

# Guia para Avaliação de Impacto Socioambiental para Utilização em Investimentos de Impacto

Guia Geral com Foco em  
Verificação de Adicionalidade



São Paulo  
Terceira revisão, janeiro de 2017

## Sobre o Insper Metricis

Núcleo de Medição para Investimentos de Impacto Socioambiental do Insper, realiza estudos sobre investimentos e estratégias empresariais de alto impacto socioambiental. Ênfase especial é dada ao desenvolvimento de ferramentas para avaliação e remuneração por impacto, bem como para a validação de medidas realizadas por organizações e governos.

Com procedimentos de avaliação claros e objetivos, validados por uma organização independente, é possível estimular não somente o crescimento dos investimentos e negócios de impacto, como também impulsionar novas formas de captação e financiamento destes projetos.

Além disso, com o aprendizado gerado pelos projetos, espera-se um contínuo registro e aprimoramento de melhores práticas por meio de casos e pesquisas ligadas aos projetos de alto impacto.

## Sobre o Insper

O Insper é uma instituição independente e sem fins lucrativos, dedicada ao ensino e à pesquisa nas áreas de Administração, Economia, Direito e Engenharia. Tem como missão ser um centro de referência, explorando as complementariedades dessas áreas.

Suas atividades de ensino abrangem cursos para várias etapas de uma trajetória profissional: graduação (Administração, Economia e Engenharias), pós-graduação lato e stricto sensu (Certificates, MBAs, programas da área de Direito, Mestrados Profissionais e Doutorado) e Educação Executiva (programas de curta e média duração, e customizados de acordo com as necessidades das empresas).

No âmbito da produção de conhecimento, a instituição atua por meio cátedras e centros de pesquisa que reúnem pesquisadores em estudos e projetos dirigidos a políticas públicas (CPP), finanças (CeFi) e negócios (CENeg). A instituição ainda conta com centros que promovem o empreendedorismo (CEMP) e a liderança e inovação (CLI). Tem as certificações de qualidade da AACSB, AMBA e Anamba. O Insper Metricis é um núcleo do Centro de Estudos em Negócios (CENeg).

# Guia para Avaliação de Impacto Socioambiental para Utilização em Investimentos de Impacto

Documento originalmente elaborado em 2013 por Sérgio G. Lazzarini, Leandro S. Pongeluppe, Pui Shen Yoong e Nobuiuki Costa Ito a partir de contribuições do grupo técnico de discussão envolvendo as seguintes pessoas: Angélica Rontondaro, Célia Cruz, Franco Veludo, Frederik Kuonen, Simon Locher, Sandro Cabral, Tatiana Fonseca, Rafaella Ziegert e Raquel Costa. Agradecemos o apoio financeiro do Latin America Impact Economy Innovations Fund (Rockefeller Foundation, Avina e Omidyar), em proposta coordenada pelo ICE (Célia Cruz e Maria Amélia Sampaio). Agradecemos também as sugestões de Naercio Menezes Filho, Luis Fernando Guedes Pinto e Maurício Voivodic, além dos comentários detalhados de Fernando Carnaúba.



## INTRODUÇÃO

Investimentos de impacto pretendem gerar resultado socioambiental positivo além de graus variados de retorno financeiro<sup>1</sup>. Embora existam métodos claros e estabelecidos para mensurar o retorno financeiro dos investimentos, ainda há muita discussão sobre como medir o impacto de organizações e projetos com foco social ou ambiental. Uma possível abordagem envolve a análise de vários indicadores que seriam, em tese, aplicáveis a múltiplos projetos em áreas distintas. Como exemplo, há os padrões estabelecidos pelo *Impact Reporting and Investment Standards* (IRIS) e o procedimento de avaliação estabelecido pelo *Global Impact Investing Report System* (GIIRS). Apesar do uso e disseminação dessas ferramentas, em muitos casos inves-

tidores, empreendedores, governos e demais atores do ecossistema demandam abordagens de medição mais adaptadas às especificidades da sua atividade<sup>2</sup>.

Uma abordagem alternativa, que tem recebido crescente atenção, é a utilização de *poucos* indicadores *específicos* por projeto. Em tese, esses poucos indicadores deveriam representar aspectos que garantam um elevado impacto para a atividade-alvo. Por exemplo, estudos indicam que o desempenho de alunos do ensino médio e fundamental em testes padronizados têm uma elevada relação com o ganho de renda deles no futuro. Logo, o grau com que o investimento de impacto aumenta o desempenho dos alunos nesses testes padronizados ten-

1 J.P.MORGAN. Impact Investments: An emerging asset class. Report, 2010, p.7.

2 IZZO, D. Aligning Interests in Impact Investing. Stanford Social Innovation Review, p. 1-3, 2013.

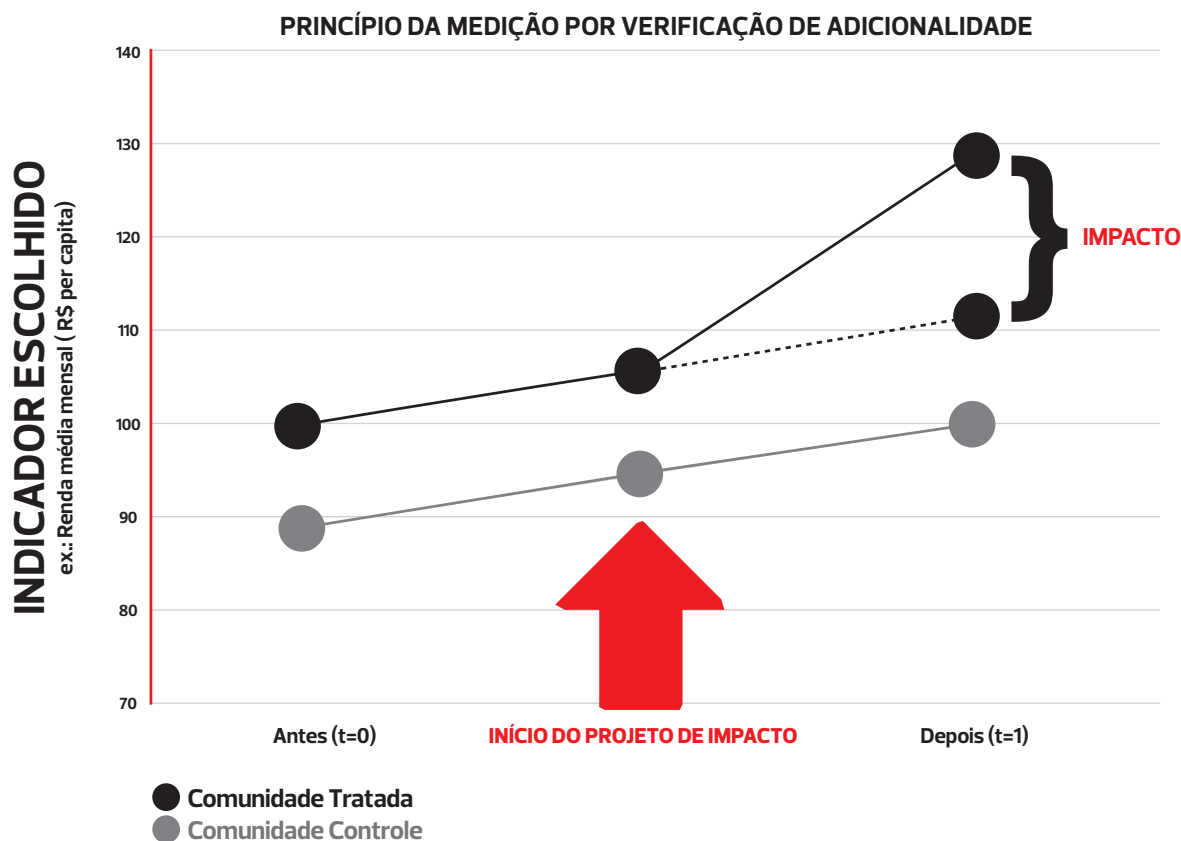
de a ser uma dimensão de alta relevância para a avaliação do impacto.

Além disso, nesse tipo de abordagem, busca-se medir não apenas o indicador em si, mas também responder a seguinte pergunta: o que teria acontecido com os indivíduos ou comunidades-alvo caso *não* tivessem sido beneficiados com o projeto? Ainda usando o exemplo em educação, a ideia é sempre comparar o grupo de alunos que recebeu o investimento com um grupo de alunos similares que não foram contemplados pelo programa educacional. Isso é importante porque, de forma simultânea ao projeto, podem ter ocorrido melhorias gerais no sistema educacional do país independente do investimento realizado.

Essa abordagem é alinhada com a chamada verificação de *adicionalidade* (*additionality*) em projetos sociais<sup>3</sup>. A ideia é similar ao desenho de um experimento. Por exemplo, no campo da medicina, elege-se um grupo de indivíduos como "grupo tratado", ou seja, aqueles que receberão o tratamento médico ou o medicamento; e o "grupo controle", que não rece-

be o referido tratamento. Dessa forma, é possível medir como o tratamento afetou a população além do que poderia ter acontecido naturalmente sem o tratamento efetuado. Esse cenário alternativo – o que teria acontecido sem o tratamento – é chamado de *contrafactual*. Em métodos de verificação mais rigorosos, como a chamada técnica de *Randomized Controlled Trial* (RCT), os grupos tratados são até mesmo escolhidos de forma aleatória, a fim de controlar o impacto de fatores externos, aumentando a confiabilidade das medidas.

A Figura 1 exemplifica essa forma de medição, baseada em adicionalidade. Imagine uma empresa que queira investir no desenvolvimento de determinadas comunidades e avaliar o impacto desses investimentos. Para tanto, esta empresa realiza medições do nível de renda dos indivíduos dos grupos de controle e tratado antes de iniciar o projeto (no momento  $t = 0$ ). Neste momento, observou-se que o nível médio de renda na comunidade de controle era de R\$ 90 e na comunidade a ser tratada (que receberá o investimento), de R\$ 100.



<sup>3</sup> BREST, B. P.; BORN, K. When Can Impact Investing Create Real Impact? Stanford Social Innovation Review. 2013, p. 24.



Após a medição, a empresa então inicia o projeto com a comunidade tratada. Após um ano do início do projeto (no momento  $t = 1$ ), a empresa realiza novamente a medição do nível de renda individual nas duas comunidades, tratada e de controle. Um analista inexperiente avaliaria que o impacto do projeto é de 30%, dada a elevação da renda de R\$ 100 para R\$ 130. Contudo, com a utilização da metodologia por verificação de adicionalidade, compara-se a evolução do grupo tratado frente ao grupo controle. Observa-se que na comunidade que não recebeu o investimento houve um aumento de renda de 10%, por outras causas que vinham afetando a localidade (como, por exemplo, melhorias naturais nas condições de vida ou programas de distribuição de renda do governo).

Dessa forma, ao considerarmos não apenas a evolução do grupo tratado no tempo, mas também a tendência natural que este grupo seguiria (evidenciada pelo grupo de controle), é possível medir o real impacto do programa. No presente caso, percebe-se que o impacto gerado pelo programa elevou em média 20% a renda dos indivíduos na comunidade. Esse valor é simplesmente a diferença entre a evolução verificada na comunidade tratada (30%) e a evolução verificada na comunidade de controle (10%).

Tomando como base essa metodologia, esse guia propõe um princípio de medição baseado em três pilares essenciais:

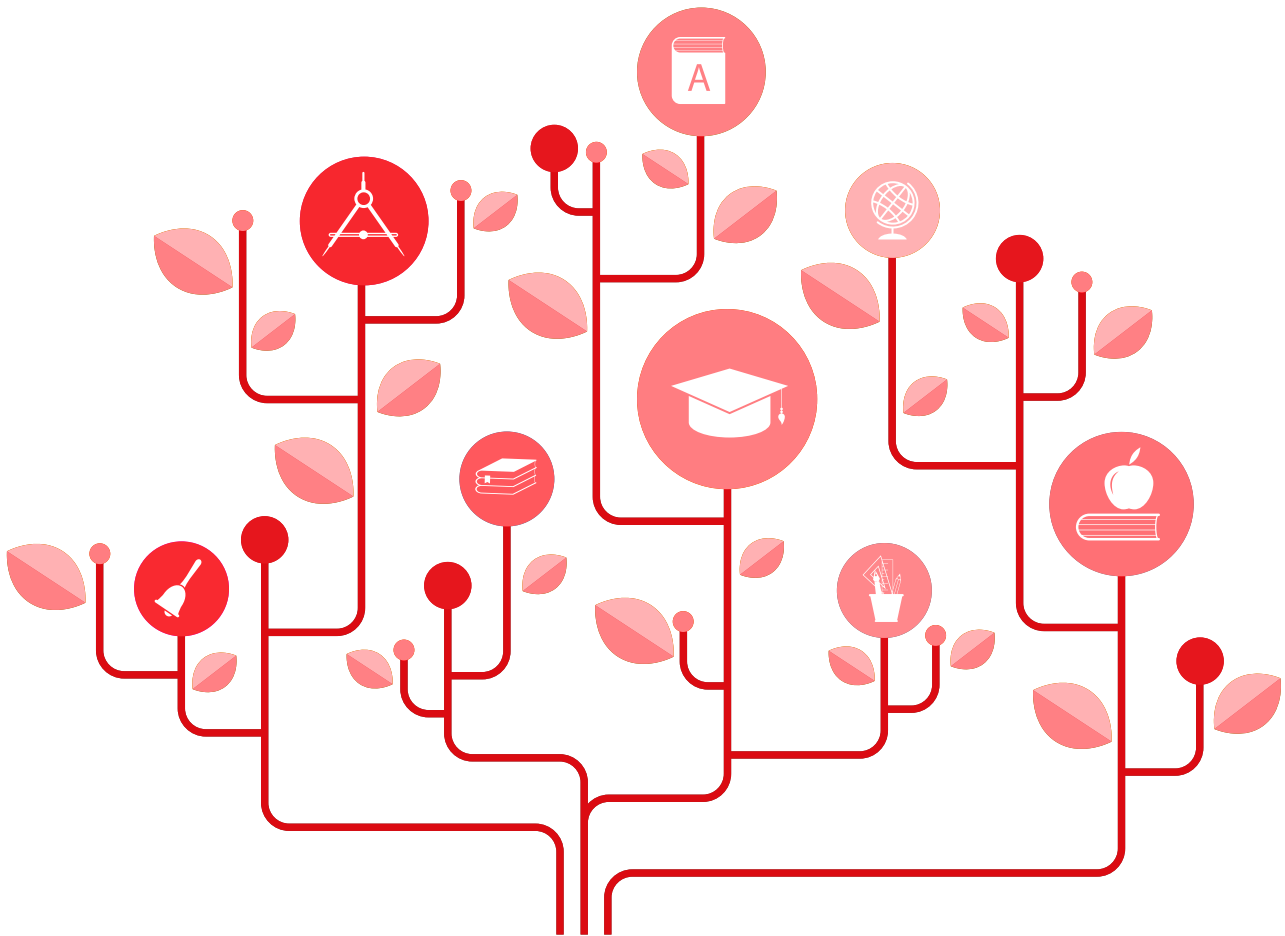
**1 Poucos indicadores com elevada relevância para aferir o impacto do projeto, isto é, as mudanças positivas esperadas nas populações-alvo;**

**2 Avaliação com o que poderia ter acontecido a essas populações sem o investimento (por exemplo, comparação com grupos de controle similares que não foram contemplados pelo projeto).**

**3 Sempre que possível, medições antes e depois do início do projeto, para comparar a evolução relativa de grupos afetados e não afetados pelo investimento.**

Dessa forma, o *Guia para Avaliação de Impacto Socioambiental para Utilização em Investimentos de Impacto* tem a finalidade de fornecer um ferramental prático que possibilite a fundos, empresas e governos avaliarem o impacto que causam nos projetos em que investem, atraindo e alavancando, dessa forma, investidores interessados em impacto socioambiental. Com métricas mais claras e estritamente ligadas aos projetos, validadas por uma terceira parte independente, deve-se facilitar o estabelecimento de contratos entre investidores de impacto e os executores dos projetos.

Fundamentado com uma base quantitativa, o presente guia estabelece um conjunto de normas metodológicas que podem ser seguidas para uma avaliação efetiva dos investimentos de impacto. Na sequência, a seção II trata do desenho do plano de medição e da coleta de dados. A seção III expõe os níveis de avaliação considerados para a medição e a seção IV, os principais cuidados a serem tomados ao se utilizar a presente técnica de medição. A seção V apresenta o fluxograma com o processo de validação do plano de medição estabelecido. Por fim, na seção VI está presente a ficha padronizada a ser entregue à terceira parte, independente, para validação do referido plano de medição.



# ELABORAÇÃO DO PLANO DE MEDIÇÃO

O plano de medição é essencial para planejar, operacionalizar, executar e controlar a medição. Com a finalidade de implementar uma medição robusta e acurada, é importante seguir as seguintes etapas:

## ETAPA 1: Delimitação do setor de impacto

O primeiro passo na definição de um plano de medição é a delimitação do setor de atuação do projeto de impacto. É necessário que o investidor tenha clareza sobre o setor em que ele irá aplicar seus recursos com a finalidade de gerar impacto além de retorno financeiro. Por exemplo, considere um projeto que tenha como objetivo impactar positivamente o acesso dos cidadãos a medicamentos e formas de tratamento efetivas de doenças crônicas a um menor custo. Define-se então como alvo do projeto

o setor de saúde com uma ênfase em tratamento médico para doenças crônicas.

Além disso, nesta etapa 1 é importante realizar a delimitação da região ou comunidade-alvo dos investimentos de impacto. Tal definição afetará o tamanho amostral necessário para a realização da medição. Seguindo com nosso exemplo do projeto relativo ao setor de saúde, é possível especificar, por exemplo, que o investimento terá foco em postos de saúde e hospitais em bairros selecionados das Zonas Norte e Leste de São Paulo.

## ETAPA 2: Benchmarking

A expansão dos investimentos de impacto depende de um esforço de aprendizado contínuo e consolidação de erros e acertos de projetos similares, bem como do uso de conhecimento prévio de medição gerado por pesquisas e estudos. Dessa forma, a análise de casos

de sucesso prévios no setor de impacto (*benchmarking*) é essencial: antes de se definir indicadores, é preciso levantar o que já foi feito ou estudado na área do projeto, sobretudo quais os indicadores foram previamente utilizados em projetos semelhantes.

Fontes de informação incluem: estudos acadêmicos, publicações de organizações internacionais, relatórios de outros investidores de impacto que realizaram aplicações no mesmo setor, dentre outros. Em estudos de impacto, tem sido também cada vez mais comum realizar "meta análises": estudos que tentam agregar os resultados de diversas avaliações prévias sobre um determinado tema.

### ETAPA 3: Definição da teoria da mudança

Antes de aprofundar a questão de como medir, primeiramente é necessário definir *o que* medir. Nessa etapa, o projeto/programa/negócio a ser avaliado deve apresentar sua *teoria da mudança*: uma forma clara e lógica de articular a conexão entre as atividades realizadas e o objetivo socioambiental pretendido. A Figura 2 a seguir apresenta as quatro etapas de uma teoria da mudança, cujo mapeamento é essencial em um plano de medição: atividades, que geram certos resultados imediatos (*outputs*); que, por sua vez, provocariam mudanças (*outcomes*); e, finalmente, geram impactos socioambientais.

Para ilustrar melhor essa ferramenta, tomamos como exemplo uma empresa fictícia cuja *atividade* seja criar jogos educacionais para melhorar os resultados acadêmicos de alunos da rede pública de

ensino, utilizando *insumos (inputs)* diversos como os jogos em si, equipes de suporte, etc.

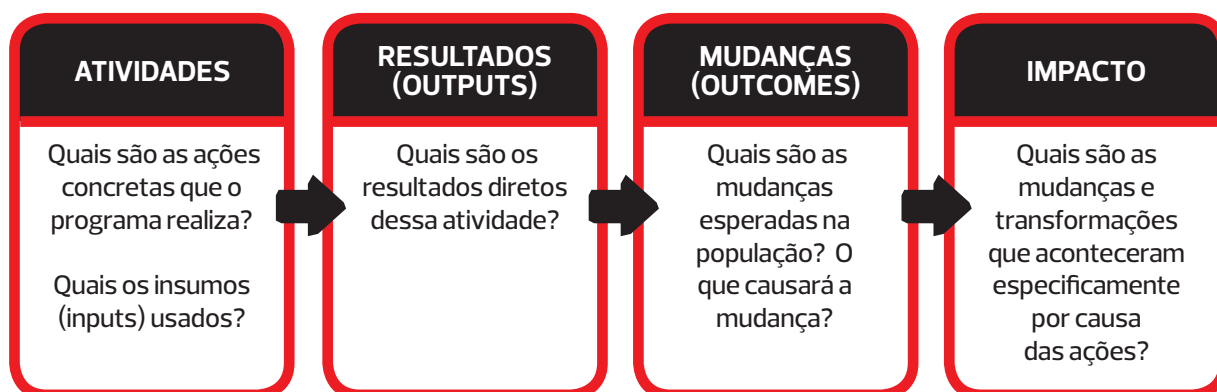
O *resultado (output)* dessa atividade seria o número de alunos/usuários dos jogos, e a *mudança esperada (outcome)* seria a melhoria da nota do aluno em uma prova padronizada. Logo em seguida, o *impacto social* da empresa seria a melhoria do aprendizado dos alunos em comparação ao que teria acontecido sem a ação da empresa (conforme indicado na seção I); e as habilidades adicionais que essas ações geram podem potencialmente aumentar o bem-estar e renda futura dos alunos. Nesse processo, é importante descrever o *mecanismo causal* que irá determinar a mudança. Por exemplo, os jogos podem contribuir para o desenvolvimento cognitivo dos alunos e aprimorar a sua capacidade de compreensão e retenção de conhecimento. Feito esse mapeamento, à luz da etapa 2, será mais fácil definir indicadores relevantes e claros para o impacto do projeto avaliado.

### ETAPA 4: Definição das Métricas (Indicadores de Impacto)

Considerando os resultados das etapas 2 e 3, devem ser propostos indicadores para o projeto. Como forma de preservar foco e facilidade de compreensão da análise, recomenda-se não mais que *três métricas* de elevada relevância para avaliar o impacto gerado. Métricas ideais apresentam as seguintes características<sup>4</sup>:

- **Têm elevada relevância em termos de geração de impacto. Isto é, representam resultados altamente desejáveis para a população-alvo.**

#### TEORIA DE MUDANÇA APLICADA A INVESTIMENTOS DE IMPACTO



Fonte: Yoong, P. S., com base em Hehenberger; Harling; Scholten (2013).<sup>5</sup>

<sup>4</sup> ROBERTS, J. Designing incentives in organizations. *Journal of Institutional Economics*, Cambridge University Press, 6(1): 125–132, 2010.

<sup>5</sup> HEHENBERGER, L.; HARLING, A.M.; SCHOLTEN, P. *A Practical Guide to Measuring and Managing Impact*, abril 2013, p.9.

- **São passíveis de serem afetados pela ação dos executores do projeto. Não podem, por exemplo, ser indicadores muito agregados cujo resultado depende de muitos fatores sem controle direto dos executores.**

- **Devem ser medidos e verificados com relativa precisão. Em geral, indicadores baseados em dados objetivos são preferíveis. Da mesma forma, os dados não podem ser passíveis de manipulação por gestores, apoiadores e outras partes interessadas em uma boa avaliação do projeto.**

- **Devem ter um baixo custo de medição – considerando, inclusive, o custo de obter dados em populações utilizadas como comparação.**

Como sugestão, ao se considerar diferentes medidas alternativas, pode-se avaliar cada métrica com base nos critérios acima definidos e escolher aquelas que mais atendem à maior parte desses critérios. Há, certamente, casos em que simplesmente não é possível medir com precisão e baixo custo o resultado almejado.

Não se propõe, aqui, que todo tipo de impacto possa ser facilmente medido com indicadores objetivos. Contudo, em vários casos, é possível propor métricas razoavelmente ajustadas ao impacto esperado, considerando outras experiências de medição já realizadas. Com o acúmulo de projetos e experiências de medição, espera-se que ocorra um aprendizado contínuo sobre como medir e avaliar o impacto socioambiental de um grande número de projetos em setores diversos.

### ETAPA 5: Definição do Método de Verificação de Adicionalidade e do Nível de Medição

Após a definição dos indicadores a serem verificados, é importante definir os grupos *tratados* (indivíduos e comunidades beneficiadas pelo projeto) e os grupos de *controle* (indivíduos e comunidades semelhantes às tratadas, mas que não foram selecionadas para o projeto).

A definição destes dois grupos é de extrema importância para a medição com base em adicionalidade (ver seção I). Esta definição irá variar de acordo com o *nível de medição* definido para o projeto, conforme detalhado na seção III. Além disso, para a definição dos grupos de tratados e de controle, algumas

medidas cautelares devem ser seguidas. A seção IV lista diversos cuidados a serem tomados para a medição. Uma vez definidos os grupos tratados e de controle, e com base nos indicadores escolhidos, é importante definir com precisão como será calculado o impacto. No exemplo da Figura 1, o impacto será medido como a diferença de variação de renda entre a comunidade tratada e a de controle, considerando os dois períodos definidos.

### ETAPA 6: Plano Amostral

A definição do tamanho amostral – quantas pessoas ou grupos são impactados pelo investimento – é outra crucial etapa do plano de medição. Com um adequado tamanho amostral, será possível verificar com maior precisão estatística a diferença entre o resultado medido antes e depois do projeto, bem como a diferença entre os grupos tratados e de controle. Em outras palavras, o tamanho da amostra ajuda a definir a confiabilidade da mensuração<sup>6</sup>.

### ETAPA 7: Execução da Medição

Sempre que possível, de forma a reduzir os custos com a coleta de dados, recomenda-se considerar bases de dados confiáveis e publicamente disponíveis, isso é, dados secundários. Havendo necessidade de coletar dados originais (primários), é imprescindível contar com o apoio de uma equipe treinada e especializada em coleta de dados. É importante que no plano de medição, o executor do projeto evidencie qual o indicador de performance, quem irá executar a medição, seu orçamento e o contato principal. Ao se buscar dados primários, entretanto, é importante lembrar que os dados devem ser coletados não apenas nas populações afetadas pelo projeto mas também nos grupos de controle definidos.

### ETAPA 8: Cronograma de Medição

Por fim, o executor do projeto e o investidor devem ter consciência de que, na maioria dos casos, a verificação de impacto ocorrerá após o período de, no mínimo, um ano desde o início do projeto. Algumas áreas de investimento requerem um horizonte ainda maior para possibilitar a real verificação do impacto. Por exemplo, no setor de educação, onde o aprendizado dos alunos requer um esforço longo e persistente, o horizonte de tempo tende a ser de dois ou três anos, pelo menos, e em diversos casos até mais.

<sup>6</sup> Uma discussão técnica sobre como definir o tamanho da amostra pode ser visto no texto de DUFLO, E.; GLENNERSTER, R.; KREMER, M. Using Randomization in Development Economics Research: a Toolkit. Mimeo, 2013. Disponível no endereço: <http://www.nber.org/papers/t0333> (acesso em 04/10/2013).



# NÍVEIS DE MEDIÇÃO E LIMITAÇÕES

**E**specialistas em avaliação com verificação de adicionalidade enfatizam que, idealmente, os grupos e indivíduos tratados – recebendo o investimento – devem ser preferencialmente escolhidos de forma aleatória<sup>7</sup>. Ao escolher de forma aleatória quem será avaliado é possível corrigir fatores externos decorrentes de diferenças pré-existentes entre os grupos. Em particular, evita-se que sejam escolhidos indivíduos a serem tratados que já, de início, tenham uma maior propensão natural a apresentar melhoria mesmo se não recebessem o tratamento.

Por exemplo, imagine que o investidor de impacto queira avaliar o efeito de um novo método pedagógico

e ofereça esse método a diversas escolas. É possível que as escolas que aceitem aderir ao programa sejam justamente aquelas com gestão mais motivada para melhorias no ensino. Se essa variável afetando a escolha (“motivação”) não for observada pelo avaliador, pode-se concluir erroneamente que possíveis melhorias nas escolas que voluntariamente aderiram ao programa resultaram do programa em si e não da propensão natural a melhorar o ensino.

No entanto, sabe-se que a aleatorização de indivíduos ou grupos é particularmente complexa e em muitos casos até mesmo inexecutável. A decisão pelo investimento em determinados projetos, seja essa tomada por fundos, empresas ou governos, segue pressupostos e planos previamente estabe-

## NÍVEIS (TIERS) DE MEDIÇÃO

**NÍVEL 3**

**Grupo tratado definido por meio de aleatorização (Randomized Controlled Trial, RCT)**

**NÍVEL 2**

**Comparação dos tratados com grupo de controle composto por indivíduos com características similares**

**NÍVEL 1**

**Comparação dos tratados com dados agregados locais ou regionais**

<sup>7</sup> É importante distinguir a definição aleatória dos grupos de tratamento e controle de uma amostra aleatória de uma população de interesse. Aqui estamos interessados no primeiro caso, isso é, quem recebe o tratamento é escolhido de forma aleatória.

lecidos. Embora isso inviabilize a aleatorização plena da amostra, conforme discutido anteriormente, para fins de análise de impacto pode-se tentar mimetizar um experimento por meio da escolha de grupos mais próximos possíveis aos selecionados<sup>8</sup>.

Devido a essas dificuldades, propõe-se que, qualquer que seja a forma de verificação adotada, seja dada total *transparência* sobre suas possíveis limitações. Especificamente, abaixo são propostos três *níveis* (*tiers*) de medição, de acordo com o nível de robustez pretendido para a avaliação de impacto, em especial no que tange ao rigor sobre a aferição do efeito *causal* do projeto de impacto. Na sequência, os níveis considerados são explicados em detalhe:

### NÍVEL 1 – Mensuração comparando o grupo tratado com dados agregados locais ou regionais.

Nesse nível, não é utilizada aleatorização e empregam-se dados agregados como forma de comparação. Tal situação ocorre quando, antes da medição, há uma definição deliberada da localidade onde o projeto será realizado e de quem são os indivíduos beneficiados. Além disso, para fins comparativos, são utilizados dados agregados já disponíveis para uma determinada região onde é feito o investimento. Devem ser buscados dados do que ocorreu com o projeto versus o agregado antes e depois da intervenção.

**Exemplo:** Em um projeto de impacto cujo indicador é a renda da comunidade, a avaliação é efetuada comparando-se a renda das famílias/indivíduos tratadas com dados sobre a renda de famílias/indivíduos no município ou microrregião onde foi feito o investimento, antes e depois da intervenção. Sendo que a definição das comunidades beneficiadas é feita de forma deliberada pelo executor do projeto (isto é, sem aleatorização).

### NÍVEL 2 – Mensuração utilizando um grupo de controle com características semelhantes às do grupo tratado.

Nesse nível, embora ainda não seja implementada aleatorização, há uma tentativa de criar um grupo de controle com indivíduos ou comunidades compa-

ráveis. Tal situação ocorre quando, antes da medição, há uma definição deliberada da localidade onde o projeto será realizado e de quem são os indivíduos beneficiados. E, posteriormente à coleta dos dados, há implementação de técnicas para construir um grupo de indivíduos de controle que sejam próximos aos indivíduos tratados. Ou seja: em vez de utilizar dados agregados de um determinado local, há uma tentativa de acompanhar grupos ou indivíduos não recebendo o investimento, porém com características similares às dos tratados.

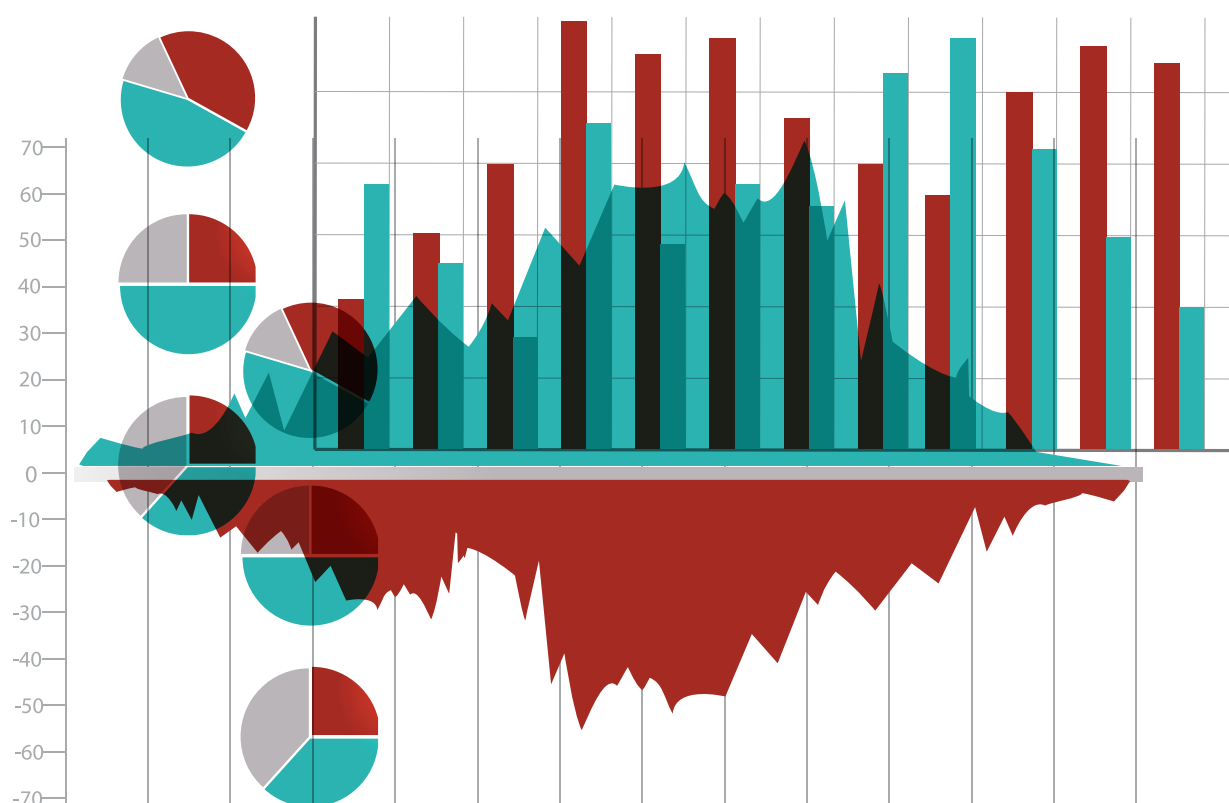
A forma mais usual de escolher indivíduos com características semelhantes entre os grupos de controle/tratados envolve a chamada técnica de *pareamento*. Quando existir competição para participar no projeto e um determinado critério de seleção, uma possibilidade é considerar grupos no *limiar* entre os participantes tratados que quase não foram aceitos e, como grupo de controle aqueles que quase foram aceitos mas, eventualmente, ficaram abaixo do limite de corte<sup>9</sup>. Recomenda-se que a metodologia utilizada para realizar o pareamento seja definida já no plano de medição, indicando claramente como os dados do grupo de controle serão pareados ao grupo de tratados.

**Exemplo (a):** Em um projeto de impacto cujo indicador é a renda da comunidade, a avaliação é efetuada comparando-se a renda das famílias/indivíduos tratados com dados sobre a renda de famílias/indivíduos no grupo de controle. A definição das comunidades beneficiadas pelo projeto é feita de forma deliberada pelo executor. Contudo, neste nível, as famílias/indivíduos de controle são selecionados com base em características próximas às dos tratados. Por exemplo, no grupo de controle podem ser selecionadas famílias com níveis iniciais de renda, escolaridade e número de filhos similares aos níveis observados nas famílias tratadas.

**Exemplo (b):** Neste mesmo projeto, a organização pretende implementar um programa de microcrédito. Pode-se convidar microempreendedores a participar do programa, pré-avaliados com base em indicadores de risco de crédito (renda, histórico em empréstimos anteriores, etc.). Havendo vagas limitadas, apenas os mais

<sup>8</sup> ANGRIST, J. e PISCHKE, J.-S.. *Mostly Harmless Econometrics*. Princeton: Princeton University Press, 2009; MENEZES FILHO, N. A. (editor) *Avaliação econômica de projetos sociais*. 1. Ed. São Paulo: Dinâmica Gráfica e Editora, 2012.

<sup>9</sup> Um exemplo de forma de implementação de técnica de pareamento pode ser vista em: ABADIE, A.; DRUKKER, D.; HERR, J. L.; IMBENS, G. W. *Implementing matching estimators for average treatment effects* *Stata. The Stata Journal*, v. 4, n. 3, p. 290–311, 2004. Para uma discussão mais geral de vários métodos alternativos, consulte ANGRIST, J. e PISCHKE, J.-S., op. cit.



bem avaliados serão selecionados. No limiar da nota de corte, pode-se supor que os indivíduos são similares. Desta forma, pode-se comparar o grupo de microempreendedores selecionados logo acima da linha de corte com aqueles que não foram selecionados, mas que ficaram logo abaixo da linha de corte.

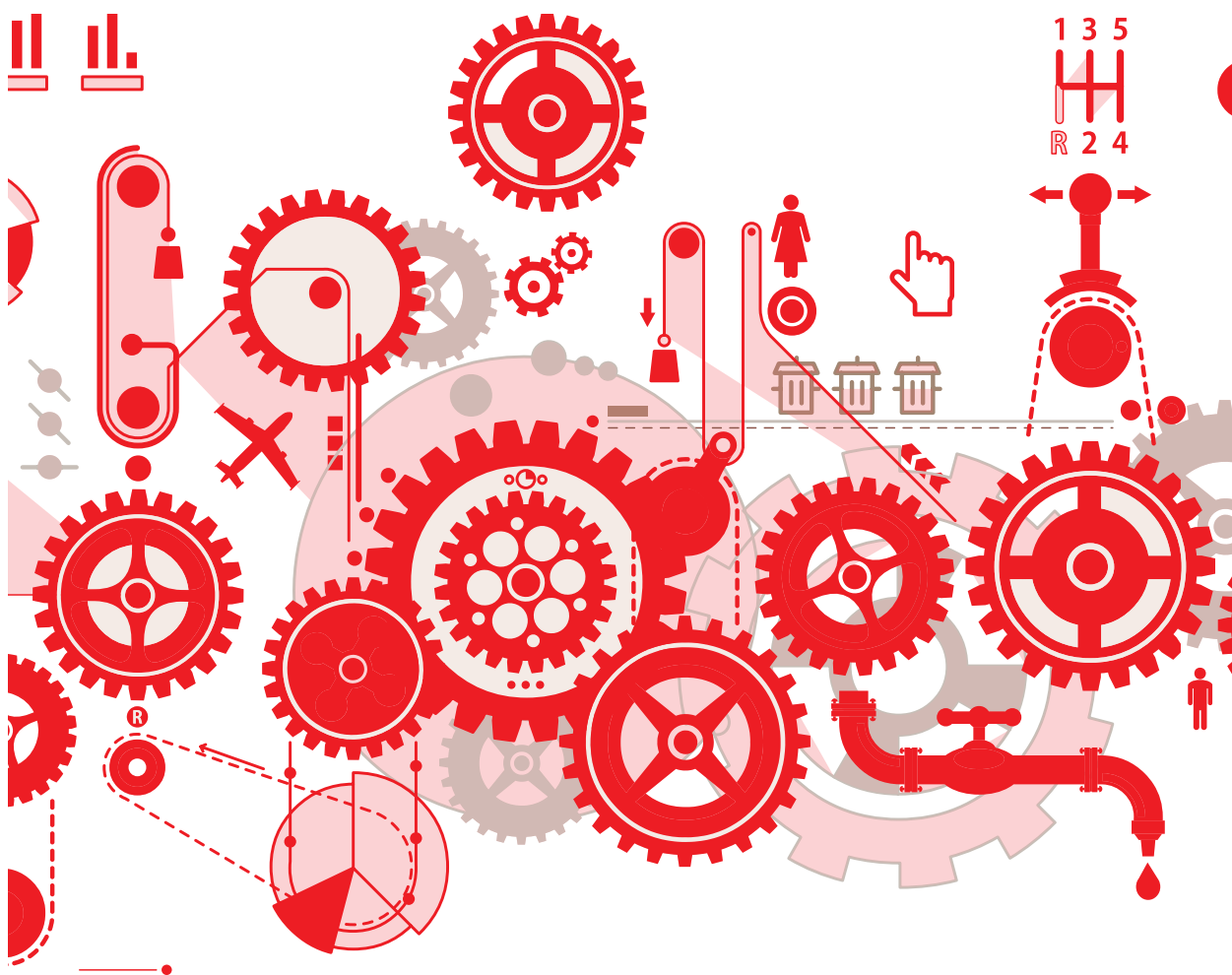
### NÍVEL 3 – Mensuração com aleatorização (RCT).

Esse é o nível de medição que permite máxima confiabilidade da estimativa de avaliação de impacto. Neste nível a escolha dos grupos tratados e de controle é realizada de forma totalmente aleatória, logo reduzindo possíveis vieses de seleção dos grupos-alvo do projeto. A escolha aleatória, se bem feita, garante que não existiam diferenças entre as características, observáveis ou não, dos grupos tratados e de controle.

**Exemplo:** Em um projeto de impacto cujo indicador é a renda da comunidade, avalia-se a evolução da renda das famílias/indivíduos, tanto no grupo de controle quanto no grupo tratado. Os beneficiários do projeto são aleatoriamente escolhidos. Por exemplo, suponha que uma organização ofereça um programa limitado de

microcrédito em um contexto onde não há recursos para apoiar todos da comunidade. A organização pode convidar interessados e definir quem vai receber o microcrédito de forma aleatória. Caso se queira implementar o programa para aqueles com bom desempenho em uma avaliação prévia de risco de crédito, pode-se aleatorizar dentro do grupo daqueles que quase foram selecionados, isto é, logo abaixo da linha de corte. Considerando esse grupo, pode-se definir de forma aleatória um subconjunto de indivíduos que receberão o programa, sendo então o grupo de controle definido por aqueles não escolhidos pelo procedimento de aleatorização.

É importante notar que a aleatorização se dá na escolha do grupo que deve receber ou não a intervenção. Por exemplo, suponha que determinado projeto de educação tenha beneficiado 30 escolas, sendo o processo de escolha das escolas definido de forma não aleatória. Então, o nível 3 não será atingido mesmo se forem selecionadas, aleatoriamente, outras 30 escolas como grupo de controle. Isso porque a seleção de quem recebeu a intervenção não foi feita de forma aleatória, podendo portanto ser afetada por fatores não observáveis pelos avaliadores no momento de calcular o impacto da intervenção.



## CUIDADOS DURANTE A MEDIÇÃO

**A**lgumas precauções são extremamente importantes para a realização satisfatória de uma medição de impacto. Abaixo serão descritos de forma pontual alguns aspectos críticos que o executor do projeto deve considerar na realização de um projeto de medição<sup>10</sup>.

### Externalidades

Um dos principais pontos a ser considerado é que deve existir o mínimo possível de efeitos cruzados (*spillovers*) entre os indivíduos de grupos diferentes. Por exemplo, considere que se pretenda avaliar o impacto de um projeto empresarial para o aumento de renda em uma determinada localidade X. Para

tanto, previamente ao projeto, se realiza um levantamento de dados nessa comunidade e na comunidade vizinha Y, que não foi considerada para o projeto. O efeito cruzado pode ocorrer se, por exemplo, a comunidade tratada X, ao receber renda advinda do investimento, puder dispendir parte dos seus recursos consumindo ou transferindo renda à comunidade Y. Se isto ocorrer, a adicionalidade de impacto do programa seria subestimada.

Esses efeitos cruzados são problemáticos para avaliações de impacto, pois eles podem distorcer o resultado pela contaminação do grupo de controle com a intervenção feita no grupo tratado. Para contornar esse problema, é importante garantir que não haja comunicação, troca de informação ou troca físi-

<sup>10</sup> Essa seção é fortemente baseada em DUFLO et al., op. cit.

ca de bens e serviços entre os indivíduos. Uma das formas de conseguir isso é buscando grupos que, por exemplo, estejam geograficamente distantes ou isolados, com menor risco de efeitos cruzados.

### Atrito

Como a medição de impacto ocorre em dois momentos distintos no tempo (um antes do projeto começar e outro depois de certo tempo de vigência), é de se esperar que ocorra um "descasamento" entre os indivíduos avaliados no momento inicial e os avaliados no momento final, tanto no grupo de controle quanto no grupo tratado.

Por exemplo: pode ocorrer que alguns indivíduos observados no momento inicial não estejam mais na localidade para avaliação final. A esta falha na coleta de resultados para alguns indivíduos que tenham se evadido, ou não se façam mais presentes no momento final, dá-se o nome de *atrito*. Vale destacar que, apesar da importância da aleatorização para a presente metodologia, ela não garante a inexistência de atrito não aleatório ao fim. Ademais, o problema do atrito reforça a importância de se definir, anteriormente à execução do projeto, uma amostra robusta e suficientemente grande para a realização das medições.

Como exemplo prático deste problema, considere um projeto de impacto para o aumento de renda em uma determinada localidade. Ainda que inicialmente os grupos tenham sido escolhidos de forma aleatória, ao fim do projeto pode-se deparar com atrito não aleatório decorrente da evasão de indivíduos que não tenham recebido o aumento de renda, frente a indivíduos que tenham recebido e continuem na localidade. Nesse caso, o atrito gera um viés na medição de impacto, pois haverá uma perda não aleatória de dados dos indivíduos do grupo de controle (pela evasão de indivíduos de um determinado perfil) frente ao grupo tratado.

É sempre importante expor as taxas de atrito presentes tanto no grupo tratado quanto no de controle de forma transparente. Ou seja, ao final da medição é preciso evidenciar quantos indivíduos medidos no primeiro momento não foram encontrados para a medição no segundo momento, tanto no grupo de controle quanto no grupo tratado.

### Efeito Hawthorne e Efeito John Henry

Por fim, outra das principais limitações existentes em uma medição de impacto é a mudança de comportamento dos indivíduos tanto do grupo tratado ("efeito Hawthorne"), quanto do grupo de controle ("efeito John Henry")<sup>11</sup>.

O efeito Hawthorne pode ocorrer pela percepção dos indivíduos tratados de que estão sob algum tipo de intervenção e, assim, há uma mudança de comportamento em resposta. Em projetos de educação, por exemplo, pode haver um esforço superior dos professores e dos alunos pela conscientização destes indivíduos de que seu grupo está sendo beneficiado pela intervenção. Já o efeito comportamental no grupo de controle, "efeito John Henry," pode ocorrer quando professores do grupo controle se sentem desafiados e começam a competir com os do grupo tratado para mostrar que também teriam condições de serem beneficiados pelos programas; ou, pelo contrário, por não serem tratados, os professores se desmotivam, decorrendo em uma superestimação do impacto do projeto. Nos dois casos, o efeito do programa pode ser distorcido.

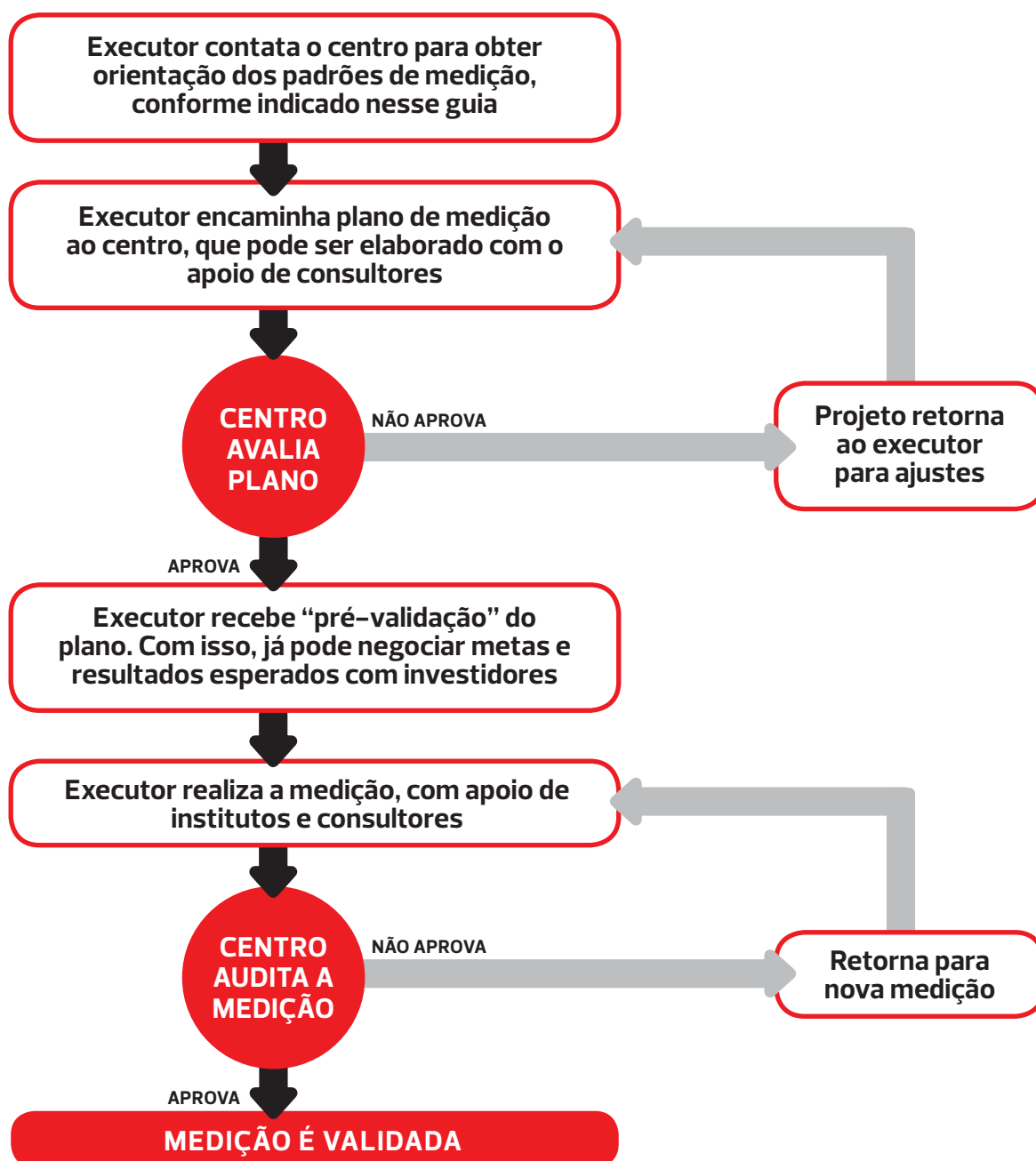
Ainda que respostas comportamentais sejam sempre complexas, há formas de desenho de medição que possibilitam minimizar os efeitos Hawthorne e John Henry. Por exemplo, pode-se utilizar três grupos: tratado, controle e *placebo*, neste último se estabelece um grupo que será observado, ou que manterá relações com o executor do projeto social, sem contudo ser efetivamente beneficiado pelo projeto. Considere, por exemplo, um projeto no qual uma grande empresa deseja melhorar a renda de determinadas comunidades por meio da compra de insumos locais. Assim, a empresa não deve apenas considerar uma comunidade de atuais fornecedores (tratada) e uma comunidade comparável de não fornecedores (controle). A empresa pode também monitorar uma comunidade placebo de não fornecedores com as quais tem alguma forma de relação (i.e. a empresa pode enviar gestores para conversar com estas comunidades, mas sem estabelecer contratos de suprimento). O impacto medido no grupo tratado deve ser superior ao verificado nos grupos de controle e placebo.

<sup>11</sup> O efeito Hawthorne leva esse nome em função da experiência homônima ocorrida em 1927 na fábrica West Electric Company (Chicago, EUA), onde se verificou que os indivíduos mudam seu comportamento de acordo com alterações percebidas no ambiente de trabalho. Já o efeito John Henry é assim chamado pois foi descoberto na John Henry Company, onde se introduziu, ainda em fase de testes, uma nova ferramenta que visava aumentar a produtividade. Na ocasião, os trabalhadores que foram mantidos usando a ferramenta antiga se sentiram desafiados a trabalharem mais, de forma a se mostrarem tão bons quanto os demais.

# FLUXOGRAMA

**N**a sequência segue detalhado todo o fluxo para a realização do projeto de avaliação de impacto. Nesse processo, um centro de avaliação de impacto social (como o Insper Metricis) pode atuar como suporte na validação tanto do plano quanto da medição final. Ao pactuarem uma forma de medir impacto, gestores e avaliadores podem se valer da ficha padronizada descrita na seção VI.

FLUXOGRAMA DO PROCESSO DE AVALIAÇÃO



# FICHA PADRONIZADA PARA AVALIAÇÃO DE IMPACTO

<b>SETOR DO PROJETO</b>			
<b>NOME DO PROJETO</b>			
<b>BREVE DESCRIÇÃO DO PROJETO COM SUA TEORIA DA MUDANÇA</b> (descrição das atividades, resultados, mudanças e impacto pretendidos)			
<b>MÉTRICA(S) ESCOLHIDA(S)</b>			
<b>COMO IMPACTO SERÁ MEDIDO</b>			
<b>NÍVEL DE MEDIÇÃO UTILIZADO</b>	<b>NÍVEL 1</b>	<b>NÍVEL 2</b>	<b>NÍVEL 3</b>
<b>JUSTIFICATIVA DA ESCOLHA DO NÍVEL DE MEDIÇÃO</b>			

## FICHA PADRONIZADA PARA AVALIAÇÃO DE IMPACTO

<b>PERÍODO DA MEDIÇÃO INICIAL (T=0)</b>		
<b>PERÍODO DA MEDIÇÃO FINAL (t=1)</b>		
<b>DESCRIÇÃO PLANO AMOSTRAL (tamanho amostral e indicação de quem são os grupos controle e tratado)</b>	<b>GRUPO TRATADO</b>	
	<b>GRUPO CONTROLE (e placebo caso aplicável)</b>	
<b>EMPRESA EXECUTORA DO PROJETO</b>		
<b>INVESTIDOR DE IMPACTO</b>		
<b>ENDEREÇO DA EMPRESA EXECUTORA DO PROJETO</b>		
<b>RESPONSÁVEL PELO PROJETO</b>		
<b>TELEFONE E E-MAIL</b>		
<b>ÁREA ABAIXO PARA PREENCHIMENTO DO CENTRO DE AVALIAÇÃO DE IMPACTO</b>		
<b>RESPONSÁVEL PELA ANÁLISE</b>	<b>DATA</b>	<b>CARIMBO CENTRO</b>