

Aportes para avaliação de programas de Cooperação Sul-Sul:

Tendências e Perspectivas

Melissa Andrade e Joaquim José Soares Neto

Resumo:

Este artigo trata das tendências e lacunas na área de avaliação da Cooperação Sul-Sul (CSS). Este tipo de cooperação tem crescido, mas ainda muito carente de mecanismos mais eficazes de gestão, o que inclui esforços de monitoramento e avaliação. Existe um consenso na comunidade internacional sobre a necessidade de tornar este tipo de cooperação mais estruturada, transparente e focada em resultados. Uma volta pelo campo da avaliação revela que os ramos afeitos à CSS estão no aspecto do uso da avaliação e do julgamento, ou seja, a CSS deve ser elaborada a partir do uso que será feito e considerar a visão apresentada pelos atores participantes (conforme definidos por Alkin, 2013). Uma análise das práticas dos organismos internacionais revela o uso da matriz lógica como prática universal e aponta para o foco mais em resultados do que em processos. No entanto, para além disso, é preciso olhar para outras práticas avaliativas que possam captar a riqueza do que é a CSS focada em prioridades políticas, valores, estratégias e relações. Os mais recentes esforços no campo da avaliação têm apontado no sentido de avaliar estratégias e não apenas programas e em considerar a complexidade das iniciativas, para além de uma modelo linear de causa e efeito. Trazer esta discussão para a CSS pode ser muito útil para fazer com que a riqueza deste tipo de cooperação possa ser melhor capturada pelos estudos avaliativos.

Palavras-chave: cooperação internacional para o desenvolvimento, avaliação, Cooperação Sul-Sul, estratégia, complexidade.

Abstract:

This article discusses the tendencies and gaps in the area of evaluation of South-South Cooperation (SSC). This type of cooperation has evolved, but still lacks more effective management tools, including monitoring and evaluation mechanisms. There is a consensus in the international community about the need of making SSC better structured, transparent and focused on results. A tour around the field of evaluation shows that the streams closer to SSC are utilization focused evaluation and the branch judgment, meaning that the evaluation of SSC needs to consider the use it will have and also the views of the participating actors (as defined by Patton, 2011 and Alkin, 2013). An analysis of the practices of international organizations shows the use of the logic matrix as an universal tool and reinforces the focus more on results than processes. However, beyond this, it is necessary to look at other evaluation practices which can capture the wealth of SCC focused on political priorities, values, strategies and institutional relations. The most recent efforts in the evaluation field have brought innovations such as assessing strategies and not only programmes and in considering the complexity of the initiatives, beyond a linear model of cause and effect. It can be very useful to bring this discussion to SCC in order to make the richness of this type of cooperation to be better captured by evaluative studies.

Key words: international development cooperation, evaluation, South-South Cooperation, strategy, complexity.

Introdução

Uma audiência global cada vez mais exigente pressiona governos a serem mais eficientes e responsivos às pressões populares e a processos de prestação de contas. A avaliação se torna assim, um imperativo para a política pública.

No caso brasileiro, a cooperação internacional tem crescido e os esforços avaliativos não têm acompanhado. Este artigo se insere na perspectiva de ampliar o esforço avaliativo e de literatura especializada em avaliação relativos a projetos e estratégias de Cooperação Sul-Sul (CSS) no contexto da crescente importância dos países emergentes.

A Cooperação Sul-Sul não é uma novidade histórica. As origens datam da década de 50 na Conferência de Bandung na Indonésia, quando líderes de 34 países africanos e asiáticos geraram um diálogo dos países não alinhados à Guerra Fria, no contexto do conflito entre os EUA e a União Soviética.

Silva (2009) faz um histórico da CSS e a divide em três fases distintas: a da mobilização inicial da década de 50, da desmobilização da década de 80, em função da crise do petróleo e o ressurgimento nos anos 2000, com o crescimento econômico das ditas economias emergentes.

O último relatório global de Desenvolvimento Humano publicado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD, 2013) teve como temática a CSS. O documento relata que o comércio Sul-Sul aumentou de 8,1% a 26,7% (UNDP, 2013) entre 1980 e 2011, como percentual do comércio internacional, apesar do progresso heterogêneo (considerando a diversidade de países).

Este ímpeto continua forte, mas perdeu força recentemente, em função da diminuição dos níveis de crescimento de estrelas emergentes como os BRICS (Brasil, Rússia, Índia, China e África do Sul). Apesar deste arrefecimento, o fato é que a dita arquitetura do desenvolvimento já mudou muito e se adaptou a um cenário Sul-Sul multipolar e apesar da diminuição do ritmo de crescimento dos emergentes, seria muito difícil voltar atrás. Um exemplo disso é que o novo Plano Estratégico do PNUD para 2014-2017 torna a CSS/Tr como parte integrante dos programas do PNUD em nível global, regional e nacional.

A Cooperação Sul-Sul é feita entre países que normalmente possuem pouca tradição avaliativa e dificuldades de gestão pública. Vários autores e organismos internacionais concordam sobre a necessidade de se aumentar as avaliações da CSS conforme será relatado a seguir.

Em abril de 2013, foi promovida uma reunião pelo órgão oficial de pesquisa do governo da Índia, o *Research and Information System for Developing Countries*. A primeira recomendação feita na síntese da conferência foi exatamente ‘fortalecer coleta sistemática e análise de evidência’ da cooperação sul-sul que foi seguida por ‘fortalecimento da avaliação de programas e projetos de cooperação sul-sul’. (RIIS, 2013).

As Nações Unidas também tem se pronunciado com frequência sobre o assunto. O Secretário Geral das Nações Unidas, Ban Ki-moon em nota sobre a CSS/Tr nas Nações Unidas, afirmou que as agências possuem um consenso sobre a importância de monitorar e avaliar (UN, 2012) as atividades de cooperação Sul-Sul nos Quadros de Assistência das Nações Unidas nos países. Mas, apesar do consenso, isso ainda vem sendo feito de forma heterogênea e insuficiente.

Na avaliação da cooperação Sul-Sul dentro do PNUD feita por consultores independentes (UNDP, 2013), foi identificado que as plataformas de compartilhamento de conhecimento e os sistemas de prestação de contas em CSS não estão gerando aprendizado adequado ou dando informação sistemática sobre desempenho.

Da mesma forma, o Comitê de Alto Nível para a Cooperação Sul-Sul, em sua reunião de maio de 2012, aprovou o Quadro de diretrizes operacionais para as Nações Unidas na área de CSS/Tr e recomendou ações de monitoramento e avaliação que ajudassem a assegurar que os resultados capturem os objetivos da CSS para prestação de contas e planejamento.

Nesta direção, a Organização para a Cooperação e o Desenvolvimento Econômico (OCDE) fez um esforço para analisar 110 casos de CSS/Tr no contexto da efetividade da ajuda.¹

¹ Cumpre mencionar que ‘ajuda’ não é termo próprio para a CSS, pois geralmente denota uma relação norte-sul, onde um país do norte oferece ajuda a um país do Sul. A mentalidade própria da CSS, conforme todo seu histórico de diálogo político, é muito mais o da ajuda mútua. Por isso, que muitos países, inclusive o Brasil, possuem dificuldade em aceitar este termo tão utilizado pelos doadores do norte. Os países do Sul não se consideram doadores de maneira geral. Uma exceção para esta posição é a Colômbia, crescente protagonista da cooperação e que aceita o discurso da ajuda e que foi o país a sugerir para a Comissão de Assistência ao Desenvolvimento da OECD, o grupo de trabalho sobre CSS.

O estudo da OCDE identificou que aplicar práticas de gestão para resultados é um desafio na CSS, especialmente em atividades de desenvolvimento de capacidades em pequena escala. Muitas vezes, os resultados são claramente identificados, mas por causa da falta de práticas avaliativas é difícil avaliar o impacto. A falta de avaliação muitas vezes é devida à falta de tempo, experiência e conhecimento quando os três países formulam um projeto triangular. O relatório argumenta ainda que a maior parte dos países ainda tem dificuldades em construir sistemas de monitoramento e avaliação e gerar informação transparente, regular e em tempo apropriado.

No caso brasileiro, há poucas avaliações externas, normalmente apenas uma avaliação de meio-termo interna feita entre os gestores dos projetos dos dois países e uma avaliação final que funciona como uma revisão interna da equipe do projeto.

Diante deste quadro, serão abordados algumas características próprias da cooperação brasileira e as dificuldades presentes neste tipo de cooperação. Posteriormente, será discutido o Estado da arte da avaliação internacional e serão apresentadas tendências e inovações da avaliação aplicadas à CSS que podem ser úteis à cooperação brasileira.

1. A Cooperação Sul-Sul (técnica) Brasileira

Apesar da ausência de um documento que explicita as diretrizes e princípios da cooperação técnica brasileira, pode-se dizer que esta seja em linhas gerais pautada por 'demanda', o que implica a solicitação do país parceiro por ajuda. Não se trata de uma demanda induzida ou provocada, como muitas vezes acontece na cooperação Norte-Sul. Ademais, ela faz uso predominantemente de técnicos e conhecimento presentes no governo brasileiro, possuindo ainda poucas parcerias com a sociedade civil. Trata-se de um diálogo horizontal e que tem por princípio a mobilização de parcerias.

Outros princípios da cooperação brasileira incluem 'considerar o elemento sul' de modo que as experiências e conhecimentos brasileiros sejam compartilhados; promover a apropriação das lideranças locais pelo conhecimento; ter objetivo de produzir 'impacto estrutural' e resultados sustentáveis e representar oportunidades de inovação em relação ao compartilhamento de experiências e a articulação de parcerias. Estes são princípios colocados pela Agência Brasileira de Cooperação nas suas apresentações públicas.

O Brasil tem se colocado cada vez mais como um parceiro 'global' para o desenvolvimento. Exemplo disso é a assinatura de protocolos entre o Brasil e países como

a Espanha, Estados Unidos e Reino Unido para atuação em outros países, notoriamente a África, no modo de Cooperação Triangular. Ou seja, o Brasil tem se aliado às potências tradicionais e também às emergentes para realizar projetos de triangulação e outras iniciativas comuns para levar ajuda a países mais pobres da comunidade internacional.

No âmbito da Cooperação Técnica, os grandes protagonistas são a Embrapa com todo o seu aporte tecnológico, o SENAI na área de formação profissional e o Ministério da Saúde, sobretudo na área de DST/AIDS, onde a experiência brasileira é referência internacional.

Dos 95 países beneficiados pela cooperação brasileira, a cooperação é mais expressiva junto aos países que são prioridade do ponto de vista da política externa brasileira, nominalmente os PALOP – Países de Língua Portuguesa e os parceiros da América Latina. Há pouca cooperação com países asiáticos, com exceção do Timor Leste.

Apesar do crescimento desta cooperação, não há um sistema de gestão desta cooperação que esteja bem definido. A Agência Brasileira de Cooperação não funciona como uma agência (é um departamento do Ministério das Relações Exteriores), pois não possui autonomia e navega em meio a uma cooperação fragmentada realizada por toda a burocracia brasileira e onde os órgãos não necessariamente reportam toda a atividade de cooperação que desenvolvem. Os problemas da cooperação brasileira são muitos, estruturais e bem conhecidos: falta de legislação adequada, falta de pessoal permanente na Agência Brasileira de Cooperação, dependência do governo em relação aos organismos internacionais para realizar a cooperação, além da atuação no exterior sendo feita de forma desarticulada e pouco estratégica (Pino, 2010, Cabral, 2010).

Existe hoje uma preocupação entre especialistas e gestores da cooperação de que toda a estrutura e a mobilização política gerada no Governo Lula sofra um retrocesso por falta de prioridade dada à cooperação. A presidência anunciou em maio de 2013 a criação de uma agência autônoma de cooperação, com forte ênfase no comércio, em linha com a nova gestão, mas esta permanece em suspenso.

De maneira mais ampla, a posição do Brasil frente à CSS pode ser caracterizada pela sua ênfase no multilateralismo; pela cooperação dentro da Região e na defesa por uma maior participação dos países emergentes no centro das decisões internacionais; pela prioridade para investimento em ciência e tecnologia; foco para a promoção da infraestrutura na Região de modo a acelerar a integração, especialmente do ponto de vista

econômico e comercial; uma intensa atividade de cooperação na área de saúde, desde bancos de leite a prevenção ao HIV/AIDS e acesso a medicamentos; a busca por novas matrizes energéticas; a abertura para o diálogo com outros blocos e países tanto no âmbito da União de Nações Sul-Americanas (UNASUL) quanto da Comunidade de Estados Latino-Americanos e Caribenhos (CELAC) e discussões mais ambiciosas no âmbito do BRICS e a necessidade de se aprofundar o diálogo sobre fontes de energia renováveis e segurança alimentar, temas também presentes no diálogo regional.

Diante do panorama brasileiro apresentado, convém abordar os desafios desta gestão da CSS que não são exclusivos do Brasil, mas permeiam esta cooperação em nível global.

2. Desafios da gestão da Cooperação Técnica Sul-Sul

São muitos os desafios da Cooperação Técnica Sul-Sul. Muitas vezes as iniciativas não possuem orçamento adequado, mecanismos de coordenação e seguimento, além de uma capacidade limitada e a falta de mecanismos apropriados de monitoramento e avaliação (Bilal, 2012).

Na avaliação da cooperação Sul-Sul dentro do PNUD feita por consultores independentes (2013) foi identificado que as plataformas de compartilhamento de conhecimento e os sistemas de prestação de contas em CSS não estão gerando aprendizado adequado ou dando informação sistemática sobre desempenho. Além disso, avaliou-se que ainda é cedo para concluir se os resultados de CSS são sustentáveis em função do contexto e da falta de sistemas de monitoramento efetivos. Estas afirmativas reforçam o olhar cauteloso e mais atento que deve ser dada à gestão da CSS. O mesmo relatório recomenda que o PNUD deveria adotar uma estratégia corporativa para apoiar a CSS/Tr de modo a facilitar o seu direcionamento, prestação de contas, avaliação do apoio dado e resultados obtidos nesta área. A proposta é que quadros mais robustos de M&A deveriam ser desenvolvidos para documentar e extrair lições de melhores práticas para posterior replicação. Este diagnóstico é muito relevante, pois é justamente o PNUD, na arquitetura das Nações Unidas, que possui significativa expertise e manuais voltados para avaliação. Se uma agência global estruturada não tem avaliado a CSS de forma suficiente, muito menos pode se esperar no âmbito dos países participantes de tal cooperação.

Um estudo também sobre a gestão da CSS financiado pela Agência de Cooperação Internacional do Japão (JICA) e o Escritório das Nações Unidas para a Cooperação Sul-Sul

- UNOSS (Vazquez, 2013) analisou vários casos de CSS/Tr (Cooperação Sul-Sul e Triangular) e chegou a uma síntese de que a melhoria das práticas de gestão da CSS e Triangular (Tr) tem sido um processo de tentativa e erro ou seja, ainda há muito improvisado e pouco sistemático.

Ademais, com o intuito de estimular a Cooperação Sul-Sul, a UNOSS promoveu em parceria com o Governo do Panamá em 2012, uma Feira de Conhecimento e selecionou as melhores práticas de cooperação Sul-Sul. Tais práticas foram analisadas (Simplicio & Andrade, 2013) e se mostraram muito diversas. Há vários arranjos institucionais, com muitos parceiros ou apenas em relações bilaterais. As modalidades de troca de conhecimento são várias e vão desde visitas e seminários a troca de experiências e casos de cooperação de longo prazo onde os especialistas ficam muito tempo no país parceiro. Os maiores desafios encontrados incluem: coordenação entre vários atores; falta de apoio político de alto nível e diálogo com estratégias nacionais; falta de dados e coletas regulares; falta de recursos; rotatividade de pessoal; falta de adaptação de materiais para outros idiomas e contextos; falta de envolvimento do grupo alvo e resistência à mudança e falta de sustentabilidade dos programas.

Neste caso, não houve um tópico que tratasse em particular de avaliação, mas é sabido que a falta de avaliação de projetos e CSS é a regra. Existe ainda uma resistência em realizar avaliações copiando os moldes da tradicional cooperação norte-sul e uma crença na importância dos dividendos políticos de tais relações - nem sempre capturadas por avaliações tradicionais.

A Cúpula sobre a arquitetura da ajuda internacional em Busan 2011 reforçou o apoio à participação dos países emergentes e da Cooperação Sul-Sul e Triangular na cooperação internacional. A declaração final mencionou a necessidade de fortalecer esforços para atingir resultados concretos e sustentáveis, o que envolve uma gestão baseada em resultados, monitoramento, avaliação, comunicação e aumento da capacidade dos países.

Em seguimento ao trabalho de Busan, em Abril de 2014, aconteceu o primeiro Encontro de Alto Nível da Parceria Global para Cooperação para o Desenvolvimento Efetivo (GPED, 2014), um grupo formado a partir da liderança do México, Malawi e Holanda. Este grupo reuniu lideranças de países do Norte e Sul e no seu comunicado final enfatizou o 'foco em resultados', colocando a importância de esforços para monitorar, avaliar e comunicar progresso da cooperação e também abordou a necessidade de 'transparência e prestação de contas entre os parceiros' onde os parceiros do Sul deveriam continuar a melhorar a disponibilidade de informação, resultados e impacto dos seus

esforços de cooperação de acordo com as metodologias que considerassem mais adequadas às suas especificidades. Este evento reforça a necessidade de melhorar os mecanismos de prestação de contas da Cooperação Sul-Sul dentro de um consenso global. Os métodos não são consensuais, mas a necessidade de prestação de contas é um ponto comum a todos os parceiros.

A CSS é uma realidade e uma necessidade, tendo em vista todos os desafios anteriormente citados, mas ainda com sérios problemas de gestão e pouquíssima avaliação. Convém agora, voltar os olhos para a literatura da avaliação para ver o que está disponível e pode ser usado de forma mais direcionada para avaliar a Cooperação Sul-Sul.

3. Uma síntese da avaliação ontem e hoje

A história da avaliação contemporânea está associada à busca pela eficiência no setor público. Nasce nos EUA e depois se desenvolve também na Europa e se dissemina como prática internacional associada à boa gestão, seja pública ou privada. O estudo marcante no desenvolvimento da disciplina de avaliação foi o trabalho de Ralph Tyler no seu 'Eight Year Study' sobre educação nos anos 1940. Aliás, educação é a área de origem das práticas avaliativas que depois se disseminaram para áreas como saúde e assistência social. O primeiro grande 'boom' de avaliação aconteceu no final da década de 1960 e 1970 nos Estados Unidos sob as administrações Kennedy e Johnson quando os programas sociais foram desenvolvidos em larga escala e apoiados por recursos federais sob as políticas de 'Guerra contra a Pobreza' e 'Grande Sociedade. (Donaldson & Lipsey, 2006).

A avaliação era inicialmente identificada apenas com técnicas estatísticas e experimentais e depois foi se ampliando para abordagens mais qualitativas e participativas. Constitui um campo extremamente diverso - uma torre de Babel do ponto de vista epistemológico e metodológico. Assim, como há uma diversidade de abordagens e metodologias no campo das ciências sociais, também é na avaliação. Por isso, o comum é que os avaliadores nas organizações e consultores, diante de cada necessidade avaliativa, adaptem métodos e abordagens de acordo com o objeto de estudo a ser analisado.

Alkin (2013) possui uma obra seminal em que organiza as perspectivas teóricas da avaliação numa árvore onde as raízes representam as motivações para avaliação e os galhos crescem em métodos relacionados a cada um dos propósitos da prática avaliativa.

A raiz da prestação de contas social (social accountability) gera métodos preocupados com o uso da avaliação, a raiz de investigação social gera abordagens quantitativas cuja preocupação é o rigor científico, enquanto a raiz da epistemologia gera métodos preocupados com a valoração dos projetos e programas avaliados, ou seja, com o julgamento de valor associado, dando preferência para métodos qualitativos. As derivações de abordagens da avaliação são divididas em métodos, julgamento e uso; ou seja, classifica os teóricos de acordo com a sua ênfase, se mais metodológica (normalmente associada a métodos quantitativos, nas origens experimentais da avaliação), na apreciação feita pelos atores envolvidos (com vieses mais participativos e subjetivos) ou no uso feito pela avaliação.

Os ramos mais afeitos à CSS são o do uso e do julgamento - são as abordagens que consideram o que vai ser feito com o uso da avaliação e a perspectiva dos vários atores envolvidos. Considerando o enfoque deste tipo de cooperação cuja origem é política, as múltiplas perspectivas devem ser consideradas. A seguir são apresentadas as influências principais dos autores mais relevantes para se estudar a CSS nestes dois ramos. Será discutido o trabalho de Jennifer Greene, Michael Patton e Lincoln & Guba.

Jennifer Greene trabalha com uma abordagem baseada em valores dentro do ramo do julgamento. Para ela todas as abordagens de avaliação promovem um conjunto de valores e a avaliação está sempre fazendo julgamentos baseados em critérios selecionados de qualidade. As influências no seu pensamento remontam a Paulo Freire e à disciplina de psicologia. Ela argumenta que a avaliação nos Estados Unidos sempre foi muito baseada em uma discussão de métodos, para se ampliar posteriormente. Foi influenciada por Ernie House. Este argumentava que a avaliação é muito impactada pelo contexto e o contexto molda as prioridades da avaliação. A avaliação pode ajudar a manter o status quo ou questioná-lo. Há sempre uma influência e por isso o avaliador deveria se questionar: a quem ele vai servir? A avaliação responsiva é outra influência na obra dela (avaliação guiada pelas necessidades dos atores envolvidos), além da dimensão cultural, pois cultura e contexto estão profundamente interligados.

Lincoln & Guba também são muito influentes e a sua obra 'As Raízes da Avaliação de Quarta Geração' é uma referência. Para eles, a quarta geração de avaliação está associada ao uso, particularmente quanto a quem deve usar os resultados da avaliação e como os usuários podem utilizar os resultados para afetar mudança. No tocante a métodos - deve haver uma variedade deles, tanto quantitativos quanto qualitativos. Trata também de valoração, ou seja, dos valores plurais que os diferentes atores trazem para os

programas e necessariamente o contexto político em que tanto programas quanto as avaliações são feitas.

Este tipo de avaliação com viés construtivista aborda múltiplas questões em epistemologia, sobretudo as questões do dualismo sujeito-objeto, o papel da causalidade e as formas possíveis de afirmações causais nas ciências sociais; a natureza de possíveis julgamentos de valor; e o papel de diversas formas de saber – ou seja – a experiência do mundo por meio da raça, classe, gênero, história e variações diversas incorporadas na vida.

A quarta geração vê os *stakeholders* engajados em gerar sentido no centro da avaliação. Para eles a causalidade mútua sugere que eventos ou pessoas não estão ligadas sequencialmente, mas estão interpenetradas e organizadas em um complexo padrão que pode ser alterado por qualquer número de atores exibindo intenções de uma forma ou outra.

Um outro autor que dialoga com este tipo de abordagem e é extremamente influente é Michael Patton, cujo enfoque está no uso. A avaliação focada no uso começa com a premissa de que as avaliações deveriam ser julgadas pela sua utilidade e uso real. Assim, os avaliadores deveriam facilitar o processo de avaliação e planejamento com cuidado em como tudo é feito, do começo ao fim, e considerando como suas ações vão afetar o uso. O uso está relacionado como pessoas reais no mundo real aplicam os resultados e experiência da avaliação. Ela não advoga nenhum tipo particular de conteúdo, modelo ou método – incluindo métodos qualitativos. A avaliação baseada no uso também não advoga um tipo particular de uso. Mas, ajuda usuários primários a selecionar os conteúdos, métodos, modelo, teoria e usos mais apropriados para sua situação em particular. A resposta da situação guia o processo interativo entre o avaliador e os usuários primários (ele enfatiza esta relação entre avaliador e atores). Ele defende que usuários intencionais terão mais possibilidade de usar a avaliação se entenderam e se sentirem ‘donos’ do processo e resultados e eles se sentirão mais ‘donos’ se forem ativamente envolvidos. Em envolver os usuários para quem a avaliação é feita, o avaliador está treinando os usuários quanto ao uso e reforçando a utilidade da avaliação ao longo de todo o processo. A base para este ramo da avaliação está na sociologia organizacional, na difusão de inovações e perspectivas sociológicas de poder e conflito.

Para além da Europa, EUA, Nova Zelândia e Austrália, Alkin, M. & Carden, F. (2012) fizeram uma análise sobre práticas de avaliação de programas de desenvolvimento no Sul Global e se propuseram a encaixá-las dentro da ‘Árvore Teórica’ desenvolvida

pelos autores. O argumento é de que prática avaliativa foi desenvolvida especialmente por acadêmicos do Norte e no Sul, de que existe muito mais prática do que teoria e a teoria desenvolvida é essencialmente baseada na prática. Além disso, as origens práticas da avaliação do Sul ainda não foram elaboradas em teorias prescritivas. O artigo sugere a necessidade de construir um conhecimento mais aprofundado em práticas de avaliação do desenvolvimento em países de renda baixa e média.

Nota-se uma abordagem muito distinta entre as perspectivas do Norte de cunho mais racional e sistemático, focadas em resultados objetivos e definidos a priori, e as abordagens do Sul que tendem a ser mais participativas, subjetivas, adaptadas a contextos dinâmicos e mutáveis. A avaliação de programas de CSS pode se beneficiar muito de uma combinação de indicadores objetivos definidos a priori que ajudam no processo de gestão e ao mesmo tempo a consideração para a perspectiva de vários atores, olhando para as mudanças mais significativas e os resultados obtidos que não foram previstos. Em programas de CSS, muito se muda no meio do caminho. É preciso analisar os resultados dentro deste contexto e não apenas ignorar o que foi feito de maneira não intencional, restringindo-se a uma matriz lógica definida no início do programa. E como avançar neste sentido, construindo avaliações que considerem as especificidades da CSS?

4. Avaliação e a Cooperação Sul-Sul

Para além de todo o histórico de avaliação de políticas públicas desenvolvido especialmente nos Estados Unidos na área educacional, há uma longa tradição de avaliação em projetos de cooperação para o desenvolvimento reforçada com a criação há 30 anos da Rede de Avaliação do Comitê de Assistência ao Desenvolvimento (DAC em inglês) da OCDE. Há que se considerar que o DAC reúne os principais países doadores que possuem larga experiência na área.

A Rede de Avaliação do DAC desenvolveu princípios que passaram a nortear a avaliação de projetos de cooperação em todo o mundo e geraram referências para ações de todos os países da rede. Os tradicionais países doadores em sua maioria possuem uma preocupação clara com avaliação, instituindo manuais e orientações expressas neste sentido.

Além disso, o Banco Mundial possui um grupo de Avaliação Independente com orçamento de US\$32.5 milhões anuais e uma equipe de quase 100 pessoas (Relatório

Anual, 2012). O Banco opta por produzir conhecimento na área, disseminando as melhores práticas e resultados das avaliações, mas neste caso, não foram identificadas diretrizes gerais para a área, mas apenas práticas comuns de avaliação de projetos com metodologias diversificadas e um crescente uso de avaliações de impacto.

O Banco Mundial possui uma ênfase em gestão baseada em resultados como uma ferramenta chave para a efetividade do desenvolvimento. O Banco tem se preocupado mais em explicitar resultados e impacto e não apenas atividades, conforme a tendência internacional.

A outra grande referência internacional é o Programa das Nações Unidas para o Desenvolvimento que possui um manual de avaliação com diretrizes claras e extensas para uso em suas operações no mundo com países parceiros.

A revisão desta literatura (OECD, 1986, 1991, 2000 e 2013 e UNDP, 2009) permite chegar à seguinte síntese preliminar:

- **Uso da matriz lógica como prática universal:**

A matriz lógica é uma ferramenta utilizada amplamente como instrumento para elaboração de projetos de cooperação internacional (e que servirão de base para as futuras avaliações), com pequenas variações. É dado como certo o seu uso por detalhar de forma objetiva o propósito macro, os objetivos, as ações, os indicadores e as condições necessárias para implementação. É a partir de uma matriz bem elaborada e um bom documento de projeto que se tem o básico para realizar qualquer ação avaliativa.

- **Foco em resultados mais que em processo:**

Os processos de reforma do Estado já na década de 90 e a matriz da Nova Gestão Pública deram a tônica para uma ação pública baseada muito mais em resultado do que na mera realização de atividades (com o foco no processo). Migrou-se de um enfoque em eficiência (relação custo-benefício) para uma visão voltada para eficácia (resultados-outputs) e impacto (outcomes). Foi assim na Inglaterra, na Austrália, no Canadá e vem avançando também no Brasil.

- **Diversidade de metodologias e instrumentos de avaliação:**

Avaliação tem basicamente a natureza de uma pesquisa. A criatividade e diversidade de metodologias e instrumentos de pesquisa se aplicam ao campo da avaliação. A construção de qualquer estratégia de avaliação deve passar por uma análise do que se quer descobrir (dimensões a serem investigadas) e recursos disponíveis (tempo, pessoas e outros insumos). Avaliação não é apenas uma decisão de gestão, é também uma decisão política e estratégica.

- **Importância da participação dos atores:** toda a literatura analisada, sem exceção, abordou a importância da participação dos atores, dos 'stakeholders' envolvidos, desde governos, parceiros diversos a comunidades beneficiadas. Dentro de um contexto de múltiplos arranjos organizacionais e um trabalho absolutamente em rede e colaborativo, a capacidade de envolver vários atores e chegar a um acordo sobre o que é relevante e como fazer a avaliação é fundamental para que o processo avaliativo gere resultados e leve à reflexão das pessoas envolvidas. Trata-se de um trabalho gerencial e de liderança.
- **Comunicação interna e externa:** na mesma linha do item anterior, onde há múltiplos atores envolvidos, a comunicação com atores e a comunicação dos resultados são estratégias para que o esforço avaliativo tenha impacto.
- **Apropriação dos resultados:** este item trata do uso da avaliação. Avaliações externas são muitas vezes rechaçadas ou simplesmente engavetadas. É necessário que as pessoas envolvidas se apropriem dos seus resultados.

Para além da prática usual de projetos de cooperação, especialmente norte-sul, vem surgindo avaliações inovadoras e abordagens que podem ajudar a trazer mais luz para a análise da CSS.

4.1. Tendências e inovações no campo da avaliação

Serão discutidas nesta seção, a crescente avaliação de estratégia e os aportes da discussão sobre complexidade aplicadas ao campo avaliativo, muito adequadas ao debate sobre CSS.

A OCDE promoveu uma avaliação independente da Declaração de Paris que teve início em 2006. A Declaração de Paris foi um documento assinado pelos países da OECD no marco da efetividade da ajuda internacional em 2005. Foi feita uma avaliação internacional que ao fim cobriu 22 países e 18 doadores/agências diretamente. Este esforço dividido em duas partes subsidiou o Encontro de alto nível sobre Efetividade da Ajuda que aconteceu em Busan, na Coreia, em 2011. Foram 5 anos de esforços de uma avaliação que custou £10,4 milhões e que em 2012 ganhou o prêmio da Associação Americana de Avaliação como 'Avaliação de Destaque'.

Tratou-se da avaliação de uma declaração, cujo cerne são princípios. Não foi uma avaliação de programa, mas uma avaliação macro de diretrizes de atuação no marco da cooperação para o desenvolvimento tradicional. A Associação Canadense de Avaliação lançou uma edição da sua revista em 2012 dedicada a esta avaliação, onde houve contribuição dos principais técnicos e pesquisadores envolvidos na gestão e pesquisa deste esforço internacional. Convém refletir sobre este esforço milionário que mobilizou toda a comunidade de países doadores.

Foi estabelecida inicialmente uma estrutura de governança e definidos os princípios que conduziriam a avaliação. As perguntas relevantes foram validadas, foi estabelecida uma teoria do programa explícita, uma matriz lógica e recomendados vários métodos que os países poderiam utilizar para responder cada pergunta. Um *paper* de 2006 diagnosticou que a série de hipóteses que davam a lógica da Declaração ou a teoria do programa não estavam plenamente articuladas. Assim, a avaliação teria que aplicar uma abordagem 'baseada em teoria' - ou seja tornar explícitas as teorias de mudança que estavam nas entrelinhas. A equipe buscou esta lógica implícita, incluindo os objetivos expressos, tanto intermediários quando de longo prazo e a descrição de como estes objetivos seriam cumpridos.

Os métodos recomendados aos países para responder às perguntas foram bem variados, desde revisão de literatura e documentação a análise quantitativa/estatística, surveys, grupos focais, análise de atores, estudos de caso e outras metodologias como *outcome mapping*, análise de decisão, *appreciative enquiry* e Análise da Mudança mais significativa.

Este estudo complexo mostra a importância de se ter clareza sobre a lógica por trás do que vai ser avaliado para poder mapear resultados e processos (foi um estudo sumativo e formativo). A diversidade de métodos mostra como a metodologia específica

não é o ponto mais importante para levantamento dos resultados, mas é definida em função do que se propõe a investigar, a conveniência do contexto e os meios disponíveis para isso.

Niels Dabelstein e Ted Kliet foram respectivamente chefe do Secretariado para Avaliação da Declaração de Paris sobre Efetividade da Ajuda e co-chefe do Grupo de Gestão da Avaliação. Os dois escreveram dois artigos (Dabelstein & Kliet, 2002a e 2002b) onde argumentam que avaliações conjuntas são difíceis. Os parceiros possuem culturas avaliativas e procedimentos administrativos diferentes, prioridades e agendas políticas próprias e alinhar parceiros é uma tarefa que demanda tempo. Este foi certamente um dos maiores desafios em se conduzir uma avaliação desta dimensão – gerar acordos em meio a tantos participantes. Este também é um desafio próprio da Cooperação Sul-Sul que envolve negociação entre muitos atores e cada vez mais com as iniciativas triangulares.

Dialogando com a avaliação de Paris, vem surgindo um tipo de avaliação que está além do programa e que se centra na estratégia da organização. A Associação Americana de Avaliação possui um periódico que se chama 'New Directions for Evaluation' em que esta temática foi abordada em 2012.

Para Patton e Patrizi (2010), estratégia é uma nova unidade de análise na avaliação, porque tradicionalmente, avaliação tem focado em projetos, programas, produtos, políticas e pessoal. Muito se fala sobre estratégia, mas não sobre avaliação de estratégia. Para eles, avaliar estratégia não é avaliar planejamento estratégico ou planos estratégicos. Trata-se de avaliar a estratégia em si mesma. Os autores destacaram uma abordagem em que a estratégia é avaliada por meio do exame de padrões de comportamento – o que a organização realmente faz – não apenas a retórica sobre estratégia e planos estratégicos.

A referência para o trabalho com estratégia é Henry Mintzberg que define estratégia como um padrão, a 'consistência em comportamento ao longo do tempo.' Para Patton e Patrizi (2010), estratégia é diferente do que os avaliadores entendem por teoria de mudança ou matriz lógica. Estratégia seria definida como um plano olhando para o futuro. Mintzberg diferencia 'estratégia buscada' de 'estratégia realizada'. Para ele, na implementação da estratégia há uma combinação entre processos deliberados e processos que não foram previstos. Existe a estratégia emergente que emana das oportunidades das situações.

Patton e Patrizi (2010) argumentam que esta perspectiva sobre implementação da estratégia no mundo real contrasta com a abordagem clássica de prestação de contas da avaliação em que a implementação e resultados do programa são medidos e avaliados baseados no que o programa planejou fazer e fez, ou seja, um programa inovador que aproveita a oportunidade e se ajusta às circunstâncias seria avaliado de forma negativa. Este argumento se alinha perfeitamente ao contexto dos programas de CSS, onde as circunstâncias são dinâmicas e muitas vezes imprevisíveis, em função da instabilidade política e econômica. Avaliar programas de cooperação não apenas a partir de uma matriz lógica seria mais efetivo e realista, considerando a estratégia realizada e reconhecendo diferentes caminhos e resultados não esperados que agregam valor ao processo de desenvolvimento de maneira mais ampla e não apenas ao estrito cumprimento de metas.

No modelo de Mintzberg, a estratégia é um processo contínuo de aprendizagem. 'Fazer' seria o precursor de 'aprender' e aprender seria o precursor de desenvolver uma visão robusta para o trabalho que vai avançar. Muitas vezes quando um programa não alcança um objetivo, a execução do programa é culpada, quando na realidade, ele argumenta que muitas vezes, um desenvolvimento inadequado de estratégia é a raiz do problema, ou seja, existe um trabalho inicial de planejamento e estratégia e não um problema na execução. A verdade é que muitos programas de CSS são super estimados e não partem de uma análise objetiva e factível da realidade.

Para Mintzberg, o contexto também importa numa estratégia. Um trabalho de planejamento tradicional, metódico e detalhado funciona mais em ambientes relativamente estáveis enquanto em ambientes mais incertos, as abordagens deveriam ser mais flexíveis.

No caso da CSS, existem princípios e valores estabelecidos que podem ser identificados no Plano de Ação de Buenos Aires (1978) que marca a Cooperação Técnica entre Países em Desenvolvimento. A UNOSS fez uma análise da implementação dos projetos IBSA com base nos princípios basilares da CSS expressos em Buenos Aires e na Conferência de Alto Nível das Nações Unidas sobre Cooperação Sul-Sul em 2009, em Nairobi (UNOSS, 2012), o que pode ser uma inspiração para futuros estudos nesta área.

As reflexões sobre estratégia e princípios parecem especialmente pertinentes no caso da CSS, onde existe pouca estratégia deliberada, mas muita ação e iniciativa e um contexto relevante de política internacional, onde princípios são importantes. E o que dizer da questão da complexidade?

4.2. Avaliação e Complexidade

O conceito de Complexidade traz um avanço no campo das ciências sociais por reconhecer que os fenômenos são dinâmicos e admitem a interação de vários fatores que se influenciam mutuamente. Nenhum processo de desenvolvimento pode ser explicado meramente pelo investimento em um setor, mas pela conjugação de vários fatores que passam pela economia, pela liderança, pelo contexto externo e pela interação entre as pessoas, entre inúmeros outros fatores. A realidade é literalmente muito 'complexa' para ser analisada simplesmente numa relação binária de causa e efeito (input-output). São muitas causas e muitos efeitos que interagem o tempo todo. A complexidade é definida em situações em que existe incerteza sobre o que fazer porque tanto o conhecimento é insuficiente quanto os atores-chaves estão em conflito. Para Patton (2011), pensamento e mapeamento sistêmico oferecem uma alternativa para a modelagem linear lógica.

Uma diferença considerável no momento de modelar a avaliação é que a avaliação mais complexa busca mapear a dinâmica e as interdependências e identificar as interconexões que surgem. A causalidade é baseada na detecção de padrões construída retrospectivamente a partir de observações. Deve-se esperar o inesperado, relações tanto lineares quanto não lineares, resultados intencionais e não intencionais, previstos e não previstos. Um modelo sistêmico de um programa para Patton (2011) colocaria todos os elementos interligados, influenciando-se mutuamente. Por exemplo, em um determinado programa, supõe-se que o aumento da renda vai estimular a presença da criança na escola e conseqüentemente, aumentar as possibilidades de emprego no futuro. Um modelo sistêmico e complexo considera que o aumento da renda estimula o emprego e o emprego influencia a presença da criança na escola e também o contrário, ou seja, há múltiplas interações que devem ser consideradas e mapeadas no momento de se fazer uma avaliação e que escapam de uma relação causal linear. Muitas vezes, os formadores de opinião tem muito mais influência no resultado de um programa do que simplesmente a execução das atividades e isso precisa ser mapeado para identificar princípios e relações que possam ser melhoradas. O princípio é de que todos os fatores estão interligados e que indivíduos afetam organizações que estes afetam os valores de uma sociedade e vice-versa.

O campo da avaliação já vem incorporando conceitos de complexidade. Michael Patton (2011) escreve um livro onde aplica conceitos de complexidade para aquilo que chama de 'Avaliação de Desenvolvimento' associada a avaliação de contextos dinâmicos

e mutáveis, onde tudo está mudando sempre e onde o avaliador contribui intensamente para o aperfeiçoamento organizacional.

Há que se considerar que o 'negócio' da CSS é inovação, é a mudança de realidade. É certo que modelos lineares e fechados não conseguem abarcar a complexidade das relações e resultados de um projeto de CSS que se dá essencialmente entre países que com frequência são instáveis.

Patton reafirma que não há uma única forma adequada de fazer uma avaliação. O desenho de uma avaliação particular depende das pessoas envolvidas e da situação em particular.

Os conceitos de complexidade que ele utiliza são:

- **Não-linearidade:** diz respeito à sensibilidade às condições iniciais em que pequenas ações podem gerar grandes reações (em função da instabilidade e incerteza inicial). Captura também resultados não previstos e não intencionais.
- **Emergência:** inovadores não podem determinar antes o que vai acontecer e por isso avaliadores não podem determinar com segurança o que medir. Precisam estar observando o que vai acontecer. Este constructo se aplica parcialmente à CSS técnica, onde a transferência de conhecimento é feita com base em experiências já testadas e bem sucedidas. A inovação se dá na transferência a um novo contexto. O constructo é válido para ajudar a capturar o que surge de forma não prevista. Subgrupos que denotam uma cultura própria também devem ser mapeados e observados, pois emergem padrões a partir da interação entre as pessoas.
- **Adaptabilidade:** elementos e agentes que interagem se adaptam uns aos outros e ao ambiente em que estão; então o que surge é uma função da adaptação tanto entre os elementos que interagem e a resposta destes elementos dinâmicos ao meio.
- **Incerteza:** a incerteza vem da turbulência no ambiente e os limites do conhecimento. Se há conhecimento suficiente sobre uma medida (e.g. eficácia de uma vacina), é muito mais fácil prever os resultados, mas se ambiente é instável e são muitos fatores afetando os resultados, a incerteza é parte do cenário.
- **Processo co-evolucionário:** está relacionado ao resultado da co-participação entre avaliador, gestores do projeto e entre os membros da equipe entre si e

com todos os envolvidos no sistema. O foco está na interação entre os atores relevantes que 'co-criam' a realidade.

Tais conceitos podem ser muito úteis para ampliar o foco de avaliação de um programa de CSS que é complexo, ocorre em ambientes instáveis, envolve muitos atores e é fortemente baseado em relações multiculturais que geram um contexto por vezes, imprevisível.

5. Uma síntese da avaliação voltada para a Cooperação Sul-Sul

Diante do que foi discutido, cabe agora fazer uma síntese dos principais pontos, visando apontar caminhos para a avaliação de projetos de Cooperação Sul-Sul. Inicialmente, foram feitas reflexões sobre a cooperação brasileira, desafios próprios da CSS e por último foram discutidas abordagens avaliativas que podem agregar valor à análise da CSS. Esta série de peculiaridades que envolve contextos políticos mais instáveis e menos previsíveis já é suficiente para se pensar em formas de avaliação próprias para este tipo de cooperação. A seguir, são consideradas as principais indicações do que seria necessário para uma avaliação da CSS:

1. A importância de combinar análise da estratégia com estudo de caso e utilizar novos indicadores

CSS trata não apenas da implementação de programas, mas de visões de desenvolvimento (na base da ideologia da CSS) que exigem o diálogo entre diferentes áreas: econômica, política e social. A avaliação restrita a programas sem visão estratégica é limitada para a riqueza do que é a CSS. Idealmente, em um escopo mais abrangente, as avaliações de vários programas de cooperação entre dois países deveriam ser feitas de forma interligada, somando para a análise das relações construídas entre as duas nações para gerar uma visão da macro estratégia da cooperação. Uma abordagem como essa tornaria a cooperação muito mais efetiva e clara, de modo a aprofundar as relações entre os países. A dimensão estratégia abrange especialmente a dimensão da macropolítica – foco da CSS quando analisada historicamente. Neste caso, a avaliação da qualidade da parceria e das relações se tornaria crucial.

Por outro lado, o estudo de caso é uma abordagem qualitativa que dá espaço para o aprofundamento do que são as nuances da CSS. Estudos de caso podem ser muito úteis

para fazer a ponte entre culturas diferentes e formas de trabalhar – dando elementos para outros programas da mesma área temática ou do mesmo país – podendo gerar muito aprendizado entre diferentes políticas ou na cooperação entre dois países. Além disso, a CSS demanda outros tipos de indicadores para aprofundar a análise estratégica para além de programas. Seria necessário considerar outras dimensões, tais como: aprofundamento dos laços, apoio mútuo e convergência de posições.

2. Ampliação do modelo da matriz lógica linear e consideração de modelagens complexas

A matriz lógica, apesar de muito útil para sistematizar objetivos, atividades, resultados, indicadores e riscos, parte de um modelo linear de *input* e *output*. Um ambiente de CSS é muito mais dinâmico, há muitos fatores inesperados, resultados não previstos e atividades substituídas. Seria necessário numa avaliação da CSS mapear a dinâmica e as interdependências e identificar as interconexões forjadas na execução dos projetos. A realidade da CSS não consegue ser espelhada por uma matriz lógica convencional.

3. O papel da cultura na implementação de projetos

Existe uma grande diferença de cultura gerencial que pode ser vista dentro de uma mesma organização e ainda mais em países diversos. A própria cultura avaliativa europeia e americana são diferentes. A avaliação de impacto, por exemplo, possui uma concepção essencialmente norte-americana, pelo seu enfoque em modelagens econométricas afins com a tradição acadêmica quantitativa naquele país. A dimensão cultural revela muito sobre o que é importante em cada país. O peso está mais nas relações ou nas atividades realizadas? Conta mais o percentual de resultado identificado ou a construção de parcerias sólidas e de longo prazo para forjar projetos políticos de desenvolvimento comum? O que está se tornando cada vez mais consensual é a importância de considerar o que cada ator deseja na hora de conduzir uma avaliação e não apenas a introdução de modelos pré-concebidos sem nenhuma reflexão prévia entre os parceiros de um projeto.

4. O impulso do Estado na avaliação

No caso dos EUA, o governo impulsionou o mercado avaliativo na década de 50, especialmente nos programas sociais e posteriormente na área educacional. Na Europa, a OCDE foi crucial para gerar consensos e uma arquitetura institucional que fomentasse e

pensasse a avaliação da cooperação para o desenvolvimento. É necessário o aporte de instituições públicas para fomentar uma cultura avaliativa séria que possua centros de formação de comunidades de prática, especialmente na área da CSS, pois este conhecimento é ainda inexistente.

5. Importância da 'cooperação baseada em evidências'

Muito se faz cooperação, mas com poucos resultados e evidências sobre o valor que possuem. Mostrar resultados será cada vez mais necessário numa era em que aumentam as pressões populares por prestação de contas públicas. Como países em desenvolvimento justificam os recursos destinados para cooperação? É necessário mostrar resultados que possam combinar recursos analíticos tradicionais com as características próprias da CSS.

6. O papel das teorias por trás dos programas

Muitos programas de CSS são fracos em teoria, possuem pouca consistência por trás da troca de experiência e por isso, mesmo quando avaliados, agregam pouco à literatura especializada. Uma visão clara de resultados a serem obtidos e da lógica por trás do programa (além da mera matriz de resultados) ajuda a gerar um diálogo mais consistente entre a teoria e a prática.

7. Avaliação compartilhada e uso dos resultados

Um dos itens que foi discutido neste artigo foi a importância do usuário da avaliação. Afinal, para quem serve a avaliação da CSS? Serve para prestação de contas numa cooperação triangular ou realmente serve para que os governos envolvidos reflitam sobre suas práticas e as aperfeiçoem? É preciso que em avaliações de CSS, haja envolvimento de equipes avaliadoras dos dois, três ou múltiplos países de modo a fornecer subsídios mais apropriados para cada contexto.

8. Foco de uma avaliação Sul-Sul e o *mix* de métodos

O foco de uma avaliação de CSS vai depender do interesse dos atores envolvidos: o que querem avaliar? Resultado ou processo? Estratégia ou implementação? Para além das opções específicas, um bom caminho pode ser a combinação de matrizes lineares com modelos complexos, de indicadores tradicionais de eficiência com indicadores voltados para CSS, de modo a capturar mais da complexidade do que é uma cooperação.

Os métodos tradicionalmente usados na avaliação são experimentais, quasi-experimentais ou testes aleatórios. Para programas de cooperação sul-sul, estes métodos parecem pouco adequados, pois enfatizam impactos objetivos de performance, mas negligenciam dimensões caras à cooperação, como a política. Não se atenta para aspectos pouco tangíveis, mas essenciais das relações entre países e do processo de aprendizagem envolvido.

Por outro lado, é preciso ter cuidado para que a avaliação não fique muito 'soft' e não goze de legitimidade ou alguma capacidade de generalização. É preciso combinar metodologias quantitativas e qualitativas que expressem valores mais alinhados com os objetivos da cooperação sul-sul.

A avaliação de CSS é muito mais do que avaliação de programas. Está relacionada à análise de políticas, de relações entre países, de estratégias de cooperação. Restringir a avaliação ao programa é reduzir o componente estratégico da cooperação sul-sul e fazer da avaliação um instrumento gerencial, mas não chave, nem estratégico. Há que avaliar eficiência do programa e da operação, mas sobretudo o impacto de ideias e relações que não podem ser capturados pela ortodoxia avaliativa que privilegia os métodos experimentais. Foi em anos mais recentes que houve uma abertura neste tipo de abordagem e se privilegiou métodos mais participativos.

Considerações finais

Este artigo se propôs a discutir tendências na área da avaliação aplicadas à CSS. Foi feita uma revisão da discussão sobre como esta temática está sendo tratada pela comunidade internacional e revisitadas algumas das principais tendências na área da avaliação que podem inspirar a busca por resultados na CSS.

Avaliar é mandatório e isso é consensual. A grande questão é: de que tipo de resultados estamos falando? Os resultados da CSS não pode ser limitados àqueles medidos por uma matriz lógica tradicional, mas dizem respeito à diálogos políticos de longo prazo, a relações e princípios. Por isso, trazer para a análise da CSS, constructos de estratégia e complexidade podem ser muito úteis para capturar melhor a riqueza deste tipo de cooperação.

É necessário amadurecer a CSS e este artigo pode ser visto como uma provocação para que mais estudos e avaliações possam surgir nesta área, ainda tão carente de estudos.

Referências

ALKIN, M. (2013). *Evaluation Roots: a wider perspective of theorists's views and influences.* Sage: Los Angeles.

ALKIN, M. & CARDEN, F. (2012). *Evaluation Roots: An International Perspective in Journal of MultiDisciplinary Evaluation*, Volume 8, Number 17, January 2012.

BILAL, S. (2012). *What is the rise of south-south relations about? Development, not Aid.* European Centre for Development Policy Management. Bélgica. **BRICS Policy**

Center (2013). *O Brasil e a Cooperação Sul-Sul: Como Responder aos Desafios Correntes,* Núcleo de Cooperação Internacional para o Desenvolvimento, BRICS POLICY CENTER.

CABRAL, L & WEINSTOCK,

J. (2010). *Brazilian technical cooperation for Development: Drivers, mechanics and future prospects.* ODI: Londres.

DABELSTEIN, N. & KLIEST, T. (2012a). *Preparing, Governing and Managing the Paris Declaration Evaluation.* In: The Canadian Journal of Program Evaluation Vol. 26, No 3, Pages 37-67, Canadian Evaluation Society.

DABELSTEIN, N. & KLIEST, T. (2012b). *Lessons learned and the contributions of the Paris Declaration Evaluation to Evaluation Theory and Practice.* In: The Canadian Journal of Program Evaluation Vol. 26, No 3, Pages 173-200, Canadian Evaluation Society.

DONALDSON, S.I. & LIPSEY, M.W. (2006). *Roles for theory in contemporary evaluation practice: Developing practical knowledge.* In I. Shaw, J.C. Greene, & M.M. Mark (Eds.), *The Handbook of Evaluation: Policies, Programs, and Practices* (pp. 56-75).

GPED (2014). *First High-Level Meeting of the Global Partnership for Effective Development Co-operation: Building Towards an Inclusive Post-2015 Development Agenda Mexico High Level Meeting Communiqué*, 16 April 2014. GPED: Mexico.

OECD (1986). *Methods and Procedures in Aid Evaluation.* OECD: Paris.

OECD (1991). *The DAC Principles for the Evaluation of Development Assistance.* OECD: Paris.

OECD (2000). *Glossary of Evaluation and Results Based Management (RBM) Terms.* OECD: Paris.

OECD (2000). *Boosting South-South Cooperation in the Context of Aid Effectiveness, Telling the Story of Partners Involved in more than 110 Cases of South-South and Triangular Cooperation.* OECD: Paris.

OECD (2013). *Evaluating Development Activities: 12 Lessons from the OECD DAC.* OECD: Paris.

PATTON, M. & PATRIZI, P.A. (2010). *Strategy as the focus for evaluation.* DIRECTIONS FOR EVALUATION, no. 128, Winter 2010 © Wiley Periodicals, Inc., and the American Evaluation Association. Published online in Wiley Online Library (wileyonlinelibrary.com) • DOI: 10.1002/ev.342

PATTON, M. (2011). *Developmental Evaluation: applying complexity concepts to enhance innovation and use.* The Guilford Press, New York.

PINO, B. (2010). *Brazilian Cooperation: a model under construction for an emerging power.* Real Instituto Elcano, Espanha.

RIS (2013). *Conference Report: Conference of Southern Providers - South-South Cooperation: Issues and Emerging Challenges - 15-16 April 2013, New Delhi.* RIS: New Delhi.

- SILVA, M. (2010).** *How did we get there? The pathways of South-South Cooperation.* In: Andrade, M. & Roy, R. (2010). *South-South Cooperation: the same old game or a new paradigm? Poverty in Focus*, N. 20, Brasília-DF.
- SIMPLÍCIO, F. & ANDRADE, M(2013).** *South-South Cooperation Regional Trends: Experiences from Latin America and the Caribbean.* UNDP: NY (no prelo).
- UN (2012).** *South-South and triangular cooperation in the United Nations system - Note by the Secretary-General, Joint Inspection Union,* New York.
- UNDP (2009).** *Handbook on Planning, Monitoring and Evaluation for Development Results.* UNDP: New York.
- UNDP (2013).** *Summary: Human Development Report 2013. The Rise of the South: Human Progress in a Diverse World.* UNDP: NY.
- UNDP (2013).** *Facilitating exchange, evaluation of UNDP contribution to south-south and triangular Cooperation.* UNDP: NY.
- UNOSS (2012).** *South-South Cooperation Principles in Practice: The IBSA Fund Experience.* UNDP: NY. (no prelo).
- WOOD, B.& BETTS, J. (2013b).** *The Paris Declaration Evaluation Process and Methods.* The Canadian Journal of Program Evaluation Vol. 27 No. 3 Pages 69-102. Canadian Evaluation Society.
- VASQUEZ, K. (2013).** *Enhancing Management Practices in South-South and Triangular Cooperation, Study on Country-Led Practices.* JICA & United Nations Office for South-South Cooperation.