original seja corretamente citado.



Introdução

Este artigo elabora uma comparação descritiva e crítica das quatro referências em diretrizes e atributos que hoje influenciam a meta-avaliação no Brasil, isto é, os da Organização para a Cooperação e Desenvolvimento Econômico (OCDE/DAC) (OECD, 2010, 2019), os atributos e padrões do *Joint Committee on Standards of Educational Evaluation* (JCSEE) (JCSEE, 1994, 2021), as Diretrizes de Avaliação para a América Latina e Caribe (DALAC) (Rodríguez Bilella et al., 2016) e as Diretrizes para a Prática de Avaliação no Brasil (DPAB) (Silva et al., 2020). As DPAB foram incluídas já que, apesar de recentemente lançadas, têm relevância e potencial de fortalecer os processos de meta-avaliação e a sua institucionalização profissional no país.

A inspiração para o artigo se deu em uma oficina de curta duração (4hs), em 2019, no evento de mobilização nacional da X Reunião da ABAVE: Avaliação Educacional no Brasil (ABAVE, 2019): O desafio da qualidade, para um grupo heterogêneo de participantes quanto à formação, isto é, um grupo multidisciplinar e com experiência diversa na prática em avaliação. Esta atividade desencadeou uma sistematização que se considerou relevante sobre os padrões, enfatizando-se comparativamente o seu uso para a prática docente e profissional.

As avaliações ocorrem em contexto históricos datados, com características políticas e econômicas específicas. Elas envolvem um processo de valoração que responde a interesses que, espera-se, sejam definidos de forma legítima e transparente. A apreciação desta valoração, um julgamento sobre ela, ancora-se em pactuações éticas, culturais e em interações dinâmicas com o conhecimento científico relacionado e disponível (Alkin & Patton, 2020; Patton, 2018).

A preocupação em estabelecer padrões de legitimidade, qualidade técnico científica e eticidade tem formalmente preocupado avaliadores desde o início da década de 60 (Worthen et al., 1997), envolvendo tanto a perspectiva dos usuários como a dos próprios avaliadores. Nas últimas décadas, as associações profissionais em avaliação, instituições internacionais com foco no desenvolvimento econômico e social e as instituições acadêmicas têm se preocupado em adaptar, construir e implementar padrões de qualidade e características de uma boa prática profissional para o campo (Silva et al., 2020). Adicionalmente, a formação de comunidades e de redes de práticas profissionais tem contribuído para que as atividades de apreciação dos processos avaliativos se tornem mais prementes.

Embora a globalização de conhecimentos e práticas, viabilizados pela web, tenham potencializado novas formas de aprender, ensinar e agir, alguns desafios para avaliação ainda continuam presentes (Fragoso et al., 2013). O primeiro deles diz respeito ao fato que dados e informações são coletados com diferentes concepções e valores e são interpretados por padrões únicos. Apesar de inúmeros esforços para se definir padrões mínimos e critérios de qualidade para as iniciativas em monitoramento e avaliação, sistemas regulatórios com jurisdições específicas não têm suficientemente contemplado a eticidade da produção e aplicação de dados e informações (Love & Russon, 2004). Nesse sentido, o segundo desafio refere-se à formação de recursos humanos comprometidos com a mobilização de informações de qualidade e do seu uso ético e incluyente.

Apesar do crescimento expressivo e dos esforços de institucionalização, o campo da avaliação no Brasil ainda possui desafios importantes (Abreu et al., 2017; Furtado & Laperrière, 2011). Um deles, abordado neste artigo, problematiza o uso de padrões na apreciação da qualidade em avaliações, isto é, de meta-avaliação. A qualidade de avaliações é um problema cada vez maior, não só no que se refere aos recursos e esforços que elas consomem e na influência de seus achados na legitimação de escolhas estratégicas, mas ainda por assegurar que se originem de boas práticas de investigação (Hartz et al., 2008; Patton, 2018; Stufflebeam & Shinkfield, 2007).

A meta-avaliação, descrita por Scriven (1991) como avaliação da avaliação, é um processo que descreve, julga e sintetiza as funções e práticas de uma avaliação, utilizando-se de padrões e critérios previamente propostos e validados. A meta avaliação, vem sendo amplamente empregada e sua história inicia-se na década de 60, onde os vários critérios vigentes foram questionados por avaliadores. Devido à falta de consenso, surge, em 1975, um projeto ambicioso para desenvolvimento de diretrizes voltadas para avaliações educacionais,



coordenadas por Stufflebeam e validadas pelo JCSEE, que foram disseminadas e empregadas em diversas áreas, não somente a educacional (JCSEE, 1994). Em termos práticos, a institucionalização da meta-avaliação permite às organizações a implementação e governança de sistemas avaliativos éticos, efetivos e eficientes.

Internacionalmente, os atributos de verificação da qualidade de avaliações se organizam em cinco núcleos norteadores. Hartz & Contandriopoulos (2008) têm reagrupado estes núcleos em dois domínios de sentido: adequação das avaliações (mérito ou qualidade) e sua capacidade de resposta às necessidades que deram origem à avaliação (pertinência e utilidade). Estes autores consideram que as meta-avaliações devam agregar uma apreciação dos princípios e diretrizes, ou seja, o alinhamento da avaliação aos fundamentos éticos profissionais e às recomendações enunciadas como regra geral para o campo. Fundamentalmente e de acordo com Stufflebeam (Stufflebeam et al., 2000; Stufflebeam, 2001) eles destacam a meta-avaliação como componente inerente às boas práticas profissionais do avaliador e essencial para a evolução do conhecimento em avaliação.

Considerando os sentidos de pertinência e utilidade, alguns autores têm explorado a relação entre os processos avaliativos, a melhoria social (Melvin et al., 2000) e a sociedade do conhecimento (Abreu, 2018). Nestas, os processos avaliativos ocorrem vinculados a produção e utilização do conhecimento voltados para o funcionamento e o bem-estar societário. As relações entre o conhecimento científico e a prática profissional são bem mais que operações de disseminação, transferência e tradução de conhecimentos e práticas e podem ser compreendidas como translacionais (Bezerra et al., 2019), isto é, ocorrem em redes sociotécnicas, científicas e de comunidades de práticas, mediadas por relações entre atuantes humanos e não humanos (Latour, 2011).

Nesta perspectiva translacional, a compreensão das mediações entre os atuantes de redes científicas e comunidades laicas nos remete a exploração do papel do avaliador como um ator intermediário que potencialmente faz com que o conhecimento produzido cientificamente circule. A perspectiva translacional implica em uma coprodução do conhecimento e das práticas entre os atores humanos e não humanos envolvidos em que se toma as conexões como unidades de análise (Abreu et al., 2017; Clavier et al., 2012; Oliveira et al., 2019). Desta forma, retoma a tensão descrita entre as “grandes teorias” e o “empiricismo cotidiano” (Merton, 1967). Pawson (2010) discutindo essas interações como mecanismos, enfatiza a pouca atenção que os estudos avaliativos conferem às teorias que possam mediar estas instâncias, isto é, as *middle range theories*.

A proposta de (Clavier et al., 2012) para a apreensão das operações de translação envolvendo atores sociais intermediários baseia-se na teoria do ator-rede e nas operações de translação de Bruno Latour (2012). Para estes autores, as operações de translação, ou seja, a problematização, a sensibilização, a motivação e a construção de alianças ocorrem em três domínios de práticas: cognitivo (circulação do conhecimento, significados e conteúdos); estratégico (conjunto de atividades voltadas para manter o interesse das partes envolvidas) e logístico (conjunto de ações para garantir as conexões entre os parceiros).

A aplicação das operações de translação para teorização das ações de atores intermediários em processos avaliativos já tem sido discutida por alguns autores (Hartz et al., 2008; Santos et al., 2019). Para eles, elas possibilitam a criação de processos pedagógicos imersos nas avaliações que permitem apreender e negociar a pluralidade de identidades e interesses dos diversos atores envolvidos, mobilizando conhecimentos e práticas, no sentido de promover uma ação transformadora. Neste artigo, esta abordagem, sintetizada no Quadro 1, descrito na metodologia, orientou a análise comparativa.

Metodologia

A elaboração deste artigo compreendeu quatro etapas, sejam elas: a avaliação da oficina realizada, a revisão bibliográfica e a comparação dos atributos de forma independente por dois autores e validada criticamente por três outros. A oficina compreendeu uma reflexão em pequenos grupos sobre os critérios de qualidade de avaliações, por meio da aplicação



dos mesmos a quatro avaliações selecionadas. Na avaliação da oficina, foram identificados os seus pontos positivos e suas fragilidades. Entre estas últimas, percebeu-se a dificuldade dos grupos em identificar e manusear critérios de qualidade nos textos utilizados para o exercício. As dúvidas dos participantes e as controvérsias entre as definições dos atributos de qualidade nortearam a elaboração deste texto.

A revisão bibliográfica foi realizada utilizando-se as bases de dados Medline, Lilacs e o banco de teses e dissertações disponíveis na plataforma OASIS nos últimos dez anos, incluindo textos considerados chaves na temática, mesmo se publicados fora deste período. As palavras-chave utilizadas foram meta-avaliação, avaliação e padrões de referência. Além disso, foi realizada uma busca dos principais padrões empregados em meta-avaliações por organismos nacionais e multissetoriais das associações de avaliação.

A comparação dos atributos foi realizada de forma independente e envolveu quatro conjuntos de recomendações para apreciação da qualidade de avaliações. A escolha deles baseou-se no fato que todos eles são utilizados no Brasil e espelham a influência internacional europeia, norte americana e latino-americana nas discussões sobre a qualidade de avaliações no país. Vale ressaltar que esta influência tem magnitude diferente nas diversas políticas e programas sociais, considerando a pluralidade de financiamento tanto das políticas e dos programas quanto de seus processos avaliativos. As DPAB foram incluídas porque apesar de recentes, têm potencial para direcionar as meta-avaliações futuras no Brasil. Os documentos referentes aos quatro conjuntos identificados foram então descritos e comparados analiticamente. Em dois casos, isto é, os critérios da OCDE/DAC e o dos atributos e padrões do JCSEE, procedeu-se, respectivamente, a análise comparativa de dois documentos de cada instituição, uma vez que foram observadas mudanças recentes nos mesmos (OECD, 2010, 2019; JCSEE, 1994, 2021).

Após a elaboração de um documento consolidando a comparação independente, este foi submetido a análise crítica de três especialistas em avaliação, coautores deste artigo. É importante destacar que, em virtude da variedade de terminologias, optou-se pela utilização do termo categoria para designar os descritivos de qualificação de avaliações empregados em cada documento.

A comparação entre os quatro documentos contendo as recomendações para a meta-avaliação privilegiou pontos desafiadores para a qualidade do processo avaliativo. Neste texto, presume-se a avaliação como campo de investigação disciplinada e como tal, deve ser transparente, academicamente legítima e cientificamente embasada. Em outras palavras, entende-se como desafiadores para a qualidade movimentos do processo avaliativo considerados controversos, negociáveis e que frequentemente permanecem ocultos em relatórios e publicações do campo. Desta forma, os pontos apresentados no Quadro 1, foram tomados como norteadores da análise comparativa realizada.

É importante destacar que este quadro materializa de forma bastante reduzida uma longa e disputada discussão que tem percorrido a maturação da avaliação como campo de teorias, métodos e práticas. Em outras palavras, a teleologia, a axiologia, a epistemologia e a metodologia desta trajetória (Potvin & Bisset, 2009; Potvin & Bisset, 2014).

Uma das limitações do presente estudo foi que os critérios, dimensões, diretrizes e padrões são imbricados, o que resultou em dificuldades no seu alinhamento ao quadro proposto para análise. Apesar de revisitas sucessivas aos documentos analisados e mobilização da literatura que reflete sobre eles, considera-se que o estudo lidou com conceitos e concepções sem interface bem definidas que evitasse ambiguidades na alocação. Parafraseando a máxima do movimento mundial de mulheres, pode-se afirmar que *"mexeu com um, mexeu com todos"*.

Descrição das categorias encontradas nos quatro documentos analisados: atributos, padrões, critérios, dimensões e diretrizes

Nesta seção, serão apresentados os critérios dos quatro conjuntos analisados. Além da descrição, serão acrescentados alguns comentários explicativos que justifiquem a classificação dos critérios pelas perguntas orientadoras, conforme descrito nos Quadros 2, 3 e 4.

Quadro 1. Práticas avaliativas: Pontos frequentes de tensionamento em apreciação de qualidade.

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	O desenho metodológico	O processo de valoração	Usos e influências (o valor agregado pela avaliação)
Fatores facilitadores e dificultadores que influenciam a intervenção e/ou a avaliação?	Um experimento?	Pesquisador?	Esporádico para informar e negociar pontos específicos da avaliação?	Quantitativo	Prescritivo?	Difusão e disseminação de achados
Eventos críticos que modificam a teoria de mudança ou de funcionamento?	Um tratamento?	Facilitador?	Contínuo, em permanente interação e protagonismo?	Experimental ou quase experimental?	Descritivo?	Político estratégico?
	Um sistema simples? Complexo?	Mediador?		Qualitativo?	Baseado em evidências científicas?	Logístico?
	Uma rede?	Negociador?		Integrado ou misto?	Baseado no conhecimento tácito?	Mobilização de conhecimento?
	Um mecanismo? Uma conexão? Um dispositivo?	Juiz?		Estudo de caso?		
		Mobilizador de conhecimento?		Etnográfico?		
Tradutor?						
	"Translador"?					

Fonte: Os autores (2021).

Destaca-se que, nos Quadros 2, 3 e 4, as categorias de cada documento são apresentadas em negrito e suas subdivisões descritas de maneira resumida, embora estejam numeradas e mantenham a organização e disposição dos documentos originais. Por exemplo, o JCSEE (1994, 2021) confere relevância a inclusão do contexto local na apreciação da Propriedade de uma avaliação. Assim, na coluna do contexto, aparece o atributo propriedade e o padrão a ela relacionado que diz respeito ao contexto.

Joint Committee on Standards for Educational Evaluation

O JCSEE é um Comitê misto, formado por representantes de organizações profissionais e especialistas, que têm a tarefa de revisar, atualizar, desenvolver e disseminar padrões para atender às necessidades da área de avaliação. Embora o JCSEE, criado em 1975, e, na época, coordenado por Daniel Stufflebeam tenha raízes no mundo da avaliação educacional, seus atributos e padrões são amplamente usados em vários tipos de avaliação de programas (JCSEE, 2021). Note-se que a terminologia utilizada compreende atributos como categoria abrangente e seus respectivos padrões.

Uma de suas características importantes é reunir nos atributos e padrões, desde a junção da *Evaluation Research Society* e a Rede de Avaliadores nos Estados Unidos, critérios que acomodassem diversas perspectivas e abordagens em avaliação. A mescla de uma longa tradição de pesquisa avaliativa com uma rede de avaliadores profissionais, não se processou sem problemas. Por meio de uma proposta participativa e incluyente, o Comitê conseguiu acomodar a diversidade dos estudos avaliativos e da prática profissional sob o "guarda-chuva" da qualidade. Destaca-se que, atualmente, o JCSEE reúne três conjuntos de documentos que orientam as atividades em meta-avaliação (*The Program Evaluation Standards, The Personnel Evaluation Standards, The Classroom Assessment Standards for Prek-12 Teachers* (JCSEE, 2021)). Neste artigo, será analisado o primeiro documento, uma vez que os dois últimos abordam questões específicas relacionadas as boas práticas profissionais em avaliação e à apreciação do desempenho de docentes.



Quadro 2. Apreciação de qualidade de avaliações segundo o JCSEE.

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	Os Procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
Propriedade		Utilidade 1-conduzida por pessoal qualificado	Utilidade 2-envolve todos os interessados, de modo que estes se identifiquem com os métodos e com as informações geradas no processo avaliativo	Utilidade 4-explicite e respeite os valores culturais e individuais que fundamentam os processos envolvidos em avaliações	Responsabilização 2,3-serem meta-avaliadas interna e externamente e sempre que a qualidade precisar ser verificada, com propósitos somativos ou formativos	Utilidade 3-propósitos claros e negociados, agregando valor as experiências dos stakeholders
1-responsivas e inclusivas e orientadas segundo as necessidades locais				5-levantem informações plausíveis, aceitáveis e relevantes, de acordo com as necessidades dos interessados		6-origem produtos e processos que estimulem os diferentes atores a redescobrirem, reinterpretarem ou reverem entendimentos e comportamentos
3 e 6- avaliações conduzidas e divulgadas dentro dos preceitos legais e éticos, promovendo o direito e bem-estar de todos os interessados, explicitando os conflitos de interesse, caso existam.						
Viabilidade				Viabilidade		7-comunique as informações apropriadamente
3-contextual de uma avaliação, isto é, reforça que a mesma deve reconhecer, monitorar e equilibrar os interesses e necessidades dos indivíduos e grupos envolvidos				1-procedimentos adotados em uma avaliação devem empregar de forma eficaz estratégias de gerenciamento		8-promova a utilização responsável e adaptativa dos achados, preservando-se do mal uso e consequências negativas
				2-procedimentos de uma avaliação devem ser práticos e responsivos à forma como a intervenção é operada		Propriedade
				4-devem usar os recursos de forma eficaz e eficiente		
				Propriedade		
				2-negociadas formalmente		
				4-clara, compreensível e justa		
	7-cumpra os procedimentos e processos fiscais quando voltadas para a responsabilidade fiscal					
	Acurácia					
	1,3,4- propósitos e procedimentos, além da descrição do programa, claros, confiáveis e documentados					
	2-informações levantadas de acordo com os propósitos mapeados e dar suporte a interpretações válidas	5-transparentes na divulgação dos resultados a todos os interessados				

Fonte: Os autores (2021).



Quadro 2. Continuação...

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	Os Procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
				5,6,7-levantamento de informações e desenho adequados, válidos, confiáveis e sistematicamente revisados		
				8- informações quanti a qualitativas analisadas apropriadamente, gerando conclusões justificadas e relatórios imparciais, que permitam a efetiva resposta à pergunta avaliativa		
				Responsabilização		
				1-documentada por completo em termos dos seus propósitos, desenhos implementados, procedimentos, dados e resultados		

Fonte: Os autores (2021).

Os atributos definidos pelo JCSEE e oficialmente adotados pela Sociedade Americana e pela Sociedade Canadense de Avaliação são cinco e agrupam 30 padrões (Almeida & Tanaka, 2017; Stufflebeam & Shinkfield, 2007; Yarbrough et al., 2010; Yarbrough, 2017). Eles são descritos como: utilidade (U), viabilidade (V), propriedade (P), acurácia (A) e responsabilização (R). O atributo de utilidade, considerado o mais importante pelo JCSEE, é composto por oito padrões (U1 a U8) que pretendem assegurar a avaliação como dispositivo relevante para atender as necessidades de informações de seus usuários potenciais (JCSEE, 1994, 2021). Apesar de uma revisão recente desenvolvida por (Yarbrough, 2017) identificando 33 padrões, neste artigo optou-se por considerar aqueles identificados no site do JCSEE (2021).

Para o JCSEE, uma avaliação é útil quando: é conduzida por pessoal qualificado (U1); caracteriza o modo como todos os indivíduos e grupos envolvidos se identifiquem com suas questões, métodos e compreendam a influência possível dos achados gerados pelo processo avaliativo (U2); tem propósitos claros e negociados, agregando valor às experiências dos *stakeholders* (U3); explicita e respeita os valores culturais e individuais que fundamentam os procedimentos envolvidos em avaliações (U4); coletam informações plausíveis, aceitáveis e relevantes, de acordo com as necessidades dos interessados (U5); origina produtos e processos que estimulem os diferentes atores a redescobrirem, reinterpretarem ou reverem entendimentos e comportamentos (U6); comunica as informações apropriadamente (U7); e promove a utilização responsável e adaptativa dos achados, preservando-se do mal uso e consequências negativas (U8).

É destacado que para a utilização dos achados de uma avaliação é necessário a identificação das forças que atuam sobre os mesmos, tais como crenças, opiniões, atitudes, valores e comportamentos dos *stakeholders*. Para o uso responsável e adaptável dos processos e resultados gerados em uma avaliação é imprescindível à reflexão continuada sobre a adequação das práticas locais e sobre o conjunto dos padrões adotados.

Já o atributo da viabilidade (*Feasibility*) é composto por quatro padrões (F1, F2, F3, F4). Eles dizem respeito a necessidade de assegurar que as avaliações sejam realistas, negociadas e executadas de forma prudente e atenta às flutuações do contexto. Os procedimentos



Quadro 3. Apreciação de qualidade de avaliações segundo as Diretrizes para América Latina e Caribe.

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	Os procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
Avaliação rigorosa	Avaliação rigorosa	Princípios éticos e jurídicos	Avaliação rigorosa	Avaliação rigorosa	Relevância/utilidade	Avaliação rigorosa
1-importância de uma avaliação contextualizada social, cultural, política e economicamente	2- necessidade da descrição clara e detalhada das ações, produtos e propósitos do objeto sob avaliação	2-autonomia dos avaliadores de forma a não afetar a imparcialidade e transparência dos processos	3- clareza e detalhamento dos propósitos e perguntas avaliativas para os principais atores;	4- procedimentos de coleta de dados, técnicas de análise e processo de produção e construção dos resultados de acordo com os propósitos da avaliação;	3- critérios empregados para o processo de valoração explícitos e identificados	7- comunicação dos resultados precisa e oportuna, com difusão dos resultados em relatório ou informe de acesso público, tendo a devolutiva dos resultados para os atores envolvidos;
Avaliabilidade Adequada					Compreensão cultural	
3- viabilidade contextual, social e política (reconhecendo e ponderando os interesses e necessidades dos indivíduos e grupos)			5- garantia, no sentido de reconhecimento e reflexão sobre as diferentes perspectivas, do processo de desenho e execução de uma avaliação, levando em consideração as opiniões dos principais atores envolvidos	6- importância das conclusões de uma avaliação, no sentido de serem claras, documentadas e consistentes com os propósitos envolvidos		8- recomendações úteis e factíveis, com probabilidade de serem implementadas.
				Avaliabilidade adequada	1- inclusão, igualdade e equidade	
				1- gestão efetiva da avaliação (processos orientados por estratégias eficientes e efetivas, com utilização eficiente e transparente dos recursos);		
				2- utilização de procedimentos práticos (pertinentes à forma em que se executa e opera o objeto sob avaliação);		
				4- realista (seu desenvolvimento depende da coerência entre os recursos financeiros, humanos e temporais).		
Direitos e integridade	Contextualização e valoração	Método e viabilidade:	Aprendizagem e utilização	Aprendizagem e utilização	Direitos e integridade	Aprendizagem e utilização
3-respeito as identidades culturais dos atores, especialmente à língua, aos modos de dialogar, às estruturas de vínculo, parentesco e às posições sociais e geracionais	2- conhecimento sobre o objeto a ser avaliado, em suas dimensões teóricas, lógicas, políticas, técnicas e operacionais	7- coloca o papel da gestão de uma avaliação, no sentido de envolverem profissionais adequados, respeitando as etapas, prazos e entregas pactuadas e assegurando uso sustentável e transparente dos recursos	1- inclusão das perspectivas dos diferentes interessados na avaliação, garantindo a inclusão dos grupos excluídos e marginalizados	2- clareza para os atores envolvidos quanto ao seu sentido e finalidade	2- diversidade e equidade, enfocando a necessidade de tratamento equânime a todos os envolvidos, com critérios que respeitem as suas condições étnico-raciais, de gênero, econômica ou cultural	4- geração de processos e resultados capazes de ampliar a compreensão do objeto avaliado

Fonte: Os autores (2021).



Quadro 3. Continuação...

Contexto	A concepção do avaliado	O papel do avaliador	O envolvimento de stakeholders	Os procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
<p>7-legalidade das avaliações, no sentido de estarem respaldadas em leis, normas, regulamentações e boas práticas vigentes</p>			<p>Direitos e integridade</p> <p>1-respeito aos valores e condições culturais, sociais, políticas, institucionais, geracionais, físicas e subjetivas das pessoas envolvidas numa avaliação</p>	<p>3- perguntas avaliativas que gerem aprendizados e favoreçam os processos decisórios</p> <p>5- conclusões claras e coerentes, capazes de responder às suas finalidades</p> <p>Direitos e integridade</p> <p>4- reciprocidade nas avaliações, devendo ser realizada de modo a destacar as assimetrias estruturais existentes e evitar que interesses de determinados grupos se sobreponham</p> <p>5- explicitação de interesses por pessoas, equipes ou instituições envolvidas numa avaliação de modo que os riscos sejam conhecidos e mitigados</p> <p>Direitos e integridade</p> <p>3-critérios defensáveis, provenientes da literatura, postulados éticos, especialistas, leis e normas, entre outros</p> <p>Método e viabilidade</p> <p>8- importância da meta-avaliação no processo avaliativo, incluindo profissionais internos e externos, numa perspectiva autocrítica e com emprego dos padrões</p>		<p>6- recomendações que respeitem as condições reais de governabilidade e factibilidade</p> <p>7- comunicação de forma precisa e oportuna, favorecendo a compreensão e aprendizados dos atores envolvidos</p>
<p>Contextualização e Valoração</p> <p>1- contexto social, cultural, político, racial, econômico onde o objeto a ser avaliado está inserido</p>			<p>6- transparência das avaliações, no sentido de terem seus resultados compartilhados com os interessados e instituições envolvidas</p> <p>Método e viabilidade</p> <p>3- métodos robustos, válidos e confiáveis para melhor resposta as perguntas avaliativas.</p>			
<p>Método e viabilidade</p> <p>1, 2- viável no contexto social, político, cultural em que são feitas, além de levarem em consideração o tempo oportuno para ocorrerem e de fato colaborarem com o momento político e institucional colocado</p>						

Fonte: Os autores (2021).



Quadro 3. COntinuação...

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	Os procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
				Contextualização e Valoração 4- técnicas e procedimentos de acordo com a sua finalidade e contexto, devendo favorecer a percepção de fenômenos de discriminação 5- métodos e técnicas explícitos e registrados, de modo a permitir futuros exames pelos atores envolvidos 6- autônomas e independentes, gerando informações livres de constrangimentos políticos		
				Contextualização e Valoração 4- explicitação dos critérios de forma consistente e válida, favorecendo a compreensão dos mecanismos de valoração empregados pelos atores envolvidos		

Fonte: Os autores (2021).



Quadro 4. Apreciação de qualidade de avaliações segundo as Diretrizes Brasileiras.

Contexto	A concepção do avaliando	O papel do avaliador	O envolvimento de stakeholders	Os Procedimentos metodológicos	O processo de Valoração	Usos e influências (o valor agregado pela avaliação?)
			Compreensão cultural	Relevância e utilidade		Relevância e utilidade
			2- respeito a identidade cultural e dignidade dos envolvidos, com utilização de protocolos de avaliação aprovados por comitê de ética, quando solicitado;	2-propósitos devem estar explícitos e documentados.		4- utilidade da avaliação segundo as necessidades dos atores envolvidos
			3-cooperação, confiança e coesão entre os participantes, garantidos por meio de relações de reciprocidade e igualdade nas trocas e relações estabelecidas.	Princípios éticos e jurídicos		5- resultados das avaliações devem contribuir para que os atores envolvidos redescubram, replanejem os avaliandos
			Relevância e utilidade	1- os procedimentos da avaliação devem respeitar os direitos dos executores e beneficiários com relação a confidencialidade da informação, crenças, costumes e dignidade das pessoas;		6- necessidade de um plano de comunicação dos achados de uma avaliação
			1-necessidade do desenho e execução de avaliações levarem em consideração a perspectiva dos diferentes atores.	3-transparência quanto a publicização dos documentos, disponíveis para os interessados envolvidos e público geral		7- toda avaliação preocupada em promover o uso responsável e pertinente dos seus resultados
			Princípios éticos e jurídicos			
			4- legalidade, por meio de um instrumento firmado entre as partes envolvidas, explicitando os diversos acordos previamente estabelecidos			

Fonte: Os autores (2021).



adotados em uma avaliação devem empregar de forma eficaz estratégias de gerenciamento (F1). O segundo padrão (F2) recomenda que os procedimentos de uma avaliação devem ser práticos e responsivos à forma como a intervenção é operada. O terceiro padrão (F3) diz respeito a viabilidade contextual de uma avaliação, isto é, reforça que a mesma deve reconhecer, monitorar e equilibrar os interesses e necessidades dos indivíduos e grupos envolvidos. O último padrão reforça que as avaliações devem usar os recursos de forma eficaz e eficiente (F4).

O atributo propriedade engloba sete padrões. Eles asseguram que as avaliações serão conduzidas e divulgadas dentro dos preceitos legais e éticos, promovendo a autonomia, os direitos e o bem-estar de todos os interessados, explicitando e considerando os conflitos de interesse, caso existam (P3, P6). Para isso devem ser responsivas e inclusivas e orientadas segundo as necessidades locais (P1); negociadas formalmente (P2); clara, compreensível e justa (P4); transparentes na divulgação dos resultados a todos os interessados (P5); cumprindo os procedimentos e processos fiscais quando voltadas para a responsabilidade fiscal (P7).

O atributo acurácia ou precisão envolve 8 padrões, que contribuem para a construção de evidências cientificamente válidas e socialmente legítimas. Neste sentido, os propósitos e procedimentos, além da descrição do programa, devem ser claros, confiáveis e documentados (A1, A3, A4) e até meta-avaliados. As informações coletadas devem estar de acordo com os propósitos mapeados e dar suporte a interpretações válidas (A2); com desenhos, manejo e armazenamento de dados adequados, válidos, legítimos e sistematicamente revisados (A5, A6, A7). Além disso, as informações quanti e qualitativas devem ser analisadas apropriadamente, gerando conclusões justificadas e relatórios imparciais, que permitam a efetiva resposta à pergunta avaliativa (A8).

Por fim, o atributo responsabilização (*Evaluation Accountability*) possui três padrões que englobam a necessidade de que uma avaliação seja documentada por completo em termos dos seus propósitos, desenhos implementados, procedimentos, dados e resultados (E1); além de serem meta-avaliadas interna e externamente (E2, E3) e sempre que a qualidade precisar ser verificada, com propósitos somativos ou formativos (Quadro 2).

Vale lembrar que a inclusão e detalhamento deste atributo é relativamente recente, sendo considerado anteriormente um padrão incluído no atributo propriedade (JCSEE, 1994, 2021; Worthen et al., 1997). De maneira geral, ocorreu um rearranjo envolvendo diversos atributos. O padrão meta-avaliação, contido dentro do atributo acurácia na versão de 1994, foi alocado dentro do novo atributo responsabilização, subdividindo-se em meta-avaliação interna e externa (R2, R3). Além disso, na versão atual, novos padrões foram incorporados como propósitos negociados (U3) e produtos e processos significativos (U6) no atributo utilidade. Outra modificação a ser notada envolveu a alocação de gestão do projeto (V1) no atributo viabilidade. Observou-se também a inclusão do padrão documentação da avaliação (R1) no atributo responsabilização, além da retirada dos padrões interação humana e avaliação completa do atributo propriedade.

Outro ponto que merece destaque é que, apesar de reiterada referência a processos de negociação com interessados e com fatores do contexto, os critérios do JCSEE não explicitam claramente o papel do avaliador, que ora parece incorporar a prática de um negociador e ora é referido como um construtor de evidências científicas e legitimadas socialmente.

Os critérios OCDE/DAC

A Organização para a Cooperação e Desenvolvimento Econômico (OCDE/DAC) foi criada no período pós segunda guerra mundial, em 1961, sendo composta por 30 países-membros, incluindo o Brasil. Seu principal objetivo é a promoção de padrões internacionais para resolução de problemas ou situações relacionadas ao desenvolvimento econômico, financeiro, comercial, social e ambiental (OECD, 2010).

A apreciação da qualidade de avaliações pela OCDE/DAC é recomendada por meio de critérios que fornecem um guia de boas práticas para o desenvolvimento de avaliações construídos por meio de consenso internacional, destinando-se a servirem como incentivo e inspiração para



melhoria da prática em avaliação. Os critérios podem ser usados em diferentes estágios do processo avaliativo, seja para a prestação de contas, o apoio no processo de aprendizagem institucional, no monitoramento e gestão dos resultados e na avaliação da qualidade das avaliações (meta-avaliação) (OECD, 2010, 2019). A apreciação incluía a relevância, eficiência, eficácia, impacto e sustentabilidade. Em 2019, os critérios foram atualizados com a introdução da coerência (OECD, 2019).

Para a OCDE/DAC, a relevância relaciona-se a adequação das intervenções às necessidades e prioridades da população-alvo, país ou instituição. A eficiência mede a possibilidade de a intervenção transformar os recursos/insumos alocados em efeitos (quanti e qualitativos), de forma econômica e ajustada as demandas do contexto em que a intervenção ocorre. A eficácia diz respeito ao grau de alcance dos objetivos e resultados pretendidos/planejados a partir de uma demanda motivadora. O impacto relaciona-se as mudanças, transformações positivas ou negativas, diretas ou indiretas, intencionais ou não, produzidas por uma intervenção. Por fim, a sustentabilidade versa sobre a probabilidade de manutenção dos benefícios provenientes da realização de uma intervenção na sociedade, ou seja, o valor e mérito da intervenção manter-se tomando em consideração os custos correntes da mesma. A coerência diz respeito a quanto a intervenção sob análise está em sinergia, complementa, agrega valor a outras intervenções existentes e é consistente com normas e padrões internacionais, evitando duplicação de esforços (OECD, 2019).

Observa-se que os critérios sofreram alterações em suas definições. Por exemplo, a relevância sofreu adição de termos como “desenho” para captura de outros elementos da qualidade; “prioridade dos beneficiários”; “necessidade”; e incorporou a dimensão tempo (anteriormente incluída em uma nota). Além disso, houve a exclusão de termos como “desenvolvimento” e “doadores”, presentes na antiga definição.

A nova definição de eficiência teve como foco a cadeia de resultados, referindo-se a aspectos da implementação e gestão. Houve adição dos termos “em comparação com alternativas viáveis no contexto” e “oportunidade”, reforçando a ideia do valor contextualizado da intervenção e da pertinência de uma análise de custo-benefício.

No critério eficácia, observa-se a adição dos termos “resultados” e “resultados diferenciais”, justificando-se a necessidade do olhar diferenciado quanto a distribuição dos resultados entre os grupos envolvidos. Embora reconhecendo que nem toda intervenção tem como propósito modificar perfis de desigualdade, o documento atual encoraja avaliadores a incluir esta questão especialmente em contextos de reconhecida desigualdade, seja ela socioeconômica, étnico-racial ou de gênero.

A coerência captura uma perspectiva não abordada pelos critérios anteriores, incentivando os avaliadores a compreenderem o papel de uma intervenção dentro de um determinado sistema (organização, setor, área temática, país). Observa-se a sua relação com o critério eficácia.

O critério impacto sofreu adição do termo “níveis mais elevados” para capturar a importância, o escopo e a natureza transformadora dos efeitos obtidos e remoção dos termos “direto e indireto” e “primário e secundário”, já que seriam termos abordados dentro do critério eficácia.

A antiga definição de sustentabilidade era centrada no doador e focada no financiamento externo. A nova versão deixa claro que a sustentabilidade possui várias dimensões (financeira, econômica, social e ambiental). Observa-se exclusão dos termos “benefícios continuados de longo prazo” e “resiliência ao risco” e manutenção do termo “benefícios líquidos”.

Uma observação interessante é que na definição dos critérios o documento considera de qualidade avaliações que abordem características específicas da intervenção. Assim, uma avaliação cumpre os critérios de qualidade se ela aprecia a relevância e as outras características de uma intervenção. Dessa forma, o documento não discute aspectos metodológicos e operacionais do processo avaliativo, colocando com certa ambiguidade o que está em foco, ou seja, a intervenção ou a avaliação.



Diretrizes para a avaliação para a América Latina e Caribe

As diretrizes, organizadas em 2016 pela Rede de Monitoramento, Avaliação e Sistematização da América Latina e do Caribe (ReLAC), em colaboração com o FOCEVAL, MIDEPLAN e DEval (Rodríguez Bilella et al., 2016), enfatizam que os resultados das avaliações devem trazer subsídios para a tomada de decisões. Esses resultados devem contribuir para a aprendizagem de equipes e organizações, acrescentando conhecimento aos problemas que são objeto da ação pública.

As cinco diretrizes foram organizadas com alguns objetivos, dentre eles o de contribuir como um marco comum de referência a ser usado na obtenção de avaliações com altos padrões de qualidade. A terminologia utilizada no documento considera que as diretrizes se subdividem em 26 dimensões as quais incluem aspectos que promovam a aprendizagem e geração de conhecimentos a partir da prática profissional em avaliação (Rodríguez Bilella et al., 2016).

A elaboração das diretrizes se inspirou no trabalho realizado por outras associações, sociedades e redes regionais e nacionais de avaliação, que utilizaram as dimensões definidas pelo *Program Evaluation Standards* (PES) do *Joint Committee on Standards for Educational Evaluation* (JCSEE) (JCSEE, 2021). Foram propostas cinco diretrizes: avaliação rigorosa, avaliabilidade adequada, avaliação conduzida com princípios éticos e jurídicos, compreensão cultural adequada, relevância e utilidade (Rodríguez Bilella et al., 2016).

A diretriz avaliação rigorosa desdobra-se em oito dimensões que versam sobre: 1) a importância de uma avaliação contextualizada social, cultural, política e economicamente; 2) a necessidade da descrição clara e detalhada das ações, produtos e propósitos do objeto sob avaliação; 3) a clareza e o detalhamento dos propósitos e perguntas avaliativas para os principais atores; 4) os procedimentos de coleta de dados, técnicas de análise e processo de produção e construção dos resultados de acordo com os propósitos da avaliação; 5) a garantia, no sentido de reconhecimento e reflexão sobre as diferentes perspectivas, do processo de desenho e execução de uma avaliação, levando em consideração as opiniões dos principais atores envolvidos; 6) a importância das conclusões de uma avaliação, no sentido de serem claras, documentadas e consistentes com os propósitos envolvidos; 7) a comunicação dos resultados ser precisa e oportuna, com difusão dos resultados em relatório ou informe de acesso público, tendo a devolutiva dos resultados para os atores envolvidos; e 8) as recomendações serem úteis e factíveis, com probabilidade de serem implementadas.

A avaliabilidade adequada considera quatro dimensões: 1) gestão efetiva da avaliação (processos orientados por estratégias eficientes e efetivas, com utilização eficiente e transparente dos recursos); 2) utilização de procedimentos práticos (pertinentes à forma em que se executa e opera o objeto sob avaliação); 3) viabilidade contextual, social e política (reconhecendo e ponderando os interesses e necessidades dos indivíduos e grupos); e 4) realista (seu desenvolvimento depende da coerência entre os recursos financeiros, humanos e temporais).

Quanto aos princípios éticos e jurídicos, são relacionadas quatro dimensões, que destacam: 1) os procedimentos da avaliação, no sentido de respeitar os direitos dos executores e beneficiários com relação a confidencialidade da informação, crenças, costumes e dignidade das pessoas; 2) autonomia dos avaliadores de forma a não afetar a imparcialidade e transparência dos processos; 3) transparência quanto à publicização dos documentos, disponíveis para os interessados envolvidos e público geral; e 4) legalidade, por meio de um instrumento firmado entre as partes envolvidas, explicitando os diversos acordos previamente estabelecidos.

A compreensão cultural apresenta três dimensões que valorizam: 1) a inclusão, igualdade e equidade; 2) respeito a identidade cultural e dignidade dos envolvidos, com utilização de protocolos de avaliação aprovados por comitê de ética, quando solicitado; 3) cooperação, confiança e coesão entre os participantes, garantidos por meio de relações de reciprocidade e igualdade nas trocas e relações estabelecidas.

A última diretriz, que fala sobre a relevância e utilidade das avaliações, desdobra-se em sete dimensões. A primeira dimensão mostra a necessidade do desenho e execução de avaliações



levarem em consideração a perspectiva dos diferentes atores. A segunda aponta que os seus propósitos devem estar explícitos e documentados. A terceira ressalta que os critérios empregados para o processo de valoração devem estar explícitos e identificados.

A utilidade da avaliação segundo as necessidades dos atores envolvidos é destacada na quarta dimensão. A quinta dimensão enfatiza que os resultados das avaliações devem ser úteis e auxiliar os atores envolvidos a redescobrirem e revisarem seu “olhar” sobre o objeto avaliado. A sexta coloca a necessidade de um plano de comunicação dos achados de uma avaliação. A última dimensão apresenta que toda avaliação deve estar preocupada em promover o uso responsável e pertinente dos seus resultados.

Conforme pode-se verificar no Quadro 3, as DALAC destacam a interação entre a relação avaliadores, stakeholders e desenho metodológico da avaliação. A sobreposição do desenho avaliativo à interação com stakeholders reforça a importância da escolha de processos avaliativos participativos, uma vez que a reciprocidade de interesses entre a equipe de avaliação e participantes requer a construção de pontes entre os diferentes grupos. Outro ponto relacionado diz respeito as relações entre desenho metodológico e utilidade dos achados da avaliação, mais uma vez reforçando a importância da interação entre os atores.

Diretrizes para a prática da avaliação no Brasil

As diretrizes representaram um primeiro esforço da Rede Brasileira de Monitoramento e Avaliação (RBMA) na proposição de marcos comuns para as avaliações no país (Silva et al., 2020; Joppert, 2021). Elaborada em 2020, contou com cerca de 200 participantes, entre eles professores, pesquisadores, consultores, gestores, ativistas, técnicos de instituições públicas, privadas e de cooperação internacional. Vários especialistas explicitaram a importância da iniciativa como sinal de amadurecimento do tema, favorecendo pensamentos críticos e orientando práticas e comportamentos dos diversos atores envolvidos.

A participação dos interessados no processo avaliativo (governos, sociedade civil, sindicatos e partidos) foi estimulada, baseando-se no pressuposto que estes atores demandam, avaliam e meta avaliam. A prática destes profissionais é bastante diversa, compreendendo a leitura crítica de avaliações, a realização de avaliações internas, a elaboração de termos de referência e a contratação e implementação de avaliações externas. Constituem-se assim em um grupo certamente interessado na qualidade de suas entregas.

O documento brasileiro é organizado em quatro dimensões e 26 diretrizes. Verifica-se que, ao contrário do observado no documento da América Latina e Caribe (Rodríguez Bilella et al., 2016), o brasileiro inverte a terminologia utilizando o termo dimensão como o mais abrangente e o subdividindo em diretrizes. As quatro dimensões que o compõem são: aprendizagem e utilização; direitos e integridade; contextualização e valoração; método e viabilidade (Silva et al., 2020).

A dimensão aprendizagem e utilização é composta por 7 diretrizes que qualificam uma avaliação, isto é: 1) Níveis de participação adequados aos interessados, que inclui as perspectivas dos diferentes interessados na avaliação, garantindo a inclusão dos grupos excluídos e marginalizados; 2) Propósitos acordados mutuamente, ou seja, avaliações devem ser claras para os atores envolvidos quanto ao seu sentido e finalidade; 3) Perguntas relevantes, ou seja, apresentar perguntas avaliativas que gerem aprendizados e favoreçam os processos decisórios na avaliação; 4) Resultados úteis, isto é, ser capaz de gerar processos e resultados de forma a ampliar a compreensão do objeto avaliado pelos interessados; 5) Conclusões fidedignas, ou seja, ter conclusões claras e coerentes, capazes de responder às suas finalidades; 6) Recomendações contextualizadas, isto é, apresentar recomendações que respeitem as condições reais de governabilidade e factibilidade; e 7) Comunicação cuidadosa, ou seja, a avaliação (processos e resultados) deve ser comunicada de forma precisa e oportuna, favorecendo a compreensão e aprendizados dos atores envolvidos.

A dimensão direitos e integridade inclui 7 diretrizes. A primeira delas fala sobre o respeito aos direitos das pessoas, se referindo aos valores e condições culturais, sociais, políticas, institucionais, geracionais, físicas e subjetivas das pessoas envolvidas numa avaliação.



A segunda versa sobre a diversidade e equidade, enfocando a necessidade de tratamento equânime a todos os envolvidos, com critérios que respeitem as suas condições étnico-raciais, de gênero, econômica ou cultural. A terceira diretriz enfoca o respeito às identidades culturais dos atores, especialmente à língua, aos modos de dialogar, às estruturas de vínculo, parentesco e às posições sociais e geracionais. A quarta ressalta o aspecto da reciprocidade nas trocas entre os atores envolvidos nas avaliações, devendo ser realizada de modo a destacar as assimetrias estruturais existentes e evitar que interesses de determinados grupos se sobreponham. A quinta diretriz assinala que os aspectos referentes a explicitação de interesses por pessoas, equipes ou instituições envolvidas numa avaliação, devem ocorrer de modo que os riscos sejam conhecidos e mitigados. A sexta coloca a transparência das avaliações, no sentido de terem todos os seus processos compartilhados com os interessados e instituições envolvidas. A última diretriz, Legalidade, apresenta a necessidade de esclarecer as bases e os contratos legais da avaliação proposta, ou seja, elas devem ser respaldadas em leis, normas, regulamentações e boas práticas vigentes.

A dimensão da contextualização e valoração contempla quatro diretrizes, que reforçam a importância de as avaliações levarem em consideração: 1) o contexto social, cultural, político, racial, econômico onde o objeto a ser avaliado está inserido; 2) a descrição do objeto a ser avaliado, em suas dimensões teóricas, lógicas, políticas, técnicas e operacionais; 3) o uso de critérios válidos e defensáveis, provenientes da literatura, postulados éticos, especialistas, leis e normas, entre outros; 4) a explicitação destes critérios de forma consistente e válida, favorecendo a compreensão dos mecanismos de valoração empregados pelos atores envolvidos.

A dimensão denominada método e viabilidade abarca oito diretrizes. As duas primeiras diretrizes, viabilidade e tempo oportuno, ressaltam que a realização de avaliações deve ser viável no contexto social, político, cultural em que são feitas, além de levarem em consideração o tempo oportuno para ocorrerem e de fato colaborarem com o momento político e institucional colocado. A terceira diretriz enfatiza o emprego de métodos consistentes, isto é, robustos, válidos e confiáveis para melhor resposta às perguntas avaliativas. A quarta apresenta que as técnicas e os procedimentos empregados em uma avaliação devem ser válidos e estar de acordo com a sua finalidade e contexto, devendo favorecer a percepção de fenômenos de discriminação. A quinta diretriz coloca que os métodos e técnicas empregados devem estar explícitos e registrados, de modo a permitir futuros exames pelos atores envolvidos. A sexta mostra que as avaliações devem ser autônomas e independentes, gerando informações livres de constrangimentos políticos. A sétima diretriz coloca a qualidade necessária a gestão de uma avaliação, no sentido de envolverem profissionais adequados, respeitando etapas, prazos e entregas pactuadas e assegurando uso sustentável e transparente dos recursos. A última diretriz fala da importância da meta-avaliação no processo avaliativo, incluindo profissionais internos e externos, numa perspectiva autocrítica.

O Quadro 4 descrito abaixo, evidencia um núcleo importante de preocupação e o alinhamento das DPAB a processos avaliativos participativos e enfaticamente reitera a preocupação com a transparência daqueles referentes a valoração. Esta característica pode ser observada na sistemática referência a processos avaliativos inclusivos ao longo de todos os “pontos de tensão” analisados.

Síntese reflexiva da descrição dos documentos analisados

Antes de analisar comparativamente as categorias utilizadas, deve-se ressaltar que os critérios empregados pela OCDE/DAC (OECD, 2010, 2019) abordam a qualidade de uma avaliação pela apreensão detalhada do avaliando. Este aspecto dá um caráter particular a este conjunto em relação aos outros três e, portanto, não foi incluído na síntese reflexiva. As categorias OCDE/DAC não têm comparabilidade com a dos outros três documentos, os quais se referem especificamente ao processo avaliativo. Como ele se refere a qualidade como abrangência de características da intervenção pela avaliação, os seus critérios não analisam aspectos metodológicos e operacionais do processo avaliativo. Assim, eliminam a especificidade do “labor” em avaliação, que, como qualquer outro, deve ser submetido à apreciação de qualidade.



Por exemplo, pelo critério OCDE/DAC, não se aprecia a eficiência da avaliação, mas se considera que ela é de qualidade se aborda a eficiência da intervenção avaliada.

Para esclarecer melhor este ponto, o documento do JCSEE (1994) qualificava uma avaliação como eficiente se ela respondia ao padrão de custo efetividade. Atualmente (JCSEE, 2021), esse padrão de qualidade é atendido sob a análise do padrão de uso de recursos em uma avaliação, mantendo o significado, ou seja, para o JCSEE, é a avaliação que deve ser custo efetiva ou ter um custo-benefício mensurável.

O Quadro 5 evidencia um resumo da terminologia das categorias e subcategorias identificadas nos quatro documentos.

Destaca-se que, apesar dos documentos para a América Latina e Caribe e o do Brasil incluírem o termo “diretriz” em seus títulos, em seu conteúdo utilizam a mesma expressão para designar diferentes categorias de classificação. No caso latino-americano (Rodríguez Bilella et al., 2016), as diretrizes são categorias abrangentes que se subdividem em dimensões e, no brasileiro (Silva et al., 2020), são subdivisões de grandes categorias nomeadas como dimensões.

Além da terminologia, os três documentos apresentam diversidade, sobreposição e similaridades entre os significados das categorias. Observa-se que a categoria relevância aparece apenas no documento da América Latina e Caribe (Rodríguez Bilella et al., 2016), e está associada à utilidade. Esta última é presente como atributo no documento do JCSEE (2021). Nas dimensões brasileiras (Silva et al., 2020), a utilidade é citada como utilização e associada a aprendizagem, não estando de forma explícita relacionada à relevância. De maneira geral, percebe-se que as categorias e suas subdivisões não têm limites definidos entre si e algumas são condições necessárias para a conformidade à outras.

Por exemplo, as referências sobre a dimensão Utilização e Aprendizagem nas DPAB (Silva et al., 2020) têm complementariedade em diretrizes (subdivisões) de outras dimensões. Na dimensão Direitos e Integridade, a sexta diretriz coloca a importância da transparência das avaliações, no sentido de terem seus resultados compartilhados com os interessados e instituições envolvidas, condições que parecem necessárias para utilização. Este ponto fica mais claro na exploração das informações identificadas referentes aos pontos de tensionamento na apreciação de qualidade. A síntese reflexiva dos documentos será apresentada e comentada seguindo a ordem de apresentação do Quadro 1.

Quadro 5. Resumo das terminologias empregadas nos diferentes documentos analisados.

	Categorias mais abrangentes	Subcategorias
JCSEE	Cinco atributos: Utilidade; Viabilidade; Propriedade; Precisão; Responsabilização	Trinta padrões
OCDE	Cinco critérios: Relevância; Eficiência; Eficácia; Impacto e sustentabilidade; e Coerência	
LAC	Cinco diretrizes: Avaliação rigorosa; Avaliabilidade adequada; Condução alinhada aos princípios ético-jurídicos; Compreensão cultural adequada; e Relevância e utilidade	Vinte e seis dimensões
BR	Quatro Dimensões: Aprendizagem e utilização; Direitos e integridade; Contextualização e valoração; Método e viabilidade	Vinte e seis diretrizes

Fonte: Os autores (2021).



Em relação à descrição do contexto e suas implicações, tanto para a intervenção como para a avaliação, elas estão presentes nos três conjuntos de documentos analisados. A influência do contexto é referida em inúmeras categorias e subcategorias, entretanto, dois aspectos se destacam: a) qualidade da avaliação referida a sua sensibilidade cultural e política, e b) o alinhamento dos processos avaliativos relacionados aos aspectos legais específicos.

Entre atributos e padrões do JCSEE (1994, 2021), a ênfase ao contexto é dada no atributo Utilidade, referida aos padrões de Respeito aos Valores Culturais e no atributo Propriedade, padrões de Valores Éticos-Legais e de Responsividade. Já nas DALAC (Rodríguez Bilella et al., 2016), o contexto é relacionado seja como condicionante da conformidade à diretriz de Preceitos Éticos Jurídicos, seja como modulador dos contratos de avaliação realizados até o respeito e a dignidade dos sujeitos envolvidos, abordado na diretriz Compreensão Cultural. As DPAB (Silva et al., 2020) enfatizam a importância do contexto jurídico e sociopolítico, mas ressaltam a importância do contexto institucional com implicações para a escolha do desenho metodológico, da qualidade e tempestividade das entregas.

A concepção do avaliando é ponto de convergência entre os documentos avaliados. Todos abordam a necessidade de clareza, conhecimento e descrição detalhada do avaliando, embora não coloquem claramente que tipo de objeto seria, por exemplo: sistemas, dispositivos ou redes. Nos três documentos ainda é incipiente a concepção, tanto do avaliando quanto da avaliação, como sistemas dinâmicos complexos e situados, predominando a abordagem de ambos como sistemas lineares, nos quais a interação entre componentes e destes com o contexto é pouco explorada para além de uma lógica linear (Patton, 2020).

Este ponto pouco explorado revela de maneira indireta a concepção sobre o papel do avaliador que percorre os três conjuntos de documentos. Há várias referências a autonomia dos avaliadores em relação aos financiadores e relações de poder, destacando-se o seu alinhamento às evidências científicas como salvaguarda de independência. Ressalta-se que, de forma indireta, identificada a partir de diferentes referências sobre relações formais entre avaliadores e demandantes da avaliação, predomina o papel do avaliador como um ator externo a intervenção e a seu ambiente. Ou seja, a imagem mental do avaliador, presente nos documentos analisados é a de um ator externo ou pesquisador, confirmado pela citação explícita no documento do JCSEE (1994, 2021) à necessária qualificação da equipe de avaliação. Isto é particularmente interessante, uma vez que, tanto na DALAC (Rodríguez Bilella et al., 2016) como nas DPAB (Silva et al., 2020), o avaliador já se configura como um mediador ou negociador, no que diz respeito aos interesses dos stakeholders e na transparência dos processos de valoração e de construção de evidências.

Em relação ao papel dos avaliadores, destacam-se duas vinculações de sentido: a função e as competências. Nos três documentos predominam as funções do avaliador como pesquisador e, embora de forma explícita, os documentos Latino-Americano e Brasileiro (Rodríguez Bilella et al., 2016; Silva et al., 2020) já apontem a função e competências de gestão. A função de negociação e as respectivas competências aparecem, ainda que de forma implícita, nas DPAB (Silva et al., 2020) quando pontuam a conveniência de composições de equipes mistas e com inúmeras referências a abordagens responsivas (Quadro 4 - Processo de Valoração-Método e Viabilidade).

A apropriação do papel dos avaliadores em todos os documentos é expressa pela expectativa de atributos do processo avaliativo, isto é, legitimidade, transparência, uso de métodos e técnicas cientificamente reconhecidas. No material analisado, tanto a função de pesquisador quanto a de negociador não consideram a relevância do avaliador/equipe de avaliação nos processos de produção e mobilização do conhecimento, além de não articular as mudanças esperadas, seja no programa ou em suas consequências sociais. Não se verifica, em quaisquer dos documentos analisados, a vinculação do avaliador à função de tradutor ou mesmo de “translador” (Oliveira et al., 2017; Abreu et al., 2017), não se discutindo a sua influência nos processos de valoração.

As relações entre avaliadores, os envolvidos na intervenção e na avaliação são abordadas nos três documentos. Predominam referências a inclusão dos participantes em certos estágios do processo avaliativo, tais como: elaboração de perguntas avaliativas e compartilhamento do



andamento do processo avaliativo de forma a promover o uso de seus achados. A necessidade da inclusão de interesses de grupos sem voz, estigmatizados e discriminados por diferenças étnico-raciais, gênero ou idade é uma referência comum aos três documentos. Entretanto, ela se torna mais evidente nas DALAC (Rodríguez Bilella et al., 2016) e permeia todos os pontos analisados na diretriz brasileira (Silva et al., 2020). Pode-se, assim, identificar a inclusão dos interessados potenciais na avaliação como um espectro que parte das considerações contextuais socioculturais, inclui o processo de valoração e alcança o compromisso com a utilização e influência da avaliação de forma a responder a diversidade e desigualdades que possam existir no grupo de participantes.

Com relação aos procedimentos metodológicos, observa-se que eles envolvem cinco grupos de conteúdo/sentido: 1) o embasamento das evidências e das análises na metodologia científica; 2) a adaptação de métodos e técnicas ao contexto (responsividade); 3) a transparência dos procedimentos; 4) a gestão do manuseio e garantia de qualidade dos processos e informações, aos quais se somam 5) os procedimentos que garantam a *accountability* financeira, técnica e ética da avaliação.

Dessa forma, diferentes atributos do JCSEE (1994, 2021) são relacionados aos procedimentos metodológicos tais como a viabilidade, a propriedade, a acurácia e a responsabilização. Envolvem padrões que dizem respeito desde a gestão de logística e qualidade de evidências, alinhamento ao método científico, devido registro e documentação.

As DALAC (Rodríguez Bilella et al., 2016) enfatizam o compromisso entre a metodologia e a ética, reforçando a confidencialidade e o respeito à dignidade das pessoas envolvidas, além de destacar a importância de conclusões baseadas em evidências documentadas e consistentes. Esse documento considera de grande relevância a gestão efetiva do processo avaliativo e sua coerência realista aos limites de recursos financeiros, humanos e temporais.

As DPAB (Silva et al., 2020) reforçam a utilização de métodos cientificamente robustos e válidos, associados as técnicas pertinentes e com o devido registro para posterior verificação, se necessária. Enfatizam a transparência e a utilização de metodologias que destaquem as simetrias de poder e participação, com a finalidade de evitar que interesses específicos se sobreponham durante o processo avaliativo, modelando os seus resultados.

De forma geral, os três documentos identificam a avaliação como um campo de práticas científicas, entretanto, o compromisso com processos metodológicos inclusivos é explícito com maior clareza nas DALAC (Rodríguez Bilella et al., 2016) e DPAB (Silva et al., 2020).

A análise realizada identificou que a qualidade da valoração é um dos pontos menos explorados pelos documentos, seja como padrão, atributo, dimensão ou diretriz. A opção por classificar os atributos que se referem à meta-avaliação como alinhados à questão da valoração foi decorrência da compreensão que o requisito prioritário de qualidade de uma avaliação é a clareza referente a seus processos de valoração.

Neste sentido, os documentos do JCSEE (1994, 2021) enfatizam a meta-avaliação tanto interna como externa como mecanismos para garantir a qualidade da avaliação, subentendida como alinhamento do processo avaliativo aos seus propósitos. Em outras palavras, a avaliação é valorada positivamente pelo critério de *accountability* e, portanto, não se refere aos critérios de valor agregado pela intervenção, apenas afirmando o atributo de responsabilização. De forma implícita, a valoração para o JCSEE está vinculada a qualidade metodológica da avaliação, uma vez que seu critério de valoração é referido as evidências baseadas cientificamente.

Comparativamente, as DALAC (Rodríguez Bilella et al., 2016) destacam que a valoração deve contemplar dois aspectos. O primeiro é a transparência de critérios e processos. O segundo considera as questões da igualdade e equidade sem negligenciar as evidências embasadas cientificamente e o alinhamento aos propósitos da avaliação.

O documento brasileiro (Silva et al., 2020), além de recomendar a transparência, é enfático em definir que os processos de valoração devem abordar o tratamento equânime e critérios que respeitem as condições de diversidade e desigualdade. Esta orientação reflete-se



nas recomendações para os procedimentos metodológicos ressaltando a importância de processos avaliativos interativos, adaptativos e inclusivos.

Os usos e influências da avaliação são tratados nos três documentos. Os usos e utilidade da avaliação são abordados nas diretrizes do JCSEE (1994, 2021), incluindo um padrão que requer do processo avaliativo a promoção dos usos de seus achados.

Nas DALAC (Rodríguez Bilella et al., 2016) predominam como condições para utilização dos achados a legitimidade dada pelas boas práticas, a inclusão dos diferentes interesses identificados, a transparência e um plano de comunicação efetivo. As DPAB (Silva et al., 2020) enfatizam a importância das conclusões claras e coerentes, com recomendações realistas, favorecendo a compreensão e aprendizado dos atores envolvidos. Vale ressaltar nesta última a nítida abordagem da avaliação como um processo reflexivo pedagógico, aspecto também elaborado com menor ênfase pelas DALAC (Rodríguez Bilella et al., 2016) e apenas implícito nas diretrizes do JCSEE (1994, 2021).

Um aspecto a ser registrado é que, apesar da importância das ações de comunicação durante e sobre os achados do processo avaliativo continua referida a interface entre produção de evidências científicas e disseminação. A função do avaliador como ator intermediário nos processos de transferência, tradução e translação do conhecimento não é problematizada, esvaziando a potência dos processos avaliativos como pedagógicos e fomentadores de mudança.

Lições aprendidas

Para que uma avaliação seja considerada de qualidade ela deve seguir os padrões estabelecidos nacional e globalmente, considerando os diferentes contextos e valores. A função principal deste artigo foi a de problematizar uma das dificuldades observadas durante o ensino sobre qualidade de avaliações na formação de avaliadores. Este estudo evidenciou que esta problematização deve ultrapassar a camada das diferenças semânticas, embora elas devam ser exploradas para a clareza norteadora de processos de averiguação de qualidade, neste caso, da avaliação da avaliação. Os profissionais, ao mobilizarem as categorias utilizadas em meta-avaliação, devem estar atentos às suas equivalências e diferenças e à dinâmica de suas mudanças no tempo.

Apesar de enfatizarem o caráter situado dos Atributos e das Diretrizes é possível identificar uma tensão interna entre categorias universais e locais em todos os documentos, exceto nos critérios empregados pela OCDE/DAC (OECD, 2019). Esta optou por organizar seus critérios em princípios, assumindo como pré-requisito a competência da equipe de avaliação em empreender uma apreciação holística da intervenção, para a qual os critérios são descritos. Este é um ponto interessante para discussão no ambiente docente e discente.

As análises realizadas apontaram para as dificuldades de transferência dos padrões de uma cultura para outra, já que acabam se consolidando em categorias temáticas gerais, com detalhamento que pode ser considerado prescritivo por alguns avaliadores. Entretanto, em todos os documentos, não se identificam nos atributos, critérios, dimensões e diretrizes aspectos que possam cercear a prática avaliativa e seus avanços em teorias e métodos.

A apreciação do contexto combina a sua concepção como cenário a ser considerado, no que diz respeito a legitimidade e legalidade do processo avaliativo com a sua relevância referente ao processo de valoração. Nesse sentido, o contexto é determinante como fator de desigualdades e diferenças sociais, bem como um cenário que influencia a viabilidade e eticidade dos processos avaliativos.

Predomina, em todos os documentos, o perfil do avaliador pesquisador, especialmente nas diretrizes do JCSEE (1994, 2021). Entretanto, já emerge nos documentos analisados, especialmente no Latino-Americano e Brasileiro (Rodríguez Bilella et al., 2016; Silva et al., 2020), o avaliador especialista, capaz de escuta diferenciada à diversidade e à desigualdade dos atores envolvidos e do compromisso da avaliação com a redução das assimetrias de direitos e poderes.



As referências à responsividade ao contexto e ao papel dos participantes, apesar de presentes, ainda são incipientes nas questões de métodos e controle da gestão dos processos avaliativos. A concepção da avaliação como modo de pensar e agir para captura de uma mudança intencional (processo reflexivo para ação) contraposta a avaliação como um projeto de pesquisa científica tradicional implicam na necessidade de recolocar em contexto a discussão sobre a especificidade e identidade do campo e a institucionalização dos processos pertinentes. Estes são pontos essenciais que emergem como críticos para a formação de avaliadores e a qualificação das práticas avaliativas.

Uma decorrência dessa reflexão é a pouca referência explícita, nos três documentos, do trabalho avaliativo como essencialmente de equipe e multidisciplinar. As competências dos avaliadores, das equipes de avaliação e da organização das boas práticas coletivas necessitam ser revistas a luz das mudanças recentes nas funções de avaliações.

Um aspecto interessante e pouco explorado na formação do avaliador é seu papel como ator intermediário e “ponte” entre o conhecimento e as práticas avaliativas. Abordado como mediador e/ou facilitador da construção de conexões em redes de conhecimento e práticas, avaliadores podem ter funções relevantes principalmente em processos participativos. Nesta situação, podem se tornar construtores ou engenheiros de conexões. Dito de outra forma, para análise de mérito eles têm importância não só na problematização de controvérsias oriundas das relações entre o contexto, as intervenções, seus atores humanos e não humanos que dão origem a pergunta avaliativa. São protagonistas ativos, imersos no híbrido que se origina em intervenções em avaliação. Assim, podem contribuir para potencializar conexões de mobilização, formação de alianças no sentido de solução das controvérsias que eventualmente emergem destas relações, sejam elas de mudanças individuais, institucionais ou contextuais. Este ponto é particularmente importante para avaliações de implementação.

Apesar dos documentos apresentarem, de forma recorrente, a inserção da avaliação em contextos que devem ser explorados, a vinculação destes processos com a corrente globalização, associado a políticas de austeridade, é pouco presente. Esta articulação tem requerido de avaliadores atenção redobrada aos processos globais de valoração, mas esta atenção ainda é bastante tímida no alinhamento da qualidade da avaliação aos compromissos da sustentabilidade planetária e da justiça social, corporificados nos 17 Objetivos para o Desenvolvimento Sustentável- Agenda 2030.

Fonte de financiamento

Não há.

Conflito de interesse

Não há.

Agradecimentos

Aos participantes da oficina e a Faculdade Cesgranrio pela infraestrutura disponibilizada.

Referências

Abreu, Dolores Maria Franco de, Santos, Elizabeth Moreira dos, Cardoso, Gisela Cordeiro Pereira, & Artmann, Elizabeth. (2017). Usos e influências de uma avaliação: Translação de conhecimento? *Saúde em Debate*, 41(spe), 302-316. <http://dx.doi.org/10.1590/0103-11042017s22>

Abreu, Dolores Maria Franco de. (2018). Usos e influências de uma avaliação de desempenho do Programa de Controle da Tuberculose em três municípios brasileiros (Tese de doutorado). Fundação Oswaldo Cruz, Rio de Janeiro.

Alkin, Marvin C, & Patton, Michael Q. (2020). The birth and adaptation of evaluation theories. *Journal of Multidisciplinary Evaluation*, 16(35), 1-13.



- Almeida, Cristiane Andrea Locatelli de, & Tanaka, Oswaldo Yoshimi. (2017). Meta-avaliação de uma avaliação participativa desenvolvida por gestores municipais de saúde com foco no padrão "utilidade". *Cadernos de Saude Publica*, 33(2), <http://dx.doi.org/10.1590/0102-311x00105615>
- Associação Brasileira de Avaliação Educacional – ABAVE. (2019). X Reunião da ABAVE: Avaliação educacional no Brasil. Recuperado em 6 julho, 2021, de <https://abave.org.br/x-reuniao-da-abave/>
- Bezerra, Luciana, Felisberto, Eronildo, Costa, Juliana, & Hartz, Zulmira. (2019). Translação do conhecimento na qualificação. *Physis: Revista de Saúde Coletiva*, 29(1), e290112.
- Clavier, Carole, Sénéchal, Yan, Vibert, Stéphane, & Potvin, Louise. (2012). A theory-based model of translation practices in public health participatory research. *Sociology of Health & Illness*, 34(5), 791-805. PMID:21929647. <http://dx.doi.org/10.1111/j.1467-9566.2011.01408.x>
- Fragoso, Suely, Recuero, Raquel, & Amaral, Adriana. (2013). O panorama de estudos de internet. In Suely Fragoso, Raquel Recuero & Adriana Amaral (Eds.), *Métodos de pesquisa para internet* (pp. 27-54). Porto Alegre: Sulina.
- Furtado, Juarez Pereira, & Laperrière, Hélène. (2011). A avaliação da avaliação. In Rosana Onocko Campos & Juarez Pereira Furtado (Eds.), *Desafios da avaliação de programas e serviços de saúde: Novas tendências e questões emergentes* (pp. 19-39). Campinas: Unicamp.
- Hartz, Zulmira, & Contandriopoulos André-Pierre. (2008). Do quê ao pra quê da meta-avaliação em saúde. In Zulmira Hartz, Eronildo Felisberto & Lígia Vieira-da-Silva (Eds.), *Meta-avaliação da atenção básica à saúde: Teoria e prática* (pp. 27-46). Rio de Janeiro: Fiocruz.
- Hartz, Zulmira, Santos, Elizabeth, & Matida, Álvaro. (2008). Promovendo e analisando o uso e a influência das pesquisas avaliativas: Desafios e oportunidades ao se institucionalizar a avaliação em saúde. In Zulmira Hartz, Eronildo Felisberto & Lígia Vieira-da-Silva (Eds.), *Meta-avaliação da atenção básica em saúde: Teoria e prática* (pp. 325-340). Rio de Janeiro: Fiocruz.
- Joppert, Márcia Paterno. (2021). Fatores para a profissionalização e a importância das diretrizes para a prática da avaliação no Brasil. *Revista Brasileira de Avaliação*, 10(1), 1-14.
- Latour, Bruno. (2011). *Ciência em ação: Como seguir cientistas e engenheiros sociedade afora* (2. ed.). São Paulo: Editora Unesp.
- Latour, Bruno. (2012). Reagregando o social: Uma introdução à teoria do Ator-Rede. Salvador: EDUFBA. Recuperado em 26 de agosto de 2021, de <http://www.edufba.ufba.br/2012/11/reagregando-o-social-uma-introducao-a-teoria-do-ator-rede-2/>
- Love, Arnold, & Russon, Craig. (2004). Evaluation standards in an international context. *New Directions for Evaluation*, 2004(104), 5-14. <http://dx.doi.org/10.1002/ev.132>
- Melvin, Mark, Gary, Henry, & George, Julnes. (2000). Understanding social betterment and the role of values in evaluation. In George Julnes, Gary Henry & Melvin Mark (Eds.), *Evaluation an integrated framework for understanding, guiding and improving policies and programs* (pp. 18-49). San Francisco: Jossey-Bass.
- Merton, Robert. (1967). On sociological theories of the middle range. In Robert Merton (Ed.), *On theoretical sociology: Five essays old and new* (pp. 39-72). New York: Free Press. <https://doi.org/10.2307/2574724>
- Oliveira, Euglébia, Cardoso, Gisela C., Santos, Elizabeth M, Oliveira, Margareth Martins, & Cruz, Marly Marques. (2017). O apoiador local como ator estratégico na implementação do QualiSUS-Rede: Engenheiros de conexão? *Saúde em Debate*, 41(spe), 275-289. <http://dx.doi.org/10.1590/0103-11042017s20>
- Oliveira, Margarete Martins de, Shimizu, Helena Eri, Santos, Elizabeth Moreira dos, & Silva, Everton Nunes da. (2019). Análise estratégica do Projeto QualiSUS-Rede: Contribuições para avaliação em saúde pública. *Saúde em Debate*, 43(123), 987-1002. <http://dx.doi.org/10.1590/0103-1104201912301>
- Organisation for Economic Co-operation and Development & Development Assistance Committee – OECD. (2010). *Quality standards for development evaluation*. Paris: OECD Publishing. <https://doi.org/10.1787/9789264167100-ar>
- Organisation for Economic Co-operation and Development & Development Assistance Committee – OECD. OECD/DAC Network on Development Evaluation. (2019). *Better criteria for better evaluation: Revised evaluation criteria definitions and principles for use*. Paris: OECD Publishing.
- Patton, Michael Quinn. (2018). Evaluation science. *The American Journal of Evaluation*, 39(2), 183-200. <http://dx.doi.org/10.1177/1098214018763121>
- Patton, Michael Quinn. (2020). Blue marble evaluation webinar: Evaluation criteria for transformation. Recuperado em 6 julho, 2021, de https://www.youtube.com/watch?v=5S_nV8yjZgs&t=23s
- Pawson, Ray. (2010). Middle range theory and program theory evaluation: From provenance to practice. In Jos Vaessen & Frans Leeuw (Eds.), *Mind the gap: Perspectives on policy evaluation and the social sciences* (Vol. 16, pp. 171-202). New Brunswick: Transaction Publishers. Recuperado em 26 de agosto de 2021, de <https://www.taylorfrancis.com/https://www.taylorfrancis.com/books/edit/10.4324/9781315124537/mind-gap-jos-vaessen-frans-leeuw>
- Potvin, Louise, & Bisset, Sherri. (2009). There is more to methodology than method. In Louise Potvin & David McQueen (Eds.), *Health promotion evaluation practices in Americas: Values and research* (p. 63-80). New York: Springer-Verlag. <https://doi.org/10.1007/978-0-387-79733-5>
- Potvin, Louise, & Bisset, Sherri. (2014). Há mais na metodologia do que o método. In Zulmira Hartz, Louise Potvin & Regina Bodstein (Eds.), *Avaliação em promoção da saúde: Uma antologia comentada da parceria*



entre o Brasil e a Cátedra de abordagens comunitárias e iniquidades em saúde (CACIS), da Universidade de Montreal de 2002 a 2012. Brasília: CONASS. (Tradução do texto de 2009).

Rodríguez Bilella, Pablo, Martinic Valencia, Sergio, Soberón Alvarez, Luis, Klier, Sarah, Guzmán Hernández, Ana, Tapella, Esteban. (2016). *Diretrizes para avaliação para a América Latina e o Caribe* (1. ed.). Ciudad Autónoma de Buenos Aires.

Santos, Elizabeth M, Cardoso, Gisela C.P, & Abreu, Dolores F. (2019). E-learning and problematizing pedagogies: A brazilian experience in monitoring and evaluation teaching. In André Pereira Neto & Matthew Flynn (Eds.), *The Internet and health in Brazil: Challenges and trends* (p. 275-298). Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-99289-1>

Scriven, Michael. (1991). *Evaluation Thesaurus* (4th ed., 228 p.). California: Sage.

Silva, Rogério Renato, Joppert, Marcia Paterno, & Gasparini, Max Felipe Vianna. (2020). *Diretrizes para a prática de avaliação no Brasil* (1. ed.). Rede Brasileira de Monitoramento e Avaliação. Recuperado em 26 de agosto de 2021, de https://drive.google.com/file/d/11dGgY6SG5juPp1b4L6fuuuxOIT4ArDR/view?usp=sharing&usp=embed_facebook

Stufflebeam, Daniel, & Shinkfield, Anthony. (2007). Metaevaluation evaluating evaluations. In Daniel Stufflebeam & Anthony Shinkfield (Eds.), *Evaluation theory, models & applications* (1st ed., pp. 647-684). San Francisco: Jossey-Bass.

Stufflebeam, Daniel, Madaus, George, & Kellaghan, Thomas. (2000). *Evaluation models: Viewpoints on educational and human services evaluation* (2nd ed). Boston: Kluwer Academic.

<jrn>Stufflebeam, Daniel. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7-98. <http://dx.doi.org/10.1002/ev.3>

The Joint Committee on Standards for Educational Evaluation – JCSEE. (1994). *The program evaluation standards: How to assess evaluations of educational programs* (2nd ed.). Thousand Oaks: Sage Publications.

The Joint Committee on Standards for Educational Evaluation – JCSEE. (2021). Recuperado em 26 de agosto de 2021, de <https://evaluationstandards.org/>

Worthen, B. R., Sanders, J. R., & Fitzpatrick, J. L. (1997). *Program evaluation: Alternative approaches and practical guidelines*. White Plains: Longman.

Yarbrough, Donald. (2017). Developing the program evaluation utility standards: Scholarly foundations and collaborative processes. *The Canadian Journal of Program Evaluation*, 31(3), <http://dx.doi.org/10.3138/cjpe.349>

Yarbrough, Donald, Shula, Lyn, Hopson, Rodney, & Caruthers, Flora. (2010). *The program evaluation standards: A guide for evaluators and evaluation users* (3rd ed.). Thousand Oaks, CA: Corwin Press.