

Uma Cartilha para o Desenvolvimento de Medidas de Efetividade

Major Shon McCormick, Exército dos EUA

PARA COMPREENDER O nível operacional da guerra, os estudantes precisam entender a doutrina conjunta mais recente. Na Escola de Comando e Estado-Maior dos EUA, em Forte Leavenworth, Kansas, um importante conceito doutrinário é o chamado *design operacional* [*operational design*, no original em inglês — N. do T.]. Como instrutor na escola, observei que as “medidas de efetividade” constituem um aspecto do planejamento operacional difícil de ser entendido pelos alunos. Visto que eu próprio não tinha pleno conhecimento do conceito, conduzi uma pesquisa sobre o tema, vasculhando a doutrina combinada em vigor e perguntando a outros, na escola. Só tive sucesso ao consultar os integrantes do Centro de Análise do Comando de Instrução e Doutrina do Exército (*Training and Doctrine Command — TRADOC*) do Exército dos EUA, que me forneceram um manual da OTAN sobre as melhores práticas para avaliar os sistemas de comando e controle¹. Esse manual é informativo sobre as medidas de efetividade, embora não no contexto original do *design*, mas é demasiadamente técnico e especializado para a maioria dos oficiais de Estado-Maior.

A falta de conhecimento em relação às medidas de efetividade vai além do corpo discente da Escola. Essa conclusão não visa a depreciar quaisquer pessoas ou instituições, mas a salientar a falta de entendimento generalizada em relação ao conceito de medidas de efetividade. Este artigo reflete meus esforços para descrever um método para desenvolver medidas de efetividade que seja prático, mas preciso, e que qualquer pessoa não especialista possa empregar. Este artigo irá:

- examinar a utilidade dos métodos para avaliar resultados em geral;
- analisar como a atual doutrina combinada retrata a avaliação da efetividade como parte dos elementos de planejamento operacional;
- explorar como outros ramos de atividade externos às Forças Armadas lidam com conceitos como medidas de efetividade;
- explorar esses ramos para encontrar perspectivas de como preencher algumas lacunas na doutrina militar atual;
- prover observações sobre as implicações dos meus achados na doutrina de *design* do Exército e nos conceitos a ela relacionados.

Medidas

Por que devemos nos importar com as medidas de efetividade? Porque é o que diz a doutrina conjunta vigente. Esse é um argumento genérico, contudo, e a pergunta merece uma resposta melhor. Mesmo que não tivessem outra razão qualquer, os líderes militares pragmáticos deveriam se importar com as medidas de efetividade porque os representantes do povo norte-americano no Congresso se preocupam com elas. As rotineiras convocações para depor no Congresso sobre o progresso das operações no Iraque e no Afeganistão são exemplos disso. Um artigo de Patrick Cronin afirma que membros dos dois partidos, no Congresso, indicaram que só continuarão a apoiar os esforços no Iraque se houver “evidências concretas de um progresso militar tangível”².

Além disso, uma série de recentes estudos do Comando das Forças Combinadas dos EUA reforça a utilidade de ferramentas de

O Major Shon McCormick, do Exército dos EUA, está atualmente servindo no Afeganistão. Antes disso, serviu como instrutor do Departamento de Operações Combinadas, Interações

e Multinacionais na Escola de Comando e Estado-Maior do Exército, no Forte Leavenworth, Kansas. É bacharel pela Academia Militar dos EUA e mestre pela Kansas State University.

avaliação como as medidas de efetividade. *Joint Operations: Insights and Best Practices* (“Operações Combinadas: Ideias e Melhores Práticas”, em tradução livre) cita o emprego das medidas de avaliação como “uma importante prática, cuja necessidade é reforçada cada vez mais nos quartéis-generais operacionais”³. O estudo afirma que as medidas de efetividade são especialmente importantes no complexo ambiente operacional de hoje, o que representa um desafio à habilidade dos planejadores em prever, com precisão, o resultado de seus planos⁴.

Doutrina Vigente

Um ponto lógico para se começar é pesquisar a doutrina vigente na busca de alguma instrução de como desenvolver medidas de efetividade. As principais referências doutrinárias às medidas de efetividade são as Publicações Combinadas 5-0, *Planejamento Operacional Combinado* (JP 5-0, *Joint Operation Planning*), e seu manual complementar, *Operações Combinadas* (JP 3-0, *Joint Operations*)⁵. Esses

dois manuais combinam medidas de efetividade com medidas de desempenho sob o título geral de *Medidas de Avaliação* e orientam os estados-maiores a desenvolverem-nas durante a análise da missão. Fora isso, a doutrina conjunta não proporciona perspectiva alguma sobre os processos atuais para o desenvolvimento de medidas de efetividade aceitáveis.

Na falta de orientação doutrinária, uma pesquisa além das publicações militares se torna necessária. A esse respeito, irei explorar três campos:

- os princípios básicos da metodologia de pesquisa da ciência social. Há muito tempo esse campo tem lidado com os mesmos temas que o elemento “medidas de efetividade” do *design* busca abordar.
- a avaliação de política e de programa, que abrange as mesmas ideias como medidas de efetividade no campo de política pública interna.
- a emergente doutrina do Exército sobre *design* e como seus conceitos lidam com o desafio da avaliação.



(Foto do Exército dos EUA, Sgt André Reynolds)

Um shura (conselho tribal) com anciãos na aldeia de Meinashin, no Afeganistão, 05 Set 05.

Ciência Social

Primeiro, faz-se necessário distinguir a ciência social das ciências físicas. Em poucas palavras, a ciência social envolve o comportamento humano. Uma característica do estudo da ciência social é a

...líderes militares pragmáticos deveriam se importar com as medidas de efetividade porque os representantes do povo norte-americano no Congresso se preocupam com elas.

inviabilidade de realizar pesquisas em ambientes experimentais controlados. Não é possível conduzir pesquisas sociais em um ambiente onde se possam controlar todas as influências. Considerações práticas e éticas nos impedem de realizar experimentação em grupos de seres humanos da mesma forma como fazemos com cobaias. Em consequência, aceitamos como inevitável certa margem de erro na ciência social, cuja ocorrência é tanto aleatória quanto sistemática⁶.

Embora a doutrina vigente muitas vezes considere causa e correlação como conceitos equivalentes, a ciência social trata os dois de maneira bastante distinta. Correlação significa que dois eventos tendem a ocorrer juntos com alguma frequência. Um exemplo clássico disso é um galo que canta ao amanhecer. Pode-se afirmar que quase todas as vezes que o sol nasce, um galo canta. Os dois eventos demonstram um alto grau de correlação. Contudo, essa correlação não tem correspondência com causa. Atribuir causa, nesse caso, é uma *falácia post hoc*. Com base apenas na nossa observação do sol e do galo, não podemos determinar se o canto do galo leva ao nascimento do sol, se o nascimento do sol leva o galo a cantar ou até mesmo se há qualquer relação causal entre os dois eventos. Como determinar o grau de causa e efeito entre duas atividades correlacionadas é o cerne da ciência. Após muito tempo debatendo sobre causa e efeito, as ciências físicas e sociais

chegaram a um conjunto de conhecimentos geralmente aceito. Embora ele seja vasto, alguns de seus pontos-chave são relevantes aqui:

- correlação não equivale a causa;
- uma causa somente pode ser determinada com o emprego de uma hipótese; e
- nunca será possível determinar uma causa de forma absoluta; só será possível reduzir a incerteza.

Hipótese

Uma hipótese nada mais é do que uma proposta de relação causal entre duas atividades que pode ser testada. Por exemplo, o conceito da “escalada de tropas” (*surge*) no Iraque foi essencialmente uma hipótese de duas fases que pôs à prova se um aumento de soldados da coalizão em Bagdá reduziria a violência insurgente, e se essa redução em violência levaria à reconciliação entre as facções políticas sunita e xiita⁷. Nunca poderemos ter certeza se o aumento de efetivo militar realmente levou a uma redução da violência. Em vez disso, podemos somente reduzir nossa incerteza com a aplicação de algumas técnicas para determinar se outros fatores causais estavam presentes. Os procedimentos necessários para isso estão além do escopo deste artigo.

Sendo necessária uma hipótese para testar a relação de causa e efeito, o desafio seguinte é desenvolvê-la. Nossa doutrina é ambígua nesse sentido, mas a ciência oferece três alternativas: empregar uma hipótese desenvolvida por outra pessoa para circunstâncias semelhantes, desenvolver a própria hipótese ou empregar uma combinação dos dois métodos anteriores. A forma mais simples de encontrar hipóteses existentes é

Nunca poderemos ter certeza se o aumento de efetivo militar realmente levou a uma redução da violência.

consultar o conjunto de conhecimentos atual sobre o assunto de interesse. Se o objetivo é determinar a conexão entre a presença de Forças de segurança e a violência insurgente, um bom começo seria

estudar as pesquisas feitas por outros sobre esse tema. No entanto, não há duas situações que sejam exatamente iguais. Mesmo as circunstâncias mais semelhantes podem ter omitido fatores importantes. Devemos armar um aldeão afegão com as mesmas armas que demos ao Conselho do Despertar no Iraque? Será que o que funcionou no Iraque funcionará no Afeganistão, dadas as histórias e graus de desenvolvimento muito diferentes entre as duas nações?

Se não existirem hipóteses aceitáveis, deve-se criar uma a partir de uma pesquisa original. Em termos simples, a formulação de uma hipótese exige que alguém especule sobre uma relação causal entre duas atividades ou duas variáveis. A fonte dessa relação causal pode

ser simplesmente uma intuição ou qualquer outra forma de perspectiva. Formulada a hipótese inicial, o pesquisador deve pô-la à prova contra casos históricos apropriados. Isso é difícil, porque na história não há dois casos que sejam completamente iguais ou que tenham os mesmos fatores causais. A meta do pesquisador é determinar, nos vários casos, quais fatores são gerais e quais são específicos. Independentemente da origem da hipótese, o passo seguinte é empregá-la para prever eventos futuros. Em termos simples: se uma ideia parece ter funcionado no passado, ela deverá funcionar no futuro, caso ocorram circunstâncias relativamente semelhantes. À luz da já estabelecida advertência de que o passado não é um indicador perfeito para o futuro, nossa hipótese proporciona, no melhor dos casos, uma “conjectura fundamentada” sobre um resultado que é desconhecido. Já que devemos aceitar que nossa hipótese possui um grau de erro, nossa tarefa será determinar quando ela tiver falhado, ou quando se tornar falsa. Infelizmente, pode ser que já tenhamos nosso plano em execução antes que possamos chegar a uma conclusão sobre nossa hipótese. É nesse momento em que se tornam importantes as medidas de efetividade.

Para melhor explorar o papel das medidas de efetividade na prova de hipóteses, nos afastamos da ciência social e entramos no campo de avaliação de programas. Uma busca do termo “avaliação de programas” na internet mostra uma ampla disciplina com um grande conjunto de pesquisas. Agências civis têm lidado formalmente com métodos para avaliar a efetividade de vários programas desde meados dos anos 40. Programas recentes como o “*No Child Left Behind*” (“Nenhuma criança deixada para trás”, em tradução livre) ou até mesmo o conjunto de estímulos econômicos do Presidente Obama nada mais são do que esforços para influenciar sistemas de modo positivo.

Programas

Antes de explorar o campo da avaliação de programas, são necessárias algumas definições:

- um *programa* é um “conjunto de recursos direcionados a uma ou mais metas comuns”⁸, ou uma hipótese que, “se for usada, ocorrerão os resultados esperados”⁹.



Foto do Departamento de Defesa, Cb Artur Shvartsberg, CFN dos EUA

Um capitão do Corpo de Fuzileiros Navais dos EUA conversa com anciãos durante um shura no Distrito de Nawa na Província de Helmand, no Afeganistão, 10 Ago 09.



Foto do CFN dos EUA, Sgt. Ezekiel R. Klandwe

O General Comandante do 215º Corpo do Exército Nacional Afegão e seu intérprete se dirigem aos líderes locais e a fuzileiros navais dos EUA durante uma reunião de segurança regional no Camp Dwyer, no Afeganistão, 13 Mai 10.

- um *insumo* do programa é simplesmente algo que se insere nele.

- uma *saída* do programa consiste nos “produtos, bens ou serviços” que dele advêm e que são, então, proporcionados aos beneficiários pretendidos.

- por último, o *resultado* é uma “mudança ou benefício resultante das saídas do programa”¹⁰.

A definição dos elementos de avaliação de programas é semelhante aos termos doutrinários militares das medidas de desempenho e das medidas de efetividade. Uma medida de saída é semelhante a uma medida de desempenho e uma medida de resultado é análoga a uma medida de efetividade. À luz dessa semelhança, as medidas de resultados da teoria de programas devem ser úteis para ajudar a explicar as medidas de efetividade da doutrina militar.

Os “modelos lógicos”, ou “modelagem” são conceitos centrais ao campo de avaliação de programas. Esclarecem a relação entre os insumos, as saídas e os resultados de um programa. Implícita no modelo lógico está sua

teoria de programa, a hipótese causal que vincula os elementos do modelo. As teorias de programa predizem resultados no desenvolvimento do programa e determinam relações causais entre insumos e resultados depois da implantação dele¹¹.

A comparação de resultados exige sua medição. A operacionalização é o processo de criar medidas de insumos, saídas e resultados. Alguns resultados se prestam à medição mais facilmente que outros. São resultados facilmente quantificáveis, por exemplo, os custos financeiros e as baixas em um conflito militar.

Medindo Variáveis Problemáticas

No entanto, nem todos os resultados podem ser medidos tão facilmente. Exemplos de variáveis mais problemáticas relacionadas às Forças Armadas são resultados como segurança e democracia. No caso desses conceitos mais abstratos, o pesquisador deve empregar indicadores ou algo que os represente. Embora pareça simples, a escolha de indicadores é complexa. Por exemplo, como se pode medir a

democracia? Frequentemente, a dificuldade em se desenvolver medidas válidas de resultados abstratos impõe uma análise da literatura de pesquisa existente e a consulta a peritos e a especialistas no campo de interesse¹².

Retornar à metodologia da ciência social é útil, neste ponto. Gary King, Robert O. Keohane e Sidney Verba aconselham que os pesquisadores determinem tantas “consequências observáveis” de suas hipóteses quanto possível, de modo a criar mais casos para pô-las à prova. O entendimento é de que um maior número de testes sobre um maior número de consequências previsíveis, irá, provavelmente, evidenciar possíveis falhas na seleção dos indicadores¹³.

Insumos, saídas e resultados. Uma vez iniciado o programa, a comparação entre insumos, saídas e resultados passa a alertar seu gestor sobre a validade da lógica adotada por ele. Se a lógica do programa for falha, o gestor deve reexaminar e possivelmente aperfeiçoar o modelo. Nesse ponto, um estado-maior militar, buscando empregar a teoria de lógica de programas, precisaria de alguém treinado em análise estatística para determinar qual dos elementos é o defeituoso. Dois métodos comuns são: os *experimentos aleatórios* e a *quase-experimentação*¹⁴. A diferença básica entre os dois métodos envolve o grau de controle que o analista mantém sobre o ambiente. As mecânicas atuais para conduzir experimentos aleatórios ou a quase-experimentação estão além do objetivo deste artigo.

Faz-se necessário um exemplo militar do desenvolvimento de medidas de efetividade. Dadas as operações em andamento no Iraque e no Afeganistão e o atual interesse das Forças militares em operações de estabilidade, este exemplo enfoca o desafio de estabelecer segurança em um ambiente pós-conflito.

No cenário deste exemplo, consideremos um estado-maior responsável pela execução de uma operação de estabilidade em um ambiente pós-conflito. O ambiente operacional em questão está inaceitavelmente violento, ameaçando a capacidade da débil nação anfitriã de estabelecer sua autoridade. O problema que o estado-maior deve enfrentar é como reforçar a capacidade do governo da nação anfitriã para que ele possa controlar efetivamente seu território.

(Observe-se que o problema não é simplesmente reduzir a violência. Definir o problema dessa forma é supor que há uma relação causal entre a redução da violência e o aumento da capacidade de governança da nação anfitriã). Para este exemplo, suponha que o estado-maior determinou que a redução de violência é o objetivo, ou seja, que é o resultado a ser buscado. O próximo desafio do estado-maior é determinar quais são os recursos disponíveis e como empregá-los para alcançar o resultado desejado. Esse passo exige a adoção de uma teoria de programas que proponha uma relação causal entre os insumos, as saídas e os resultados. O estado-maior escolheu não conduzir sua própria pesquisa independente devido a limitações de tempo e, portanto, teve de confiar nas pesquisas existentes. Uma fonte lógica é a doutrina. No entanto, como Christopher Paparone observou, um dos problemas da doutrina é que ela nunca cita suas fontes. Por exemplo, o Manual de Campanha 3-07, *Operações de Estabilidade* (FM 3-07, *Stability Operations*), fornece boa orientação geral sobre como conduzir a tarefa de estabilização de “Garantir a Segurança Civil”, mas os leitores não sabem que casos históricos específicos influenciaram essa generalização¹⁵. Já que a doutrina é insuficiente, o estado-maior deve ampliar sua pesquisa.

Uma fonte potencial de hipóteses causais é *Winning the Peace: An American Strategy for Post-Conflict Reconstruction* (“Ganhar a Paz: uma Estratégia Americana para Reconstrução Pós-Conflito”, em tradução livre). Esse livro serviu como base para a “Matriz de Tarefas Essenciais para Reconstrução Pós-Conflito” (“*Post Conflict Reconstruction Essential Tasks Matrix*”, no original, em inglês) do Departamento de Estado, a qual influenciou o Manual de Campanha 3-07. Segundo o capítulo de Scott Feil sobre o aperfeiçoamento das capacidades de segurança, estabelecer segurança é um pré-requisito para qualquer atividade de desenvolvimento ou de reconstrução¹⁶. Os esforços de segurança bem-sucedidos consistem em uma combinação de atividades de proteção defensiva e ofensiva que “retirem a capacidade de grupos e indivíduos se engajarem em atos de violência ilegítima”. Com respeito às medidas defensivas, a população é um elemento que exige proteção¹⁷. À luz da

pesquisa de Feil, consideramos a hipótese de que a proteção da população geral leva a uma redução da violência ilegítima. Contudo, essa hipótese não nos indica como proteger a população em geral e, portanto, devemos continuar nossa pesquisa. Em *The Quest for Viable Peace* (“Em Busca da Paz Viável”, em tradução livre), Ben Lovelock afirma que o aumento de patrulhas a pé em áreas povoadas foi uma técnica bem-sucedida para proteger o povo em Kosovo, nos anos 90¹⁸.

A combinação das hipóteses de Feil e de Lovelock produz a seguinte narrativa lógica: Se uma organização aumentar patrulhas a pé (insumos), a população geral estará mais segura (saída). Se a população estiver mais segura, a violência ilegítima deve diminuir (resultado).

Elementos do Modelo de Programa Tendo determinado a teoria lógica de programa e criado um modelo de lógica, o próximo passo é determinar as medidas de efetividade para os vários elementos do modelo de programa. Mensurar patrulhas a pé é relativamente fácil. Nesse caso, seriam apropriadas as medidas de desempenho derivadas da doutrina do Exército. As medidas devem incluir o número e a duração das patrulhas, bem como a área coberta por elas. A mensuração da saída — a segurança da população — é algo mais desafiante, porque “segurança” é mais abstrata. Como resultado, dependemos de indicadores ou algo que os representem. Se nem Feil nem Lovelock — as fontes para nossa teoria de programas — fornecem indicadores, então faz-se necessária pesquisa adicional. Uma obra que aborda indicadores de segurança é o recente *Guidebook for Supporting Economic Development in Stability Operations* (“Guia para Apoiar o Desenvolvimento Econômico em Operações de Estabilidade”, em tradução livre) da RAND [organização sem fins lucrativos que estuda e propõe soluções para temas da agenda internacional — N. do T.], a qual cita alguns indicadores para a segurança da população, como o número de pessoas abandonando suas casas¹⁹, por exemplo.

A medição do resultado, representado pela redução da violência ilegítima, pode incluir o registro de crimes e de mortes violentas. Nesse caso, o analista dependeria mais da intuição do que da teoria existente para escolher o indicador. No entanto, o guia da RAND oferece

algumas orientações sobre o emprego de dados de criminalidade como indicadores. Como, por exemplo, a advertência de que a fonte dessa informação é, provavelmente, o banco de dados policial, que somente reflete a violência reportada. Além disso, o guia adverte que, em geral, os projetos de reconstrução bem-sucedidos representam alvos compensadores para os insurgentes e podem, na verdade, levar a um aumento de violência. O analista que desejasse medir a violência com precisão teria de aceitar a imprecisão decorrente dos atos de violência não relatados e do aumento de violência perto de projetos de reconstrução.

Tendo o analista criado e empregado um modelo aceitável de lógica de programa, o estado-maior precisa determinar a efetividade do programa de segurança. O primeiro passo é conduzir as medições. O *United States Institute for Peace* (Instituto da Paz dos EUA) oferece quatro métodos principais para coletar dados de medição. Esses métodos são:

- análise de conteúdo dos produtos da mídia local,
- consulta a um painel de especialistas,
- análise estatística e
- uso de pesquisas de opinião pública e enquetes²⁰.

O próximo passo é estabelecer a relação entre as medições. Segundo a teoria do programa, um aumento em patrulhas a pé deve, no final, levar a uma diminuição da violência reportada, a qual indica um aumento na estabilidade geral. Se essa cadeia de hipóteses não for verdadeira, os analistas devem reexaminar o modelo de lógica. Nós nos baseamos em modelos de eventos passados, e não há garantia alguma de que nossa lógica de programa seja completamente válida no ambiente atual. Talvez o aumento de patrulhas a pé tenha servido somente para aumentar ainda mais o antagonismo da população e para ampliar a percepção de que as Forças da coalizão são uma Força de ocupação. Se esse for o caso, a situação exigirá uma teoria de programa mais adequada. Pode ser que a teoria de programa e o modelo de lógica sejam válidos, mas os indicadores de segurança não. Nesse caso, os analistas devem desenvolver melhores indicadores.

Para um exemplo prático da aplicação dos princípios abordados nesta seção, consulte

o documento de trabalho *Guidance on Evaluation Conflict and Peacebuilding Activities* (“Orientação sobre Avaliação de Conflitos e Atividades de Construção da Paz”, em tradução livre) da Organização para a Cooperação e Desenvolvimento Econômico. O documento é um guia de fácil leitura para usuários governamentais e não governamentais que empreguem a teoria de lógica de programas durante atividades de estabilidade e reconstrução. No Anexo 6, “Entendendo e Avaliando Teorias de Mudança”, há um sumário tabulado com as principais teorias adequadas para usar como lógica de programa²¹.

O Futuro

A complexidade do ambiente operacional atual levou a um número de iniciativas para melhorar o planejamento militar por meio do conceito de *design*. O Exército incorporou os princípios de “*design* operacional sistêmico” com a publicação do Manual de Campanha 5-0, *O Processo de Operações* (FM 5-0, *The Operations Process*)²².

A linguagem do *design* expressa no FM 5-0 é muito parecida à da teoria de lógica de programas, existente desde os anos quarenta. Ambas as construções aceitam que soluções iniciais talvez não sejam válidas. As duas enfocam hipóteses explícitas que vinculam

insumos, saídas e resultados. Os métodos para criar medidas de efetividade, sob o enquadramento do novo conceito de *design*, são semelhantes àqueles encontrados neste artigo e na comunidade da Área Funcional 49 (*Functional Area 49*) — Analistas de Sistemas de Pesquisa de Operações. Os analistas de sistemas do Exército vêm empregando medidas em ambientes complexos há muito tempo e podem fornecer insumos úteis aos processos de planejamento que estão surgindo. À medida que o Exército abandonar os métodos de planejamento mecânicos e determinísticos — associados à obsoleta “abordagem baseada em efeitos” — e incorporar princípios de *design* em sua doutrina, ele não deve negligenciar a existência desses blocos de conhecimento.

A doutrina que está sendo desenvolvida indica que as medidas de efetividade e os conceitos associados ao *design* irão perdurar. Enquanto o Exército adotar os elementos do *design*, um entendimento básico do que são e de como criar medidas de efetividade permanecerá sendo uma habilidade fundamental a comandantes e seus estados-maiores. O conceito de medidas de efetividade não nos deve intimidar. Quase todas as facetas de sua construção estão dentro das capacidades normais de um oficial superior. **MR**

REFERÊNCIAS

1. NATO. *Code of Best Practice for Command and Control Assessment* (Washington, DC: DOD Command and Control Research Program, 2002).
2. CRONIN, Patrick M. “Irregular Warfare: New Challenges for Civil-Military Relations,” *Strategic Forum* 234 (out. 2008): p. 4, disponível em: <www.ndu.edu/inss>.
3. JOINT WARFIGHTING CENTER, UNITED STATES JOINT FORCES COMMAND, *Joint Operations: Insights and Best Practice* 2d ed. (jul. 2008), Gary Luck, Mike Findlay e a JWFC Joint Training Division, p. 33.
4. *Ibid.*, p. 35.
5. Joint Publication (JP) 3-0, *Joint Operations*, incorporando Chg. 1 (Washington, DC: U.S. Government Printing Office [GPO], February 2008) e JP 5-0 *Joint Operation Planning* (Washington, DC: GPO, dez. 2006).
6. KING, Gary; KEOHANE, Robert O; VERBA, Sidney. *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton: Princeton University Press, 1994), p. 8-9.
7. KAGAN, Frederick W. *Choosing Victory: A Plan for Success in Iraq*, Phase I Report, American Enterprise Institute Iraq Study Group (jan. 2007), p. 1.
8. NEWCOMER, Kathryn E.; HATRY, Harry P.; WHOLEY, Joseph S. “Meeting the Needs for Practical Evaluation: An Introduction”, *Handbook of Practical Program Evaluation*, 2d ed. (San Francisco: Jossey-Bass, 2004), xxxiv.
9. MCLAUGHLIN, John A.; JORDAN, Gretchen B. “Using Logic Models,” *Handbook of Practical Program Evaluation*, 2d ed. (San Francisco: Jossey-Bass, 2004), p. 7.
10. *Ibid.*, p. 9.
11. *Ibid.*, p. 10-11.
12. LOVE, Arnold. “Implementation Evaluation”, *Handbook of Practical Program Evaluation* 2d ed. (San Francisco: Jossey-Bass, 2004), p. 76-78.
13. KING, KEOHANE e VERBA, p. 28.
14. REICHHART, Charles S.; MACK, Melvin M. “Quasi-Experimentation,” e ST. PIERRE, Robert G. “Using Randomized Experiments,” from *Handbook of Practical Program Evaluation*, 2d ed. (San Francisco: Jossey-Bass, 2004).
15. Esse comentário aplica à análise proporcionada por Stephen Biddle no seu estudo do FM 3-24 em “The New U.S. Army/Marine Corps Counterinsurgency Field Manual as Political Science and Political Praxis”, *Perspectives on Politics* 6, vol. 2 (jun. 2008): p. 347-50.
16. FEIL, Scott. “Laying the Foundation: Enhancing Security Capabilities”, *Winning the Peace: An American Strategy for Post-Conflict Resolution*, ed. Robert C. Orr (Washington, DC: The CSIS Press, 2004), p. 40.
17. *Ibid.*, p. 42.
18. LOVELOCK, Ben. “Securing a Viable Peace: Defeating Militant Extremists—Fourth-Generation Peace Implementation”, *The Quest for Viable Peace: International Intervention and Strategies for Conflict Transformation*, ed. Jock Covey, Michael Dziedzic, and Leonard R. Hawley (Washington DC: United States Institute of Peace, 2005), p. 134.
19. CRANE, Keith et al. *Guidebook for Supporting Economic Development in Stability Operations* (RAND: Santa Monica, 2009).
20. United States Institute of Peace, *Measuring Progress In Conflict Environments (MPICE): Metrics Framework For Assessing Conflict Transformation and Stabilization* (fev. 2008).
21. ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT, *Guidance on Evaluation Conflict and Peacebuilding Activities*. Documento de trabalho para o período de aplicação (2008), disponível em: <www.oecd.org/document/35/0,3343,en_21571361_34047972_31779555_1_1_1_1,00.html>, Acesso em 24 fev. 2009.
22. FM 5-0, *The Operations Process* (Washington, DC: GPO, mar. 2010)