



DOI: 10.31560/pimentacultural/2020.867.12-44

Tatiana de Sousa Ribeiro
Yves de Carvalho Carabajal
Ovidio Orlando Filho

**Relatório de avaliação
do projeto
futebol - uma nova
visão do jogo:
meta-avaliação**

Meta-avaliação: uma visão geral

Nos tempos atuais a avaliação se tornou um aspecto primordial no sentido de se aferir o nível de qualidade de ações praticadas por indivíduos e organizações públicas ou privadas. Os resultados alcançados a partir da implementação de projetos, programas e serviços desenvolvidos nos campos empresarial, social, da saúde, educação, entre outros, podem ser aperfeiçoados por meio da realização de um processo avaliativo. A todo momento a avaliação está presente, tendo como intencionalidade a emissão de um julgamento sobre determinada realidade, trazendo subsídios para tomada de decisões inteligentes.

Destaca-se, nesse ponto, que na realização de processos avaliativo alguns aspectos devem ser observados como, por exemplo: (a) termos éticos, direcionando-a para onde ela é mais necessária; (b) termos pragmáticos, uma vez que produtos e serviços de qualidade deficitária podem custar vidas e saúde; (c) termos intelectuais que aprimoram estruturas mentais; e (d) termos pessoais, que permitem a elevação da autoestima na contínua reavaliação dos processos e resultados.

Considerando-se a avaliação como um processo que favorece a tomada de decisão, Gimenes (2007, p. 227) afirma que “os diferentes processos avaliativos não se encerram nas conclusões de seus relatórios ou nas ações realizadas a partir de seus resultados”. Assim, depreende-se que uma avaliação também deve ser aferida em relação à sua qualidade, ou seja, deve ser avaliada, pois, como afirmam Worthen, Sanders e Fitzpatrick (2004), existem avaliações bem-feitas e avaliações malfeitas. Sendo que, neste último caso, existe a possibilidade de o processo avaliativo posto em prática causar mais malefícios que benefícios aos seus interessados.

Conforme afirmou Scriven (1991) o ato de se avaliar uma avaliação é denominado meta-avaliação. Segundo a definição de avaliação de Stufflebeam (2001 apud ELLIOT, 2011, p. 943) em congruência com as dimensões abordadas pelo *Joint Committee on Standards for Educational Evaluation* (2019), meta-avaliação é conceituada como:

O processo de delinear, obter e aplicar informação descritiva e de julgamento — sobre a utilidade, a viabilidade, adequação e precisão de uma avaliação e sua natureza sistemática, competente conduta, integridade/ honestidade, respeitabilidade e responsabilidade social — para orientar a avaliação e divulgar publicamente seus pontos fortes e fracos.

A meta-avaliação coopera para a melhoria e o aperfeiçoamento das políticas e ações institucionais por meio de um acompanhamento e também revisão permanente e sistemática do objeto que está sendo avaliado. Sendo assim, uma meta-avaliação tem como objetivo assegurar o mérito e o valor do processo avaliativo.

O *Joint Committee*, comitê inicialmente responsável pela criação das diretrizes destinadas à avaliação educacional nos Estados Unidos da América e Canadá, foi fundado em 1975. Além de estabelecer padrões de avaliação, também é responsável pelas suas revisões e atualizações; capacitação de formuladores de políticas, avaliadores e educadores; e atua como uma câmara de compensação na literatura de padrões de avaliação (JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION, 2019).

Com este Comitê foram originados 30 padrões agrupados, inicialmente, em quatro categorias a saber: utilidade, viabilidade, adequação e precisão deu-se o nome de *Standards* ou Padrões de Avaliação, incluindo-se posteriormente, a categoria responsabilização em sua última revisão, apresentada pelo *Joint Committee*, em 2011 (YARBROUGH et al, 2011).

Elliot (2011), sustentada em pensamentos formulados por Berends e Roberts (2003), apontou duas razões para a realização de uma meta-avaliação: a) possibilitar a verificação da aplicação de um código de comportamento adequado no decorrer de todo o estudo, ou seja, a ética adotada pelos avaliadores, se ela existe e como foi praticada; e b) a oferta de meios de incorporação de novos conhecimentos advindos não somente dos meta-avaliadores, como também dos avaliadores, à medida que tomam ciência da análise realizada. Logo, define-se meta-avaliação como uma prática de reflexão sobre todos os procedimentos utilizados na avaliação.

Objetivo do estudo

Por meio de projetos sociais, a administração pública vem assumindo a responsabilidade de criar e implementar políticas públicas. Esses projetos necessitam ser avaliados, dada a forte relevância de suas atividades no atual cenário socioeconômico do país. Os projetos sociais são implementados em diferentes áreas da ação social dos governos, como nas áreas da educação, agricultura, habitação, meio ambiente, lazer entre outras. Eles possuem funções essenciais, atuando no auxílio às populações de diferentes regiões, faixa etária e situação econômica. Todos esses fatores reafirmam a importância de avaliá-los, no sentido de que os mesmos alcancem de fato e com a devida qualidade os objetivos pelos quais foram implementados.

Sobre esse aspecto, a relevância da avaliação de projetos sociais, Tavares (2009, p. 3) afirma:

As avaliações de projetos [sociais] tornam-se fundamentais para a sua continuidade no sentido de evidenciar a justificativa dos esforços empenhados e a relevância da sua atuação para a comunidade. [...] A importância deste tipo de avaliação é o de prestar contas para a comunidade, sobre a qualidade do trabalho que vem

sendo desenvolvido por estes projetos. Verificar se as instituições aplicam bem os recursos que recebem e se procuram melhorar a atuação no sentido de atingir seus propósitos sociais.

A citada autora aborda senão um dos pontos de maior significância na avaliação de projetos sociais: a prestação de contas. Observa-se que a receita dos projetos sociais terá variação conforme a extensão de seus domínios. Um projeto social localizado em uma cidade ribeirinha terá um fluxo de caixa completamente diferente de um projeto social financiado por um renomado banco estrangeiro. Ainda assim, a prestação de contas deverá ser realizada, tenha o projeto social um pequeno ou grande investidor.

Logo, a relevância de se avaliar projetos sociais deve-se por dois motivos: o primeiro, devido ao fato de que a avaliação auxilia esses projetos no julgamento do seu mérito, proporcionando assim o surgimento de gestões consolidadas e responsáveis que agregam valor à sua existência; e o segundo, devido ao fato de justificar os “esforços empenhados e a relevância da sua atuação para a comunidade” (TAVARES, 2009, p. 3).

Assim sendo, o objetivo deste estudo é apresentar o desenvolvimento de uma meta-avaliação somativa, realizada em um relatório final de uma avaliação já encerrada, produzido pela Fundação Itaú Social (FUNDAÇÃO EPROCAD, 2016), o Relatório de Avaliação Econômica do Projeto Futebol - uma nova visão de jogo, realizado utilizando a metodologia de avaliação econômica da citada Fundação. Assim sendo, esse estudo procurou, por meio de uma meta-avaliação, verificar se os objetivos traçados pela avaliação do projeto social, objeto daquele estudo, foram atingidos, assim como, identificar os pontos fortes e fracos do processo avaliativo.

O objeto do estudo

As avaliações são instrumentos essenciais para qualificar o investimento social realizado pelas organizações públicas e privadas, servindo também como meio de verificação da efetividade das iniciativas e seu impacto social. Esses instrumentos orientam as escolhas da Fundação Itaú Social em relação ao investimento social e proporcionam respostas objetivas e mensuráveis a respeito dos recursos investidos frente aos resultados alcançados. Devido à tamanha importância dada ao tema da avaliação de projetos sociais, a Fundação Itaú Social (2015):

Realiza o Programa de Avaliação Econômica de Projetos Sociais que tem como objetivo difundir a cultura e a prática deste tipo de avaliação por meio de cursos, seminários, oficinas e eventos direcionados a gestores de projetos sociais de organizações da sociedade civil e órgãos públicos.

Assim, o objeto deste estudo é o Relatório de Avaliação Econômica de um projeto social intitulado Projeto Futebol - Uma Nova Visão de Jogo, avaliado pela Fundação Itaú Social em 2016. Desenvolvido pela Fundação Esportiva Educacional Pró-Criança e Adolescente (EPROCAD). O Projeto Futebol iniciou suas atividades no ano de 2008, no município de Santana do Parnaíba, no estado de São Paulo.

A iniciativa de se empregar o esporte como complemento de formação de crianças nesse município teve início em 1986, utilizando-se uma outra nomenclatura conhecida como Esportes Pró-Parnaíba (EPROPAR). A versão atual do projeto tem como público-alvo crianças e adolescentes de seis a 17 anos, estudantes da rede escolar do município, e conta com cerca de 350 vagas. O projeto valoriza a permanência de seus integrantes, de forma que os veteranos têm prioridade no processo de inscrição. Dentre seus principais objetivos, está o de melhorar a autoestima de seus

participantes, a partir de valores como respeito, cooperação, solidariedade, honestidade, justiça, responsabilidade e prudência (FUNDAÇÃO EPROCAD, 2016).

Tais atributos, por serem de difícil mensuração, levaram o projeto a buscar compatibilizá-los com medidas validadas na literatura especializada: (a) medida de sociabilidade a partir do questionário de capacidades e dificuldades (SDQ); (b) medida de impulsividade por meio da escala de impulsividade de domínio específico de Angela Duckworth; (c) medida de Grit usando-se de Angela Duckworth; (d) medida de Lócus de Controle a partir da escala de Rotter; (e) medida de autoestima com o uso da escala Rosenberg de autoestima; e (f) medida de honestidade por meio da escala HEXACO-PI-R de honestidade (FUNDAÇÃO EPROCAD, 2016).

Como apresentado pela *streetfootballworld*, mentora da metodologia baseada em *Fútbol Callejero*, ou futebol de rua, aplicada pela EPROCAD, “o futebol é usado para transformar crianças com baixa autoestima e auto percepção negativa em cidadãos e cidadãs proativos e comprometidos a mudar a realidade que vivem” (FUNDAÇÃO EPROCAD, 2016).

De acordo com Stottili (2013 apud LEAL, 2015, p. 123):

A metodologia do Fútbol Callejero (ou de Rua) mostra que o futebol pode ter outra cara, a cara do diálogo, da mediação de conflitos e da garantia dos direitos humanos. [...] a iniciativa é democrática, solidária, respeita as diversidades, as diferenças e mobiliza a todos. É uma oportunidade para discutir os padrões dos grandes eventos, como em relação às mulheres e negros, por exemplo.

A metodologia adotada na avaliação realizou-se por meio de uma comparação entre participantes de anos anteriores e indivíduos que nunca participaram do programa. Foram utilizadas três ferramentas metodológicas para estimar o efeito do EPROCAD:

Mínimos Quadrados Ordinários (MQO) com controles, MQO sem controles e Pareamento por Escore de Propensão (P-SCORE).

A questão avaliativa do relatório, proposta pela Fundação Itaú Social (FUNDAÇÃO EPROCAD, 2016) que se buscou responder com a metodologia foi:

Qual o efeito de participar do projeto da EPROCAD sobre indicadores de autoestima, impulsividade e honestidade dos participantes?

A escolha dos participantes foi feita a partir da divulgação nas escolas da região municipal de Santana de Parnaíba, onde os interessados se inscreveram e concorreram às vagas. Para a seleção dos participantes, além do número de vagas, existiram outros elementos que priorizavam a vulnerabilidade socioeconômica da família, o fato do aluno ser veterano, ter carta-convite (indicação de participantes de anos anteriores) e o encaminhamento do Conselho Tutelar, Assistência Social e Secretaria da Educação.

Para responder à questão avaliativa foi desenvolvido, em conjunto com o grupo gestor da EPROCAD, um questionário como instrumento de coleta de dados. Nele foram colocadas questões associadas aos indicadores de interesse descritos anteriormente e questões que mapeiam informações, como convivência dos indivíduos com seus familiares, cotidiano da vida escolar, condições de moradia e características socioeconômicas dos indivíduos.

Os respondentes foram divididos em dois grupos: (a) grupo de tratamento, que são os participantes do projeto em 2014 e anos anteriores (veteranos) que continuaram a participar em 2015 e (b) grupo de controle, que são os participantes apenas em 2015 (novatos) e não participantes (colegas de turma dos participantes).

Devido às particularidades das medidas utilizadas e a dificuldade de se criar questões que acessassem os indicadores propostos

de pessoas entre 6 a 17 anos, o questionário foi construído e aplicado para indivíduos que tinham entre 10 e 15 anos, totalizando a quantidade de 624 respondentes (FUNDAÇÃO EPROCAD, 2016).

O Relatório de Avaliação da Fundação Itaú Social (FUNDAÇÃO EPROCAD, 2016) aponta que o programa EPROCAD gera um aumento de 0,112 (3%) pontos no nível de autoestima de seus participantes, resultado esperado e condizente com o projeto. Observou-se também que o programa seleciona indivíduos com maiores níveis de honestidade, sendo estes níveis aumentados aos indivíduos que participam há mais tempo. Além disso, para indivíduos com idades entre 14 e 15 anos, o programa mostrou diminuir seus níveis de impulsividade em 0,245 (13%).

A partir de estimativas obtidas pela literatura acadêmica, os avaliadores detectaram que um aumento médio de 10% no nível de autoestima leva a um aumento de 4,8% no salário dos indivíduos. Transportando esse resultado para o efeito do projeto, foi detectado que o projeto EPROCAD aumenta em 1,5% o salário de seus participantes.

Considerando que o investimento contábil do programa em 2014, por aluno, foi de R\$ 3.266,93, tendo em vista custos com trabalho voluntário (presidente, psicólogos, conselho), aluguel da unidade centro, doações, vale-transporte e alimentação. Já as estimativas de retorno econômico mostram que, no pior dos cenários, a taxa interna de retorno do projeto é de 6,8% a.a. e o valor presente líquido por participante é de R\$ 14.076,56. No melhor dos cenários, a taxa interna de retorno obtida foi de 11,1% a.a. e o valor presente líquido por participante, de R\$ 31.420,06. Para cada R\$ 1,00 gasto no projeto, o programa gera em termos de salários no ciclo de vida de seus participantes, de R\$ 5,31 a R\$ 10,62.

Finalmente, com os resultados obtidos por meio do Relatório de Avaliação, concluiu-se que o projeto da EPROCAD é uma

possibilidade favorável para melhorar a autoestima de crianças do município de Santana de Parnaíba e que o investimento social é considerado válido, com base no cálculo do retorno econômico.

Metodologia da meta-avaliação

A metodologia eleita para nortear este estudo foi pautada nos padrões do *Joint Committee on Standards for Evaluation*. Segundo Yarbrough et al. (2011), os padrões de avaliação do *Joint Committee* compõem cinco categorias: utilidade, exequibilidade, adequação, precisão e responsabilização. No entanto, devido à necessidade de se delimitar o estudo, somente as categorias Utilidade e Precisão foram escolhidas para serem contempladas na meta-avaliação do Relatório de Avaliação Econômica do Projeto Futebol - uma nova visão de jogo. Foram escolhidos, assim, os dois padrões que melhor se adequam, considerando os propósitos do processo avaliado por meio da metodologia da Fundação Itaú Social.

A primeira delas, a Categoria Utilidade é composta por oito padrões, abarcando o conceito da:

Preocupação com as necessidades de informação dos usuários, interessados na avaliação. [...] a avaliação deve ser conduzida por um avaliador ou equipe competente e isto deve servir como garantia de sua credibilidade; dar atenção a todos os interessados (*stakeholders*) na avaliação em foco; negociar os propósitos da avaliação e selecionar informação relevante de modo a atender às necessidades desses interessados; indicar os valores que embasam os propósitos, processos e a interpretação dos resultados; apresentar comunicações e relatórios que descrevam claramente o que foi realizado e que sejam entregues nos prazos estabelecidos; considerar as possíveis consequências e influência que os resultados da avaliação podem vir a ter. (ELLIOT, 2011, p. 951).

O instrumento aplicado para a coleta e análise das categorias e seus respectivos padrões foram as listas de verificação

composta por cada categoria e seus respectivos padrões. Existem várias definições para o que seja uma lista de verificação e deve-se ter em mente que ela pode ser empregada com inúmeras possibilidades, sempre em conformidade com o objetivo que se deseja atingir (LEITE, 2012).

Para que os padrões de Utilidade e Precisão da avaliação fossem atendidos no processo avaliativo, foram utilizadas as Listas de verificação de atendimento dos padrões expressas nos Quadros 1 e 2:

Quadro 1 – Lista de verificação de atendimento aos padrões de utilidade

Padrões	Condição de atendimento
U1 - Credibilidade do avaliador	As avaliações devem ser conduzidas por pessoal qualificado que estabeleça e mantenha credibilidade no contexto da avaliação.
U2 - Atenção aos interessados	As avaliações devem dedicar atenção a todos os indivíduos e grupos envolvidos no programa e por elas afetados.
U3 - Propósitos negociados	Os propósitos da avaliação devem ser identificados e continuamente negociados com base nas necessidades dos <i>stakeholders</i> .
U4 - Explicitação de valores	As avaliações devem esclarecer e especificar os valores culturais e individuais que fundamentam objetivos, processos e julgamentos.
U5 - Informação relevante	A informação obtida por meio da avaliação deve servir às necessidades identificadas e emergentes dos <i>stakeholders</i> .
U6 - Produtos e processos significativos	As avaliações devem construir atividades, descrições e julgamentos, de modo que os participantes se sintam estimulados a redescobrir, reinterpretar ou rever seus entendimentos e comportamentos.
U7 - Comunicação e relatórios apropriados e no prazo	As avaliações devem atender às constantes necessidades de informação de seus vários públicos.

U8 - Preocupação com consequências e influências	As avaliações devem promover o uso responsável e adequado de seus resultados, além de prevenir consequências negativas indesejáveis e má utilização.
--	--

Fonte: Os autores (2019) adaptado de YARBROUGH et al. (2011).

Em contrapartida, os oito padrões da Categoria Precisão do *Joint Committee* (2011b) (Quadro 2) estão relacionados “à clareza da descrição do objeto avaliado, bem como o contexto, apresentando o planejamento da avaliação desde os procedimentos até suas conclusões. Asseguram assim, confiabilidade, validade e fidedignidade das informações” (SANTOS, 2017, p. 92-93).

Quadro 2 - Lista de verificação de atendimento aos padrões de precisão

Padrões	Condição de atendimento
P1 - Conclusões e decisões justificadas	As conclusões e decisões da avaliação devem ser explicitamente justificadas nas culturas e nos contextos onde têm consequências.
P2 - Informação válida	A informação da avaliação deve servir aos propósitos pretendidos e sustentar interpretações válidas.
P3 - Informação fidedigna	Os procedimentos da avaliação devem gerar informações suficientemente confiáveis e consistentes, de modo que sirvam aos usos pretendidos.
P4 - Explicitação das descrições do programa e do contexto	As avaliações devem documentar os programas e seus contextos com apropriado detalhe e escopo para os propósitos estabelecidos.
P5 - Gerenciamento da informação	A avaliação deve empregar métodos sistemáticos de coleta, revisão, verificação e armazenamento das informações.
P6 - Análises e delineamentos (<i>designs</i>) sólidos	A avaliação deve empregar delineamentos tecnicamente adequados e análises apropriadas aos propósitos estabelecidos.

P7 - Explicitação do raciocínio da avaliação	A fundamentação da avaliação, desde informações e análises até resultados, interpretações, conclusões e julgamentos, deve ser clara e completamente documentada.
P8 - Comunicação e relatório	As comunicações da avaliação devem ter escopo adequado, evitando concepções errôneas, vieses, distorções e erros.

Fonte: Os autores (2019) adaptado de YARBROUGH et al. (2011).

Para cada categoria, foi elaborado um quadro para a análise do atendimento, ou não, do relatório avaliado à luz de cada padrão. Para cada padrão, foram definidas quatro opções de resposta: *Atende (A)*, *Atende Parcialmente (AP)*, *Não Atende (N)* e *Não se Aplica (NA)*. Todo juízo de valor emitido pela meta-avaliação possui caráter qualitativo, ou seja, subjetivo, tendo sido formulado em função da forma de se ver a realidade pelos autores do estudo. Os resultados de julgamento de cada padrão são apresentados por meio de evidências extraídas do próprio relatório de avaliação, seguidas de análises construídas pelos meta-avaliadores, autores deste estudo.

Na busca do estabelecimento de critérios que contemplassem o objetivo apresentado por este estudo, optou-se pela utilização dos parâmetros constantes da lista de verificação fundamentada nos estudos do *Joint Committee*, expostos nos Quadros 1 e 2 (YARBROUGH et al., 2011). Para cada padrão, foi realizado um estudo no conteúdo do Relatório, no sentido de se verificar seu atendimento ou não. Posteriormente, foram realizadas análises e apontados os resultados da aplicação de cada padrão ao relatório de avaliação.

Resultados da categoria de utilidade

Os padrões correspondentes às características de utilidade levam a ajuizar até que ponto uma avaliação se prestará às necessidades de seus usuários. Em conformidade com essa descrição, Elliot (2001, p. 948) afirma que:

Uma avaliação precisa ser útil para aqueles que a encomendaram e que nela têm interesse. Assim, a utilidade de uma avaliação é determinada pela necessidade que determinado grupo de interessados possui em relação a seus resultados.

Segundo Yarbrough et al (2011), os julgamentos sobre a utilidade da avaliação são elaborados baseados na medida em que os interessados encontram processos de avaliação valiosos no atendimento de suas necessidades. As descrições, as evidências de atendimento aos padrões da categoria Utilidade identificadas no Relatório e os pareceres avaliativos são apresentados a seguir:

Atendimento ao padrão credibilidade do avaliador

As avaliações devem ser conduzidas por pessoas tanto fidedignas quanto competentes para realizar o estudo e para que as descobertas da avaliação tenham o máximo de credibilidade e aceitação.

Segundo a Fundação Itaú Social, em texto extraído de seu Relatório de avaliação (FUNDAÇÃO EPROCAD, 2016):

Uma das organizações responsáveis por realizar o investimento social do Banco Itaú, certificada pelo Conselho Municipal de Assistência Social (COMAS) nº 1052/2013, tem como atividades centrais a formulação, a implantação e a disseminação de metodologias voltadas à melhoria de políticas públicas na área educacional e à avaliação econômica de projetos sociais. A atuação da

Fundação abrange todo o território brasileiro, desenvolve-se em parceria com governos, setor privado e organizações da sociedade civil e é orientada pelo foco em resultados, buscando contribuir de maneira efetiva para a transformação social. Os projetos selecionados recebem assessoria técnica de especialistas para realizar a avaliação econômica.

Assim sendo, a composição da equipe responsável pela avaliação é feita por profissionais qualificados, capacitados e com experiência na aplicação da metodologia utilizada. Esses profissionais avaliadores são certificados para aplicarem a metodologia de Avaliação Econômica de Projetos Sociais da Fundação Itaú Social. Segundo Yarbrough et al. (2011), a marca registrada de avaliadores e equipes de avaliação confiáveis é a capacidade de atender às necessidades dos interessados e, concomitantemente seguir padrões e diretrizes profissionais. Assim, considera-se que o Relatório de Avaliação *atendeu* ao padrão designado.

Atendimento ao padrão atenção aos interessados (*stakeholders*)

Considerando que as avaliações devem dedicar-se a atender as necessidades de todos os indivíduos e grupos identificados em um programa e por elas afetados, o Relatório de Avaliação aponta que o Programa Futebol:

Tem como público-alvo crianças e adolescentes de 6 a 17 anos estudantes da rede escolar do município e conta com cerca de 350 vagas. O projeto valoriza a permanência de seus integrantes, de forma que veteranos têm prioridade no processo de inscrição [...] devido a peculiaridade das medidas utilizadas e a dificuldade de se criar questões que acessem os indicadores propostos de pessoas entre 6 e 17 anos, resolvemos construir e aplicar o questionário para indivíduos que tenham entre 10 e 15 anos. (FUNDAÇÃO EPROCAD, 2016, p. 9-10).

Conforme expressou Elliot (2011, p. 949) “o avaliador não deve ter preferências por qualquer grupo de *stakeholders*, nem apresentar julgamentos tendenciosos”, e de acordo com Yarbrough et al. (2011), os interessados são identificados por meio de uma análise da abrangência do programa. Uma vez identificados, o desafio é trabalhar para que revelem suposições, interesses, valores, comportamentos e dúvidas com relação ao programa.

Apesar de não ter ocorrido o envolvimento de participantes em cargo ou função de liderança, a avaliação se dirigiu a determinados *stakeholders*, os quais, somente estes, teriam o interesse em participar do processo avaliativo. Pode-se verificar que não houve a preocupação do avaliador em contemplar todas as partes interessadas no processo avaliativo, apesar de terem sido identificadas claramente no Relatório de Avaliação. A partir da evidência mencionada, pode-se verificar que o padrão foi *parcialmente atendido*.

Atendimento ao padrão propósitos negociados

Considera-se que os propósitos de uma avaliação devem ser identificados e negociados permanentemente, priorizando as necessidades e interesses dos *stakeholders* para que possam trazer respostas pertinentes ao programa. Ressalta-se que o Relatório expressou, de forma clara, seu objetivo por meio de uma avaliação econômica do Projeto Futebol - Uma Nova Visão do Jogo, organizado e desenvolvido pela Fundação EPROCAD. O objetivo do relatório foi o de avaliar o impacto de se participar do projeto sobre indicadores de habilidades não cognitivas, proposto pelos gestores da própria ONG. Esta avaliação de impacto, por sua vez, propôs identificar se o projeto teve ou não efeito sobre seus beneficiários, tendo como público-alvo pessoas entre seis e 17 anos, participantes

de 2014, e anos anteriores com indivíduos que nunca participaram ou novatos de 2015 (FUNDAÇÃO EPROCAD, 2016).

Conforme a proposta definida pelos gestores do EPROCAD, e de acordo com a afirmação de Yarbrough et al. (2011), “o alcance da descrição do programa, a coleta de dados, as estratégias de análise, e os métodos de divulgação e relatório serão determinados com base em sua contribuição para os propósitos estabelecidos”. Logo, ficou claro o *atendimento* ao padrão.

Atendimento ao padrão explicitação de valores

Nos processos avaliativos, os valores culturais e individuais que fundamentam objetivos, processos e julgamentos usados para interpretar as descobertas, devem ser esclarecidos e especificados de forma minuciosa pelas avaliações para que as bases dos juízos de valor sejam cuidadosamente compreendidas. Nesse aspecto, cabe se observar a declaração da organização *Streetfootballworld* (FUNDAÇÃO EPROCAD, 2016), ao afirmar que o futebol é usado para transformar crianças com baixa autoestima e auto percepção negativa em cidadãos proativos e comprometidos a mudar a realidade em que vivem.

Nesse ponto, na gestão do EPROCAD “busca-se desenvolver nos participantes atributos como respeito, cooperação, solidariedade, honestidade, justiça, responsabilidade e prudência, como elencado na Escala de Valores, documento criado pelos gestores do programa” (FUNDAÇÃO EPROCAD, 2016, p. 8).

Logo, os avaliadores contemplaram todas as fontes relevantes de valores para interpretação dos resultados da avaliação, considerando as necessidades sociais, as necessidades dos inte-

ressados, missão institucional e metas do programa. Segundo Yarbrough et al. (2011), os valores também determinam a negociação dos propósitos da avaliação por estarem no centro de todas as decisões e julgamentos resultantes de uma avaliação. Portanto, houve *atendimento* do Relatório de Avaliação ao que é recomendado pelo padrão.

Atendimento ao padrão informação relevante

As informações obtidas por meio da avaliação devem ser relevantes, adequadas e em conformidade às necessidades identificadas e emergentes dos interessados. Nesse aspecto, destaca-se no Relatório meta-avaliado que para avaliar o efeito de participação no programa foram construídas questões associadas aos indicadores de medidas de sociabilidade, de impulsividade, a medida de Grit (“garra”), de locus de controle e medida de autoestima. Foram, também, elaboradas questões que mapeiam informações, como convivência dos indivíduos com seus familiares, cotidiano da vida escolar, condições de moradia e características socioeconômicas dos indivíduos.

Contribuindo com esta evidência, é esclarecido no texto do documento que:

As perguntas associadas às medidas de impulsividade, locus de controle, autoestima e honestidade produziram escalas razoavelmente coerentes. A medida de locus de controle, contudo, teve um problema alternativo: nela estamos interessados em diagnosticar como um indivíduo relaciona seus resultados na vida com questões internas (esforços nos estudos, por exemplo) e externas (sorte, por exemplo). [...] por conta de imprevistos não foi possível fazer uma avaliação de impacto de fato. (FUNDAÇÃO EPROCAD, 2016, p. 10-11).

Assim, considerando a complexidade mencionada na fase de coleta dos dados e tendo em vista que, segundo Yarbrough et al. (2011), o avaliador precisa ajudar os *stakeholders* a entenderem o valor dos procedimentos alternativos para coleta e análise da informação, além das razões pelas quais determinada abordagem é recomendada, conclui-se que o padrão foi *parcialmente atendido*.

Atendimento ao padrão produtos e processos significativos

As atividades, descrições, resultados e julgamentos construídos a partir de uma avaliação devem estimular os participantes a redescobrirem, reinterpretarem ou reverem seus entendimentos e comportamentos. Conforme apresentado no Relatório de Avaliação, o acordo foi avaliar o impacto de se participar do programa sobre indicadores de habilidades não cognitivas, como proposto pelos gestores da Fundação EPROCAD. Os resultados obtidos por meio de questionários aplicados com a escala Rosenberg de Autoestima, possibilitou considerar que o Projeto Futebol - Uma Nova Visão de Jogo afeta positivamente a autoestima de seus participantes.

Foi esclarecido também que, para indivíduos mais velhos (14-17 anos), o projeto reduz a impulsividade. Contudo, apesar de as crianças que participam do projeto terem maiores níveis de honestidade, não se pôde atribuir esta medida ao programa. O resultado, por meio do estudo aplicado e disponibilizado ao público por meio da internet, aponta que o programa é mais eficaz que outros projetos de Santana de Parnaíba em melhorar a autoestima das crianças e conseqüentemente aumenta o salário de seus participantes em 1,5%. Conclui-se que o projeto EPROCAD é um investimento social com bom retorno econômico (FUNDAÇÃO EPROCAD, 2016).

Desse modo, de acordo com as evidências apresentadas, é possível perceber o planejamento de atividades que contribuem para resultados significativos, auxiliando os *stakeholders* a entenderem melhor a complexidade das operações, a analisarem a própria contribuição para o programa, além de como intensificá-la. Portanto, o Relatório de Avaliação *atende* ao que preconiza o padrão.

Atendimento ao padrão comunicação e relatórios apropriados e no prazo

Para que os propósitos de uma avaliação sejam atingidos é fundamental que os relatórios descrevam de forma clara o programa avaliado, inclusive seu contexto, os propósitos, os procedimentos e conclusões da avaliação. Com isso, é possibilitada a produção de informações essenciais que atendam de modo oportuno e contínuo os *stakeholders* e que sejam facilmente entendidas e utilizadas no momento adequado. Então, cabe observar que o estudo no Relatório meta-avaliado mostrou a existência de preocupação, no processo avaliativo, de se gerar uma documentação que possibilitasse o entendimento pelos interessados, não ocorrendo uso excessivo de teor técnico ou acadêmico. O Relatório encontra-se disponível para *download* no site da Rede Itaú Social de Avaliação, possibilitando o acesso por parte dos gestores do projeto EPROCAD e consequente tomada de decisões (FUNDAÇÃO EPROCAD, 2016).

Assim sendo, em conformidade com Yarbrough et al. (2011), o avaliador demonstrou responsabilidade em garantir o acesso imediato pelos interessados às informações importantes, disponibilizando-as de tal forma que atendam adequadamente às suas necessidades. Nesse sentido, observa-se a *conformidade* com o padrão mencionado.

Atendimento ao padrão preocupação com influências e consequências

As avaliações devem ser planejadas, realizadas e apresentadas de forma a promover o uso responsável e adequado de seus resultados, evitando impactos e consequências negativas e sua utilização indevida. Nesse aspecto, a avaliação construiu atividades, descrições e julgamentos, de modo que os participantes se sentissem estimulados a redescobrir, reinterpretar ou rever seus entendimentos e comportamentos. Não foram identificadas evidências de que os avaliadores condicionaram, às partes interessadas, oportunidade de testar suas suposições e ajustar seus entendimentos correspondente ao objetivo e as consequências dos resultados do processo avaliativo.

Logo, a não conformidade foi percebida por não terem sido encontradas no Relatório de Avaliação evidências do *atendimento* do padrão Preocupação com Influências e Consequências.

Resultados da categoria de precisão

As avaliações devem buscar a precisão, ou seja, a clareza na transcrição dos resultados, dos pareceres avaliativos e na redação do relatório de avaliação. A redução de inconsistências, distorções e concepções equivocadas aumentam a precisão das avaliações (YARBROUGH et al., 2011).

Atendimento ao padrão conclusões e decisões justificadas

De acordo com Yarbrough et al. (2011), as avaliações devem conter conclusões e decisões precisas sendo estas explicitamente

justificadas nas culturas e nos contextos onde têm consequências. Nesse aspecto, no relatório meta-avaliado encontra-se exposto que:

Foram utilizadas três metodologias para estimar o efeito do programa: Mínimos Quadrados Ordinários, Mínimos Quadrados Ordinários com controles e Pareamento por Escore de Propensão. As estimativas do efeito do programa mostraram resultados positivos e estatisticamente significativos sobre o nível de Autoestima das crianças. [...] ao analisar os subgrupos dos participantes, vimos que indivíduos que participam há mais tempo do programa têm ainda maiores níveis de Honestidade. Além disso, para indivíduos com idades entre 14 e 15 anos o programa mostrou diminuir seus níveis de Impulsividade em 0,245 (13%). (FUNDAÇÃO EPROCAD, 2016, p. 4).

Considerando que todo o processo avaliativo deve evitar qualquer tipo de omissão que interfira no seu mérito e na fidedignidade de suas informações, assim, a explicitação do processo de coleta de dados expressa no relatório meta-avaliado deve ser relevante e essencial para a precisão de uma avaliação. De acordo com o exposto, considera-se que o padrão foi *atendido*.

Atendimento ao padrão informação válida

Considerando a afirmação de Yarbrough et al. (2011) de que os procedimentos de coleta de informações devem ser gerados e implementados de tal maneira que sirva aos propósitos pretendidos e assegurem interpretações válidas dos resultados, e de acordo com a afirmativa expressa no relatório: “Devido à particularidade das medidas utilizadas e a dificuldade de se criar questões que acessem os indicadores propostos de pessoas entre 6 e 17 anos, resolvemos construir e aplicar o questionário para indivíduos que tenham entre 10 e 15 anos” (FUNDAÇÃO EPROCAD, 2016, p. 10). Além de observar que:

Foram utilizadas *cinco* medidas: *i.* Medida de sociabilidade: Questionário de Capacidades e Dificuldades (SDQ); *ii.* Medida de impulsividade: Escala de Impulsividade de domínio específico de Angela Duckworth; *iii.* Medida de Grit (“garra”): Escala de Angela Duckworth; *iv.* Medida de Lócus de Controle: Escala Rotter; *v.* Medida de Autoestima: Escala Rosenberg de Autoestima; *vi.* Medida de Honestidade: Escala HEXACO-PI-R de honestidade. (FUNDAÇÃO EPROCAD, 2016, p. 8).

Considera-se que o padrão foi *atendido parcialmente*, pois, segundo Yarbrough et al. (2011), a validade, em seus vários aspectos importantes integra todos os padrões de precisão. Nesse ponto, o relatório contemplou o padrão Propósitos Negociados, o qual a validade das interpretações está relacionada com os sistemas de pensamento e de sua relação com os propósitos da avaliação, constantes no padrão Comunicação e Relatório.

Atendimento ao padrão informação fidedigna

As avaliações devem se pautar em informações suficientemente confiáveis servindo de base para a validade e qualidade de seus resultados. Nesse ponto, o Relatório de Avaliação assinala que, devido às peculiaridades das medidas utilizadas e a dificuldade de se criar questões que acessem os indicadores propostos de pessoas entre 6 a 17 anos, para coleta de dados utilizou-se como instrumento o questionário aplicado para indivíduos que tinham entre 10 e 15 anos, totalizando a quantidade de 624 respondentes.

Os avaliadores documentaram e relataram as ocorrências geradoras de informações obtidas e incluíram os instrumentos de coleta de dados utilizados em um apêndice técnico no relatório. Mesmo estando em congruência com Yarbrough et al. (2011), que recomendam aos avaliadores descreverem os procedimentos adotados para fundamentar a consistência, oferecendo estimativas

adequadas de fidedignidade, a evidência aponta que o procedimento de coleta de dados não contemplou todos os participantes em função da dificuldade da criação de indicadores precisos. Assim, considerou-se o *atendimento parcial* ao padrão.

Atendimento ao padrão explicitação das descrições do programa e do contexto

O contexto de um programa deve ser examinado e documentado de forma suficientemente detalhada, possibilitando aos interessados uma compreensão adequada dos propósitos estabelecidos. Nesse aspecto a Fundação Itaú Social (FUNDAÇÃO EPROCAD, 2016, p. 9) expressa em seu Relatório que:

O programa Futebol – Uma nova visão do jogo acontece anualmente no município de Santana de Parnaíba e existe desde 2008. Ele tem como público-alvo crianças e adolescentes de 6 a 17 anos estudantes da rede escolar do município e conta com cerca de 350 vagas. O projeto valoriza a permanência dos participantes, de forma que veteranos têm prioridade no processo de inscrição.

O propósito deste padrão está na descrição detalhada do programa, a fim de que os avaliadores compreendam o contexto em que será realizada a avaliação. Por meio da leitura do relatório de avaliação, é possível observar o esclarecimento do que exatamente foi acordado para a execução do processo avaliativo. Logo considera-se que o padrão foi *atendido*.

Atendimento ao padrão gerenciamento da informação

As informações produzidas nas avaliações devem ser sistematicamente coletadas, revisadas, verificadas e armazenadas,

visando assegurar sua validade, acessibilidade e confiabilidade. Nesse ponto, observa-se que o Relatório de Avaliação, em sua metodologia, aponta que, para responder à questão avaliativa, o Relatório foi dividido em dois grupos aos respondentes do questionário: de tratamento, sendo os participantes do projeto em 2014 e anos anteriores; e de controle, formado por participantes apenas do ano de 2015. Um dos recursos metodológicos adotados foi o Pareamento por Escore de Propensão (*P-score*), que buscou separar os indivíduos em subgrupos a partir das variáveis observadas (idade, sexo, escolaridade da mãe).

Conforme Yarbrough et al. (2011 apud SANTOS, 2017)

Um dos principais objetivos da realização de uma avaliação é o de oferecer aos *stakeholders* as informações de que precisam e, para isso, existem diferentes métodos que podem ser aplicados ao processo avaliativo, de forma a garantir a precisão dessas informações.

Desse modo, as informações podem ser obtidas por meio da comunicação oral, dados numéricos, observação, aplicação de instrumento de coleta de dados, entre outros. Portanto, houve *atenção* do relatório ao que é recomendado pelo padrão.

Atendimento ao padrão análises e delineamentos sólidos

A avaliação deve adotar delineamentos tecnicamente adequados que facilitem uma estrutura de raciocínio capaz de resultar em conclusões justificáveis e permitir análises apropriadas aos propósitos e questões estabelecidas. Nesse aspecto, segundo a Fundação Itaú Social:

A medida de lócus de controle, contudo, teve um problema alternativo: nela estamos interessados em diagnosticar como um indivíduo relaciona seus resultados na vida com questões internas (esforço nos estudos, por exemplo) e externas (sorte, por exemplo). O ideal

é obrigar o respondente a escolher qual das duas questões é mais relevante. Em nosso caso, não fizemos isso e, portanto, os indivíduos acabaram relatando que tudo é importante: fatores externos e internos. (FUNDAÇÃO EPROCAD, 2016, p. 11).

Os avaliadores esboçaram um planejamento exequível do ponto de vista de uma avaliação de um projeto social. Conforme apresentado no relatório, durante o processo de coleta de dados, os avaliadores não hesitaram em modificar os delineamentos referente às questões construídas a partir da medida de locus de controle. De acordo com as evidências, percebe-se a *conformidade* do relatado ao padrão.

Atendimento ao padrão explicitação do raciocínio da avaliação

A fundamentação da avaliação, desde informações e análises até resultados, interpretações, conclusões e julgamentos, deve ser clara e completamente documentada. Assim, se pode considerar que a justificativa do atendimento do padrão está presente no relatório de avaliação no trecho que fala sobre a construção do questionário, no qual “foram colocadas questões que mapeiam informações, como convivência dos indivíduos com seus familiares, cotidiano da vida escolar, condições de moradia e características socioeconômicas dos indivíduos” (FUNDAÇÃO EPROCAD, 2016, p. 10).

Para a construção do instrumento de coleta de dados foi considerado um conhecimento apropriado sobre os interessados do processo avaliativo, considerando sua cultura e seus valores. Por meio desta explicitação, considera-se que o padrão foi *atendido*.

Atendimento ao padrão comunicação e relatório

Yarbrough et al. (2011) afirmam que o registro dos procedimentos, para que seja construída uma comunicação precisa nas avaliações, deve ter escopo adequado, evitando concepções errôneas, vieses e distorções intencionais e não intencionais. Seguindo essa visão, a Fundação Itaú Social (FUNDAÇÃO EPROCAD, 2016, p. 8) esclarece que:

Tais atributos são de difícil mensuração. Buscou-se, portanto, compatibilizá-los com medidas validadas na literatura especializada. Foram utilizadas cinco medidas: i. Medida de sociabilidade: Questionário de Capacidades e Dificuldades (SDQ); ii. Medida de impulsividade: Escala de Impulsividade de domínio específico de Angela Duckworth; iii. Medida de Grit ("garra"): Escala de Angela Duckworth; iv. Medida de Lócus de Controle: Escala Rotter; v. Medida de Autoestima: Escala Rosenberg de Autoestima; vi. Medida de Honestidade: Escala HEXACO-PI-R de honestidade.

De acordo com as evidências, é possível perceber que, na apresentação das medidas utilizadas no relatório de avaliação foram redigidos cinco medidas. Entretanto, foram listadas seis. Embora esta falha não comprometa o processo avaliativo e os resultados, pelo fato de não apresentar uma comunicação precisa, conforme orientam os padrões do *Joint Committee*, considera-se que o Relatório de Avaliação Econômica *atendeu parcialmente* o que recomenda o padrão.

Conforme mostrado no Quadro 3, considerando a análise das evidências e comentários, percebe-se que dos oito padrões de utilidade, cinco foram *plenamente atendidos* no Relatório avaliado, dois foram *atendidos de forma parcial* e apenas o padrão Preocupação com Influências e Consequências *não foi atendido* mediante a suas recomendações. Não há evidências no Relatório deste padrão, que está diretamente relacionado à atenção aos interessados e sua participação no processo avaliativo.

Conclui-se neste sentido que, apesar de dois padrões não terem alcançado a totalidade em seu atendimento, o estudo avaliativo, por meio de seu relatório, *atende à categoria Utilidade*, evidenciando que a metodologia aplicada na avaliação de retorno econômico do programa EPROCAD necessita ser revista no aspecto relacionado à Preocupação com Influências e Consequências.

Ainda no Quadro 3, considerando a análise das evidências e comentários, percebe-se que dos oito padrões de precisão, cinco foram *atendidos* no Relatório avaliado e três padrões foram *atendidos de forma parcial*. Logo, conclui-se que apesar de três padrões não terem alcançado a totalidade em seu atendimento, o estudo meta-avaliativo *atende à categoria Precisão*, contribuindo para a avaliação de projetos sociais.

Quadro 3 – Nível de atendimento dos padrões de utilidade e precisão

Padrões	Atendimento
U1 - Credibilidade do avaliador	A
U2 - Atenção aos interessados	AP
U3 - Propósitos negociados	A
U4 - Explicitação de valores	A
U5 - Informação relevante	AP
U6 - Produtos e processos significativos	A
U7 - Comunicação e relatórios apropriados e no prazo	A
U8 - Preocupação com consequências e influências	N
P1 - Conclusões e decisões justificadas	A
P2 - Informação válida	AP
P3 - Informação fidedigna	AP
P4 - Explicitação das descrições do programa e do contexto	A
P5 - Gerenciamento da informação	A

P6 - Análises e delineamentos (<i>designs</i>) sólidos	A
P7 - Explicitação do raciocínio da avaliação	A
P8 - Comunicação e relatório	AP

Legenda: A - Atende. N - Não Atende.

AP - Atende Parcialmente. NA - Não se Aplica.

Fonte: Os autores (2019).

Considerações finais

O presente estudo objetivou meta-avaliar o Relatório de Avaliação Econômica de um programa social sob o enfoque dos padrões de Utilidade e Precisão do *Joint Committee on Standards for Educational Evaluation*, procurando destacar o mérito e relevância da avaliação original, sem a pretensão de desconsiderar a importância das demais categorias não abordadas neste estudo.

Justifica-se a escolha dos autores pelos atributos Utilidade e Precisão pelo fato de terem uma conexão especialmente forte entre si, compartilhando diversos tópicos, tais como a interação entre o avaliador e interessados, colaboração entre avaliadores e outros *stakeholders*. Não obstante, os padrões de Utilidade enfocam, em primeiro lugar, as qualidades que preparam os *stakeholders* para utilizar os processos, descrições, conclusões, julgamentos e recomendações da melhor maneira que atenda às suas necessidades. Os padrões de Precisão enfocam as informações e a qualidade dos procedimentos empregados para sua seleção, coleta, análise, divulgação e interpretação (YARBROUGH et al., 2011).

Os padrões de Utilidade apresentados neste estudo contribuem para fundamentar uma boa utilização da avaliação. O objetivo desses padrões é aumentar as chances de a avaliação ter consequências positivas e influência significativa, toda vez que surgirem oportunidades e necessidades no seu curso.

Os padrões de Precisão abordados nesta meta-avaliação orientam como diminuir três fatores capazes de comprometer a precisão das avaliações: (a) inconsistências que impedem as informações serem fidedignas; (b) distorções causadas por negligências como omissões, desatenções, incompetência, parcialidade ou conduta imprópria; e (c) concepções equivocadas que não se baseiam em evidências e informações, e podem resultar de ideias preexistentes, apelos emocionais ou tentativas de persuasão.

Difícilmente uma avaliação consegue eliminar todas as fontes ocasionais ou sistemáticas de imprecisão. Entretanto o objetivo principal dos padrões de Precisão é minimizar, na medida do possível, a ocorrência e o impacto de tais fatores.

A meta-avaliação não é considerada uma atividade trivial e sugere que seja efetuado um processo crítico e com pleno conhecimento para sua realização. Não obstante, uma meta-avaliação possibilita aos interessados identificarem os pontos fortes e limitações do estudo avaliativo de programas e serviços, o que caracteriza a sua função somativa.

Como pontos fortes do relatório de avaliação, salienta-se a intensificação da importância das atividades esportivas na educação como método de integração, o alinhamento dos valores éticos e morais na prática desportiva e o incentivo à maior participação nas atividades escolares das crianças e adolescentes envolvidos no projeto.

Duas limitações foram observadas no processo avaliativo apresentado no Relatório: (a) não foram identificadas evidências de que os avaliadores condicionaram às partes interessadas a oportunidade de testar suas suposições e de ajustar seus entendimentos correspondentes ao objetivo e às consequências dos resultados do processo avaliativo; e (b) a falta de evidências correspondente a

recomendações direcionadas aos interessados após a obtenção dos resultados da avaliação.

Como recomendação é importante considerar reflexões sobre: (a) a importância dos esforços por parte dos avaliadores para envolver o maior número de interessados durante a fase de planejamento até a conclusão, uma vez que estes são de alguma forma afetados, além de serem potenciais usuários dos resultados da avaliação e, (b) o empenho dos avaliadores em mencionar as devidas recomendações aos interessados após a obtenção dos resultados do processo avaliativo. Tais reflexões podem elucidar que os objetivos do projeto da EPROCAD não sejam atingidos em sua plenitude.

Com base na análise crítica do Relatório, conclui-se que a adoção dos padrões estabelecidos pelo *Joint Committee* é capaz de promover uma reflexão sobre boas práticas de avaliação de programas e projetos sociais existente no Brasil por intermédio de investimento social privado. Assim, verificou-se que a meta-avaliação foi positiva, tendo em vista que 10, dos 16 padrões, foram considerados como aspectos úteis e aplicáveis para o processo avaliativo, confirmando a relevância do relatório. No entanto, foi possível identificar algumas limitações que podem ser objeto de reflexões mais aprofundadas em avaliações futuras.

Referências

BERENDS, Lynda; ROBERTS, Bridget. Evaluation standards and their application to indigenous programs in Victoria, Australia. *Evaluation Journal of Australasia*, Canberra, v. 3, n. 2, p. 54-59, 2003.

ELLIOT, Lígia Gomes. Meta-avaliação: das abordagens às possibilidades de aplicação. *Ensaio: Avaliação e Políticas Públicas em Educação*, Rio de Janeiro, v. 19, n. 73, p. 941-964, out./dez. 2011.

EHRHARDT, Giovani. *Avaliar a avaliação: percurso para a meta-avaliação na educação superior em países latino-americanos: os casos da Argentina, Brasil, Chile e México*. Perafita: TDP Ltda. Edições, 2012.

FUNDAÇÃO EPROCAD. *Relatório de avaliação econômica: projeto futebol: uma nova visão do jogo*. São Paulo: Fundação Itaú Social, 2016. Disponível em: <https://www.redeitausocialdeavaliacao.org.br/wp-content/uploads/2016/03/Relat_EPROCAD_20160317.pdf>. Acesso em: 25 out. 2018.

FUNDAÇÃO ITAÚ SOCIAL; FUNDAÇÃO ROBERTO MARINHO; MOVE (Org.). *Avaliação do investimento social privado: estratégia organizacional*. São Paulo: Fundação Santillana, 2015. Disponível em: <<https://fundacao-itausocial-producao.s3.amazonaws.com/files/s3fs-public/biblioteca/documentos/publicacao-estrategia-organizacional-062015.pdf>>. Acesso em: 25 out. 2018.

GIMENES, Nelson Antônio Simão. Estudo meta-avaliativo do processo de autoavaliação em uma instituição de educação superior no Brasil. *Estudos em Avaliação Educacional*, São Paulo, v. 18, n. 37, maio/ago. 2007.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. *The program evaluation standards: a guide for evaluators and evaluations users*. Los Angeles, CA: Sage, 2011a. Parte I: Padrões de Utilidade. Tradução Neuza Capelo; revisão técnica: Lígia Gomes Elliot.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. *The program evaluation standards: a guide for evaluators and evaluations users*. Los Angeles, CA: Sage, 2011b. Parte IV: Padrões de Precisão. Tradução: Neuza Capelo; revisão técnica: Christina Marília Teixeira da Silva.

JOINT COMMITTEE ON STANDARDS FOR EDUCATIONAL EVALUATION. Program evaluation standards statements. *Pages*, 2019. Disponível em: <<http://www.jcsee.org/programevaluation-standards/program-evaluation-standards-statements>>. Acesso em: 27 out. 2018.

LEAL, Paulo Henrique. *Processos educativos construídos com participantes do projeto vivências em atividades diversificadas de lazer*. 2015. 196 f. Tese (Doutorado em Educação)-Universidade Federal de São Carlos, São Carlos, 2015.

LEITE, Lígia S. Lista de verificação. In: ELLIOT, Lígia G. (Org.). *Instrumentos de avaliação e pesquisa: caminhos para construção e validação*. Rio de Janeiro: WAK Editora, 2012. p. 101-147.

MENEZES FILHO, Naercio Aquino; PINTO, Cristine Campos de Xavier (Org.). *Avaliação econômica de projetos sociais*. 3. ed. São Paulo: Fundação Itaú Social, 2017. Disponível em: <https://www.itausocial.org.br/wp-content/uploads/2018/05/avaliacao-economica-3a-ed_1513188151.pdf>. Acesso em: 2 nov. 2018.

SANTOS, Hylisson Mataruna dos. *A metodologia de avaliação econômica de projetos sociais da Fundação Itaú Social: um estudo meta-avaliativo*. 2017. 118 f. Dissertação (Mestrado Profissional em Avaliação)-Faculdade Cesgranrio, Rio de Janeiro, 2017.

SCRIVEN, Michael. *Evaluation thesaurus*. 4. ed. Newbury Park, CA: Sage, 1991. Disponível em: <http://books.google.co.nz/books/about/Evaluation_Thesaurus.html?id=koL0Fs_ZSvQC>. Acesso em: 20 out. 2018.

STUFFLEBEAM, Daniel L.; SHINKFIELD, Anthony J. *Evaluación sistemática: guía teórica y práctica*. Madrid: Paidós, 1987.

TAVARES, Marialva. *Análise de projetos sociais: caminho para melhorar o ensino*. Rio de Janeiro: Fundação Carlos Chagas, 2009. Disponível em: <http://www.fcc.org.br/conteudos especiais/difusaoideias/pdf/entrevista_marialva.pdf>. Acesso em: 23 out. 2018.

WORTHEN, Blaine R.; SANDERS, James R.; FITZPATRICK, Jody L. *Avaliação de programas: concepções e práticas*. São Paulo: Ed. Gente, 2004.

YARBROUGH, Donald B. et al. *The program evaluation standards: a guide for evaluators and evaluations users*. 3. ed. Thousand Oaks, Califórnia: Sage, 2011.